Preface

This is a book that is started out as a reading note of the book Royden (1988) and the MATH 441 & 442 notes by Professor Peter Leob of University of Illinois at Urbana-Champaign. I am inspired by Luenberger (1969) on the wide applicability of function analysis tools and its unification of classical results in decision theory, optimization theory, control theory, and numerical approximation. My background is in control theory, and I find that the tools in Luenberger (1969) requires deep knowledge in Lebesgue integral and therefore measure theory. So, I started with Royden (1988) to study measure theory and general integration with the goal of studying measure theoretic calculus. Another objective of mine is to set a solid foundation for my control theory research. Thus, I started out on this decade long study that leads to this book. During this study, I consulted a number of books that are the guiding light that shine upon my way. In Chapter 6, I have included some material from the book Maunnder (1996). The Chapters 6–10 include a significant amount of material from the book Luenberger (1969). Chapter 9 also referenced Bartle (1976). The book Bartle (1976) is solid in its mathematical rigor and serves as the starting point of this endeavor. The Global Inverse Function Theorem in Section 9.6 is inspired by Ambrosetti and Prodi (1993). Chapter 11 includes significant amount of self-developed material due to lack of reference on this subject. The proof of Radon-Nikodym Theorem 11.169 was adapted from the MATH 442 notes by Prof. Peck of University of Illinois at Urbana-Champaign. The book Royden (1988) offers clues on how to invent the wheel. The calculation of measure (see Section 11.7) is a main contribution of this book. As I understand from this study, the definition of a measure, calculating new measures based on already defined measures, and working with the cumulative distribution functions of these measures are essential characteristic in the application of measure theory. In Chapter 11, I finally resolve the dual of the space of continuous functions of a compact space to a Banach space. I further correctly defined the space $C_c(X, Y)$ and find its dual. I could further define $C_c(X', Y)$ for $X'$ that are open subset of a finite-dimensional Banach space. I chose to stop at here and leave this as an exercise for the reader. In Chapter 12, differentiation and integration finally comes together. The Tonelli’s Theorem is standard. The Fubini’s Theorem comes
in two flavors: one for general (separable) Banach space, and the other for finite-dimensional Banach space. The one for the finite-dimensional Banach space is as simple as it can get: for absolutely integrable functions of $\mathbb{R}^m$, the integral can be calculated as repeated integrals, as long as one set the intermediate integral to the null vector whenever the integral is not absolutely integrable. The concept of isomeasure is a great discovery. It allows me to equate integrations on different measurable spaces. This has direct application in the Change of Variable Theorem 12.91. There, all I need to do is to show that certain induced measure is absolutely continuous with respect to the standard Borel measure on $\mathbb{R}^m$. The connection between Riemann-Stieltjes integral and Lebesgue-Stieltjes integral is established in Proposition 12.100. I attempted to define Sobolev space in a correct fashion, which is different from that of Zeidler (1995) when the dimension of the domain of the function is greater than 1. I feel that my definition is the correct way to go, due to the characterization of absolutely continuous functions on multi-dimensional domains in the Fundamental Theorem of Calculus II. In Chapter 13, standard projection theorems are established as well as the dual of a Hilbert space. Examples of complete orthonormal sequences are obtained. The Hermitian adjoint of a linear operator is introduced, which then lead to the definition of positive definite and positive semi-definite operators (including complex versions). Standard results on pseudoinverse of a linear operator between Hilbert spaces is proved. Then, the spectral theory of compact Hermitian linear operator is established. The singular value decomposition is obtained for all compact linear operators between Hilbert spaces. In Chapter 14, standard results are obtained. I consulted my MATH 451 & 452 notes of Professor Burkholder of University of Illinois at Urbana-Champaign, Stark and Woods (1986), Williams (1991), and Billingsley (1995). ... (To be continued).

After I took my Math 441 & 442 on Real Analysis based on the book Royden (1988) in University of Illinois at Urbana-Champaign, there is something fuzzy about the theory. I am confident that the results are correct, but the arbitrary assignment of $0 \times \infty = 0$ hangs like a dark cloud in the horizon; and the lack of Change of Variable Theorem in the theory for general measure spaces keeps me wanting for more results in this field. I didn’t bother then to study measure theory myself, and pressed on with my PhD study in control theory without solidifying my background. That proves to be a major mistake. After I lost my teaching job for good, I finally settled down to study the basics that underpins my past research. My wife got a secure job and I became an at-home dad. I have all the time in the world. I really don’t want to waste my brain away, and pressed on with research to keep my mind sharp and my resumé competitive. God put the thought in me that I could work on foundational mathematics rather than waiving my hands at some advanced control problems. I could do both, but the urge to stand on my own feet finally win over my attention. I put down some fruitful control theory research and focused myself on the
foundational mathematics that underpins my earlier research. I expected that it might take a few years, but it actually started me on this decade long pursuit. In retrospect, I am happy with my decision of working in mathematics, which is a topic that fascinated me since my youth. The end result is very satisfying for me and I look back at this journey as something that I can be proud of.

I want to thank my Uncle Desheng Pan and Aunt Deyin Pan for my upbringing. Uncle Sheng gave me a rigorous mathematical mind when I was ten years old that he drilled me to study Euclidean geometry. I wasn’t happy then, but it gave me a sharp mind ever since. All those classes and lectures in later life worth much more to me than to any other person, since I constantly think about the why behind everything I learned. Aunt Yin and Uncle Sheng shouldered most of the life’s burdens when I was in China and gave me and my sister the freedom to study and excel. I thank Aunt Yin especially for all those chores that she did for us, all those clothes that she washed without a washing machine, and all those meals she cooked. One has to remember she did all these in addition to a six day per week job. I want to express my gratitude for University of Illinois at Urbana-Champaign that gave me a world class education. The friendships I formed during my graduate study will last a lifetime. Most importantly, I am grateful for Professor Tamer Ba¸sar for his friendship throughout my last 28 years, first as my advisor and later as my friend. I want to thank my wife for all the help she gave me. I want to thank Professor Wen-ben Jone for remaining my friend after I lost my last job.

Zigang Pan
Mason, Ohio
July, 2018
# Contents

Preface iii

1 Introduction 1

2 Set Theory 23
   2.1 Axiomatic Foundations of Set Theory . . . . . . . . . . . . 23
   2.2 Relations and Equivalence . . . . . . . . . . . . . . . . . . 24
   2.3 Function . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
   2.4 Set Operations . . . . . . . . . . . . . . . . . . . . . . . . 26
   2.5 Algebra of Sets . . . . . . . . . . . . . . . . . . . . . . . . 27
   2.6 Partial Ordering and Total Ordering . . . . . . . . . . . . . 28
   2.7 Basic Principles . . . . . . . . . . . . . . . . . . . . . . . . 30

3 Topological Spaces 39
   3.1 Fundamental Notions . . . . . . . . . . . . . . . . . . . . . . 39
   3.2 Continuity . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
   3.3 Basis and Countability . . . . . . . . . . . . . . . . . . . . . 45
   3.4 Products of Topological Spaces . . . . . . . . . . . . . . . . 46
   3.5 The Separation Axioms . . . . . . . . . . . . . . . . . . . . 51
   3.6 Category Theory . . . . . . . . . . . . . . . . . . . . . . . . 52
   3.7 Connectedness . . . . . . . . . . . . . . . . . . . . . . . . . 54
   3.8 Continuous Real-Valued Functions . . . . . . . . . . . . . . . 59
   3.9 Nets and Convergence . . . . . . . . . . . . . . . . . . . . . 63

4 Metric Spaces 77
   4.1 Fundamental Notions . . . . . . . . . . . . . . . . . . . . . . 77
   4.2 Convergence and Completeness . . . . . . . . . . . . . . . . 79
   4.3 Uniform Continuity and Uniformity . . . . . . . . . . . . . . 82
   4.4 Product Metric Spaces . . . . . . . . . . . . . . . . . . . . . 84
   4.5 Subspaces . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
   4.6 Baire Category . . . . . . . . . . . . . . . . . . . . . . . . . 90
   4.7 Completion of Metric Spaces . . . . . . . . . . . . . . . . . 91
   4.8 Metrization of Topological Spaces . . . . . . . . . . . . . . . 96
   4.9 Interchange Limits . . . . . . . . . . . . . . . . . . . . . . . 97
## 5 Compact and Locally Compact Spaces

- 5.1 Compact Spaces ........................................ 101
- 5.2 Countable and Sequential Compactness ............... 107
- 5.3 Real-Valued Functions and Compactness .............. 109
- 5.4 Compactness in Metric Spaces .......................... 111
- 5.5 The Ascoli-Arzelá Theorem ............................. 113
- 5.6 Product Spaces ......................................... 115
- 5.7 Locally Compact Spaces ................................ 118
  - 5.7.1 Fundamental notion ................................. 118
  - 5.7.2 Partition of unity ................................. 121
  - 5.7.3 The Alexandroff one-point compactification .... 124
  - 5.7.4 Proper functions ................................. 125
- 5.8 σ-Compact Spaces .................................... 127
- 5.9 Paracompact Spaces .................................. 128
- 5.10 The Stone-Čech Compactification .................... 132

## 6 Vector Spaces

- 6.1 Group .................................................. 135
- 6.2 Ring .................................................... 137
- 6.3 Field ................................................... 138
- 6.4 Vector Spaces .......................................... 138
- 6.5 Product Spaces ........................................ 140
- 6.6 Subspaces .............................................. 141
- 6.7 Convex Sets ........................................... 145
- 6.8 Linear Independence and Dimensions ................ 148

## 7 Banach Spaces

- 7.1 Normed Linear Spaces ................................ 151
- 7.2 The Natural Metric .................................... 157
- 7.3 Product Spaces ........................................ 160
- 7.4 Banach Spaces ......................................... 162
- 7.5 Compactness ............................................ 167
- 7.6 Quotient Spaces ....................................... 169
- 7.7 The Stone-Weierstrass Theorem ....................... 171
- 7.8 Linear Operators ...................................... 179
- 7.9 Dual Spaces ........................................... 183
  - 7.9.1 Basic concepts .................................... 183
  - 7.9.2 Duals of some common Banach spaces ............ 184
  - 7.9.3 Extension form of Hahn-Banach Theorem ........ 188
  - 7.9.4 Second dual space ............................... 196
  - 7.9.5 Alignment and orthogonal complements ........... 198
- 7.10 The Open Mapping Theorem ......................... 203
- 7.11 The Adjoints of Linear Operators .................. 207
- 7.12 Weak Topology ...................................... 210
8 Global Theory of Optimization 221
  8.1 Hyperplanes and Convex Sets ........................... 221
  8.2 Geometric Form of Hahn-Banach Theorem ................. 224
  8.3 Duality in Minimum Norm Problems ....................... 226
  8.4 Convex and Concave Functionals ......................... 229
  8.5 Conjugate Convex Functionals ........................... 233
  8.6 Fenchel Duality Theorem ............................... 240
  8.7 Positive Cones and Convex Mappings ..................... 247
  8.8 Lagrange Multipliers ................................. 249

9 Differentiation in Banach Spaces 257
  9.1 Fundamental Notion .................................. 257
  9.2 The Derivatives of Some Common Functions .............. 261
  9.3 Chain Rule and Mean Value Theorem ..................... 265
  9.4 Higher Order Derivatives ............................. 271
    9.4.1 Basic concept .................................. 271
    9.4.2 Interchange order of differentiation .............. 277
    9.4.3 High order derivatives of some common functions .. 282
    9.4.4 Properties of high order derivatives ............. 286
  9.5 Mapping Theorems ................................... 292
  9.6 Global Inverse Function Theorem ....................... 304
  9.7 Interchange Differentiation and Limit .................. 311
  9.8 Tensor Algebra .................................... 315
  9.9 Analytic Functions .................................. 317

10 Local Theory of Optimization 329
  10.1 Basic Notion ..................................... 329
  10.2 Unconstrained Optimization .......................... 335
  10.3 Optimization with Equality Constraints ............... 338
  10.4 Inequality Constraints ............................. 343

11 General Measure and Integration 353
  11.1 Measure Spaces .................................. 353
  11.2 Outer Measure and the Extension Theorem ............... 359
  11.3 Measurable Functions ................................ 372
  11.4 Integration ....................................... 387
  11.5 General Convergence Theorems ........................ 402
  11.6 Banach Space Valued Measures ......................... 418
  11.7 Calculation With Measures ........................... 450
  11.8 The Radon-Nikodym Theorem ........................... 481
  11.9 \( L_p \) Spaces .................................... 497
  11.10 Dual of \( C(X, Y) \) and \( C_c(X, Y) \) ............... 513
12 Differentiation and Integration 543
  12.1 Carathéodory Extension Theorem .......................... 543
  12.2 Change of Variable ........................................ 549
  12.3 Product Measure ............................................ 554
  12.4 Functions of Bounded Variation ............................ 581
  12.5 Absolute and Lipschitz Continuity ......................... 605
  12.6 Fundamental Theorem of Calculus ......................... 622
  12.7 Representation of \((C_0(\Omega,Y))^*\) ..................... 651
  12.8 Sobolev Spaces ............................................. 660

13 Hilbert Spaces 667
  13.1 Fundamental Notions ....................................... 667
  13.2 Projection Theorems ........................................ 671
  13.3 Dual of Hilbert Spaces .................................... 673
  13.4 Hermitian Adjoints ........................................ 676
  13.5 Approximation in Hilbert Spaces ......................... 678
  13.6 Other Minimum Norm Problems ............................. 688
  13.7 Positive Definite Operators on Hilbert Spaces .......... 692
  13.8 Pseudoinverse Operator ................................... 697
  13.9 Spectral Theory of Linear Operators ..................... 701

14 Probability Theory 721
  14.1 Fundamental Notions ....................................... 721
  14.2 Gaussian Random Variables and Vectors ................... 741
  14.3 Law of Large Numbers ..................................... 743
  14.4 Martingales Indexed by \(\mathbb{Z}_+\) ....................... 744
  14.5 Banach Space Valued Martingales Indexed by \(\mathbb{Z}_+\) ... 752
  14.6 Characteristic Functions .................................. 757
  14.7 Convergence in Distribution .............................. 763
  14.8 Central Limit Theorem .................................... 768
  14.9 Uniform Integrability and Martingales ................. 772
  14.10 Existence of the Wiener Process ......................... 776
  14.11 Martingales With General Index Set .................... 791
  14.12 Stochastic Integral ....................................... 800

15 Numerical Methods 803
  15.1 Newton’s Method ........................................... 803

A Elements in Calculus 807
  A.1 Some Formulas ........................................... 807
  A.2 Convergence of Infinite Sequences ....................... 807

List of Figures 809

Bibliography 811
CONTENTS

Notations 812
Index 821
Chapter 1

Introduction

When I started on this mathematics endeavor, I knew I had to start with set theory. Various paradoxes hang in the air but I knew this subject is beyond my ability. I have to start with a set axioms that are solid and extensive. I thought that I studied so much math, and did so many problems, I never have to question my own mathematics understanding. So, what I know by heart should be correct. I just have to find a set of axiomatic system that conforms with my understanding that avoids the various paradoxes that says warning in certain directions. I find the ZFC axiomatic set theory from my Chinese award book that I received during my high school years Mathematics Handbook Editors Group (1979). This axiomatic system claims to be consistent. It agrees with my understanding in set theory. Most importantly, I see that it avoids the obvious pitfalls of the paradoxes. I took a leap of faith here and based my Chapter 2 on this axiom system. My understanding here is that in my entire life, all the sets I encounter are defined via these axioms except Axiom 4. Axiom 4 is something that I am not familiar with, and is the source of the paradoxes when violated. I feel that I am safe that I define sets only using the rest of the axioms. So, there it is. I had the understanding that in set definition using Axiom 5, one must specify clearly the set A always and then the paradoxes are avoided. I hope that I am correct with my decision. I am confident that I am correct. Throughout the book, I never had those weird set definitions that calls for suspicion. So, even if this axiomatic system is not entirely consistent, I feel that my work will not be of void. The major theorem in Chapter 2 is the equivalence of the Axiom of Choice, the Hausdorff’s Maximal Principle, the Zorn’s Lemma, and the Well-Ordering Principle. The proof of this equivalence (Theorem 2.18) is based on the teaching notes of Professor Peter Leob on Real Analysis. There is no more leap of faith in the rest of the book. Everything follows from reasoning, and Axiom 4 is never invoked.

Chapter 3 is a study of topological spaces. A topological space is a set
together with a collection of subsets of the set which is called the topology on the set. Any set in the topology, it is considered an open set. Any set whose complement is an open set is considered a closed set. Then, we can talk about set closure, interior, boundary. Continuity is the key concept for a function between topological spaces. Fundamental notions of a basis of a topological space, a basis at a particular point in the topological space, the first and second countability of a topological space, the various classes of topological spaces depending on the separability axioms, the category theory of topological spaces, the connectedness concepts are important building blocks for this chapter. In the end of the chapter, the key notion of net is introduced that is a generalization of the concept of a sequence. A sequence has only one direction: forward. A net is like the water flow in a river, it has many vortexes locally, but in the big picture, it has only one direction: forward. With the definition of a net, we can talk about limits and convergence. If we have a net of real (or extended real) numbers, one can talk about limit superior and limit inferior. The concept of continuity is fundamentally linked to the concept of net in topological spaces.

Metric spaces are studied in Chapter 4. A metric space is a set together with a function that defines the distance (metric) between each pair of points in the set. It induces a topology on the set, and therefore we can talk about open and closed sets, continuity and convergence. The concept of completeness is essential in metric spaces. A complete metric space is one in which every Cauchy sequence converges. In addition to these topological concepts, we can further talk about uniform continuity, uniform convergence, equi-continuity, etc, in general uniformity. The fundamental difference between a complete metric space and an incomplete one is delineated in Baire Category Theorem 4.41. Because of the additional properties of a metric spaces instead of a topological space, we are very interested in the question when a topological space is metrizable. This question is solved in Urysohn Metrization Theorem 4.53. Because of the ability to talk about uniform convergence, we give the conditions when two limit operations in series can be interchanged without changing the final answer. This is the topic of Section 4.9. The limit operation in series can be interchanged if the joint limit exists. The joint limit exists if certain convergence is uniform.

The concept of compactness of a set is pervasive in analysis. Chapter 5 devotes to the study of compact spaces and locally compact spaces. The concept of compactness involves actually four concepts: countably compact, Bolzano-Weierstrass property, sequentially compact, and compact for general topological spaces. A countably compact space is equivalent to the space has the Bolzano-Weierstrass property. Otherwise, these compactness properties are distinct in general topological spaces. Here, metric space again demonstrates its superiority over topological spaces, where all these compactness notions are equivalent in metric spaces as demonstrated in Borel-Lebesgue Theorem 5.37. Ascoli-Arzelá Theorem 5.44 then says that for metric space valued equi-continuous functions, pointwise convergence
implies uniform convergence implies that the limit function is continuous. In some sense, this is a sequential compactness theorem for equi-continuous family of functions. In Section 5.6, Tychonoff Theorem 5.47 states that the arbitrary product of compact spaces is compact. The chapter then turns to study locally compact spaces in Section 5.7. A main result for locally compact Hausdorff spaces is that a compact set in such a space admits partition of unity with respect to any open covering among class of functions that satisfy certain standard properties. Another main result for locally compact Hausdorff spaces is that it admits Alexandroff one-point compactification. Here, the concept of proper functions is introduced. The existence of a proper continuous function on a locally compact Hausdorff space is equivalent to $\sigma$-compactness of the space. This chapter ends with a discussion of Stone-Cech compactification.

Starting in Chapter 6, we bring binary operations into our study of spaces. A set with a binary operation (addition) is called a group if the binary operation is associative, admits a unit element (0), and every element in the set admits an additive inverse. If the group operation is commutative, then we have an abelian group. On the other hand, a semi-group is something less than a group. It will become a group if it admits inverses for every element in the set. A step further of the concept of a group, is a ring. A ring is an abelian group together with another binary operation defined on the set: multiplication. On a ring, the multiplication must be associative, and distributive with respect to addition (left and right). A ring can be commutative if the multiplication operation is commutative. It can have a 1 element if the multiplication operation admits a identity element. A commutative ring with an identity element (for multiplication) is further called a field, if every element in the set except 0 has a multiplicative inverse. The concept of field leads us to the familiar entities: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. A vector space over a field is a set together with vector addition and scalar multiplication defined on the set, that further satisfies a list of conditions for scalar multiplication and vector addition. A vector space reminds us of $\mathbb{R}^n$ or $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$ (but with less technical capabilities). We can define linear operators of vector space to vector space. The concept of linear combination, subspaces, linear variety, linear independence, and dimension are introduced in vector spaces. For real or complex vector spaces (vector spaces over the field $\mathbb{R}$ or $\mathbb{C}$), the concept of convexity can be defined.

A vector space is a space with vector operations but without a topology. A normed linear space is a vector space together with a nonnegative real-valued function defined on the set: which is called the norm. The norm is like the concept of the length of an arrow. The arrow has a direction and a length. With the definition of a norm, the space admits a metric space structure, where the distance between two points is simply the length of the difference of the two points (vectors). Thus, we have a topology on the space, and can talk about open and closed sets, continuity, convergence. Since it is a metric space, we can talk about uniform convergence and
uniformity in general. A complete normed linear space is called a Banach space. (Remember that completeness is a concept introduced for metric spaces). The Banach space is the focus of the study in Chapter 7. In $\mathbb{R}^n$, compactness of a set is equivalent to the set being closed and bounded. For finite dimensional Banach spaces, compactness of a set is again equivalent to the set being closed and bounded. But, for infinite dimensional Banach spaces, compactness is hard to come by. But, a lot can be said for a set that is compact in a Banach space. (Then, it is countably compact, has Bolzano-Weierstrass property, sequentially compact. Remember Borel-Lebesgue Theorem). Since a Banach space has the vector space structure, we can talk about the quotient space with respect to a given subspace. The quotient space is a Banach space if the given subspace is closed and we define norm on the quotient space in a specific way. The given subspace can be the nullspace of a pseudo-norm (which is less than a norm in the sense that the pseudo-norm can be 0 for non-null vectors in the vector space). The vector space with a pseudo-norm, can become a normed linear space if we take the quotient of the vector space with respect to the nullspace of the pseudo-norm. The bounded linear operators of a normed linear space to a Banach space forms a Banach space. In particular, bounded linear operators of a Banach space to its associated field $\mathbb{K}$, (stands for $\mathbb{R}$ or $\mathbb{C}$) are called functionals on the original Banach space, form a Banach space of special interest: it is called the dual of the original Banach space. A Banach space together with its dual allows many optimization problems to be solved. A fundamental result concerning a Banach space and its dual is the Hahn-Banach Theorem. In this chapter, we discuss the extension form of the Hahn-Banach Theorem. In simple words, Hahn-Banach Theorem guarantees the existence of an extension of a bounded linear functional defined on a subspace of the Banach space to a bounded linear functional on the entire Banach space without increasing the norm of the functional. This key existence result is proved via a direct application of the Axiom of Choice (or its equivalent). It allows the proof of existence of many optimization problem formulated in the dual space. With this theorem, we know that not all formulations of an optimization problem are equal. A crafty formulation allows the solution to be found, while a naive formulation leads to nowhere. The dual of the dual of a Banach space in general contains the original Banach space as a subspace. It should be noted that there are some Banach spaces whose dual is unknown to us (even though it exists in theory). We can not comprehend such a dual space. The dual of the dual of a Banach space may be the original Banach space, in which case, we say the original Banach space is reflexive. Reflexive Banach space are simpler to deal with. We will restrict ourselves to such spaces if certain proof doesn’t go through for general Banach spaces. In the space of $\mathbb{R}^n$, we have a powerful theorem called the projection theorem. Minimum norm problem from a vector to a subspace can be found by the projection theorem. In Banach space, the concept of orthogonality can be generalized, except that,
for a set in the Banach space, its orthogonal complement is a subset in the
dual of the Banach space. For a subset in the dual of a Banach space, its
orthogonal complement is a subset of the dual of the dual of the Banach
space. To simplify things, we define the pre-orthogonal complement of this
set in the dual to be a set in the original Banach space. With these concepts,
minimum norm problems of a point to a subspace can be characterized
using the orthogonal complement or the pre-orthogonal complement (see
Propositions 7.97 and 7.99). A strong result concerning a bijective bounded
linear operator between two Banach spaces is that its inverse is a bounded
linear operator. This is a deep result that is a consequence of the Baire
Category Theorem. This theorem demonstrates the additional advantage
in working in a Banach space, rather than a normed linear space. For a
bounded linear operator $A$ of Banach space $X$ to Banach space $Y$, it induces
a bounded linear operator $A': Y^* \to X^*$, which is called the adjoint of $A$
that maps of the dual of $Y$ to the dual of $X$. In terms of matrices, the
adjoint of a matrix is its transpose. The range spaces and null spaces
of $A$ and $A'$ are closely related as delineated in Propositions 7.112 and
7.114. These are the infinite dimensional counterparts of the matrix fact
$\dim(R(A)) = n - \dim(N(A))$, where $A$ is an $m \times n$-dimensional matrix,
but are strictly weaker conclusions. Chapter 7 concludes with a discussion
of weak topology on a Banach space and the weak$^*$ topology on the dual
of a Banach space. On a Banach space, the norm induce a topology on the
space, which is called the strong topology. Then, there is the weak topology
which is the weakest topology such that all bounded linear functionals of
the Banach space are continuous. A weak open set is a strong open set,
and a weak closed set is a strong closed set. But, a strong open set may
or may not be a weak open set, and a strong closed set may or may not
be a weak closed set. Remember the topology is the collection of subset
of the space that we call them open sets. Weak topology is a subset of
the strong topology. With the definition of weak topology on the Banach
space, we can talk about continuity and convergence in the weak topology.
It is true that vector addition and scalar multiplication are continuous in
the weak topology. For finite dimensional Banach spaces, weak topology
is the same as the strong topology. So, the study of weak topology is for
infinite dimensional Banach space only. On the dual of a Banach space,
(remember it is also a Banach space) we have the weak topology, and we
will define a possibly even weaker topology: the weak$^*$ topology: it is
the weakest topology on the dual such that all the element of the Banach
space viewed as a bounded linear functional on the dual are continuous.
When the Banach space is reflexive, then the weak$^*$ topology coincides
with the weak topology on the dual of the Banach space. Again, it is true
that vector addition and scalar multiplication are continuous in the weak$^*$
topology. The main reason why we study these topologies is the Alaoglu
Theorem 7.122 that says that the closed unit ball of the dual of the Banach
space is a weak$^*$ compact set. Don’t rush to the wrong conclusion that
every sequence in the closed unit ball has a weak* convergent subsequence! Remember this is weak* topology, not a metric space, so weak* compactness doesn’t imply weak* sequential compactness. The right conclusion can be made here is that every sequence in the closed unit ball has a weak* cluster point (Bolzano-Weierstrass property). Another piece of useful result in the weak topology is that a bounded linear operator $A$ between Banach spaces is continuous as a operator between the weak topologies of these Banach spaces.

Chapter 8 studies the convex optimization problems in real Banach spaces. The concept of a hyperplane is defined to be a maximal proper linear variety in the normed linear space. Every hyperplane can be identified by a set $\{ x \in X \mid f(x) = c \}$, where $f$ is a non-null bounded linear functional in the dual of $X$ and $c \in \mathbb{R}$. The direct consequence of the Hahn-Banach Theorem is Mazur’s Theorem 8.7 that says that given a linear variety and convex set with nonempty interior in a real normed linear space, if the linear variety doesn’t intersect the interior of the convex set, then there is a hyperplane containing the linear variety such that the convex set is contained in exactly one of the two closed half spaces associated with the hyperplane. Another direct consequence of Hahn-Banach Theorem is Eidelheit Separation Theorem 8.9 that says given two nonempty convex sets $K_1$ and $K_2$ in a real normed linear space $X$, if the interior of $K_1$ is nonempty, and the interior of $K_1$ doesn’t intersect $K_2$, then we may find a hyperplane that separates $K_1$ and $K_2$. A direct consequence of these theorems is that a closed convex subset of a real normed linear space is weakly closed (remember the weak topology). This combined with Alaoglu Theorem yields that, for a reflexive real Banach space $X$, a closed bounded convex subset in the dual of $X$ is weak* compact. The infimization of the norm of a point to a nonempty convex set in a real normed linear space is equal to the maximization problem involving the support functional of the convex set over the support of the convex set. The same problem, if formulated in the dual of a reflexive real Banach space then admits a minimizer. For more general problems where the optimizing functional is convex but not necessary a norm expression, we introduce the concept of conjugate convex functional. For convex functionals defined on the dual of a real normed linear space, we introduce the concept of pre-conjugate convex functional. This culminates in Propositions 8.32 and 8.33, where the original function is equal to the supremization over bounded linear functional in the conjugate convex set of a functional involving the conjugate convex functional. Under stronger assumptions, the supremization becomes maximization in Proposition 8.34. A step further in this line of study is the Fenchel Duality Theorem 8.35. The key game theory result is obtained in Proposition 8.39. For constrained convex optimization problems, we first introduce the concept of positive cone and its dual concept of positive conjugate cone. The chapter ends with presentation of two flavor of the Lagrange multiplier theorems, as well as sufficient conditions of optimality.
In a Banach space, we can do vector addition and scalar multiplication and has the concept of distance. All the necessary ingredients are present for the study of calculus on the space. In Chapter 9, we embark on the study of differentiation on Banach space. This topic is relatively well understood and has significant body of presentation in Luenberger (1969). I follow closely the results on Fréchet derivative for functions between Banach spaces. I take extra care in the results for functions with general domains (not simply open domains). There, an extra condition is needed for the shape of local shape of the domain at the point of differentiation. For high order derivatives, (including partial derivatives) we inevitably deal with tensor operators. A function \( f \) is \( C_k \) at \( x_0 \in X \) and \( \text{dom}(f) \) is locally convex implies that the \( k \)th order derivative \( f^{(k)}(x_0) \) is a \( k \)-fold symmetric operator on \( X \). The calculation of high order derivatives of a function can be carried out according to Proposition 9.27, once it is shown that it is sufficiently many times differentiable. Interchange order of differentiation can be carried out according to Propositions 9.31 and 9.32. These results are standard (see Bartle (1976)) if the domain of the function is open. When this assumption is not valid, then the technical assumptions on the domain that allow the interchange of order of differentiation are delineated in these propositions. The derivatives of useful functions are obtained as well as some basic properties governing the high order derivatives are obtained. A key take away from Section 9.4 is that \( C_k \) function composed with \( C_k \) functions yields \( C_k \) functions. In particular, this result holds at each point in the domain (see Proposition 9.45). For a point in the interior of the domain, the function is \( C_k \) at the point if all partial derivatives up to \( k \)th order are continuous at the point (see Proposition 9.47). The study of high order derivative ends with the well-known Taylor’s Theorem 9.48. The usefulness of the differentiation in Banach space is because of various mapping theorems Section 9.5 and analytic functions Section 9.9. In Section 9.5, we present the Contraction Mapping Theorem, Injective Mapping Theorem, Surjective Mapping Theorem, Open Mapping Theorem, Inverse Function Theorem, and two versions of Implicit Function Theorems (for continuous and differentiable cases). These results are motivated by Luenberger (1969). The global version of the Inverse Function Theorem is motivated by Ambrosetti and Prodi (1993). The result is that for a continuous and countably proper function \( F \) between Hausdorff topological spaces, if it is locally continuously invertible at each point in the domain, \( \text{dom}(F) \) is arcwise connected and the range \( \text{range}(F) \) is simply connected, then \( F \) is a homeomorphism between \( \text{dom}(F) \) and \( \text{range}(F) \). (A continuous function is said to be countably proper if the inverse image of any compact set is countably compact). The technicality all go away once we deal with metric spaces. In Section 9.7, we present sufficient conditions under which the limit operation and differentiation operation in series can be interchanged. The result becomes standard (Bartle, 1976, see) if the domain of the function is open set. In Section 9.8, we give the definition and properties of basic
tensor operations. The chapter ends with the section on analytic functions that gives all essential ingredients of analytic function on Banach space: the definition of an analytic function, the Taylor series expansion of an analytic function, the composition of analytic functions yields an analytic function, the inversion of bounded linear operators between Banach spaces as an analytic function, the analytic versions of Inverse Function Theorem and Implicit Function Theorem. It also includes the discussion of the analytic function: the exponentiation of a bounded linear operator on a Banach space.

In Chapter 10, we derive the local necessary, as well as sufficient conditions for the unconstrained optimization of a functional, the constrained optimization of a functional subject to equality constraints, as well as inequality constraints, on real Banach spaces. We define the concept of positive definite, positive semi-definite, negative definite, and negative semi-definite operators on a real Banach space. We derive the necessary and sufficient conditions for Fréchet differentiable functionals to be a convex functional. In constrained optimization problems, the constraints are mapping between Banach spaces, and the inequality constraints are in terms of positive cones. Lagrange multiplier theory is presented for constrained optimization problems. The generalized Kuhn-Tucker Theorem for optimization problems with infinite dimensional inequality constraints is presented. The results are standard and inspired by Luenberger (1969).

The first five sections of Chapter 11 is dedicated to the study of measure theory, measurable functions on a measure space, general integration theory on a measure space, and the interchange of integration and limit via the Lebesgue Dominated Convergence Theorems. As our study proceeded, we decided that it is unfruitful to deal with complete measure spaces, rather we focus on Borel measure spaces, when dealing with measures on a topological space. The foundation of measure theory is the Carathéodory Extension Theorem. In the definition of the Lebesgue measure on \( \mathbb{R} \), we observe the key feature of Lebesgue measure: any Lebesgue measurable set \( E \) can be contained by some open set \( U \) with arbitrarily small Lebesgue measure \( \mu_L(U \setminus E) \); or equivalently, it can contain some closed set \( F \) with arbitrarily small Lebesgue measure \( \mu_L(E \setminus F) \). I restrict my consideration to only those Borel measurable sets \( B_B(\mathbb{R}) \). The restriction of the Lebesgue measure to these Borel measurable sets is defined to be the Borel measure on \( \mathbb{R} \), \( \mu_B \). This measure clearly inherits the key feature from Lebesgue measure. This leads me to define topological measure space \( X \) on a topological space \( (X, \mathcal{O}) \) to be a the space \( ((X, \mathcal{O}), \mathcal{B}, \mu) \) where \( \mu \) is a measure on the measurable space of \( (X, \mathcal{B} := B_B(X)) \), where \( \mathcal{B} \) is the \( \sigma \)-algebra generated by the open subsets of \( X \), such that any (Borel) measurable set \( E \in \mathcal{B} \) can be contained by some open set \( U \in \mathcal{O} \) with arbitrarily small measure \( \mu(U \setminus E) \). A metric measure space is a topological measure space on top of a metric space, and Banach measure space is a topological measure space on top of a Banach space, and so on and so forth. One simple relationship due
to this definition is that any set $E \in \mathcal{B}$, let $(E, \mathcal{O}_E)$ be the topological subspace of $X$, and $(E, \mathcal{B}_E, \mu_E)$ be the measure subspace of $(X, \mathcal{B}, \mu)$, then $\mathcal{B}_E$ is the $\sigma$-algebra generated by $\mathcal{O}_E$ and $\mu_E$ satisfies the preceding condition such that $E := ((E, \mathcal{O}_E), \mathcal{B}_E, \mu_E)$ form a topological measure space. So, I call $E$ the topological measure subspace of $X$. On a topological measure space, all measurable subset admits a topological measure structure. With the focus on Borel measurable sets, a function is measurable if the inverse image of any open set (and hence any Borel measurable set) is Borel measurable. These concepts are not changed. But, the concept of two function are almost everywhere equal to each other is significantly changed. We say a property holds almost everywhere if the set on which the property doesn’t make sense or doesn’t hold is a Borel measurable set and has measure zero. The burden here is to guarantee that this negligible set is (Borel) measurable, which can be ensured by the measurability of the functions on both sides of the equality and the range of the functions is a separable metric space. The pointwise limit of measurable functions is again measurable. But for almost everywhere limit of measurable functions, one has to be careful: the limiting function is measurable but not necessarily defined everywhere. To get a measurable function that is defined on the entire domain as an almost everywhere limit of the sequence of measurable functions, one can simply set the function to an arbitrary constant whenever the pointwise limit doesn’t exist. By taking caution in proof arguments, the standard machinery of measure theory carries through to Borel measurable spaces. The Littlewood’s Three Principles are cleanly established and appropriately generalized to Borel measurable spaces. The integration of a measurable function on a measure space is carefully defined. So, $0 \times \infty = 0$ is completely banished from the theory. First, we define a directed system, which is called the integration system, on the range of the integrand (which is a normed linear space). First, integration on finite measure space is defined. This directed system results in a net for the corresponding integrand, whose convergence determines the integrability of the integrand. Special care is taken for cases when the integrand is possibly unbounded on the domain of integration, where the integration system has special rules for the sample point that represents a measurable set in the range of the integrand. Based on this definition, we further define the integration of a measurable function on a measure space of $\infty$ measure to be the limit of the net of integrals on finite measure subspaces of the measure space with the partial ordering being set containment for the finite measure subspaces. Then, we prove that these definitions lead to expected results for the integration of simple functions. Then, Bounded Convergence Theorem 11.77 is established for general integrands over finite measure spaces. Then, we prove that our definition leads to the same integration formula for nonnegative real-valued measurable functions over any measure space Proposition 11.78 as Royden (1988). Now, we are fully sure that our definition of integration is correct. We then prove Monotone Convergence Theorem and establish
all well-known properties for Lebesgue integral of nonnegative real-valued functions. Then, we continue in Section 11.5 to establish Lebesgue Dominated Convergence Theorem for Banach space valued measurable functions, which culminates in Theorem 11.91. Thus, under the assumption that the integrands are absolutely integrable over the measure space and the Banach space being separable, all important properties of integration for Banach space valued functions are established in Proposition 11.92. The integral version of the Jensen’s Inequality is established in Theorem 11.98. With these success in integration theory, I couldn’t help to wonder whether a measure can be Banach space valued. The integrations on measurable subset over a measure space of a Banach space valued measurable function seems to generate a $\sigma$-finite Banach space valued measure on the original measurable space. The question is how to defined such a measure. The definition is relatively straightforward if for every measurable set, the Banach space valued measure generates a vector in the Banach space. This leads to definition of pre-measure on a measurable space. For the pre-measure to be countably additive, one must assume that the measures on a countable collection of disjoint measurable sets must be absolutely summable. Taking this idea to the limit when the collection of disjoint sets are small in size but covers the entire measurable space, we inevitably stumble upon the total variation of the Banach space valued pre-measure. This total variation must be a finite measure on the measurable space. Now, the intuition is obtained and the rest is just labor. To define a Banach space valued measure $\mu$ on a measurable space, there must be first a measure $\nu$ on the measurable space, which is known as the total variation of $\mu$, to be denoted by $P \circ \mu$. Whenever a measurable set $E$ has $\nu(E) = \infty$, $\mu(E)$ is undefined; on the other hand, when $\nu(E) < \infty$, then $\mu(E)$ is a point in the Banach space, and satisfies the countable additivity for its subsets. $\nu(E)$ must then relate to $\mu(E)$ as the set of $E$ is breakdown to ever smaller disjoint pieces $(E_i)_{i=1}^\infty$ and the sum of the $\| \mu(E_i) \|$’s must give us back $\nu(E)$. This definition is very different from its obvious counterpart in Royden (1988), the signed measure. But, this is the correct path forward. We confirm the intuition that integrations on measurable subset over a measure space of an measurable function $f(\cdot)$ indeed generate a $\sigma$-finite Banach space valued measure on the original measurable space, where the total variation is simply the integrations of measurable subset over the measure space of the measurable function $\| f(\cdot) \|$ (see Proposition 11.116). To generate a $\sigma$-finite Banach space valued measure space, we may piece together countably many finite Banach space valued measure spaces as long as they are compatible on common domain of definition. The resulting $\sigma$-finite Banach space valued measure space admits all of the finite Banach space valued measure space as measure subspaces. This is so-called generation process, and the resulting $\sigma$-finite Banach space valued measure space is furthermore unique. I will have a few opportunities to apply the generation process to yield desired $\sigma$-finite Banach space valued measure spaces. Then, I define
integration over Banach space valued measure space along the same line as for measure spaces. The integral of simple functions over a Banach space valued measure space is determined and is as expected. Then, Lebesgue Dominated Convergence Theorem 11.131 is established. We define a function to be absolutely integrable over a Banach space valued measure space if its norm function is integrable over the total variation of Banach space valued measure. Under the assumption of absolutely integrability of integrands, most important properties of integration holds for integration over Banach space valued measure spaces (see Proposition 11.132). Now, we have Banach space valued measure spaces. For a fixed measurable space $\mathcal{X} := (X, \mathcal{B})$ and a fixed Banach space $Y$, the collection of all finite $Y$-valued measures on $\mathcal{X}$ form a Banach space $\mathcal{M}_f(\mathcal{X}, Y)$, where the norm of each finite $Y$-valued measures is the total variation of the measure of $\mathcal{X}$. The key here is the definition of vector addition and scalar multiplication for finite $Y$-valued measures. The vector addition and scalar multiplication are defined for $\sigma$-finite $Y$-valued measures as well via the generation process. Then, the collection of all $\sigma$-finite $Y$-valued measures on $\mathcal{X}$ form a vector space $\mathcal{M}_\sigma(\mathcal{X}, Y)$. We can further define a topology on $\mathcal{M}_\sigma(\mathcal{X}, Y)$ whose restriction on $\mathcal{M}_f(\mathcal{X}, Y)$ is exactly the one induced by the norm on $\mathcal{M}_f(\mathcal{X}, Y)$. Then, $\mathcal{M}_\sigma(\mathcal{X}, Y)$ is a vector space with a topology, we can then talk about continuity and convergence in $\mathcal{M}_\sigma(\mathcal{X}, Y)$. Then, we prove the Lebesgue Dominated Convergence Theorem 11.151 in the general setting of $\mathcal{M}_\sigma(\mathcal{X}, Y)$, when the underlying measure on the measurable space are also changing, but converges to some $\sigma$-finite $Y$-valued measure $\mu$ according the topology on $\mathcal{M}_\sigma(\mathcal{X}, Y)$. The result of Proposition 11.116 is further generalized to $Y$-valued measure spaces. The integrations on measurable subset over $Y$-valued measure space of an measurable function $f(\cdot)$ generate a $\sigma$-finite Banach space valued measure on the original measurable space, where the total variation of the new measure is upper bounded by the integrations of measurable subset over the total variation of $Y$-valued measure space of the measurable function $\| f(\cdot) \|$ (see Proposition 11.153). The upper bound of the total variation is exactly the total variation if we restrict $Y$ to be $K$ (see Proposition 11.154). Until now, I have only the following: give me a measurable function over a ($K$-valued) measure space, I can integrate the function to generate a $Z$-valued measure on the measurable space, where $Z$ is a Banach space which contains the range of the measurable function. The inverse question of the above is interesting and is resolved by Radon-Nikodym Theorem. The fundamental condition that allows this inverse operation (Radon-Nikodym derivative) is that the $Z$-valued measure must be absolutely continuous with respect to the $K$-valued measure. Once the Radon-Nikodym derivative is known, integration with respect to the $Z$-valued measure can be easily carried out with respect to $K$-valued measure. This study is carried out in Section 11.8, which culminates in Theorems 11.169 and 11.171, where the first one is for measures and the second one is for Banach space valued measures. Section 11.9 stud-
and introduce the concept of isomeasure: there are two measurable spaces. In Section 12.2, I have to derive a generalized Carathéodory Extension Theorem 12.4 to handle these measure spaces. The main objective is to characterize the dual space of $L_p(\mathcal{X}, \mu)$, where $\mu$ is a measure space, and $\mathcal{Y}$ is a separable Banach space. I present the Minkowski’s Inequality and Hölder’s Inequality, which then guarantees that $L_p(\mathcal{X}, \mu)$ is a pseudo-normed linear space. If we take quotient space of $L_p(\mathcal{X}, \mu)$ with respect to the nullspace of the pseudo-norm, we have the normed linear space $L_p(\mathcal{X}, \mu)$, which a Banach space if $\mathcal{Y}$ is a Banach space. The main objective is to characterize the dual space of $L_p(\mathcal{X}, \mu)$. This leads to the Riesz Representation Theorem 11.186, which says that $(L_p(\mathcal{X}, \mu))^* = L_q(\mathcal{X}, \mu^*)$, where $p \in [1, \infty]$ and $q \in (1, \infty)$ with $1/p + 1/q = 1$. 

Chapter 12 is the climax of this mathematical journey. Finally, the I must somehow link integration, Radon-Nikodym derivative, and Fréchet derivative together that allow us to go smoothly from one thing to the other, solve the puzzle of the connection between multi-dimensional Riemann integral with Lebesgue integral that we have defined, and derive the Change of Variable formula that we are so familiar for Riemann integration. Because I have introduced Banach space valued measures, to obtain product measure spaces, I have to derive a generalized Carathéodory Extension Theorem 12.4 to handle these measure spaces. In Section 12.2, I introduce the concept of isomeasure: there are two measurable spaces $\mathcal{X}$ and $\mathcal{X}$ and each with a (Banach space valued) measure $\mu$ and $\mu$, respec-
The two measures are said to be isomeasurable if there is a bijective mapping $g: \mathcal{X} \rightarrow \bar{\mathcal{X}}$ such that the inverse image of any measurable set in $\bar{\mathcal{X}}$ is measurable in $\mathcal{X}$, the image of any measurable set in $\mathcal{X}$ is measurable in $\bar{\mathcal{X}}$, the two total variations of the two measures agree (equal to each other) for any measurable set in $\bar{\mathcal{X}}$ and its inverse image under $g$ in $\mathcal{X}$, and the two measures agree (equal to each other whenever one of them is defined, or both remain undefined when the common total variation is $\infty$) for any measurable set in $\bar{\mathcal{X}}$ and its inverse image under $g$ in $\mathcal{X}$. Two (Banach space valued) measure spaces are isomeasurable implies that they are equivalent up to relabeling of elements. Isomeasure preserves the completeness, finiteness, and $\sigma$-finiteness of measure spaces. A step further, I introduce the concept of homeomorphic isomeasure between two (Banach space valued) topological measure spaces, these basically are bijective mappings that are isomeasure and a homeomorphism at the same time. Two topological measure spaces are equivalent up to relabeling of elements if they are homeomorphically isomeasurable. Given a topological measure space, a Borel measurable space on a topological space, and homeomorphism between the two topological spaces, the induced measure on the measurable space based on the bijective mapping and the topological measure space, is such that the topological measure space and the topological measurable space with the induced measure are homeomorphically isomeasurable under the given mapping. It is straightforward to prove that integration on one of the measure space can be easily computed via integration on the other measure space after the integrand is composed with the bijective mapping. As a prelude to the ultimate Change of Variable Theorem, we obtain Theorem 12.16: we have two measure spaces and a bijective mapping between them; the mapping induces a measure on the second measure space based on the first measure space. Now, if we know that the induced measure admits a Radon-Nikodym derivative with respect to the measure on the second measure space, then we may easily calculate integral on the first measure space by calculating the integral on the second measure space with respect to the induced measure, which can be expressed as an integration over the second measure space as long as we multiply the Radon-Nikodym derivative into the integrand. The ultimate theorem on change of variables will delineate the conditions such that the induced measure admits the Radon-Nikodym derivative as desired. This task is postponed until Section 12.6. In Section 12.3, we established the existence of the product ($\mathbb{K}$-valued) measure space of finitely many ($\mathbb{K}$-valued) measure spaces. These results are motivated by the Math 442 notes of Professor Peck of University of Illinois at Urbana-Champaign. Special care is taken to deal with Borel measure spaces, rather than Lebesgue measure spaces. Here, things are actually simpler. A measurable set in the product measure space has its sections in the multiplier measure spaces to be all measurable. Then a measurable function in the product measure space has all of its section in the multiplier measure spaces to be measurable. Tonelli’s Theorem 12.29 allows the inte-
Integral in product measure spaces of nonnegative real-valued integrands to be written as repeated integrals, where we have to allow integrals of nonnegative extended real-valued functions, which is defined in Definition 11.79. Tonelli's Theorem is usually used to prove a function to be absolutely integrable over the product measure space. Under the assumption that the integrand is absolutely integrable over the product measure space, Fubini's Theorem then gives conditions that allow the joint integral to be computed iteratively. It is observed that general integration can be done if the range of the integrand is a subset of a separable Banach space. But, for product measure spaces, if the integration system is set up on a separable Banach space, there is no guarantee that iterative integrals will make sense, that first integrating with respect to one multiplier measure space may not render a measurable function for the next integration. Here, we can say that the product integral can be written as repeated integrals if we appropriately define the intermediate integrands so that they are measurable and equals to the integral with respect to first multiplier measure space almost everywhere whenever the first iterative integral is absolutely integrable. Theorem 12.30 promises that such expression as iterative integrals always exist, and if one can make the integrand for the second iterative integration step to be measurable and equal to the integral with respect to the first multiplier measure space almost everywhere, then you can write the product integral as repeated integrals using such an integrand. The picture for integrands whose range is a subset of a finite-dimensional Banach space is much simpler. All you need to do (see Fubini’s Theorem 12.31) is that you write the product integration as repeated integral, and in the integration with respect to the first multiplier measure space, just set the integration result to be the null vector if the integrand is not absolutely integrable (with respect to the first multiplier measure space), then automatically the second integration has an absolutely integrable integrand. In this section, we further prove that the finite product measure space of second countable topological measure spaces is a second countable topological measure space (with respect to the product topological space). Things are nice here. Same can be said when the multiplier topological measure spaces are $\mathbf{K}$-valued topological measure space. Fubini’s Theorem 12.35 generalizes the Theorem 12.30 to $\mathbf{K}$-valued product measure spaces. We see that (Banach space valued) measures are important objects to deal with. The question is how do we keep track of them, even when their domain is $\mathbb{R}^m$. It is known in 1-dimensional case, a measure can be represented by it cumulative distribution function, which are functions of bounded variation when the measure is finite. In Section 12.4, we embark on a study to characterize so called functions of locally bounded variation, which are functions on $\mathbb{R}^m$ that will correspond to the cumulative distribution function of (Banach space valued) measures on $\mathbb{R}^m$. First observation is that for a measure on the interval $(1,2]$, the cumulative distribution function must live on the interval $[1,2]$ unless the rule of the cumulative distribution function is modified.
at end points of the interval. This modification seems to be difficult to characterized for measures on $\mathbb{R}^m$. Also, things will be weird if the domain of the cumulative distribution function is too strange. As motivated by 1-dimensional case, we must define these cumulative distribution functions to be continuous on the right. (Continuous on the left works as well, if we insist that cumulative distribution function is continuous on the left on every point in the domain. We just have to pick one and stick with it.) Then, I proceed first to define the notion of a region $\Omega$, of which the cumulative distribution functions are defined, and it principal $P(\Omega)$ on which the measure is defined. The region $\Omega$ is the countable union of closed rectangles in $\mathbb{R}^m$; while the principal $P(\Omega)$ is the countable union of the corresponding semi-open rectangles in $\mathbb{R}^m$, that further satisfies the assumption that for any closed rectangle contained in $\Omega$, its corresponding semi-open rectangle must be contained in $P(\Omega)$. As defined, there is a unique $P(\Omega)$ for each region $\Omega$, but not vise versa. It is easy to see that a region and its principal are (Borel) measurable sets. The principal enlarges if the region enlarges. For a rectangle in $\mathbb{R}^m$, it must be a region, and its principal is a rectangle that can be easily characterized. The countable union of semi-open rectangles in $\mathbb{R}^m$ must be a region, whose principal is itself. Product of regions is a region, whose principal is the product of the principals of the regions. An open subset of $\mathbb{R}^m$ is a region whose principal is itself. Then, we proceed to define the function of locally bounded variation on a region. These are functions that are continuous on the right first of all. We define the increment of such a function on a semi-open rectangle (whose closure is a subset of the region) to be an alternate sum of the values of the function on vertices of the rectangle. We define the total variation of such a function on a semi-open rectangle (whose closure is a subset of the region) to be supre- mum of the sum of the norm of the increments of the function on semi-open rectangles whose disjoint union is the given semi-open rectangle. Second of all, for a function to be of locally bounded variation, it must have finite total variation on any semi-open rectangle whose closure is a subset of the region. Lastly, on any closed rectangle that is a subset of the domain (a region), we may define the total variation function of the semi-open rectangle with opposite corners consisting the lower right corner of the rectangle and any point in the rectangle, this function must be continuous from the right. So, three conditions for a function to be of locally bounded variation. The total variation on semi-open rectangles reminds us of the total variation of the Banach space valued measure, and the last condition on the right continuity of the total variation function, reminds us of the fact that the total variation must be a measure, and therefore must have a right continuous cumulative distribution function. The function is said to be of bounded variation if it total variation on the domain (region) is finite. We hope that we have nailed this definition. To prove this, we define a cumulative distribution function for a $\sigma$-finite Banach space valued measure on the principal of the region to be one such that for any semi-open rectangle
whose closure is a subset of the region, the measure of the semi-open rectangle equals to the increment of the cumulative distribution function on the semi-open rectangle. In Theorem 12.50, we establish the result that any Banach space \((Y)\) valued function of locally bounded variation defined on a region that is a subset of \(\mathbb{R}^m\), there exists a unique \(\sigma\)-finite \(Y\)-valued measure on the principal of the region such that the function is a cumulative distribution function of the \(Y\)-valued measure, and the total variation of measure on any semi-open rectangle whose closure is a subset of the region is equal to the total variation of the function on the semi-open rectangle, and the total variation of the measure on the entire principal of the region is equal to the total variation of the function. This gives one way of the correspondence. For any finite \(Y\)-valued Borel measure on the principal of the region, we can always find a function of locally bounded variation defined on the region that is the cumulative distribution function of the measure (see Proposition 12.51). So, the opposite way is valid for finite measures. For \(\sigma\)-finite \(Y\)-valued measures, a cumulative distribution function may not exist and is not unique. The converse holds (see Proposition 12.52) if the region is a rectangle \(\Omega\) (which must be a region by Proposition 12.39 as we have discussed) and the \(Y\)-valued measure is on the principal of \(\Omega\), if the total variation of the measure on any semi-open rectangle, whose closure is a subset of the rectangle (region), is finite, then for any point \(x_0\) in the region, we may find the cumulative distribution function of the measure with origin \(x_0\) (defined on the region), which is a function of locally bounded variation; for the total variation of the \(Y\)-valued measure, we may find the cumulative distribution function of the total variation with origin \(x_0\), and this function is also of locally bounded variation and its value on any semi-open rectangle whose closure is a subset of the region equals to the total variation of the cumulative distribution function for the \(Y\)-valued measure on the semi-open rectangle. For more general regions, a cumulative distribution function may not be unique or exist since the increment of a function has many degree of freedom that may not be captured by just specifying the origin \(x_0\). With these converses done, I ponder on the question whether a function of locally bounded variation or a cumulative distribution function of a \(Y\)-valued measure is measurable on \(\mathbb{R}^m\). Propositions 12.54 and 12.55 shows that the cumulative distribution function on a region defined for finite \(Y\)-valued measures, and cumulative distribution function with origin \(x_0\) defined on a rectangle for \(\sigma\)-finite \(Y\)-valued measures are measurable. For arbitrary function of locally bounded variation, it should be clear that 1-dimensional case is OK. For higher dimensional domain, there is too much freedom in the function to allow a positive conclusion. An example on \(\mathbb{R}^m\), which is the \(m\)-dimensional product of the Borel measure space \(\mathbb{R}\), demonstrates that the \(m\)-dimensional outer measure induced by the \(m\)-dimensional Borel measure of a set \(E\) is equal to infimum over sum of \(m\)-dimensional Borel measures of any countable open covering of \(E\). The measurable sets with respect to this outer measure
(m-dimensional Lebesgue outer measure) are the m-dimensional Lebesgue measurable sets. Since the Borel measure \( \mu_B \) on \( \mathbb{R} \) is translational invariant, so is the m-dimensional Borel measure \( \mu_{Bm} \), as well as the m-dimensional Lebesgue measure \( \mu_{Lm} \). The section ends with the definition of Lebesgue-Stieltjes Integral: the integral on a region in \( \mathbb{R}^m \) of a measurable function with respect to some function of locally bounded variation is defined to be the integral on the principal of the region of the measurable function with respect to the unique \( \sigma \)-finite Banach space valued measure that is prescribed in Theorem 12.50. In Section 12.5, I investigate the property of a cumulative distribution function for a Banach space valued measure that is absolutely continuous (admits Radon-Nikodym derivative) with respect to the m-dimensional Borel measure \( \mu_{Bm} \). In essence, I seek to define absolutely continuous functions on \( \mathbb{R}^m \). In \( \mathbb{R} \), the definition has been widely known. On higher dimensional spaces, the m-dimensional integral of the Radon-Nikodym derivative still allows additional degree of freedom in the result in terms of integration constant (functions) in the lower dimensional surface of the region. A function can be the cumulative distribution function for an absolutely continuous measure (with respect to \( \mu_{Bm} \)) without being continuous. This is obviously an undesirable aspect. So, I define a function on a rectangle to be absolutely continuous at a point \( x_0 \) if freeze any subset of the m-coordinates yields a absolutely continuous function in a lower dimension; and in the m-dimensional space, \( \forall \epsilon > 0, \), there exists a \( \delta > 0, \) for any pairwise disjoint semi-open rectangles in a vicinity of \( x_0, \) their \( \mu_{Bm} \) measure sums to less than \( \delta, \) implies that the norm of the increment of the function on the semi-open rectangles sums to less than \( \epsilon. \) This is an obvious try at the definition of absolutely continuous functions on higher dimensions. The reason why I insist on the domain being a rectangle is that the \( \epsilon-\delta \) language does not make sense if the domain is full of holes, or has weird boundaries. A possible alternative to my choice is to consider open domain, which should work just as well. But, I wanted to piece together absolutely continuous functions on closed rectangles, which will result in an absolutely continuous function on the joint rectangle. The rectangle domain allows me to investigate absolute continuity on the boundary of a set. Then, I think that one just need to have a big set on which the function is absolutely continuous. A smaller region in the big set might be investigated for physical phenomenon. So, I settle in this formulation. It is proved a function is absolutely continuous at a point implies that it is continuous at the point. Vectorizing two absolutely continuous functions on the same domain yields an absolutely continuous function. Absolute continuity is preserved under restriction of the domain. Piecing together absolutely continuous functions on two smaller rectangles yields an absolutely continuous function on the joint rectangle if the smaller rectangles are both relatively open or both relatively closed. Closely related to absolutely continuous functions on intervals, is the Lipschitz continuous functions. A Lipschitz continuous function composed with an absolutely continuous function on
an interval is again an absolutely continuous function. This fact is the key to the integrability of nonlinear ordinary differential equations. There is no counterpart of this result in higher dimensions, which determines the nonsolvability of general partial differential equations. The product of two absolutely continuous functions on intervals remains absolutely continuous. This is the reason why Integration by Part Theorem exists. Again, there is no counterpart of this result in higher dimensions. The composition of bounded linear operator with a absolutely continuous function on $\mathbb{R}^m$ is again absolutely continuous. The key word here is linear, it does not work for nonlinear operators. An absolutely continuous function is automatically of locally bounded variation. Therefore, there exists a unique $\sigma$-finite $Y$-valued measure that admits the function as a cumulative distribution function (see Proposition 12.73). This $Y$-valued measure is absolutely continuous with respect to $\mu_{Bm}$, and therefore admits Radon-Nikodym derivative $f$ if $Y$ is reflexive and separable with $Y^*$ being separable. Then, the $Y$-valued measure is such a measure with kernel $f$ over $\mathbb{R}^m$. Then, the increment of the function on a semi-open rectangle is equal to the integral of $f$ over the semi-open rectangle with respect to $\mu_{Bm}$. The converse direction of the preceding discussion works like this: start with a rectangle in $\mathbb{R}^m$ and a $\sigma$-finite $Y$-valued measure on its principal (as a topological subspace of $\mathbb{R}^m$), assume that $Y$-valued measure is absolutely continuous with respect to $\mu_{Bm}$ on the principal, then the cumulative distribution function of the $Y$-valued measure with any origin $x_0$ in the rectangle is absolutely continuous. So far, we have been only working with integration and Radon-Nikodym derivative. Section 12.6 studies the line between integration and Fréchet derivative. The key here is the Vitali’s Lemma. For higher dimensional space (beyond 1), the key is to adapt the Vitali’s Lemma to rectangles rather than balls. Then, we may define Lebesgue point of an open set in $\mathbb{R}^m$ and show that the set of Lebesgue points is Lebesgue measurable (not Borel), and its complement in the original open set has $(m$-dimensional) Lebesgue measure 0. Given a right continuous function on $\mathbb{R}^m$, it directly leads to that the partial derivatives with respect to coordinate variables are $\mu_{Bm}$-measurable, and the Fréchet derivative of the function is $\mu_{Bm}$-measurable. The question is when does these derivatives exists almost everywhere, and whether it allows the interchange of integration and differentiation. Standard result in Riemann integral case directly carries over to Lebesgue integral (only in 1-dimension, see Theorems 12.82 and 12.83. In Fundamental Theorem of Calculus I, Theorem 12.86, we show that the function defined as a $m$-dimensional integral of a measurable function that is absolutely integrable over any semi-open rectangles must admit all partial derivatives along all coordinate variables almost everywhere $\mu_{Bm}$. The partial derivatives equals to (almost everywhere $\mu_{Bm}$) the $(m - 1)$-dimensional integral subject to measurability concerns. If the function is finite-dimensional Banach space valued, then the partial derivatives along coordinate variables is equal to (almost everywhere $\mu_{Bm}$) the
\( (m - 1) \)-dimensional integral whenever the integral is absolutely integrable and set to the null vector whenever it is not. In Fundamental Theorem of Calculus II, Theorem 12.88, a \( Y \)-valued absolutely continuous function defined on an open rectangle of \( \mathbb{R}^m \) can be expressed as the sum of \( 2^m - 1 \) integrals of the stream functions of the original function with respect to a point \( x_0 \) in the domain. These stream functions are unique up to a set of measure 0 on their domain of definition. The partial derivatives along coordinate variables is equal to (almost everywhere \( \mu_B \)) some one-dimensional stream function with respect to some point. If the stream functions are locally bounded, then the Fréchet derivative of the original function exists almost everywhere in the open rectangle. The stream functions of the original function with respect to any point \( \bar{x} \) can be obtained form the stream functions with respect to \( x_0 \). When \( Y \) is finite-dimensional, this transformation of stream functions are calculated whenever the integrands in the transformation are absolutely integrable, but set to null vector of \( Y \) if any of the integrands is not absolutely integrable over its domain of integration. Integration by Parts Theorem 12.89 allows for absolutely continuous functions on intervals instead of requiring \( C^1 \) functions. The Change of Variable Theorem 12.91, specifies the conditions under which the induced measure is absolutely continuous with respect \( \mu_B^m \) and proved the usual formula, which allows for absolutely integrable integrand functions and homeomorphisms that may not be \( C^1 \). In Section 12.7, I present some incomplete characterization of the dual of \( C^k(X, Y) \). Then, I present the specialization of the Change of Variable Theorem to 2-dimensional and 3-dimensional common domain transformations, and 1-dimensional general domain transformation. I also present a result that states the equivalence of Riemann-Stieltjes and Lebesgue-Stieltjes integrals for bounded measurable integrand and monotone integrator. The surface measure on the unit sphere in \( \mathbb{R}^m \) is defined. In the last section of the chapter, I present a new definition of Sobolev space that agrees with the classical definition (Zeidler, 1995) in 1-dimensional case. My understanding here is that for functions with generalized derivatives, one must look at absolutely continuous functions. These functions can be expressed as sum of integrals of the stream functions with respect to some point \( x_0 \) in the domain exactly. If we defined norm of the absolutely continuous function to include the \( L_p \) norm of the function as well as of its stream functions, then it allows proof that the space is a Banach space. It is a space of (absolutely) continuous functions with all integral norms that is complete. When \( p = 2 \), this space is a Hilbert space with integral inner products. It has the smell of heaven that all these smooth functions, projection operators, Fréchet derivatives, and eigenvectors and eigenvalues. Applicable tools abound in this space. The question is how relevant is this space to Mathematical Physics formulation. We will see.

In Chapter 13, I present key results that are specific to Hilbert spaces. Pre-Hilbert space is a vector space together with the binary operation inner
product that induces the norm on the space. A Hilbert space is a pre-Hilbert space that is further complete. So, it is a special kind of Banach space. All of the results for Banach spaces are valid in Hilbert spaces. The special feature of Hilbert space is that the Projection Theorem 13.13 holds. The dual of a Hilbert space is easily characterized in Riesz-Fréchet Theorem 13.15, and a Hilbert space is reflexive. The dual is isometrically isomorphic to itself if the underlying field is real. When the underlying field is complex, then the dual is isometric to the Hilbert space. There exists a conjugate linear mapping from the Hilbert space to its dual, which is further invertible. With this mapping, we can define two vectors in the Hilbert space to be orthogonal if one of the vector is orthogonal (as defined for Banach space and its dual) to the bounded linear functional that is the image of the other vector under the conjugate map. In Section 13.4, I define the Hermitian adjoint of a bounded linear operator between Hilbert spaces, and obtain its properties. In Section 13.5, I present the Gram-Schmidt procedure and other results related to orthonormal sequences in Hilbert spaces. The Legendre polynomials is presented in Example 13.34. The $L_2$ theory for Fourier series is presented in Example 13.36. The Projection Theorem is revisited in Section 13.6, and I present the minimum norm problem for convex sets. Then, I define positive-definite and positive semi-definite operators among Hermitian operators on a Hilbert space, and derive their properties in Section 13.7. Section 13.8 is devoted to the study of projection operators on Hilbert space, and the pseudoinverse of bounded linear operators between Hilbert spaces. The last section of this chapter, Section 13.9, presents the study of eigenvector and eigenvalues for a bounded linear operator on a Hilbert space. In general, eigenvector or eigenvalue may not exist. But, once we restrict ourselves to compact Hermitian linear operators, then this concept is very useful, and generalizes common understandings for symmetric matrices. In Spectral Theory, Theorem 13.52, the singular value decomposition of a compact linear operator between Hilbert spaces is obtained and the eigenstructure of a compact Hermitian linear operator is obtained. This topic is motivated by Zeidler (1995) and the well-known theory for matrices. It is shown in Example 13.53 that linear integral operators with continuous kernel is a compact operator between $L_2$ spaces. In Proposition 13.55, certain matrix like properties is established for compact linear operators. For compact Hermitian linear operators, Fredholm Alternative Proposition 13.56 establishes certain matrix like properties for these operators. In Example 13.57, I establish that first order Sobolev space $W_{2,1,\mathcal{E}}(X, Y)$ is a Hilbert space when $X$ is a $\sigma$-finite metric measure subspace of $\mathbb{R}^m$ and $Y$ is separable Hilbert space with $Y^*$ being separable.

In Chapter 14, I studied the probability theory. These results are inspired mainly by Williams (1991). The fundamental notions in probability theory is introduced in Section 14.1: random variable, Banach space valued random variable, the expectation, independence of $\sigma$-algebras, and independence of random variables, conditional expectation, the law of a
random variable, and its probability density function. The conditional expectation of a random variable given a sub-$\sigma$-algebra, is defined to be the Radon-Nikodym derivative of appropriated defined measures (see Definition 14.7). The major properties are listed in Proposition 14.11, which is motivated by Williams (1991). I then prove the Fundamental Theorem of Modeling, Theorem 14.13, which allows all probability problems to be set up as measure theoretic problems. After a section on Gaussian random variables and vectors (Section 14.2), I present the strong Law of Large Numbers, Theorem 14.23, in Section 14.3. Then, I directly turn to the study of Martingale theory in Section 14.4. I prove the Doob’s Optional Stopping Theorem 14.32, Doob’s Upcrossing Lemma 14.35, and Doob’s Forward Convergence Theorem 14.37. I then turn in Section 14.5 to study Banach space valued Martingales. I present the generalized Doob’s Optional Stopping Theorem 14.41, Kolmogorov 0-1 Law, Theorem 14.43, and some results on the tail $\sigma$-algebra Proposition 14.44. Then, I present results (especially Lévy’s Inversion Formula Theorem 14.48) related to the characteristic function for a random variable in Section 14.6. In Section 14.7, I discuss the mode of convergence, and formally apply the Definition 11.210 in the probability measure space Definition 14.49. Then, the Modes of Convergence Theorem 14.50 follows readily from our definition. We introduce the concept of uniformly growth boundedness in Definition 14.51, which is also known as the tightness condition in Williams (1991). The convergence in distribution notion we defined is weaker than that of Williams (1991). Proposition 14.53 then gives the conditions to bridge the classical notion of convergence in distribution and our notion of convergence in distribution. The uniformly growth bounded condition plays key role in this result, which is dependent on Skorokhod Representation Theorem 14.52. In Section 14.8, I proceed to present the Central Limit Theorem 14.56, which are dependent on Helly’s Lemma 14.54 and Lévy’s Convergence Theorem 14.55. Results on the characteristic function and its inversion formula have direct application to Fourier transform. In Section 14.9, I present some results on uniform integrable Martingales. One of my goals in this mathematics study is to establish the existence of Wiener process so as to provide a solid foundation for my past control research. It was sad that I was not able to do this before late July 2018, when I opened my MATH 452 notes of Professor Burkholder of University of Illinois at Urbana-Champaign and found that the proof for Brownian motion is right there, which was not there the last time I checked the notes. So, finally, God is opening the door to probability theory for me. I was going to finish the book here and publish it by starting to write this introduction section. Now, I am continuing the research on Wiener processes and will not end until I fully prove Girsonov’s Theorem...

(To be continued.)
Chapter 2

Set Theory

2.1 Axiomatic Foundations of Set Theory

We will list the nine axioms of ZFC axiom system. The ninth axiom, which is the Axiom of Choice, will be introduced in Section 2.7. Let $A$ and $B$ be sets and $x$ and $y$ be objects (which is another name for sets).

Axiom 1 (Axiom of Extensionality) $A = B$ if $\forall x \in A$, we have $x \in B$; and $\forall x \in B$, we have $x \in A$.

Axiom 2 (Axiom of Empty Set) There exists an empty set $\emptyset$, which does not contain any element.

Axiom 3 (Axiom of Pairing) For any objects $x$ and $y$, there exists a set $\{x, y\}$, which contains only $x$ and $y$.

Axiom 4 (Axiom of Regularity) Any nonempty set $A \neq \emptyset$, there exists $a \in A$, such that $\forall b \in A$, we have $b \notin a$.

Axiom 5 (Axiom of Replacement) $\forall x \in A$, let there be one and only one $y$ to form an ordered pair $(x, y)$. Then, the collection of all such $y$’s is a set $B$.

Axiom 6 (Axiom of Power Set) The collection of all subsets of $A$ is a set denoted by $^2A$.

Axiom 7 (Axiom of Union) For any collection of sets $(A_\lambda)_{\lambda \in \Lambda}$, where $\Lambda$ is a set, then $\bigcup_{\lambda \in \Lambda} A_\lambda$ is a well defined set.

Axiom 8 (Axiom of Infinity) There exists a set $A$ such that $\emptyset \in A$ and $\forall x \in A$, we have $\{\emptyset, x\} \in A$. 
By Axiom 2, there exists the empty set \(\emptyset\), which we may call 0. Now, by Axiom 3, there exists the set \(\{\emptyset\}\), which is nonempty and we may call 1. Again, by Axiom 3, there exists the set \(\{\emptyset, \{\emptyset\}\}\), which we will call 2. After we define \(n\), we may define \(n + 1 := \{0, n\}\), which exists by Axiom of Pairing. This allows us to define all natural numbers. By Axiom 8, these natural numbers can form the set, \(\mathbb{N} := \{1, 2, \ldots\}\), which is the set of natural numbers. Furthermore, by Axiom 6, we may define the set of all real numbers, \(\mathbb{R}\).

For any \(x \in A\) and \(y \in B\), we may apply Axiom 3 to define the ordered pair \((x, y) := \{\{\{x\}\}, \{\{y\}\}, 2\}\). Then, the set \(A \times B\) is defined by \(\bigcup_{x \in A} \bigcup_{y \in B} \{(x, y)\}\), which is a valid set by Axiom 7. By Axiom 5, any portion \(A_x\) of a well-defined set \(A\) is again a set, which is called a subset of \(A\), we will write \(A_x \subseteq A\). Thus, the formula

\[
\{ x \in A \mid p(x) \text{ is true.} \}
\]
defines a set as long as \(A\) is a set and \(p(x)\) is unambiguous logic expression.

### 2.2 Relations and Equivalence

**Definition 2.1** Let \(A\) and \(B\) be sets. A relation \(R\) from \(A\) to \(B\) is a subset of \(A \times B\). \(\forall x \in A, \forall y \in B, \text{ we say } x \sim y \text{ if } (x, y) \in R.\) We will say that \(R\) is a relation on \(A\) if it is a relation from \(A\) to \(A\). We define

\[
\begin{align*}
\text{dom}(R) & := \{ x \in A \mid \exists y \in B, \text{ such that } x \sim y \} \\
\text{range}(R) & := \{ y \in B \mid \exists x \in A, \text{ such that } x \sim y \}
\end{align*}
\]

which are well-defined subsets.

**Definition 2.2** Let \(A\) be a set and \(R\) be a relation on \(A\). \(\forall x, y, z \in A,\)

1. \(R\) is reflexive if \(x \sim x\).
2. \(R\) is symmetric if \(x \sim y\) implies \(y \sim x\).
3. \(R\) is transitive if \(x \sim y\) and \(y \sim z\) implies \(x \sim z\).
4. \(R\) is a equivalence relationship if it is reflexive, symmetric, and transitive, and will be denote by "\(\equiv\)".
5. \(R\) is antisymmetric if \(x \sim y\) and \(y \sim x\) implies \(x = y\).

Let \(\equiv\) be an equivalence relationship on \(A\), then, it partitions \(A\) into disjoint equivalence classes \(A_x := \{ y \in A \mid x \equiv y \}, \forall x \in A.\) The collection of all equivalence classes, \(A/ \equiv := \{ A_x \subseteq A \mid x \in A \}\), is called the quotient of \(A\) with respect to \(\equiv\).
2.3 Function

**Definition 2.3** Let $X$ and $Y$ be sets and $D \subseteq X$. A function $f$ of $D$ to $Y$, denoted by $f : D \to Y$, is a relation from $D$ to $Y$ such that $\forall x \in D$, there is exactly one $y \in Y$, such that $(x, y) \in f$; we will denote that $y$ as $f(x)$. The graph of $f$ is the set graph $(f) := \{(x, f(x)) \in X \times Y \mid x \in D\}$. The domain of $f$ is $\text{dom}(f) = D$. $\forall A \subseteq X$, the image under $f$ of $A$ is $f(A) := \{y \in Y \mid \exists x \in A \cap D \text{ such that } f(x) = y\}$, which is a subset of $Y$. The range of $f$ is $\text{range}(f) = f(D) = \{y \in Y \mid \exists x \in D \text{ such that } f(x) = y\}$, which is a subset of $Y$. The collection of all functions of $X$ is $\mathcal{P}(X)$.

Sufficiency Let $\phi : X \to Y$; then $\exists \psi : Y \to X$, $\forall x \in X$, $\phi \circ \psi(y) = y$, which implies that $y \in \text{range}(\phi)$, and hence, $\phi$ is surjective. Suppose that $\phi$ is not injective, then $\exists x_1, x_2 \in X$ with $x_1 \neq x_2$ such that $\phi(x_1) = \phi(x_2)$. Then, we have $x_1 = \text{id}_X(x_1) = \psi_2(\phi(x_1)) = \psi_2(\phi(x_2)) = \text{id}_X(x_2) = x_2$ which is a contradiction. Hence, $\phi$ is injective. This proves that $\phi$ is bijective.

Necessity Let $\phi$ be bijective. Then, $\phi^{-1} : X \to X$ exists. $\forall x \in X$, let $y = \phi(x)$, then $x = \phi^{-1}(y)$, hence $\phi^{-1}(\phi(x)) = x$. Therefore, we have $\phi^{-1} \circ \phi = \text{id}_X$. $\forall y \in Y$, let $x = \phi^{-1}(y)$, then $y = \phi(x)$, hence $\phi(\phi^{-1}(y)) = y$. Therefore, we have $\phi \circ \phi^{-1} = \text{id}_Y$. Hence, $\psi_1 = \psi_2 = \phi^{-1}$. The inverse function is denoted by $\phi^{-1} : Y \to X$. We will say that $f$ is a function from $X$ to $Y$.

Let $f : D \to Y$ and $g : Y \to Z$ be functions, we may define a function $h : D \to Z$ by $h(x) = g(f(x))$, then $h$ is called the composition of $g$ with $f$, and is denoted by $g \circ f$. Let $A \subseteq X$. We may define a function $l : A \cap D \to Y$ by $l(x) = f(x)$, $\forall x \in A \cap D$. This function is called the restriction of $f$ to $A$, and denoted by $f_{|A}$. Let $f : D \to Y$, $g : Y \to Z$, and $h : Z \to W$, we have $(h \circ g) \circ f = h \circ (g \circ f)$. Let $f : D \to D$ and $k \in \mathbb{Z}_+$, we will write $f^k := f \circ \cdots \circ f$, where $f^0 := \text{id}_D$.

A function $f : X \to Y$ is a subset of $X \times Y$. Then, $f \in X^Y$. The collection of all functions of $X$ to $Y$ is then a set given by

$$ Y^X := \{ f \in X^Y \mid \forall x \in X, \exists! y \in Y \ni (x, y) \in f \} $$

We have the following result concerning the inverse of a function.

**Proposition 2.4** Let $\phi : X \to Y$, where $X$ and $Y$ are sets. Then, $\phi$ is bijective if, and only if, $\exists \psi_1 : Y \to X$, $i = 1, 2$, such that $\phi \circ \psi_1 = \text{id}_Y$ and $\psi_2 \circ \phi = \text{id}_X$. Furthermore, $\phi^{-1} = \psi_1 = \psi_2$.

**Proof**

**Sufficiency** Let $\psi_i : Y \to X$, $i = 1, 2$, exist. $\forall y \in Y$, $\phi \circ \psi_1(y) = y$, which implies that $y \in \text{range}(\phi)$, and hence, $\phi$ is surjective. Suppose that $\phi$ is not injective, then $\exists x_1, x_2 \in X$ with $x_1 \neq x_2$ such that $\phi(x_1) = \phi(x_2)$. Then, we have $x_1 = \text{id}_X(x_1) = \psi_2(\phi(x_1)) = \psi_2(\phi(x_2)) = \text{id}_X(x_2) = x_2$ which is a contradiction. Hence, $\phi$ is injective. This proves that $\phi$ is bijective.

**Necessity** Let $\phi$ be bijective. Then, $\phi^{-1} : Y \to X$ exists. $\forall x \in X$, let $y = \phi(x)$, then $x = \phi^{-1}(y)$, hence $\phi^{-1}(\phi(x)) = x$. Therefore, we have $\phi^{-1} \circ \phi = \text{id}_X$. $\forall y \in Y$, let $x = \phi^{-1}(y)$, then $y = \phi(x)$, hence $\phi(\phi^{-1}(y)) = y$. Therefore, we have $\phi \circ \phi^{-1} = \text{id}_Y$. Hence, $\psi_1 = \psi_2 = \phi^{-1}$. The inverse function is denoted by $\phi^{-1} : Y \to X$. We will say that $f$ is a function from $X$ to $Y$.
Let $\psi_1$ and $\psi_2$ satisfy the assumption of the proposition, and $\phi_{\text{inv}}$ be the inverse function of $\phi$. Then, we have
\[
\psi_1 = \text{id}_X \circ \psi_1 = (\phi_{\text{inv}} \circ \phi) \circ \psi_1 = \phi_{\text{inv}} \circ (\phi \circ \psi_1) = \phi_{\text{inv}} \circ \text{id}_Y = \phi_{\text{inv}} \\
\psi_2 = \psi_2 \circ \text{id}_Y = \psi_2 \circ (\phi \circ \phi_{\text{inv}}) = (\psi_2 \circ \phi) \circ \phi_{\text{inv}} = \text{id}_X \circ \phi_{\text{inv}} = \phi_{\text{inv}}
\]

This completes the proof of the proposition.

For bijective functions $f : X \to Y$ and $g : Y \to Z$, $g \circ f$ is also bijective and $(g \circ f)_{\text{inv}} = f_{\text{inv}} \circ g_{\text{inv}}$.

2.4 Set Operations

Let $X$ be a set, $X^2$ is the set consisting of all subsets of $X$. $\forall A, B \subseteq X$, we will define
\[
A \cup B := \{ x \in X \mid x \in A \text{ or } x \in B \} \\
A \cap B := \{ x \in X \mid x \in A \text{ and } x \in B \} \\
\tilde{A} := \{ x \in X \mid x \notin A \} \\
A \setminus B := \{ x \in A \mid x \notin B \} = A \cap \tilde{B} \\
A \triangle B := (A \setminus B) \cup (B \setminus A)
\]

We have the following results.

**Proposition 2.5** Let $A, B, D, A_\lambda \in X^2, f : D \to Y, C, E, C_\lambda \in Y^2$, where $X$ and $Y$ are sets, $\lambda \in \Lambda$, and $\Lambda$ is an index set. Then, we have

1. $A \cup B = B \cup A$ and $A \cap B = B \cap A$;
2. $A \subseteq A \cup B$ and $A = A \cup B$ if, and only if, $B \subseteq A$;
3. $A \cup \emptyset = A, A \cap \emptyset = \emptyset, A \cup X = X$, and $A \cap X = A$;
4. $\emptyset = X, \tilde{A} = A, A \cup \tilde{A} = X, A \cap \tilde{A} = \emptyset$, and $A \subseteq B$ if, and only if, $\tilde{B} \subseteq \tilde{A}$;
5. The De Morgan’s Laws:
\[
(\bigcup_{\lambda \in \Lambda} A_\lambda)^\sim = \bigcap_{\lambda \in \Lambda} \tilde{A}_\lambda; \quad (\bigcap_{\lambda \in \Lambda} A_\lambda)^\sim = \bigcup_{\lambda \in \Lambda} \tilde{A}_\lambda
\]
6. $B \cup \left( \bigcap_{\lambda \in \Lambda} A_\lambda \right) = \bigcap_{\lambda \in \Lambda} (B \cup A_\lambda)$ and $B \cap \left( \bigcup_{\lambda \in \Lambda} A_\lambda \right) = \bigcup_{\lambda \in \Lambda} (B \cap A_\lambda)$;
7. $f \left( \bigcup_{\lambda \in \Lambda} A_\lambda \right) = \bigcup_{\lambda \in \Lambda} f(A_\lambda)$ and $f \left( \bigcap_{\lambda \in \Lambda} A_\lambda \right) \subseteq \bigcap_{\lambda \in \Lambda} f(A_\lambda)$;
2.5. **ALGEBRA OF SETS**

8. $f_{\text{inv}} \left( \bigcup_{\lambda \in \Lambda} C_\lambda \right) = \bigcup_{\lambda \in \Lambda} f_{\text{inv}}(C_\lambda)$ and $f_{\text{inv}} \left( \bigcap_{\lambda \in \Lambda} C_\lambda \right) = \bigcap_{\lambda \in \Lambda} f_{\text{inv}}(C_\lambda)$;

9. $f_{\text{inv}}(C \setminus E) = f_{\text{inv}}(C) \setminus f_{\text{inv}}(E)$, $f(f_{\text{inv}}(C)) = C \cap \text{range} \,(f)$, and $f_{\text{inv}}(f(A)) \supseteq A \cap \text{dom} \,(f) = A \cap D$.

The proof of the above results are standard and is therefore omitted.

### 2.5 Algebra of Sets

**Definition 2.6** A set $X$ is said to be finite if it is either empty or the range of a function of $\{1, 2, \ldots, n\}$, with $n \in \mathbb{N}$. In this case, $\text{card} \,(X)$ denotes the number of elements in $X$. It is said to be countable if it is either empty or the range of a function of $\mathbb{N}$.

**Definition 2.7** Let $X$ be a set and $A \subseteq X^2$. $A$ is said to be an algebra of sets on $X$ (or a Boolean algebra on $X$) if

(i) $\emptyset, X \in A$;

(ii) $\forall A, B \in A$, $A \cup B \in A$ and $\sim A \in A$.

$A$ is said to be a $\sigma$-algebra on $X$ if it is an algebra on $X$ and countable unions of sets in $A$ is again in $A$.

Let $M \subseteq X^2$, where $X$ is a set, then, there exists a smallest algebra on $X$, $A_0 \subseteq X^2$, containing $M$, which means that any algebra on $X$, $A_1 \subseteq X^2$, that contains $M$, we have $A_0 \subseteq A_1$. This algebra $A_0$ is said to be the algebra on $X$ generated by $M$. Also, there exists a smallest $\sigma$-algebra on $X$, $A \subseteq X^2$, containing $M$, which is said to be the $\sigma$-algebra on $X$ generated by $M$.

**Proposition 2.8** Let $X$ be a set, $E$ be a nonempty collection of subsets of $X$, $A$ be the algebra on $X$ generated by $E$, and $\bar{A}$ be the algebra on $X$ generated by $E$, and

$$
\bar{A} := \left\{ A \subseteq X \mid \exists n, m \in \mathbb{N}, \forall i_1, \ldots, i_{2n} \in \{1, \ldots, m\}, \exists F_{i_1, \ldots, i_{2n}} \subseteq X \right. \\
\left. \text{with } F_{i_1, \ldots, i_{2n}} \in E \text{ or } (F_{i_1, \ldots, i_{2n}})' \in E, \text{ such that } \right. \\
A = \bigcup_{i_1=1}^{m} \bigcap_{i_2=1}^{m} \cdots \bigcup_{i_{2n}=1}^{m} F_{i_1, \ldots, i_{2n}} \right\}
$$

Then, $A = \bar{A}$.

**Proof** $\forall E \in E$, let $n = 1$, $m = 1$, and $F_{1,1} = E$. Then, $E = \bigcup_{i_1=1}^{1} \bigcap_{i_2=1}^{1} F_{i_1, i_2} \in \bar{A}$. Hence, we have $E \subseteq \bar{A}$. It is clear that $\bar{A} \subseteq A$. All we need to show is that $A$ is an algebra on $X$. Then, $A \subseteq \bar{A}$ and the result follows.
Fix $E \in \mathcal{E} \neq \emptyset$. $E \subseteq X$. Let $n = 1$, $m = 2$, $F_{1,1} = E$, $F_{1,2} = \bar{E}$, $F_{2,1} = E$, and $F_{2,2} = \bar{E}$. Then, $\emptyset = \bigcup_{i_1=1}^2 \bigcap_{i_2=1}^2 F_{i_1,i_2} \in \mathcal{A}$. Let $n = 1$, $m = 2$, $F_{1,1} = E$, $F_{1,2} = \bar{E}$, $F_{2,1} = E$, and $F_{2,2} = \bar{E}$. Then, $X = \bigcup_{i_1=1}^2 \bigcap_{i_2=1}^2 F_{i_1,i_2} \in \mathcal{A}$.

$\forall A, B \in \mathcal{A}, \exists n_A, m_A \in \mathbb{N}, \forall i_1, \ldots, i_{2n_A} \in \{1, \ldots, m_A\}, \exists F_{i_1, \ldots, i_{2n_A}} \subseteq X$ with $F_{i_1, \ldots, i_{2n_A}} \in \mathcal{E}$ or $(F_{i_1, \ldots, i_{2n_A}})^{\sim} \in \mathcal{E}$ such that $A = \bigcup_{i_1=1}^{m_A} \bigcap_{i_2=1}^{m_A} \ldots \bigcap_{i_{2n_A}=1}^{m_A} F_{i_1, \ldots, i_{2n_A}}$, and $\exists n_B, m_B \in \mathbb{N}, \forall i_1, \ldots, i_{2n_B} \in \{1, \ldots, m_B\}, \exists F_{i_1, \ldots, i_{2n_B}} \subseteq X$ with $F_{i_1, \ldots, i_{2n_B}} \in \mathcal{E}$ or $(F_{i_1, \ldots, i_{2n_B}})^{\sim} \in \mathcal{E}$ such that $B = \bigcup_{i_1=1}^{m_B} \bigcap_{i_2=1}^{m_B} \ldots \bigcap_{i_{2n_B}=1}^{m_B} F_{i_1, \ldots, i_{2n_B}}$.

Note that $\bar{A} = \bigcap_{i_1=1}^{m_A} \bigcap_{i_2=1}^{m_A} \ldots \bigcap_{i_{2n_A}=1}^{m_A} (F_{i_1, \ldots, i_{2n_A}})^{\sim}$. Let $n = n_A + 1$, $m = m_A, \forall i_1, \ldots, i_{2n} \in \{1, \ldots, m\}, G_{i_1, \ldots, i_{2n}} = (F_{i_1, \ldots, i_{2n-1}})^{\sim}$. Then, $\bar{A} = \bigcup_{i_1=1}^{m} \bigcap_{i_2=1}^{m} \ldots \bigcap_{i_{2n}=1}^{m} G_{i_1, \ldots, i_{2n}} \in \mathcal{A}$.

Without loss of generality, assume $n_A \geq n_B$. Let $n = n_A$ and $m = m_A + m_B$. Define $i = 1 + (i \mod m_A)$ and $i = 1 + (i \mod m_B)$, $\forall i \in \mathbb{N}$, $\forall i_1, \ldots, i_{2n} \in \{1, \ldots, m\}$, let

$$G_{i_1, \ldots, i_{2n}} = \begin{cases} F_{i_1, i_2, \ldots, i_{2n}} & \text{if } i_1 \leq m_A \\ F_{i_{m_A+1}, i_2, \ldots, i_{2n}} & \text{if } i_1 > m_A \end{cases}$$

Then, it is easy to check that $A \cup B = \bigcup_{i_1=1}^{m} \bigcap_{i_2=1}^{m} \ldots \bigcap_{i_{2n}=1}^{m} G_{i_1, \ldots, i_{2n}} \in \mathcal{A}$. Hence, $\mathcal{A}$ is an algebra on $X$.

This completes the proof of the proposition. \hfill $\square$

### 2.6 Partial Ordering and Total Ordering

**Definition 2.9** Let $A$ be a set and $\preceq$ be a relation on $A$. $\preceq$ will be called a partial ordering if it is reflexive and transitive. It will be called a total ordering if it is an antisymmetric partial ordering and satisfies $\forall x, y \in A$ with $x \neq y$, we have either $x \preceq y$ or $y \preceq x$ (not both).

As an example, the set containment $\subset$ is a partial ordering on any collection of sets; while $\subseteq$ is a total ordering on any subset of $\mathbb{R}$.

**Definition 2.10** Let $A$ be a set with a partial ordering $\preceq$.

1. $a \in A$ is said to be minimal if, $\forall x \in A$, $x \preceq a$ implies $a \preceq x$;

2. $a \in A$ is said to be the least element if, $\forall x \in A$, $a \preceq x$, and $x \preceq a$ implies that $x = a$.

3. $a \in A$ is said to be maximal if, $\forall x \in A$, $a \preceq x$ implies $x \preceq a$;
2.6. PARTIAL ORDERING AND TOTAL ORDERING

4. \( a \in A \) is said to be the greatest element if, \( \forall x \in A, x \leq a, \) and \( a \leq x \) implies that \( x = a. \)

**Definition 2.11** Let \( A \) be a set with a partial ordering \( \preceq \), and \( E \subseteq A. \)

1. \( a \in A \) is said to be an upper bound of \( E \) if \( x \preceq a, \forall x \in E. \) It is the least upper bound of \( E \) if it is the least element in the set of all upper bounds of \( E; \)

2. \( a \in A \) is said to be a lower bound of \( E \) if \( a \preceq x, \forall x \in E. \) It is the greatest lower bound of \( E \) if it is the greatest element in the set of all lower bounds of \( E. \)

We have the following results.

**Proposition 2.12** Let \( A \) be a set with a partial ordering \( \preceq. \) Then, the following holds.

(i) If \( a \in A \) is the least element, then it is minimal.

(ii) There is at most one least element in \( A. \)

(iii) Define a relation \( \succeq \) by \( \forall x, y \in A, x \succeq y \) if \( y \preceq x. \) Then, \( \succeq \) is a partial ordering on \( A. \) Furthermore, \( \succeq \) is antisymmetric if \( \preceq \) is antisymmetric.

1. \( a \in A \) is the least element for \((E, \preceq)\) if, and only if, it is the greatest element for \((E, \succeq)\).

2. \( a \in A \) is minimal for \((E, \preceq)\) if, and only if, it is maximal for \((E, \succeq)\).

(iv) If \( a \in A \) is the greatest element, then it is maximal.

(v) There is at most one greatest element in \( A. \)

(vi) If \( \preceq \) is antisymmetric, then \( a \in A \) is minimal if, and only if, there does not exist \( x \in A \) such that \( x \leq a \) and \( x \neq a. \)

(vii) If \( \preceq \) is antisymmetric, then \( a \in A \) is maximal if, and only if, there does not exist \( x \in A \) such that \( a \leq x \) and \( x \neq a. \)

(viii) If \( \preceq \) is antisymmetric, then it is a total ordering if, and only if, \( \forall x_1, x_2 \in A, \) we have \( x_1 \preceq x_2 \) or \( x_2 \preceq x_1. \)

**Proof**

(i) is straightforward from Definition 2.10.

For (ii), let \( a_1 \) and \( a_2 \) be least elements of \( A. \) By \( a_1 \) being the least element, we have \( a_1 \preceq a_2. \) By \( a_2 \) being the least element, we then have \( a_1 = a_2. \) Hence, the least element is unique if it exists.

For (iii), \( \forall x, y, z \in A. \) Since \( x \preceq x \) implies \( x \succeq x, \) then \( \succeq \) is reflexive. If \( x \succeq y \) and \( y \succeq z, \) we have \( y \preceq x \) and \( z \preceq y, \) which implies \( z \preceq x, \) and hence
$x \succeq z$. This shows that $\succeq$ is transitive. Hence, $\succeq$ is a partial ordering on $A$.

When $\preceq$ is antisymmetric, $x \succeq y$ and $y \succeq x$ implies that $x \preceq y$ and $y \preceq x$, and therefore $x = y$. Hence, $\succeq$ is also antisymmetric.

For 1, “only if” let $a \in A$ be the least element in $(E, \preceq)$. $a \preceq x$ implies $x \preceq a$, $\forall x \in A$. $\forall x \in A$ with $a \preceq x$, we have $x \preceq a$, by $a$ being the least element in $(E, \preceq)$, we have $x = a$. Hence, $a$ is the greatest element in $(E, \preceq)$. The “if” part is similar to the “only if” part.

For 2, “only if” let $a \in A$ be a minimal element for $(E, \preceq)$. Then, $\forall x \in A$ with $a \preceq x$ implies $x \preceq a$ and hence $a \preceq x$, which yields $x \preceq a$. Hence, $a$ is a maximal element for $(E, \succeq)$. The “if” part is similar to the “only if” part.

(iv) is straightforward from Definition 2.10.

For (v), Let $a_1$ and $a_2$ be greatest elements of $A$. By $a_1$ being the greatest element, we have $a_2 \preceq a_1$. By $a_2$ being the greatest element, we then have $a_1 = a_2$. Hence, the greatest element is unique if it exists.

For (vi), “if”, $\forall x \in A$ with $x \preceq a$, then, we have $a = x$, which means that $a \preceq x$; hence $a$ is minimal. “Only if”, suppose that $\exists x \in A$ such that $x \preceq a$ and $x \neq a$. Note that $a \preceq x$ since $a$ is minimal. Then, $a = x$, since $\preceq$ is antisymmetric, which is a contradiction.

For (vii), “if”, $\forall x \in A$ with $a \preceq x$, then, we have $a = x$, which means that $x \preceq a$; hence $a$ is maximal. “Only if”, suppose that $\exists x \in A$ such that $a \preceq x$ and $x \neq a$. Note that $x \preceq a$ since $a$ is maximal. Then, $a = x$, since $\preceq$ is antisymmetric, which is a contradiction.

For (viii), “if”, $\forall x_1, x_2 \in A$ with $x_1 \neq x_2$, we must have $x_1 \preceq x_2$ or $x_2 \preceq x_1$. They can not hold at the same time since, otherwise, $x_1 = x_2$, which is a contradiction. “Only if”, $\forall x_1, x_2 \in A$, when $x_1 = x_2$, then $x_1 \preceq x_2$; when $x_1 \neq x_2$, then $x_1 \preceq x_2$ or $x_2 \preceq x_1$; hence, in both cases, we have $x_1 \preceq x_2$ or $x_2 \preceq x_1$.

This completes the proof of the proposition. □

2.7 Basic Principles

Now, we introduce the last axiom in ZFC axiom system.

Axiom 9 (Axiom of Choice) Let $(A_\lambda)_{\lambda \in \Lambda}$ be a collection of nonempty sets, $\Lambda$ is a set, (this collection is a set by Axiom 5), then, there exists a function $f : \Lambda \to \bigcup_{\lambda \in \Lambda} A_\lambda$, such that, $\forall \lambda \in \Lambda$, we have $f(\lambda) \in A_\lambda$.

With Axioms 1–8 holding, the Axiom of Choice is equivalent to the following three results.

Theorem 2.13 (Hausdorff Maximal Principle) Let $\preceq$ be a partial ordering on a set $E$. Then, there exists a maximal (with respect to set containment $\subseteq$) subset $F \subseteq E$, such that $\preceq$ is a total ordering on $F$. 
Theorem 2.14 (Zorn’s Lemma) Let \( \preceq \) be an antisymmetric partial ordering on a nonempty set \( E \). If every nonempty totally order subset \( F \) of \( E \) has an upper bound in \( E \), then, there is a maximal element in \( E \).

Definition 2.15 A well ordering of a set is a total ordering such that every nonempty subset has a least element.

Theorem 2.16 (Well-Ordering Principle) Every set can be well ordered.

To prove the equivalence we described above, we need the following result.

Lemma 2.17 Let \( E \) be a nonempty set and \( \preceq \) is an antisymmetric partial ordering on \( E \). Assume that every nonempty subset \( S \) of \( E \), on which \( \preceq \) is a total ordering, has a least upper bound in \( E \). Let \( f : E \rightarrow E \) be a mapping such that \( x \preceq f(x), \forall x \in E \). Then, \( f \) has a fixed point on \( E \), i.e., \( \exists w \in E, f(w) = w \).

Proof Fix a point \( a \in E \), since \( E \neq \emptyset \). We define a collection of “good” sets:

\[
B = \{ B \subseteq E \mid (i) a \in B \quad (ii) f(B) \subseteq B \quad (iii) \forall F \subseteq B, F \neq \emptyset \}
\]

\( F \) is totally ordered with \( \preceq \) implies that the least upper bound of \( F \) belongs to \( B \).

Consider the set

\[
B_0 := \{ x \in E \mid a \preceq x \}
\]

Clearly, \( B_0 \) is nonempty since \( a \in B_0 \) and

\[
f(B_0) = \{ f(x) \in E \mid a \preceq x \preceq f(x) \} \subseteq B_0
\]

since \( f \) satisfies \( x \preceq f(x), \forall x \in E \).

For any \( F \subseteq B_0 \), such that \( F \) is totally ordered with \( \preceq \) and \( F \neq \emptyset \). Let \( e_0 \) be the least upper bound of \( F \) in \( E \). Then, \( \exists x_0 \in F \) such that \( a \preceq x_0 \preceq e_0 \), and therefore \( e_0 \in B_0 \).

This shows that \( B_0 \in B \), and \( B \) is nonempty.

The following result holds for the collection \( B \).

Claim 2.17.1 Let \( \{ B_\alpha \mid \alpha \in \Lambda \} \) be any nonempty subcollection of \( B \), then \( \bigcap_{\alpha \in \Lambda} B_\alpha \in B \).

Proof of claim:

(i) \( a \in B_\alpha, \forall \alpha \in \Lambda \). This implies \( a \in \bigcap_{\alpha \in \Lambda} B_\alpha \).

(ii) By Proposition 2.5, we have \( f(\bigcap_{\alpha \in \Lambda} B_\alpha) \subseteq \bigcap_{\alpha \in \Lambda} f(B_\alpha) \subseteq \bigcap_{\alpha \in \Lambda} B_\alpha \), where the last \( \subseteq \) follows from the fact \( f(B_\alpha) \subseteq B_\alpha, \forall \alpha \in \Lambda \).

(iii) Let \( F \subseteq \bigcap_{\alpha \in \Lambda} B_\alpha \), which is totally ordered by \( \preceq \) and \( F \neq \emptyset \).
For any $\alpha \in \Lambda$, $F \subseteq B_\alpha$ implies that the least upper bound of $F$ is an element of $B_\alpha$. Therefore, the least upper bound of $F$ is in the intersection $\bigcap_{\alpha \in \Lambda} B_\alpha$.

This establishes $\bigcap_{\alpha \in \Lambda} B_\alpha \in \mathcal{B}$, and completes the proof of the claim. \qed

The claim shows that the collection $\mathcal{B}$ is closed under arbitrary intersection, as long as the collection is nonempty. Define $A := \bigcap_{B \in \mathcal{B}} B$. By the above claim, we have $A \in \mathcal{B}$, i.e., $A$ is the smallest set in $\mathcal{B}$.

Hence, $A \subseteq B_0$, i.e., the set $A$ satisfies, in addition to (i) – (iii), (iv) $\forall x \in A$, $a \preceq x$.

Define the relation $\prec$ on $E$ as $\forall x, y \in E$, $x \prec y$ if, and only if, $x \preceq y$ and $x \neq y$.

Define the set $P$ by

$$P = \{ x \in A \mid \forall y \in A, \ y \prec x \Rightarrow f(y) \preceq x \}$$

Clearly, $a \in P$, since there does not exist any $y \in A$ such that $y \prec a$, by $\preceq$ being antisymmetric. Therefore, $P$ is nonempty.

We claim that

Claim 2.17.2 (v) $\forall x \in P$, $\forall z \in A$, then $z \preceq x$ or $f(x) \preceq z$.

**Proof of claim:** Fix $x \in P$, and let

$$B := \{ z \in A \mid z \preceq x \} \cup \{ z \in A \mid f(x) \preceq z \}$$

We will show that $B \in \mathcal{B}$.

(i) $a \in A$, $x \in P \subseteq A$, by (iv), $a \preceq x$, which further implies that $a \in B$.

(ii) $\forall z \in B \subseteq A$, then $f(z) \in A$ since $A \in \mathcal{B}$. There are three exhaustive scenarios. If $z \preceq x$, since $x \in P$ and $z \in B \subseteq A$, then $f(z) \preceq x$. This implies that $f(z) \in B$. If $z = x$, then $f(x) \preceq f(x) = f(z)$. This implies that $f(z) \in B$. If $f(x) \preceq z$, then $f(x) \preceq z \preceq f(z)$. This again implies that $f(z) \in B$. Hence, in all three scenarios, we have $f(z) \in B$. Then, $f(B) \subseteq B$ by the arbitrariness of $z \in B$.

(iii) Let $F \neq \emptyset$ be any totally ordered subset of $B$ and $e_0 \in E$ be the least upper bound of $F$. Since $F \subseteq B \subseteq A$ and $A \in \mathcal{B}$, then $e_0 \in A$. There are two exhaustive scenarios. If there exists $y \in F$ such that $f(x) \preceq y$, then, $f(x) \preceq y \preceq e_0$. This implies $e_0 \in B$. If, for any $y \in F$, $y \preceq x$, then $F \subseteq \{ z \in A \mid z \preceq x \}$. This implies that $x$ is an upper bound of $F$ and $e_0 \preceq x$, since $e_0$ is the least upper bound of $F$. Therefore, $e_0 \in B$. In both of the cases, we have $e_0 \in B$.

This establishes that $B \in \mathcal{B}$. By $A$ being the smallest set in $\mathcal{B}$, we have $A = B$. Therefore, the claim is proven. \qed

Now, we show that $P \in \mathcal{B}$.

(i) $a \in P$ and therefore $P \neq \emptyset$.

(ii) Fix an $x \in P \subseteq A$. Then, $f(x) \in A$. $\forall y \in A$ such that $y \prec f(x)$. We need to show that $f(y) \preceq f(x)$, which then implies $f(x) \in P$. By (v),
there are two exhaustive scenarios. If \( y \leq x \), then \( y \leq x \). If \( f(x) \leq y \), then \( f(x) \leq y < f(x) \) form a contradiction by \( \leq \) being antisymmetric. Therefore, we must have \( y \leq x \), which results in the following two exhaustive scenarios. If \( y < x \), then \( f(y) \leq x \) since \( x \in P \). This implies that \( f(y) \leq x \leq f(x) \). If \( y = x \), then \( f(y) = f(x) \leq f(x) \). In both cases, we have \( f(y) \leq f(x) \). By the arbitraryness of \( y \), we have \( f(x) \in P \), which further implies \( f(P) \subseteq P \) by the arbitraryness of \( x \in P \).

(iii) Let \( F \neq \emptyset \) be a totally ordered subset in \( P \). Let \( e_0 \in E \) be the least upper bound of \( F \). We have \( F \subseteq A \) implies that \( e_0 \in A \) by \( A \in B \). \( \forall z \in A \) with \( z < e_0 \), implies that \( z \) must not be an upper bound of \( F \). Therefore, \( \exists x_0 \in F \) such that \( x_0 \notin z \). By (v), we have \( z < x_0 \). Hence, by \( x_0 \in F \subseteq P \), \( z \in A \), and \( z < x_0 \), we have \( f(z) \leq x_0 \). Therefore, \( f(z) \leq e_0 \) since \( e_0 \) is an upper bound of \( F \). This further implies that \( e_0 \in P \) by the arbitraryness of \( z \).

This proves that \( P \in B \).

Since \( P \subseteq A \) and \( A \) is the smallest set in \( B \), then, \( P = A \).

The set \( A \) satisfies properties (i) – (v).

For any \( x_1, x_2 \in A \), by (v), there are two exhaustive scenarios. If \( x_1 \leq x_2 \), then \( x_1 \) and \( x_2 \) are related through \( \leq \). If \( f(x_2) \preceq x_1 \), then \( x_2 \preceq f(x_2) \preceq x_1 \), which implies that \( x_1 \) and \( x_2 \) are related through \( \preceq \). Therefore, \( x_1 \) and \( x_2 \) are related through \( \preceq \) in both cases. Then, by Proposition 2.12 (viii), \( A \) is totally ordered by \( \leq \) and nonempty. Let \( w \in E \) be the least upper bound of \( A \). Then, \( w \in A \), since \( A \in B \).

Therefore, \( f(w) \in A \) by \( f(A) \subseteq A \), which implies that \( f(w) \leq w \). This coupled with \( w \leq f(w) \) yields \( f(w) = w \), since \( \leq \) is antisymmetric.

This completes the proof of the lemma.

\[ \square \]

**Theorem 2.18** Under the Axioms 1–8, the following are equivalent.

1. **Axiom of Choice**
2. **Hausdorff Maximum Principle**
3. **Zorn’s Lemma**
4. **Well-ordering principle**

**Proof**

1. \( \Rightarrow \) 2. Define

\[ E := \{ A \subseteq E \mid \preceq \text{ defines a total ordering on } A \} \]

Clearly, \( \emptyset \in E \), then \( E \neq \emptyset \). Define a partial ordering on \( E \) by \( \subseteq \), which is set containment. This partial ordering \( \subseteq \) is clearly reflexive, transitive, and antisymmetric.

\( \forall A \in E \), define a collection

\[ A_A := \begin{cases} \{ B \in E \mid A \subset B \} & \text{if } \exists B \in E \text{ such that } A \subset B \\ \{ A \} & \text{otherwise} \end{cases} \]
Clearly, \( A_A \neq \emptyset \). By Axiom of Choice, \( \exists T : E \to E \) such that \( T(A) = B \in A_A, \forall A \in E \).

We will show that \( T \) admits a fixed point by Lemma 2.17. Let \( B \subseteq E \) be any nonempty subset on which \( \subseteq \) is a total ordering. Let \( C := \bigcup_{B \in B} B \).

Clearly, \( C \subseteq E \). We will show \( \preceq \) is a total ordering on \( C \). Since \( \preceq \) is a partial ordering on \( E \), then it is a partial ordering on \( C \). \( \forall x_1, x_2 \in C \), \( \exists B_1, B_2 \in B \) such that \( x_1 \in B_1 \) and \( x_2 \in B_2 \). Since \( \subseteq \) is a total ordering on \( B \), then, we may without loss of generality assume \( B_1 \subseteq B_2 \). Then, \( x_1, x_2 \in B_2 \). Since \( B_2 \subseteq B \subseteq E \), then \( \preceq \) is a total ordering on \( B_2 \), which means that we have \( x_1 \preceq x_2 \) or \( x_2 \preceq x_1 \). Furthermore, if \( x_1 \preceq x_2 \) and \( x_2 \preceq x_1 \), then \( x_1 = x_2 \) by \( \preceq \) being antisymmetric on \( B_2 \). Therefore, by Proposition 2.12, \( \preceq \) is a total ordering on \( C \). Hence, \( C \in E \). This shows that \( B \) admits least upper bound \( C \) in \( E \) with respect to \( \subseteq \). By the definition of \( T \), it is clear that \( A \subseteq T(A) \), \( \forall A \in E \). By Lemma 2.17, \( T \) has a fixed point on \( E \), i.e., \( \exists A_0 \in E \) such that \( T(A_0) = A_0 \).

By the definitions of \( T \) and \( A_{A_0} \), there does not exist \( B \in E \) such that \( A_0 \subseteq B \). Hence, by Proposition 2.12 (vii), \( A_0 \) is maximal in \( E \) with respect to \( \subseteq \).

2. \( \Rightarrow \) 3. Let \( E \) be a nonempty set with an antisymmetric partial ordering \( \preceq \). By Hausdorff Maximum Principle, there exists a maximal (with respect to \( \subseteq \)) totally ordered (with respect to \( \preceq \)) subset \( F \subseteq E \). We must have \( F \neq \emptyset \), otherwise, let \( x_0 \in E \) (since \( E \neq \emptyset \)), \( F \subseteq \{x_0\} \subseteq E \) and \( \{x_0\} \) is totally ordered by \( \preceq \), which violates the fact that \( F \) is maximal (with respect to \( \subseteq \)). Then, \( F \) has an upper bound \( e_0 \in E \).

Claim 2.18.1 \( \exists e_0 \in F \).

Proof of claim: Suppose \( e_0 \notin F \). Define \( A := F \cup \{e_0\} \subseteq E \). Clearly, \( F \subseteq A \) and \( F \neq A \). We will show that \( \preceq \) is a total ordering on \( A \). Clearly, \( \preceq \) is an antisymmetric partial ordering on \( A \) since it is an antisymmetric partial ordering on \( E \). \( \forall x_1, x_2 \in A \), we will distinguish 4 exhaustive and mutually exclusive cases: Case 1: \( x_1, x_2 \in F \); Case 2: \( x_1 \in F, x_2 = e_0 \); Case 3: \( x_1 = e_0, x_2 \in F \); Case 4: \( x_1 = x_2 = e_0 \). In Case 1, we have \( x_1 \preceq x_2 \) or \( x_2 \preceq x_1 \) since \( \preceq \) is a total ordering on \( F \). In Case 2, we have \( x_1 \preceq x_2 = e_0 \) since \( e_0 \) is an upper bound of \( F \). In Case 3, we have \( x_2 \preceq x_1 \preceq e_0 \). In Case 4, we have \( x_1 \preceq e_0 \preceq x_2 \). Hence, \( \preceq \) is a total ordering on \( A \). Note that \( F \subseteq A \) and \( F \neq A \). By Proposition 2.12 (vii), this contradicts with the fact that \( F \) is maximal with respect to \( \subseteq \). Therefore, we must have \( e_0 \in F \). This completes the proof of the claim.

\forall e_1 \in E \) such that \( e_0 \preceq e_1 \). \( \forall x \in F \), we have \( x \preceq e_0 \preceq e_1 \). Hence, \( e_1 \) is an upper bound of \( F \). By Claim 2.18.1, we must have \( e_1 \in F \). Then, \( e_1 \preceq e_0 \) since \( e_0 \) is an upper bound of \( F \). This shows that \( e_0 \) is maximal in \( E \) with respect to \( \preceq \).

3. \( \Rightarrow \) 4. Let \( E \) be a set. It is clear that \( \emptyset \subseteq E \) is well-ordered by the empty relation. Define

\[ E := \{ (A_\alpha, \preceq_\alpha) \mid A_\alpha \subseteq E, A_\alpha \text{ is well-ordered by } \preceq_\alpha \} \]
Then, $E \neq \emptyset$. Define an ordering $\preceq$ on $E$ by $\forall (A_1, \preceq_1), (A_2, \preceq_2) \in E$, we say $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$ if the following three conditions hold: (i) $A_1 \subseteq A_2$; (ii) $\preceq_2 = \preceq_1$ on $A_1$; (iii) $\forall x_1 \in A_1, \forall x_2 \in A_2 \setminus A_1$, we have $x_1 \preceq x_2$.

Now, we will show that $\preceq$ defines an antisymmetric partial ordering on $E$. $\forall (A_1, \leq_1), (A_2, \leq_2), (A_3, \leq_3) \in E$. Clearly, $(A_1, \leq_1) \preceq (A_1, \leq_1)$. Hence, $\leq$ is reflexive. If $(A_1, \leq_1) \preceq (A_2, \leq_2)$ and $(A_2, \leq_2) \preceq (A_3, \leq_3)$, we have $A_1 \subseteq A_2 \subseteq A_3$, and (i) holds; (ii) $\leq_3 = \leq_2$ on $A_2$ and $\leq_2 = \leq_1$ on $A_1$ implies that $\leq_3 = \leq_1$ on $A_1$; (iii) $\forall x_1 \in A_1, \forall x_2 \in A_3 \setminus A_1$, we have 2 exhaustive scenarios: if $x_2 \in A_2$, then $x_2 \in A_2 \setminus A_1$ which implies $x_1 \preceq x_2$ and hence $x_1 \preceq x_2$; if $x_2 \in A_3 \setminus A_2$, then we have $x_1 \preceq x_2$ and $x_1 \preceq x_2$, thus, we have $x_1 \preceq x_2$ in both cases. Therefore, $(A_1, \leq_1) \preceq (A_3, \leq_3)$ and hence $\preceq$ is transitive. If $(A_1, \leq_1) \preceq (A_2, \leq_2)$ and $(A_2, \leq_2) \preceq (A_1, \leq_1)$, then $A_1 \subseteq A_2 \subseteq A_1 \Rightarrow A_1 = A_2$ and $\leq_2 = \leq_1$ on $A_1$. Hence, $(A_1, \leq_1) = (A_2, \leq_2)$, which shows that $\preceq$ is antisymmetric. Therefore, $\preceq$ defines an antisymmetric partial ordering on $E$.

Let $A \subseteq E$ be any nonempty subset totally ordered by $\preceq$. Take $A := \{ (A_\alpha, \preceq_\alpha) \mid \alpha \in \Lambda \}$ where $\Lambda \neq \emptyset$ is an index set. Define $A := \bigcup_{\alpha \in \Lambda} A_\alpha$. Define an ordering $\preceq$ on $A$ by: $\forall x_1, x_2 \in A$, $\exists (A_1, \leq_1), (A_2, \leq_2) \in A$ such that $x_1 \in A_1$ and $x_2 \in A_2$, without loss of generality, assume that $(A_1, \leq_1) \preceq (A_2, \leq_2)$ since $A$ is totally ordered by $\preceq$, then $x_1, x_2 \in A_2$, we will say that $x_1 \preceq x_2$ if $x_1 \preceq x_2$. We will now show that this ordering is uniquely defined independent of $(A_2, \leq_2) \in A$. Let $(A_3, \leq_3) \in A$ be such that $x_1, x_2 \in A_3$. Since $A$ is totally ordered by $\preceq$, then there are two exhaustive cases: Case 1: $(A_3, \leq_3) \preceq (A_2, \leq_2)$; Case 2: $(A_3, \leq_3) \preceq (A_3, \leq_3)$. In Case 1, we have $A_3 \subseteq A_2$ and $\leq_3 = \leq_2$ on $A_3$, which implies that $x_1 \preceq x_2 \Leftrightarrow x_1 \preceq x_2 \Leftrightarrow x_1 \preceq x_2$. In Case 2, we have $A_2 \subseteq A_3$ and $\leq_3 = \leq_2$ on $A_2$, which implies that $x_1 \preceq x_2 \Leftrightarrow x_1 \preceq x_2 \Leftrightarrow x_1 \preceq x_2$. Hence, the ordering $\preceq$ is well-defined on $A$.

Next, we will show that $\preceq$ is a total ordering on $A$. $\forall x_1, x_2, x_3 \in A$, $\exists (A_i, \leq_i) \in A$ such that $x_i \in A_i$, $i = 1, 2, 3$. Since $A$ is totally ordered by $\preceq$, then, without loss of generality, assume that $(A_1, \leq_1) \preceq (A_2, \leq_2) \preceq (A_3, \leq_3)$. Then, $x_1, x_2, x_3 \in A_3$. Clearly, $x_1 \preceq x_1$ since $x_1 \preceq x_3$, which implies that $\preceq$ is reflexive. If $x_1 \preceq x_2$ and $x_2 \preceq x_3$, then $x_1 \preceq x_2 \preceq x_3$, which implies $x_1 \preceq x_3$ since $\leq_3$ is transitive on $A_3$, and hence, $x_1 \preceq x_3$. This shows that $\preceq$ is transitive. If $x_1 \preceq x_2$ and $x_2 \preceq x_1$, then $x_1 \preceq x_2$ and $x_2 \preceq x_1$, which implies that $x_1 = x_2$ since $\leq_3$ is antisymmetric on $A_3$. This shows that $\preceq$ is antisymmetric. Since $\leq_3$ is a well-ordering on $A_3$, then we must have $x_1 \preceq x_2 \Leftrightarrow x_1 \preceq x_2$ or $x_2 \preceq x_3 \Leftrightarrow x_2 \preceq x_1$. Hence, $\preceq$ defines a total ordering on $A$.

Next, we will show that $\preceq$ is a well-ordering on $A$ with $B \neq \emptyset$. Fix $x_0 \in B$. Then, $\exists (A_1, \leq_1) \in A$ such that $x_0 \in A_1$. Note that $\emptyset \neq B \cap A_1 \subseteq A_1$. Since $A_1$ is well-ordered by $\leq_1$, then $\exists e \in B \cap A_1$, which is the least element of $B \cap A_1$. $\forall y \in B \subseteq A$, $\exists (A_2, \leq_2) \in A$ such that $y \in A_2$. We have 2 exhaustive and mutually exclusive cases: Case 1: $y \in A_1$; Case 2: $y \in A_2 \setminus A_1$. In Case 1, $e \leq y$ since $e$ is the least element
of $B \cap A_1$, which implies that $e \leq y$. In Case 2, since $A$ is totally ordered by $\preceq$, we must have $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$, which implies that $e \preceq_2 y$, by (iii) in the definition of $\preceq$, and hence $e \leq y$. In both cases, we have shown that $e \leq y$. Since $\preceq$ is a total ordering on $A$, then $e$ is the least element of $B$. Therefore, $\preceq$ is a well ordering on $A$, which implies $(A, \preceq) \in \mathcal{E}$.

By Zorn’s Lemma, there is a maximal element $(F, \preceq_F) \in \mathcal{E}$. We claim that $F = E$. We will prove this by an argument of contradiction. Suppose $F \subset E$, then $\exists x_0 \in E \setminus F$. Let $H := F \cup \{x_0\}$. Define an ordering $\preceq_H$ on $H$ by: $\forall x_1, x_2 \in H$, if $x_1, x_2 \in F$, we say $x_1 \preceq_H x_2$ if $x_1 \preceq_F x_2$; if $x_1 \in F$ and $x_2 \in x_0$, then we let $x_1 \preceq_H x_2$; if $x_1 = x_2 = x_0$, we let $x_1 \preceq_H x_2$. Now, we will show that $\preceq_H$ is a well ordering on $H$. \forall $x_1, x_2, x_3 \in H$. If $x_1, x_2, x_3 \in H$. If $x_1 \preceq_F x_2$, then, $x_1 \preceq_F x_1$ and $x_1 \preceq_H x_1$; if $x_1 = x_0$, then $x_1 \preceq_H x_1$. Hence, $\preceq_H$ is reflexive. If $x_1 \preceq_H x_2$ and $x_2 \preceq_H x_3$. We have 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_2 \in F$; Case 2: $x_1 \in F$ and $x_2 = x_0$; Case 3: $x_2 \in F$ and $x_1 = x_0$; Case 4: $x_1 = x_2 = x_0$. In Case 1, we must have $x_2 \in F$ and then $x_1 \preceq_F x_2$ and $x_2 \preceq_F x_3$, which implies that $x_1 \preceq_F x_3$, and hence $x_1 \preceq_H x_3$. In Case 2, we have $x_1 \preceq_H x_3$. In Case 3, we have to show that $x_0 \preceq_H x_3$, hence, this case is impossible. In Case 4, we have $x_1 \preceq_H x_3$. In all cases except that is impossible, we have $x_1 \preceq_H x_3$. Hence, $\preceq_H$ is transitive. If $x_1 \preceq_H x_2$ and $x_2 \preceq_H x_1$. We have 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_2 \in F$; Case 2: $x_1 \in F$ and $x_2 = x_0$; Case 3: $x_2 \in F$ and $x_1 = x_0$; Case 4: $x_1 = x_2 = x_0$. In Case 1, we have $x_1 \preceq_F x_2$ and $x_2 \preceq_F x_1$, which implies that $x_1 = x_2$ since $\preceq_F$ is antisymmetric on $F$. In Case 2, we have $x_0 \preceq_H x_1$, which is a contradiction, and hence this case is impossible. In Case 3, we have $x_0 \preceq_H x_2$, which is a contradiction, and hence this case is impossible. In Case 4, we have $x_1 = x_2$. In all cases except those impossible, we have $x_1 = x_2$. Hence, $\preceq_H$ is antisymmetric. When $x_1, x_2 \in F$, then, we must have $x_1 \preceq_F x_2$ or $x_2 \preceq_F x_1$ since $\preceq_F$ is a well ordering on $F$, and hence $x_1 \preceq_H x_2$ or $x_2 \preceq_H x_1$. When $x_1 \in F$ and $x_2 = x_0$, then $x_1 \preceq_H x_2$. When $x_2 \in F$ and $x_1 = x_0$, then $x_2 \preceq_H x_1$. When $x_1 = x_2 = x_0$, then $x_1 \preceq_H x_2$. This shows that $\preceq_H$ is a total ordering on $H$. $\forall B \subseteq H$ with $B \neq \emptyset$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $B = \{x_0\}$; Case 2: $B \neq \{x_0\}$. In Case 1, $x_0$ is the least element of $B$. In Case 2, $B \setminus \{x_0\} \subseteq F$ and is nonempty, and hence admits a least element $e_0 \in B \setminus \{x_0\} \subseteq F$ with respect to $\preceq_F$. $\forall x \in B$, if $x \in B \setminus \{x_0\}$, then $e_0 \preceq_F x$ and hence $e_0 \preceq_H x$; if $x = x_0$, then $e_0 \preceq_H x$. Hence, $e_0$ is the least element of $B$ since $\preceq_H$ is a total ordering on $H$. Therefore, $\preceq_H$ is a well ordering on $H$ and $(H, \preceq_H) \in \mathcal{E}$.

Clearly, $F \subset H$, $\preceq_F \preceq_H$ on $F$, and $\forall x_1 \in F$ and $\forall x_2 \in H \setminus F$, we have $x_2 = x_0$ and $x_1 \preceq_H x_2$. This implies that $(F, \preceq_F) \preceq (H, \preceq_H)$. Since $(F, \preceq_F)$
is maximal in \( E \) with respect to \( \preceq \), we must have \((H, \preceq_H) \preceq (F, \preceq_F)\), and hence, \( H \subseteq F \). This is a contradiction. Therefore, \( F = E \) and \( E \) is well ordered by \( \preceq_F \).

4. \( \Rightarrow \) 1. Let \( (A_\lambda)_{\lambda \in \Lambda} \) be a collection of nonempty sets, and \( \Lambda \) is a set. Let \( A := \bigcup_{\lambda \in \Lambda} A_\lambda \). By Well-Ordering Principle, \( A \) may be well ordered by \( \leq \). \( \forall \lambda \in \Lambda, A_\lambda \subseteq A \) is nonempty and admits the least element \( e_\lambda \in A_\lambda \). This defines a function \( f : \Lambda \to A \) by \( f(\lambda) = e_\lambda \in A_\lambda, \forall \lambda \in \Lambda \).

This completes the proof of the theorem. \( \square \)

**Example 2.19** Let \( \Lambda \) be an index set and \( (A_\alpha)_{\alpha \in \Lambda} \) be a collection of sets. We will try to define the *Cartesian (direct) product* \( \prod_{\alpha \in \Lambda} A_\alpha \). Let \( A = \bigcup_{\alpha \in \Lambda} A_\alpha \), which is a set by the Axiom of Union. Then, as we discussed in Section 2.3, \( A^\Lambda \) is a set, which consists of all functions of \( \Lambda \) to \( A \). Define the *projection* functions \( \pi_\alpha : A^\Lambda \to A, \forall \alpha \in \Lambda \), by, \( \forall f \in A^\Lambda, \pi_\alpha(f) = f(\alpha) \).

Then, we may define the set

\[
\prod_{\alpha \in \Lambda} A_\alpha := \{ f \in A^\Lambda \mid \pi_\alpha(f) \in A_\alpha, \forall \alpha \in \Lambda \}
\]

When all of \( A_\alpha \)'s are nonempty, then, by Axiom of Choice, the product \( \prod_{\alpha \in \Lambda} A_\alpha \) is also nonempty. \( \diamond \)
Chapter 3

Topological Spaces

3.1 Fundamental Notions

Definition 3.1 A topological space \((X, \mathcal{O})\) consists of a set \(X\) and a collection \(\mathcal{O}\) of subsets (namely, open subsets) of \(X\) such that

(i) \(\emptyset, X \in \mathcal{O}\);

(ii) \(\forall O_1, O_2 \in \mathcal{O}, \text{ we have } O_1 \cap O_2 \in \mathcal{O}\);

(iii) \(\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}, \text{ where } \Lambda \text{ is an index set, we have } \bigcup_{\alpha \in \Lambda} O_\alpha \in \mathcal{O}\).

The collection \(\mathcal{O}\) is called a topology for the set \(X\).

Definition 3.2 Let \((X, \mathcal{O})\) be a topological space and \(F \subseteq X\). The complement of \(F\) is \(\overline{F} := X \setminus F\). \(F\) is said to be closed if \(\overline{F} \in \mathcal{O}\). The closure of \(F\) is given by \(\overline{F} := \bigcap_{B \subseteq F \in \mathcal{O}} B\), which is clearly a closed set. The interior of \(F\) is given by \(F^\circ := \bigcup_{B \subseteq F \in \mathcal{O}} B\), which is clearly an open set. A point of closure of \(F\) is a point in \(\overline{F}\). An interior point of \(F\) is a point in \(F^\circ\). A boundary point of \(F\) is a point \(x \in X\) such that \(\forall O \in \mathcal{O} \text{ with } x \in O, \text{ we have } O \cap F \neq \emptyset\) and \(O \cap \overline{F} \neq \emptyset\). The boundary of \(F\), denoted by \(\partial F\), is the set of all boundary points of \(F\). An exterior point of \(F\) is a point in \(\overline{F}^\circ\), where \(\overline{F}^\circ\) is called the exterior of \(F\). An accumulation point of \(F\) is a point \(x \in X\) such that \(\forall O \in \mathcal{O} \text{ with } x \in O, \text{ we have } O \cap (F \setminus \{x\}) \neq \emptyset\).

Clearly, \(\emptyset\) and \(X\) are both closed and open.

Proposition 3.3 Let \((X, \mathcal{O})\) be a topological space and \(A, B, E\) are subsets of \(X\). Then,
(i) \( E \subseteq \overline{E} \), \( \overline{E} = \overline{E} \), \( E^\circ \subseteq E \), \( (E^\circ)^\circ = E^\circ \), and \( \overline{E^\circ} = \overline{E} \).

(ii) \( \forall x \in X, x \) is a point of closure of \( E \) if, and only if, \( \forall O \in \mathcal{O} \) with \( x \in O \), we have \( O \cap E \neq \emptyset \);

(iii) \( \forall x \in X, x \) is an interior point of \( E \) if, and only if, \( \exists O \in \mathcal{O} \) with \( x \in O \) such that \( O \subseteq E \);

(iv) \( A \cup B = A \cup \overline{B}, (A \cap B)^^\circ = A^\circ \cap B^\circ \);

(v) \( E \) is closed if, and only if, \( E = \overline{E} \);

(vi) \( \overline{E} = E^\circ \cup \partial E \).

(vii) \( X \) equals to the disjoint union \( E^\circ \cup \partial E \cup (\overline{E})^\circ \);

**Proof**

(i) Clearly, \( E \subseteq \overline{E} \). Then, \( \overline{E} \subseteq \overline{E} \), \( \forall C \supseteq E \) with \( \overline{C} \in \mathcal{O} \), we have \( C \supseteq \overline{E} \). Then, \( \overline{E} = \bigcap_{B \in \mathcal{O}} C \supseteq \bigcap_{B \in \mathcal{O}} C = \overline{E} \). Hence, we have \( E = \overline{E} \).

Clearly, \( E^\circ \subseteq E \). Note that

\[
\overline{E} = \left( \bigcap_{B \in \mathcal{O}} B \right)^\circ = \bigcup_{B \in \mathcal{O}} B = \bigcup_{O \in \mathcal{O}} O = \overline{E}^\circ
\]

Furthermore,

\[
(E^\circ)^\circ = \left( \overline{E} \right)^\circ = \overline{E}^\circ = \overline{E} = \left( \overline{E} \right)^\circ = E^\circ
\]

(ii) “Only if” \( \forall x \in \overline{E} \), we have \( x \in \bigcap_{B \in \mathcal{O}} B \). \( \forall O \in \mathcal{O} \) with \( x \in O \), let \( O_1 := O \cup \overline{E} \in \mathcal{O} \). Note that \( x \in O_1 \) and \( E \cap \overline{E} = \emptyset \). Suppose \( O \cap E = \emptyset \). Then, \( E \cap O_1 = \emptyset \), which further implies that \( E \subseteq O_1 \). Then, \( \overline{E} \subseteq O_1 \) and \( x \in O_1 \). This contradicts with \( x \in O \). Hence, \( O \cap E \neq \emptyset \).

“If” \( \forall x \in \overline{E} = \overline{E}^\circ \in \mathcal{O} \). Then, \( \exists O := \overline{E} \in \mathcal{O} \) such that \( E \cap O = \emptyset \). Hence, the result holds.

(iii) “Only if” \( \forall x \in E^\circ \subseteq E \), then \( E^\circ \in \mathcal{O} \).

“If” \( \forall x \in X, \exists O \in \mathcal{O} \) such that \( x \in O \subseteq E \). Then, \( x \in O \subseteq \bigcup_{B \in \mathcal{O}} B = E^\circ \). Hence, the result holds.

(iv) Let \( \mathcal{B} := \{ O \in \mathcal{O} \mid O \subseteq A \cap B \} \), \( \mathcal{B}_A := \{ O \in \mathcal{O} \mid O \subseteq A \} \), and \( \mathcal{B}_B := \{ O \in \mathcal{O} \mid O \subseteq B \} \). \( \forall O_1 \in \mathcal{B}_A \) and \( \forall O_2 \in \mathcal{B}_B \), then, \( O_1 \cap O_2 \in \mathcal{B} \). On the other hand, \( \forall O \in \mathcal{B} \), we have \( O = O \cap O \) and \( O \in \mathcal{B}_A \) and \( O \in \mathcal{B}_B \). Then,

\[
(A \cap B)^\circ = \bigcup_{O_1 \subseteq A,B \in O} O = \bigcup_{O_1 \subseteq A, O_2 \subseteq B} (O_1 \cap O_2)
\]

\[
= \left( \bigcup_{O_1 \subseteq A} O_1 \right) \cap \left( \bigcup_{O_2 \subseteq B} O_2 \right) = A^\circ \cap B^\circ
\]
We also have
\[
\overline{A \cup B} = \left( \overline{A \cup B} \right)^\circ = \left( \left( \overline{A} \cap \overline{B} \right)^\circ \right) = \left( \overline{A} \cap \overline{B} \right)^\circ = \overline{A} \cap \overline{B} = \overline{A} \cup \overline{B}
\]

(v) “If” $E$ is closed since $E = \overline{E}$ and $\overline{E}$ is closed.

“Only if” Since $E$ is closed, then $\overline{E} \subseteq E$. Then, we have $E = \overline{E}$. Hence, the result holds.

(vi) This result follows directly from (ii), (iii), and Definition 3.2.

(vii) Note that $E = \overline{E}$ relative to $(X, O)$ since $\overline{E}$ and $\partial E$ are disjoint. It is obvious that $\overline{E}$ is disjoint with $E^\circ \cup \partial E = \overline{E}$. Hence, the result holds. \qed

To simplify notation in the theory, we will abuse the notation to write $x \in \mathcal{X}$ when $x \in X$ and $A \subseteq \mathcal{X}$ when $A \subseteq X$ for a topological space $\mathcal{X} := (X, O)$. We will later simply discuss a topological space $\mathcal{X}$ without further reference to components of $\mathcal{X}$, where the topology is understood to be $O_X$. When it is clear from the context, we will neglect the subscript $\mathcal{X}$.

**Proposition 3.4** Let $(X, O)$ be a topological space and $A \subseteq X$. A admits the subset topology $O_A := \{ O \cap A \mid O \in O \}$.

**Proof** Clearly, $O_A$ is a collection of subsets of $A$. $\emptyset = \emptyset \cap A \in O_A$ and $A = X \cap A \in O_A$. $\forall O_{A1}, O_{A2} \in O_A, \exists O_1, O_2 \in O$ such that $O_{A1} = O_1 \cap A$ and $O_{A2} = O_2 \cap A$. Then, $O_1 \cap O_2 \in O$ since $O$ is a topology. Then, $O_{A1} \cap O_{A2} = (O_1 \cap O_2) \cap A \in O_A$. $\forall (O_{\alpha})_{\alpha \in \Lambda} \subseteq O_A$, where $\Lambda$ is an index set, we have $\forall \alpha \in \Lambda, \exists O_\alpha \in O$ such that $O_{A\alpha} = O_\alpha \cap A$. Then, $\bigcup_{\alpha \in \Lambda} O_\alpha \in O$ since $O$ is a topology. Therefore, $\bigcup_{\alpha \in \Lambda} O_{A\alpha} = \left( \bigcup_{\alpha \in \Lambda} O_\alpha \right) \cap A \in O_A$.

Hence, $O_A$ is a topology on $A$. \qed

Let $(X, O)$ be a topological space and $A \subseteq X$. The property of a set $E \subseteq A$ being open or closed is relative with respect to $(X, O)$, that is, this property may change if we consider the subset topology $(A, O_A)$.

**Proposition 3.5** Let $\mathcal{X}$ be a topological space, $A \subseteq \mathcal{X}$ be endowed with the subset topology $O_A$, and $E \subseteq A$. Then,

1. $E$ is closed in $O_A$ if, and only if, $E = A \cap F$, where $F \subseteq \mathcal{X}$ is closed in $O_\mathcal{X}$;

2. the closure of $E$ relative to $(A, O_A)$ (the closure of $E$ in $O_A$) is equal to $\overline{E} \cap A$, where $\overline{E}$ is the closure of $E$ relative to $\mathcal{X}$.

**Proof** Here, the set complementation and set closure operation are relative to $\mathcal{X}$.

1. “If” $A \setminus E = A \setminus (A \cap F) = A \cap A \setminus F = A \cap F$. Since $F$ is closed in $O_\mathcal{X}$, then $\overline{F} \in O_\mathcal{X}$. Then, $A \setminus E \in O_A$. Hence, $E$ is closed in $O_A$. 

---

3.1. FUNDAMENTAL NOTIONS

41
“Only if” $A \setminus E \in \mathcal{O}_A$. Then, $\exists O \in \mathcal{O}_X$ such that $A \setminus E = A \cap O$. Then, $E = A \setminus (A \setminus E) = A \cap A \cap O = A \cap \overset{\sim}{O}$. Hence, the result holds.

(2) By (1), $\overline{E} \cap A$ is closed in $\mathcal{O}_A$. Then, the closure of $E$ relative to $(A, \mathcal{O}_A)$ is contained in $\overline{E} \cap A$. On the other hand, by Proposition 3.3, if $x \in \mathcal{X}$ is a point of closure of $E$ relative to $\mathcal{X}$, then it is a point of closure of $E$ in $\mathcal{O}_A$ if $x \in A$. Then, $\overline{E} \cap A$ is contained in the closure of $E$ relative to $(A, \mathcal{O}_A)$. Hence, the result holds.

This completes the proof of the proposition. \qed

**Definition 3.6** For two topologies over the same set $X$, $\mathcal{O}_1$ and $\mathcal{O}_2$, we will say that $\mathcal{O}_1$ is stronger (finer) than $\mathcal{O}_2$ if $\mathcal{O}_1 \supset \mathcal{O}_2$, in which case, $\mathcal{O}_1$ is said to be weaker (coarser) than $\mathcal{O}_1$.

**Proposition 3.7** Let $X$ be a set and $A \subseteq X^2$. Then, there exists the weakest topology $\mathcal{O}$ on $X$ such that $A \subseteq \mathcal{O}$. This topology is called the topology generated by $A$.

**Proof** Let $\mathcal{M} := \{ \mathcal{X} \subseteq X^2 \mid \mathcal{A} \subseteq \mathcal{X} \text{ and } \mathcal{X} \text{ is a topology on } X \}$ and $\mathcal{O} = \bigcap_{\mathcal{X} \in \mathcal{M}} \mathcal{X}$. Clearly, $X^2 \in \mathcal{M}$ and hence $\mathcal{O}$ is well-defined. Then,

(i) $\emptyset, X \in \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Hence, $\emptyset, X \in \mathcal{O}$.

(ii) $\forall A_1, A_2 \in \mathcal{O}$, we have $A_1, A_2 \in \mathcal{X}$, $\forall \mathcal{X} \in \mathcal{M}$. Then, $A_1 \cap A_2 \in \mathcal{X}$, $\forall \mathcal{X} \in \mathcal{M}$. Hence, $A_1 \cap A_2 \in \mathcal{O}$.

(iii) $\forall (A_{\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}$, where $\Lambda$ is an index set, we have, $\forall \mathcal{X} \in \mathcal{M}$, $A_{\alpha} \subseteq \mathcal{X}$. Then, $\bigcup_{\alpha \in \Lambda} A_{\alpha} \subseteq \mathcal{X}, \forall \mathcal{X} \in \mathcal{M}$. Hence, we have $\bigcup_{\alpha \in \Lambda} A_{\alpha} \in \mathcal{O}$.

Therefore, $\mathcal{O}$ is a topology on $X$. Clearly, $\mathcal{A} \subseteq \mathcal{O}$ since $A \subseteq \mathcal{X}$, $\forall \mathcal{X} \in \mathcal{M}$. Therefore, $\mathcal{O}$ is the weakest topology containing $A$. \qed

### 3.2 Continuity

**Definition 3.8** Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be topological spaces, $D \subseteq X$ with the subset topology $\mathcal{O}_D$, and $f : D \to Y$ (or $f : (D, \mathcal{O}_D) \to (Y, \mathcal{O}_Y)$ to be more specific). Then, $f$ is said to be continuous if, $\forall O \in \mathcal{O}_Y$, we have $f_{\text{inv}}(O_Y) \in \mathcal{O}_D$. $f$ is said to be continuous at $x_0 \in D$ if, $\forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$, $\exists U \in \mathcal{O}_X$ with $x_0 \in U$ such that $f(U) \subseteq O_Y$. $f$ is said to be continuous on $E \subseteq D$ if it is continuous at $x$, $\forall x \in E$.

**Proposition 3.9** Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \to \mathcal{Y}$. $f$ is continuous if, and only if, $\forall x_0 \in D$, $f$ is continuous at $x_0$.

**Proof** “If” $\forall O_Y \in \mathcal{O}_Y$, $\forall x \in f_{\text{inv}}(O_Y) \subseteq D$. Since $f$ is continuous at $x$, then $\exists U_x \in \mathcal{O}_X$ with $x \in U_x$ such that $f(U_x) \subseteq O_Y$, which implies, by Proposition 2.5, that $U_x \cap D \subseteq f_{\text{inv}}(O_Y)$. Then,
3.2. CONTINUITY

\[ f_{\text{inv}}(O_Y) = \bigcup_{x \in f_{\text{inv}}(O_Y)} (U_x \cap D) = (\bigcup_{x \in f_{\text{inv}}(O_Y)} U_x) \cap D \in \mathcal{O}_D. \] Hence, \( f \) is continuous.

“Only if” \( \forall x_0 \in D, \forall O_Y \in \mathcal{O}_Y \) with \( f(x_0) \in O_Y \), let \( U = f_{\text{inv}}(O_Y) \) \( \in \mathcal{O}_D \). By Proposition 3.4, \( \exists U \in \mathcal{O}_X \) such that \( U = \bar{U} \cap D \). Then, \( x_0 \in U \).

By Proposition 2.5, \( f(\bar{U}) = f(U) \subseteq O_Y \). Hence, \( f \) is continuous at \( x_0 \).

This completes the proof of the proposition.

**Proposition 3.10** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be topological spaces and \( f : \mathcal{X} \to \mathcal{Y} \). \( f \) is continuous if, and only if, \( \forall B \subseteq \mathcal{Y} \) with \( \bar{B} \in \mathcal{O}_Y \), we have \( f_{\text{inv}}(B) \in \mathcal{O}_X \), that is, the inverse image of any closed set in \( \mathcal{Y} \) is closed in \( \mathcal{X} \).

**Proof** “If” \( \forall O \in \mathcal{O}_Y \), we have, by Proposition 2.5, \( f_{\text{inv}}(O) = f_{\text{inv}}(\bar{O}) \in \mathcal{O}_X \). Hence, \( f \) is continuous.

“Only if” \( \forall B \subseteq \mathcal{Y} \) with \( \bar{B} \in \mathcal{O}_Y \). Since \( f \) is continuous, then, by Proposition 2.5, \( f_{\text{inv}}(\bar{B}) = f_{\text{inv}}(B) \in \mathcal{O}_X \). Hence, the result holds.

This completes the proof of the proposition.

**Theorem 3.11** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be topological spaces, \( f : \mathcal{X} \to \mathcal{Y} \), and \( \mathcal{X} = X_1 \cup X_2 \), where \( X_1 \) and \( X_2 \) are both open or both closed. Let \( X_1 \) and \( X_2 \) be endowed with subset topologies \( \mathcal{O}_{X_1} \) and \( \mathcal{O}_{X_2} \), respectively. Assume that \( f|_{X_1} : X_1 \to \mathcal{Y} \) and \( f|_{X_2} : X_2 \to \mathcal{Y} \) are continuous. Then, \( f \) is continuous.

**Proof** Consider the case that \( X_1 \) and \( X_2 \) are both open. \( \forall x_0 \in \mathcal{X} \). Without loss of generality, assume \( x_0 \in X_1 \). \( \forall O \in \mathcal{O}_Y \) with \( f(x_0) \in O \).

Since \( f|_{X_1} \) is continuous, then, by Proposition 3.9, \( \exists U \in \mathcal{O}_{X_1} \) with \( x_0 \in U \) such that \( f|_{X_1}(U) \subseteq O \). Since \( X_1 \in \mathcal{O}_{X_1} \), then \( U \in \mathcal{O}_X \). Note that \( f(U) = f|_{X_1}(U) \subseteq O \), since \( U \subseteq X_1 \). Hence, \( f \) is continuous at \( x_0 \). By the arbitraryness of \( x_0 \) and Proposition 3.9, \( f \) is continuous.

Consider the case that \( X_1 \) and \( X_2 \) are both closed. \( \forall \) closed subset \( B \subseteq \mathcal{Y} \), we have \( f_{\text{inv}}(B) \subseteq X \). Then, \( f_{\text{inv}}(B) \cap X_1 = (f|_{X_1})_{\text{inv}}(B) \) is closed in \( \mathcal{O}_{X_1} \), by Proposition 3.10 and the continuity of \( f|_{X_1} \). Similarly, \( f_{\text{inv}}(B) \cap X_2 = (f|_{X_2})_{\text{inv}}(B) \) is closed in \( \mathcal{O}_{X_2} \). Since \( X_1 \) and \( X_2 \) are closed sets in \( \mathcal{O}_X \), then, \( f_{\text{inv}}(B) \cap X_1 \) and \( f_{\text{inv}}(B) \cap X_2 \) are closed in \( \mathcal{O}_X \), by Proposition 3.5. Then, \( f_{\text{inv}}(B) = (f_{\text{inv}}(B) \cap X_1) \cup (f_{\text{inv}}(B) \cap X_2) \) is closed in \( \mathcal{O}_X \). By Proposition 3.10, \( f \) is continuous.

This completes the proof of the theorem.

**Proposition 3.12** Let \( \mathcal{X}, \mathcal{Y}, \) and \( \mathcal{Z} \) be topological spaces, \( f : \mathcal{X} \to \mathcal{Y} \), \( g : \mathcal{Y} \to \mathcal{Z} \), and \( x_0 \in \mathcal{X} \). Assume that \( f \) is continuous at \( x_0 \) and \( g \) is continuous at \( y_0 := f(x_0) \). Then, \( g \circ f : \mathcal{X} \to \mathcal{Z} \) is continuous at \( x_0 \).

**Proof** \( \forall O_Z \in \mathcal{O}_Z \) with \( g(f(x_0)) \in O_Z \). Since \( g \) is continuous at \( f(x_0) \), then \( \exists O_Y \in \mathcal{O}_Y \) with \( f(x_0) \in O_Y \) such that \( g(O_Y) \subseteq O_Z \). Since \( f \) is continuous at \( x_0 \), then \( \exists O_X \in \mathcal{O}_X \) with \( x_0 \in O_X \) such that \( f(O_X) \subseteq O_Y \). Then, \( g(f(O_X)) \subseteq O_Z \). Hence, \( g \circ f \) is continuous at \( x_0 \). This completes the proof of the proposition.
Definition 3.13 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \to \mathcal{Y}$. $f$ is said to be a homeomorphism between $\mathcal{X}$ and $\mathcal{Y}$ if it is bijective and continuous and $f_{\text{inv}} : \mathcal{Y} \to \mathcal{X}$ is also continuous. The spaces $\mathcal{X}$ and $\mathcal{Y}$ are said to be homeomorphic if there exists a homeomorphism between them.

Any properties invariant under homeomorphisms are called topological properties.

Homeomorphisms preserve topological properties in topological spaces. Isomorphisms preserve algebraic properties in algebraic systems. Isometries if there exists a homeomorphism between them.

Definition 3.14 Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \to \mathbb{R}$. $f$ is said to be upper semicontinuous if $\forall a \in \mathbb{R}$, $f_{\text{inv}}((-\infty,a)) \in \mathcal{O}_D$. $f$ is said to be upper semicontinuous at $x_0 \in \mathcal{X}$ if $\forall \epsilon \in (0,\infty) \subset \mathbb{R}$, $\exists U \in \mathcal{O}_X$ with $x_0 \in U$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in U \cap D$. $f$ is said to be lower semicontinuous if $-f$ is upper semicontinuous.

Proposition 3.15 Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$ with the subset topology $\mathcal{O}_D$, and $f : D \to \mathbb{R}$. $f$ is upper semicontinuous if, and only if, $f$ is upper semicontinuous at $x_0$, $\forall x_0 \in D$.

Proof This is straightforward, and is therefore omitted. \qed

Proposition 3.16 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{X} \to \mathbb{R}$ be upper semicontinuous at $x_0 \in \mathcal{X}$, $h : \mathcal{Y} \to \mathcal{X}$ be continuous at $y_0 \in \mathcal{Y}$, and $h(y_0) = x_0$. Then, $f + g$ is upper semicontinuous at $x_0$ and $f \circ h$ is upper semicontinuous at $y_0$. Furthermore, if $f$ is also lower semicontinuous at $x_0$, then $f$ is continuous at $x_0$.

Proof $\forall \epsilon \in (0,\infty) \subset \mathbb{R}$, $\exists U_f \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_f$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in U_f$, by the upper semicontinuity of $f$. By the upper semicontinuity of $g$, $\exists U_g \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_g$ such that $g(x) < g(x_0) + \epsilon$, $\forall x \in U_g$. Then, $x_0 \in U := U_f \cap U_g \in \mathcal{O}_\mathcal{X}$ and $(f + g)(x) = f(x) + g(x) < f(x_0) + \epsilon + g(x_0) + \epsilon = (f + g)(x_0) + 2\epsilon$, $\forall x \in U$. Hence, $f + g$ is upper semicontinuous at $x_0$.

$\forall \epsilon \in (0,\infty) \subset \mathbb{R}$, $\exists U_f \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_f$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in U_f$, by the upper semicontinuity of $f$. By the continuity of $h$, $\exists U_h \in \mathcal{O}_Y$ with $y_0 \in U_h$ such that $h(y) \in U_f$, $\forall y \in U_h$. Then, $(f \circ h)(y) < f(h(y_0)) + \epsilon$, $\forall y \in U_h$. Hence, $f \circ h$ is upper semicontinuous at $y_0$.

$\forall \epsilon \in (0,\infty) \subset \mathbb{R}$, by the upper semicontinuity of $f$, $\exists U_1 \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_1$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in U_1$. By the lower semicontinuity of $f$, $\exists U_2 \in \mathcal{O}_\mathcal{X}$ with $x_0 \in U_2$ such that $f(x) > f(x_0) - \epsilon$, $\forall x \in U_2$. Then, $x_0 \in U := U_1 \cap U_2 \in \mathcal{O}_\mathcal{X}$ and $|f(x) - f(x_0)| < \epsilon$, $\forall x \in U$. Hence, $f$ is continuous at $x_0$.

This completes the proof of the proposition. \qed
3.3 Basis and Countability

Definition 3.17 Let \((X, \mathcal{O})\) be a topological space and \(\mathcal{B} \subseteq \mathcal{O}\). \(\mathcal{B}\) is said to be a basis of the topological space if, \(\forall O \in \mathcal{O}, \forall x \in O, \exists B \in \mathcal{B}\) such that \(x \in B \subseteq O\). B, x \in B, \forall B \in \mathcal{B}_x\), is said to be a basis at \(x \in X\) if \(\forall O \in \mathcal{O}\) with \(x \in O\), \(\exists B \in \mathcal{B}_x\) such that \(x \in B \subseteq O\).

If \(\mathcal{B}\) is a basis for the topology \(\mathcal{O}\), then the topology generated by \(\mathcal{B}\) is \(\mathcal{O}\) and, \(\forall O \in \mathcal{O}\), \(O = \bigcup_{x \in O} B_x\), where \(B_x \in \mathcal{B}\) is such that \(x \in B_x \subseteq O\).

Proposition 3.18 Let \(X\) be a set and \(\mathcal{B} \subseteq X^2\). Then, \(\mathcal{B}\) is a basis for the topology generated by \(\mathcal{B}, \mathcal{O}\), if, and only if, the following two conditions hold:

(i) \(\forall x \in X, \exists B \in \mathcal{B}\) such that \(x \in B\); (that is, \(\bigcup_{B \in \mathcal{B}} B = X\))
(ii) \(\forall B_1, B_2 \in \mathcal{B}, \forall x \in B_1 \cap B_2, \exists B_3 \in \mathcal{B}\) such that \(x \in B_3 \subseteq B_1 \cap B_2\).

Proof “Only if” Let \(\mathcal{B}\) be a basis for \(\mathcal{O}\). Since \(X \in \mathcal{O}\), then \(\forall x \in X\), \(\exists B \in \mathcal{B}\) such that \(x \in B \subseteq X\). So (i) is true. \(\forall B_1, B_2 \in \mathcal{B} \subseteq \mathcal{O}\), we have \(B_1 \cap B_2 \subseteq \mathcal{O}\). \(\forall x \in B_1 \cap B_2\), \(\exists B_3 \in \mathcal{B}\) such that \(x \in B_3 \subseteq B_1 \cap B_2\). Hence, (ii) is also true.

“If” Define \(\mathcal{O} := \{O \subseteq X \mid \forall x \in O, \exists B \in \mathcal{B}\) such that \(x \in B \subseteq O\}\). Clearly, \(\mathcal{B} \subseteq \mathcal{O}\) and \(\emptyset \in \mathcal{O}\). \(X \in \mathcal{O}\) since (i). \(\forall O_1, O_2 \in \mathcal{O}, \forall x \in O_1 \cap O_2\), we have \(x \in O_1\) and \(x \in O_2\). Then, by the definition of \(\mathcal{O}\), \(\exists B_1, B_2 \in \mathcal{B}\) such that \(x \in B_1 \subseteq O_1\) and \(x \in B_2 \subseteq O_2\). Then, \(x \in B_1 \cap B_2\). By (ii), \(\exists B_3 \in \mathcal{B}\) such that \(x \in B_3 \subseteq B_1 \cap B_2 \subseteq O_1 \cap O_2\). Then, \(O_1 \cap O_2 \in \mathcal{O}\).

\(\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}\), where \(\Lambda\) is an index set, let \(O = \bigcup_{O \in \Lambda} O_\alpha\). \(\forall x \in O, \exists \alpha \in \Lambda\) such that \(x \in O_\alpha\). By the definition of \(\mathcal{O}\), \(\exists B \in \mathcal{B}\) such that \(x \in B \subseteq O_\alpha \subseteq O\). Hence, \(O \in \mathcal{O}\). Therefore, \(\mathcal{O}\) is a topology on \(X\). By the definition of \(\mathcal{O}\), \(\mathcal{B}\) is a basis for \(\mathcal{O}\). Note that \(\mathcal{O} \supseteq \mathcal{O}\) since \(\mathcal{O}\) is the weakest topology containing \(\mathcal{B}\). Then, \(\mathcal{B}\) is a basis for \(\mathcal{O}\).

This completes the proof of the proposition. \(\square\)

Example 3.19 For the real line \(\mathbb{R}\), let \(\mathcal{A} := \{\text{interval } (a,b) \mid a, b \in \mathbb{R}, a < b\}\). Then, the topology generated by \(\mathcal{A}, \mathcal{O}_{\mathbb{R}}\), is the usual topology on \(\mathbb{R}\) as we know before. By Proposition 3.18, \(\mathcal{A}\) is a basis for this topology.

Definition 3.20 A topological space \((X, \mathcal{O})\) is said to satisfy the first axiom of countability if there exists a countable basis at each \(x \in X\), i.e., \(\forall x \in X, \exists \mathcal{B}_x \subseteq \mathcal{O}\), which is countable, such that \(\forall O \in \mathcal{O}\) with \(x \in O\), we have \(\exists B \in \mathcal{B}_x\) with \(x \in B \subseteq O\); and \(\forall B \in \mathcal{B}_x\), we have \(x \in B\). In this case, we will say that the topological space is first countable. The topological space is said to satisfy the second axiom of countability if there exists a countable basis \(\mathcal{B}\) for \(\mathcal{O}\), in which case, we will say that it is second countable.
Clearly, a second countable topological space is also first countable.

Example 3.21 The real line is first countable, where a countable basis at any \( x \in \mathbb{R} \) consists of intervals of the form \((x - r, x + r)\) with \( r \in \mathbb{Q} \) and \( r > 0 \). The real line is also second countable, where a countable basis for the topology consists of intervals of the form \((r_1, r_2)\) with \( r_1, r_2 \in \mathbb{Q} \) and \( r_1 < r_2 \).

When basis are available on topological spaces \( X \) and \( Y \), in Definitions 3.8 and 3.14, we may restrict the open sets \( O_Y \) and \( U \) to be basis open sets without changing the meaning of the definition. In Proposition 3.3, we may restrict \( O \) to be a basis open set and the results still hold.

Definition 3.22 A collection \((A_\alpha)_{\alpha \in \Lambda}\) of sets, where \( \Lambda \) is an index set, is said to be a covering of a set \( X \) if \( X \subseteq \bigcup_{\alpha \in \Lambda} A_\alpha \). It is an open covering if \( A_\alpha \)'s are open sets in a specific topological space.

Definition 3.23 A topological space \( X \) is said to be Lindelöf if any open covering of \( X \) has a countable subcovering.

Proposition 3.24 A second countable topological space \( X \) is Lindelöf.

Proof Let \((B_i)_{i \in \mathbb{N}}\) be a countable basis for \( X \), where \( N \) is a countable index set. Let \((O_\alpha)_{\alpha \in \Lambda}\) be an open covering of \( X \), where \( \Lambda \) is an index set. \( \forall \alpha \in \Lambda, O_\alpha = \bigcup_{i \in N_\alpha} B_i \), where \( N_\alpha \subseteq N \). Then, \( X = \bigcup_{\alpha \in \Lambda} \bigcup_{i \in N_\alpha} B_i \). We may determine a subcollection \((B_i)_{i \in \bar{N}}\) where \( \bar{N} \subseteq N \) such that these \( B_i \)'s appears at least once in the previous union. Then, \( X = \bigcup_{i \in \bar{N}} B_i \), \( \forall i \in \bar{N} \), the collection \( A_i := \{ O_\alpha \mid \alpha \in \Lambda, B_i \subseteq O_\alpha \} \) is nonempty. Then, by Axiom of Choice, there exists an assignment \((O_\alpha_i)_{i \in \bar{N}}\) such that \( O_\alpha_i \in A_i \), \( \forall i \in \bar{N} \). Then, we have \((O_\alpha_i)_{i \in \bar{N}}\) is a countable subcover. This completes the proof of the proposition.

3.4 Products of Topological Spaces

Proposition 3.25 Let \( \Lambda \) be an index set, and \((X_\alpha, O_\alpha)\) be a topological space, \( \forall \alpha \in \Lambda \). The product topology \( \mathcal{O} \) on \( \prod_{\alpha \in \Lambda} X_\alpha \) is the topology generated by

\[
\mathcal{B} := \left\{ \prod_{\alpha \in \Lambda} O_\alpha \mid O_\alpha \in O_\alpha, \forall \alpha \in \Lambda, \text{ and } O_\alpha = X_\alpha \text{ for all but a finite number of } \alpha \text{'s} \right\}
\]

The collection \( \mathcal{B} \) forms a basis for the topology \( \mathcal{O} \). We will also write \((\prod_{\alpha \in \Lambda} X_\alpha, \mathcal{O}) = \prod_{\alpha \in \Lambda}(X_\alpha, O_\alpha) \). When \((X_\alpha, O_\alpha) = (X, \mathcal{O}) =: (X, \forall \alpha \in \Lambda, \text{ we will denote } \prod_{\alpha \in \Lambda}(X_\alpha, O_\alpha) \text{ by } X^\Lambda.

3.4. PRODUCTS OF TOPOLOGICAL SPACES

Proof We will prove this by Proposition 3.18. Since $\prod_{\alpha \in \Lambda} X_{\alpha} \in \mathcal{B}$, then (i) is true. $\forall B_1, B_2 \in \mathcal{B}$, $B_1 = \prod_{\alpha \in \Lambda} O_{\alpha,1}$, $O_{\alpha,1} \in \mathcal{O}_{\alpha}$, $\forall \alpha \in \Lambda$, and $O_{\alpha,1} = X_{\alpha}$, $\forall \alpha \in \Lambda \setminus \Lambda_1$, where $\Lambda_1 \subseteq \Lambda$ is a finite set; $B_2 = \prod_{\alpha \in \Lambda} O_{\alpha,2}$, $O_{\alpha,2} \in \mathcal{O}_{\alpha}$, $\forall \alpha \in \Lambda$, and $O_{\alpha,2} = X_{\alpha}$, $\forall \alpha \in \Lambda \setminus \Lambda_2$, where $\Lambda_2 \subseteq \Lambda$ is a finite set. Let $B_3 = B_1 \cap B_2 = \prod_{\alpha \in \Lambda} (O_{\alpha,1} \cap O_{\alpha,2})$. Clearly, $O_{\alpha,1} \cap O_{\alpha,2} \in \mathcal{O}_{\alpha}$, $\forall \alpha \in \Lambda$, and $O_{\alpha,1} \cap O_{\alpha,2} = X_{\alpha}$, $\forall \alpha \in \Lambda \setminus (\Lambda_1 \cup \Lambda_2)$, where $\Lambda_1 \cup \Lambda_2 \subseteq \Lambda$ is a finite set. Hence, $B_3 \in \mathcal{B}$. Then, (ii) also holds. By Proposition 3.18, $\mathcal{B}$ is a basis for $\mathcal{O}$. This completes the proof of the proposition.

Definition 3.26 Let $X_{\alpha}$ be a topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set, and $\mathcal{Y} = \prod_{\alpha \in \Lambda} X_{\alpha}$ be the product topological space. Define a collection of projection functions $\pi_{\alpha} : \mathcal{Y} \to X_{\alpha}$, $\forall \alpha \in \Lambda$, by $\pi_{\alpha}(y) = y_{\alpha}$ for $y = (y_{\alpha})_{\alpha \in \Lambda} \in \mathcal{Y}$.

Proposition 3.27 Let $(X_{\alpha}, \mathcal{O}_{\alpha})$ be a topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set, and $(Y, \mathcal{O}) = \prod_{\alpha \in \Lambda} (X_{\alpha}, \mathcal{O}_{\alpha})$ be the product topological space. Then, $\mathcal{O}$ is the weakest topology on which the projection functions $\pi_{\alpha}$, $\forall \alpha \in \Lambda$, are continuous.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$. Case 1: $\Lambda = \emptyset$. $(Y, \mathcal{O}) = (\{\emptyset\}, \{\emptyset, \{\emptyset\}\})$. Then, $\mathcal{O}$ is the only topology on $Y$. Hence, the result is true.

Case 2: $\Lambda \neq \emptyset$. Define

$$A := \left\{ O \subseteq Y \mid O = \prod_{\alpha \in \Lambda} O_{\alpha}, \exists \alpha_0 \in \Lambda \text{ such that } O_{\alpha_0} \in \mathcal{O}_{\alpha_0} \text{ and } O_{\alpha} = X_{\alpha}, \forall \alpha \in \Lambda \setminus \{\alpha_0\} \right\}$$

Clearly, $A \neq \emptyset$ is the collection of inverse images of open sets under $\pi_{\alpha}$, $\forall \alpha \in \Lambda$. It is also clear that $A \subseteq B \subseteq \mathcal{O}$, where $B$ is the basis for $\mathcal{O}$ that we introduced in Proposition 3.25. $\forall B \in \mathcal{B}$, we have $B = \bigcap_{i=1}^k B_i$, where $k \in \mathbb{N}$ and $A_1, \ldots, A_k \in A$. Hence, the topology generated by $A$ contains $\mathcal{B}$. Then, the topology generated by $A$ equals to $\mathcal{O}$. Hence, $\mathcal{O}$ is the weakest topology containing $A$. This completes the proof of the proposition.

Proposition 3.28 Let $X_i := (X_i, \mathcal{O}_i)$ be a second countable topological space, $\forall i \in N \subseteq \mathbb{N}$, and $N$ is a countable index set. Then, the product topological space $X := (X, \mathcal{O}) := \prod_{i \in N} X_i$ is second countable.

Proof $\forall i \in N$, let $B_i \subseteq \mathcal{O}_i$ be a countable basis for $X_i$, without loss of generality, assume $X_i \subseteq B_i$. Let $B$ be the basis for the product topological space $X$ as defined in Proposition 3.25. Define

$$B := \left\{ \prod_{i \in N} B_i \subseteq X \mid B_i \in B_i, \forall i \in N, \text{ and } B_i = X_i \text{ for all but a finite number of } i \text{'s} \right\} \subseteq B \subseteq \mathcal{O}$$
Since $B_i$’s are countable, then $\bar{B}$ is countable. We will show that $\bar{B}$ is a basis for $O$. \(\forall O \in O, \forall x \in O, \exists O_B \in B\) such that $x \in O_B \subseteq O$, where $O_B = \prod_{i \in \mathbb{N}} O_{B_i}, O_{B_i} \in O_i, \forall i \in \mathbb{N}, O_{B_i} = X_i, \forall i \in N \setminus \bar{N}$, and $\bar{N} \subseteq N$ is a finite set. \(\forall i \in \bar{N}, \pi_i(x) \in O_{B_i}\). Then, \(\exists B_i \in B_i\) such that $\pi_i(x) \in B_i \subseteq O_{B_i}$. Let $B_i := X_i, \forall i \in N \setminus \bar{N}$. Let $B := \prod_{i \in \mathbb{N}} B_i \in \bar{B}$. Clearly, $x \in B \subseteq O_B \subseteq O$. Hence, $\bar{B}$ is a basis for $\mathcal{X}$. Then, $\mathcal{X}$ is second countable. This completes the proof of the proposition. \(\square\)

**Proposition 3.29** Let $X_\alpha := (X_\alpha, O_\alpha)$ be a topological space, \(\forall \alpha \in \Lambda\), where $\Lambda$ is an index set. Let $X := (X, O) := \prod_{\alpha \in \Lambda} X_\alpha$ be the product space. Assume that $A_\alpha \subseteq X_\alpha, \forall \alpha \in \Lambda$. Then, $\prod_{\alpha \in \Lambda} A_\alpha = \prod_{\alpha \in \Lambda} \bar{A}_\alpha$.

**Proof** Note that

\[ \left( \prod_{\alpha \in \Lambda} A_\alpha \right)^\sim = \bigcup_{\alpha \in \Lambda} \pi_\alpha \inv(\bar{A}_\alpha) \]

where $\pi_\alpha : \mathcal{X} \rightarrow X_\alpha$ is the projection function, $\forall \alpha \in \Lambda, \forall \alpha \in \Lambda$, by Definition 3.2, $\bar{A}_\alpha \in O_\alpha$. By Proposition 3.27, $\pi_\alpha$ is continuous. Then, $\pi_\alpha \inv(\bar{A}_\alpha) \in O$. Then, $\prod_{\alpha \in \Lambda} \bar{A}_\alpha$ is a closed set in $O$. Clearly, $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \bar{A}_\alpha$. Then, we have $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \bar{A}_\alpha$.

On the other hand, $\forall x \in \prod_{\alpha \in \Lambda} A_\alpha$, $\forall$ basis open set $O \in O$ with $x \in O$, by Proposition 3.25, we have $O = \prod_{\alpha \in \Lambda} O_\alpha$, where $O_\alpha \in O_\alpha, \forall \alpha \in \Lambda$, and $O_\alpha = X_\alpha$ for all but finitely many $\alpha$’s. $\forall \alpha \in \Lambda, \pi_\alpha(x) \in O_\alpha$ and $\pi_\alpha(x) \in \bar{A}_\alpha$. By Proposition 3.3, $O_\alpha \cap A_\alpha \neq \emptyset$. Then, by Axiom of Choice, we have $O \cap (\prod_{\alpha \in \Lambda} A_\alpha) = \prod_{\alpha \in \Lambda} (O_\alpha \cap A_\alpha) \neq \emptyset$. Thus, by Proposition 3.3, $x \in \prod_{\alpha \in \Lambda} A_\alpha$. Hence, we have $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \bar{A}_\alpha$.

Therefore, $\prod_{\alpha \in \Lambda} A_\alpha = \prod_{\alpha \in \Lambda} \bar{A}_\alpha$. This completes the proof of the proposition.

An immediate consequence of the above proposition is that if $A_\alpha \subseteq X_\alpha$ is closed, $\forall \alpha \in \Lambda$, where $X_\alpha$’s are topological spaces, then $\prod_{\alpha \in \Lambda} A_\alpha$ is closed in the product topology.

**Proposition 3.30** Let $(\Lambda_\beta)_{\beta \in \Gamma}$ be a collection of pairwise disjoint sets, where $\Gamma$ is an index set. Let $(X_\alpha, O_\alpha)$ be topological space, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$. $\forall \beta \in \Gamma, (\prod_{\alpha \in \Lambda_\beta} X_\alpha, O^{(\beta)})$ is the product space with the product topology. $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, O^{(\Gamma)})$ is the product space of product spaces with the product topology. Let $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, O)$ be the product space with the product topology. Then, $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, O^{(\Gamma)})$ and $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, O)$ are homeomorphic.

**Proof** Define a mapping $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \rightarrow \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$ by, $\forall x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta, \exists! \beta_\alpha \in \Gamma : \exists! \alpha \in \Lambda_\beta, \pi_\alpha(E(x)) = \pi_\alpha^{(\beta_\alpha)}(\pi_\alpha^{(\Gamma)}(x))$. 
∀y ∈ \prod_{α ∈ Δ_λ} X_α. ∀β ∈ Γ, define \( x_β ∈ \prod_{α ∈ Δ_β} X_α \) be such that 
\[
\pi_α^{(β)}(x_β) = π_α(y), \forall α ∈ Δ_β.
\]
Define \( x ∈ \prod_{β ∈ Γ} \prod_{α ∈ Δ_β} X_α \) by
\[
π^{(Γ)}_{α}(x) = x_β, \forall β ∈ Γ.
\]
Then, \( ∀α ∈ \bigcup_{β ∈ Γ} Δ_β, \exists β_α ∋ α \in Δ_β, \) and \( π_α(E(x)) = π_α^{(β_α)}(π^{(Γ)}_{β_α}(x)) = π_α(y) \) and hence \( E(x) = y \). Hence, \( E \) is surjective.

∀x, z ∈ \( \prod_{β ∈ Γ} \prod_{α ∈ Δ_β} X_α \) with \( E(x) = E(z) \). \( ∀α ∈ \bigcup_{β ∈ Γ} Δ_β, \exists β_α ∋ α \in Δ_β \). Let \( Γ_β := \{ β_α | α ∈ Δ_β \} \). Then, \( Γ_N \) is a finite set. \( ∀β ∈ Γ_N, \lambda_{Nβ} := λ_Ν \cap Δ_β \), which is a nonempty finite set. Then, \( λ_Ν \) equals to the disjoint union of \( \bigcup_{Γ ∈ Γ_Ν} λ_{Νβ} \).

Define \( B^{(β)} := \prod_{α ∈ Δ_β} B_α, ∀β ∈ Γ \). Then, \( ∀β ∈ Γ, B^{(β)} \) is a basis open set in \( \left( \prod_{α ∈ Δ_β} X_α, O^{(β)} \right) \). Define \( B^{(Γ)} := \prod_{β ∈ Γ} B^{(β)}, ∀β ∈ Γ \setminus Γ_N \), \( B^{(β)} = \prod_{α ∈ Δ_β} X_α \). Clearly, \( B^{(Γ)} \) is a basis open set in \( O^{(Γ)} \). \( ∀β ∈ Γ_N \), \( ∀α ∈ λ_{Νβ}, \pi_α^{(β)}(π^{(Γ)}_{β}(x_β)) = π_α(y_0) \in π_α(B) = B_α \). Then, \( x_β ∈ B^{(Γ)} \).

∀x ∈ B^{(Γ)}, ∀β ∈ Γ_N, \( ∀α ∈ λ_{Νβ}, \pi_α^{(β)}(π^{(Γ)}_{β}(x_β)) = π_α(y_0) \in π_α(B) = B_α \). Then, \( x_β ∈ B^{(Γ)} \).

∀x ∈ B^{(Γ)}, ∀β ∈ Γ_N, \( ∀α ∈ λ_{Νβ}, \pi_α^{(β)}(π^{(Γ)}_{β}(x_β)) = π_α(y_0) \in π_α(B) = B_α \). Then, \( x_β ∈ B^{(Γ)} \).

Finally, we will show \( E_{inv} \) is continuous by showing that, \( ∀y_0 ∈ \prod_{α ∈ Δ_λ} X_α, E_{inv} \) is continuous at \( y_0 \). \( ∀y_0 ∈ \prod_{α ∈ Δ_λ} X_α \). Let \( x_0 = E_{inv}(y_0) ∈ \prod_{β ∈ Γ} \prod_{α ∈ Δ_β} X_α \). For any basis open set \( B^{(Γ)} ∈ O^{(Γ)} \) with \( x_0 ∈ B^{(Γ)} \). Then, \( B^{(Γ)} = \prod_{β ∈ Γ} B^{(β)}, B^{(β)} ∈ O^{(β)}, ∀β ∈ Γ \), and \( B^{(β)} = \prod_{α ∈ Δ_β} X_α \) for all \( β ∈ Γ \) except finitely many \( β’s \), say \( β ∈ Γ_N \). \( ∀β ∈ Γ_N, π^{(Γ)}_{β}(x_0) ∈ B^{(β)} \). Then, there exists a basis open set \( B^{(β)} ∈ O^{(β)} \) such that \( π^{(Γ)}_{β}(x_0) ∈ B^{(β)} ⊆ B^{(β)} \). Then, \( B^{(β)} = \prod_{α ∈ Δ_β} B_α, B_α ∈ O_α, ∀α ∈ Δ_β, \) and \( B_α = X_α \) for all \( α ∈ Δ_λ \) except finitely many \( α’s \), say \( α ∈ λ_{Νβ} \). \( ∀α ∈ λ_{Νβ}, \pi_α(y_0) = π_α^{(β)}(π^{(Γ)}_{β}(x_0)) \in B_α \). Let \( λ_Ν = \bigcup_{Γ ∈ Γ_Ν} λ_{Νβ} \), which is a finite set and the union is pairwise disjoint.

Let \( B_α = X_α, ∀α ∈ Δ_β, ∀β ∈ Γ \setminus Γ_N \). Define \( B := \prod_{α ∈ Δ_λ} B_α \). Clearly, \( B \) is a basis open set in \( O \) and \( y_0 ∈ B \). \( ∀y ∈ B \). Let \( x = E_{inv}(y) \), then, \( y = E(x) \). \( ∀β ∈ Γ, ∀α ∈ Δ_β, \pi_α(y) = \pi_α(E(x)) = π_α^{(β)}(π^{(Γ)}_{β}(x)) \in B_α \).

Then, \( ∀β ∈ Γ_N, π^{(Γ)}_{β}(x) ∈ \prod_{α ∈ Δ_β} B_α = B^{(β)} ⊆ B^{(β)} \). Then, \( x ∈ B^{(Γ)} \).
Hence, \( E_{\text{inv}}(B) \subseteq B^{(Y)} \). Hence, \( E_{\text{inv}} \) is continuous at \( y_0 \). By the arbitrariness of \( y_0 \) and Proposition 3.9, we have \( E_{\text{inv}} \) is continuous.

This shows that \( E \) is a homeomorphism of \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}(Y) \) to \( \langle \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O} \rangle \), and completes the proof of the proposition. \( \square \)

**Proposition 3.31** Let \( \Lambda \) be an index set, \( X_\alpha := (X_\alpha, \mathcal{O}_{X_\alpha}) \) and \( Y_\alpha := (Y_\alpha, \mathcal{O}_{Y_\alpha}) \) be topological spaces, \( \alpha \in \Lambda \). Assume that \( X_\alpha \) and \( Y_\alpha \) are homeomorphic, \( \forall \alpha \in \Lambda \). Then, the product topological spaces \( X = \prod_{\alpha \in \Lambda} X_\alpha \) and \( Y = \prod_{\alpha \in \Lambda} Y_\alpha \) are homeomorphic.

**Proof** Since \( X_\alpha \) and \( Y_\alpha \) are homeomorphic, then, \( \exists H_\alpha : X_\alpha \to Y_\alpha \) such that \( H_\alpha \) is bijective and both \( H_\alpha \) and \( H_{\text{inv}} \alpha \) are continuous. Define mapping \( H : X \to Y \) by, \( \forall x \in \prod_{\alpha \in \Lambda} X_\alpha, \pi^{(Y)}_\alpha(H(x)) = H_\alpha(\pi^{(X)}_\alpha(x)), \forall \alpha \in \Lambda \).

\( \forall y \in Y, \) define \( x \in X \) by \( \pi^{(X)}_\alpha(x) = H_{\text{inv}}(\pi^{(Y)}_\alpha(y)), \forall \alpha \in \Lambda \). Then, \( H(x) = y \). Hence, \( H \) is surjective.

\( \forall x_1, x_2 \in X \) with \( x_1 \neq x_2 \). Then, \( \exists \alpha \in \Lambda \) such that \( \pi^{(X)}_{\alpha}(x_1) \neq \pi^{(X)}_{\alpha}(x_2) \). Then, \( \pi^{(Y)}_{\alpha}(H(x_1)) = H_\alpha(\pi^{(X)}_{\alpha}(x_1)) \neq H_\alpha(\pi^{(X)}_{\alpha}(x_2)) = \pi^{(Y)}_{\alpha}(H(x_2)), \) since \( H_{\alpha} \) is injective. Hence, \( H(x_1) \neq H(x_2) \). Therefore, \( H \) is injective. Therefore, \( H \) is continuous with inverse \( H_{\text{inv}} \).

Next, we show that \( H \) is continuous. For any basis open set \( O_Y \in \mathcal{O}_Y \). Then, \( O_Y = \prod_{\alpha \in \Lambda} O_{Y_\alpha} \) with \( O_{Y_\alpha} \in \mathcal{O}_{Y_\alpha}, \forall \alpha \in \Lambda \), and \( O_{Y_\alpha} = Y_\alpha \) for all \( \alpha \)'s except finitely many \( \alpha \)'s, say \( \alpha \in \Lambda_N \). Then, \( H_{\text{inv}}(O_Y) = \prod_{\alpha \in \Lambda} H_{\text{inv}}(O_{Y_\alpha}) \in \mathcal{O}_X \). Hence, \( H \) is continuous.

Finally, we show that \( H_{\text{inv}} \) is continuous. For any basis open set \( O_X \in \mathcal{O}_X \). Then, \( O_X = \prod_{\alpha \in \Lambda} O_{X_\alpha} \) with \( O_{X_\alpha} \in \mathcal{O}_{X_\alpha}, \forall \alpha \in \Lambda \), and \( O_{X_\alpha} = X_\alpha \) for all \( \alpha \)'s except finitely many \( \alpha \)'s, say \( \alpha \in \Lambda_N \). Then, \( H(O_X) = \prod_{\alpha \in \Lambda} H_\alpha(O_{X_\alpha}) \in \mathcal{O}_Y \). Hence, \( H_{\text{inv}} \) is continuous.

Hence, \( H \) is a homeomorphism. This completes the proof of the proposition. \( \square \)

**Proposition 3.32** Let \( \Lambda \) be an index set, \( X \) be a topological space, \( Y_\alpha := (Y_\alpha, \mathcal{O}_{Y_\alpha}) \) be a topological space, and \( f_\alpha : X \to Y_\alpha, \forall \alpha \in \Lambda \). Let \( Y = (Y, \mathcal{O}) := \prod_{\alpha \in \Lambda} Y_\alpha \) be the product topological space and \( f : X \to Y \) be given by, \( \forall x \in X, \pi_\alpha(f(x)) = f_\alpha(x), \forall \alpha \in \Lambda \). Then, \( f \) is continuous at \( x_0 \in X \) if, and only if, \( f_\alpha \) is continuous at \( x_0, \forall \alpha \in \Lambda \).

**Proof** “Sufficiency” \( \forall O \in \mathcal{O} \) with \( f(x_0) \in O \). By Proposition 3.25 and Definition 3.17, \( \exists \) a basis open set \( B = \prod_{\alpha \in \Lambda} B_\alpha \in \mathcal{O} \) such that \( f(x_0) \in B \subseteq \mathcal{O} \), where \( B_\alpha \in \mathcal{O}_{Y_\alpha}, \forall \alpha \in \Lambda \), and \( B_\alpha = Y_\alpha \) for all \( \alpha \)'s except finitely many \( \alpha \)'s, say \( \alpha \in \Lambda_N \). \( \forall \alpha \in \Lambda_N, f_\alpha(x_0) = \pi_\alpha(f(x_0)) \in \mathcal{O}(B) = B_\alpha \). By the continuity of \( f_\alpha \) at \( x_0, \exists \mathcal{O}_{U_\alpha} \in \mathcal{O}_X \) with \( x_0 \in U_\alpha \) such that \( f_\alpha(U_\alpha) \subseteq B_\alpha \). Let \( U := \bigcap_{\alpha \in \Lambda_N} U_\alpha \in \mathcal{O}_X \). Clearly \( x_0 \in U \). \( \forall x \in U \),
3.5. THE SEPARATION AXIOMS

∀α ∈ Λ_N, π_α(f(x)) = f_α(x) ∈ f_α(U) ⊆ f_α(U_α) ⊆ B_α. Hence, f(x) ∈ B and f(U) ⊆ B ⊆ O. Then, f is continuous at x_0.

“Necessity” ∀α ∈ Λ, the projection function π_α is continuous, by Proposition 3.27. Note that f_α = π_α ∘ f. Then, f_α is continuous at x_0 by Proposition 3.12.

This completes the proof of the proposition. □

3.5 The Separation Axioms

Definition 3.33 Let (X, O) be a topological space. It is said to be

T_1 (Tychonoff): ∀x, y ∈ X with x ≠ y, ∃O ∈ O such that x ∈ O and y ∈ O̅.

T_2 (Hausdorff): ∀x, y ∈ X with x ≠ y, ∃O_1, O_2 ∈ O such that x ∈ O_1, y ∈ O_2, and O_1 ∩ O_2 = ∅.

T_3 (regular): it is Tychonoff and, ∀x ∈ X, ∀F ⊆ X with F being closed and x ∈ F, ∃O_1, O_2 ∈ O such that x ∈ O_1, F ⊆ O_2, and O_1 ∩ O_2 = ∅.

T_4 (normal): it is Tychonoff and, ∀F_1, F_2 ⊆ X with F_1 and F_2 being closed and F_1 ∩ F_2 = ∅, ∃O_1, O_2 ∈ O such that F_1 ⊆ O_1, F_2 ⊆ O_2, and O_1 ∩ O_2 = ∅.

Note that (X, O) is Tychonoff implies that, ∀x ∈ X, the singleton set \{x\} is closed since \{x\} = \bigcup_{y \in X \setminus \{x\}} O_y, where O_y ∈ O and y ∈ O_y and x ∉ O_y. Then, it is clear that T_4 ⇒ T_3 ⇒ T_2 ⇒ T_1.

Proposition 3.34 A topological space (X, O) is Tychonoff if, and only if, ∀x ∈ X, the singleton set \{x\} is closed.

Proof “Only if” ∀x ∈ X, ∀y ∈ X with y ≠ x, we have ∃O_y ∈ O such that y ∈ O_y and x ∈ O̅_y. Then, \{x\} = \bigcup_{y \in \{x\}} O_y ∈ O. Hence, \{x\} is closed.

“If” ∀x, y ∈ X with x ≠ y, we have \{y\} ∈ O. Then, x ∈ \{y\} and y ∉ \{y\}. Hence, (X, O) is Tychonoff.

This completes the proof of the proposition. □

Proposition 3.35 Let (X, O) be a Tychonoff topological space. It is normal if, and only if, for all closed subset F ⊆ X and any open subset O ∈ O with F ⊆ U, ∃U ∈ O such that F ⊆ U ⊆ \overline{U} ⊆ O.

Proof “Necessity” Since (X, O) is normal, F, O̅ are closed, and F ∩ O̅ = ∅, then ∃O_1, O_2 ∈ O such that F ⊆ O_1, O̅ ⊆ O_2, and O_1 ∩ O_2 = ∅.

Then, we have F ⊆ O_1 ⊆ O_2 ⊆ O. Since O_2 is closed, then, O̅_1 ⊆ O̅_2. Therefore, we have F ⊆ O_1 ⊆ O̅_1 ⊆ O̅_2 ⊆ O. So, U = O_1.
“ Sufficiency ”. For any closed subsets \( F_1, F_2 \subseteq X \) with \( F_1 \cap F_2 = \emptyset \), we have \( F_1 \subseteq \overline{F_2} \). Then, \( \exists U \in \mathcal{O} \) such that \( F_1 \subseteq U \subseteq \overline{U} \subseteq F_2 \). Let \( O_1 = U \in \mathcal{O} \) and \( O_2 = \overline{U} \in \mathcal{O} \). Clearly, \( O_1 \cap O_2 = \emptyset \), \( F_1 \subseteq O_1 \), and \( F_2 \subseteq O_2 \). Hence, \( (X, \mathcal{O}) \) is normal.

This completes the proof of the proposition. \( \square \)

It is easy to show that the product topological space of Hausdorff topological spaces is Hausdorff.

### 3.6 Category Theory

**Definition 3.36** In a topological space \((X, \mathcal{O})\), a subset \( D \subseteq X \) is said to be dense if \( \overline{D} = X \). \((X, \mathcal{O})\) is said to be separable if there exists a countable dense subset \( D \subseteq X \). A subset \( M \subseteq X \) is said to be nowhere dense if \( \overline{M}^\circ = \overline{M} \) is dense in \( X \). A subset \( F \subseteq X \) is said to be of first category (or meager) if \( F \) is the countable union of nowhere dense subsets of \( X \). A subset \( S \subseteq X \) is said to be of second category (or nonmeager) if \( S \) is not meager. A subset \( H \subseteq X \) is said to be a residual set (or comeager) if \( H \) is meager. \((X, \mathcal{O})\) is said to be second category everywhere if every nonempty open subset of \( X \) is of second category.

**Proposition 3.37** Let \( \mathcal{X} \) be a topological space and \( Y \subseteq \mathcal{X} \) be dense. Then, \( \forall \mathcal{O} \in \mathcal{O}, \mathcal{O} \cap Y = \overline{\mathcal{O}} \).

**Proof** Clearly, \( \overline{\mathcal{O} \cap Y} \subseteq \overline{\mathcal{O}} \). \( \forall x \in \overline{\mathcal{O} \cap Y} \), \( \forall \mathcal{U} \in \mathcal{O} \) with \( x \in \mathcal{U} \), by Proposition 3.3, \( U \cap \mathcal{O} \neq \emptyset \). Then, \( (U \cap \mathcal{O}) \cap Y \neq \emptyset \) since \( \overline{Y} = \mathcal{X} \) and \( U \cap \mathcal{O} \in \mathcal{O} \). Hence, we have \( U \cap (O \cap Y) \neq \emptyset \). This implies that \( x \in \overline{U \cap (O \cap Y)} \), by Proposition 3.3. Hence, we have \( \overline{\mathcal{O}} \subseteq \overline{U \cap (O \cap Y)} \). Therefore, \( \overline{\mathcal{O}} = \overline{O \cap Y} \). This completes the proof of the proposition. \( \square \)

**Proposition 3.38** Let \( \mathcal{X} \) be a topological space. Then, \( \mathcal{X} \) is second category everywhere if, and only if, countable intersection of open dense subsets is dense.

**Proof** “Only if” Let \((O_n)_{n=1}^\infty\) be a sequence of open dense subsets of \( \mathcal{X} \). Let \( F_n := \overline{O_n}, \forall n \in \mathbb{N} \). Clearly, \( F_n \) is closed and nowhere dense, \( \forall n \in \mathbb{N} \), since \( \overline{F_n} = \overline{F_n} = \overline{O_n} = \mathcal{X} \). Note that \( \bigcap_{n=1}^\infty O_n = \bigcup_{n=1}^\infty F_n \). Now, suppose that \( \bigcap_{n=1}^\infty O_n \) is not dense. Then, \( \exists \mathcal{O} \in \mathcal{O}_\mathcal{X} \) which is nonempty such that \( \bigcap_{n=1}^\infty O_n \cap \mathcal{O} = \emptyset \). Then, \( \bigcap_{n=1}^\infty F_n \cap \mathcal{O} = \emptyset \). Hence, \( O = \bigcup_{n=1}^\infty (F_n \cap \mathcal{O}) \) is of first category. This contradicts with the fact that \( O \) is of second category. Therefore, \( \bigcap_{n=1}^\infty O_n \) must be dense.

The case of finite intersection can be converted to the above scenario by padding \( \mathcal{X} \) as additional open dense subsets.

“If” \( \forall \mathcal{O} \in \mathcal{O} \) with \( O \neq \emptyset \). For any countable collection \((E_\alpha)_{\alpha \in \Lambda}\) of nowhere dense subsets of \( \mathcal{X} \). \( \forall \alpha \in \Lambda, E_\alpha \) is open and dense. Then, \( \bigcap_{\alpha \in \Lambda} E_\alpha \)
3.6. CATEGORY THEORY

is dense. Then, \( O \cap \left( \bigcap_{\alpha \in \Lambda} \overline{E_\alpha} \right) \neq \emptyset \), which further implies that \( O \not\subset \bigcup_{\alpha \in \Lambda} E_\alpha \). Hence, \( O \not\subset \bigcup_{\alpha \in \Lambda} E_\alpha \). Hence, \( \mathcal{X} \) is of second category by the arbitrariness of \( (E_\alpha)_{\alpha \in \Lambda} \). Hence, \( \mathcal{X} \) is second category everywhere by the arbitrariness of \( O \).

This completes the proof of the proposition.

Proposition 3.39 Let \((X, \mathcal{O})\) be a topological space. \( F, O \subseteq X \) with \( O, \overline{F} \in \mathcal{O} \). Then,

(i) \( \overline{O} \setminus O \) and \( F \setminus F^\circ \) are nowhere dense;

(ii) if \((X, \mathcal{O})\) is second category everywhere, and \( F \) is of first category, then \( F \) is nowhere dense.

Proof

(i) \( \overline{O} \setminus O = \overline{O} \cap \overline{O} \) is closed. Note that

\[
\overline{O} \setminus O = \overline{O} \cup \overline{O} = \overline{O} \cup \overline{O} = X
\]

where we have applied Proposition 3.3 in the above. Hence \( \overline{O} \setminus O \) is nowhere dense.

\( F \setminus F^\circ = F \cap \overline{F^\circ} \) is closed. Note that

\[
\overline{F \setminus F^\circ} = F \cap \overline{F^\circ} = F \cup F^\circ = F \cup F^\circ = \overline{F^\circ} = \overline{F^\circ} = X
\]

where we have applied Proposition 3.3 in the above. Hence \( F \setminus F^\circ \) is nowhere dense.

(ii) Let \( F = \bigcup_{n=1}^\infty F_n \), where \( F_n \)'s are nowhere dense subsets of \( X \). Then, \( \overline{F_n} \) is open and dense in \((X, \mathcal{O})\). By Proposition 3.38, we have \( \bigcap_{n=1}^\infty \overline{F_n} \) is dense.

Since \( F \) is closed, then \( \overline{F} = F \), \( \forall m \in \mathbb{N} \). Note that \( F \supseteq \bigcup_{n=1}^m F_n \). Then, \( \overline{F} \supseteq \bigcup_{n=1}^m F_n = \bigcup_{n=1}^m \overline{F_n} \), by Proposition 3.3. Therefore,

\[
F = \overline{F} \supseteq \bigcup_{n=1}^\infty \overline{F_n} \supseteq \bigcup_{n=1}^\infty F_n = F
\]

This implies \( \overline{F} = \bigcup_{n=1}^\infty \overline{F_n} \). Hence, we have \( \overline{F} = \bigcap_{n=1}^\infty \overline{F_n} \), which is dense. Hence, \( F \) is nowhere dense.

This completes the proof of the proposition.

Proposition 3.40 Let \((X, \mathcal{O})\) be a topological space. Then,

(i) A closed set \( F \subseteq X \) is nowhere dense if, and only if, it contains no nonempty open subset;

(ii) A subset \( E \subseteq X \) is nowhere dense if, and only if, \( \forall O \in \mathcal{O} \) with \( O \neq \emptyset \), \( \exists U \in \mathcal{O} \) with \( U \neq \emptyset \) such that \( U \subseteq O \setminus E \).
Then, \( F \cap U = \emptyset \). This contradicts with the fact that \( \overline{F} = F = X \), by Proposition 3.3. Hence, the result holds.

“If” \( \forall x \in X, \forall U \in \mathcal{O} \) with \( x \in U \). Then, \( U \not\subseteq F \) implies \( U \cap \overline{F} \neq \emptyset \). This implies that \( x \) is a point of closure of \( \overline{F} \), by Proposition 3.3. Hence, we have \( \overline{F} = F = X \). Hence, \( F \) is nowhere dense.

(ii) “Only if” \( \forall O \in \mathcal{O} \) with \( O \neq \emptyset \). Note that \( O \) not exist nonempty open sets topological space that is second category everywhere. Let \( \mathcal{F} \) open and nonempty by the nonemptyness of \( O \) \( \mathcal{X} \) of continuous real-valued functions of \( \mathcal{X} \) Let \( \text{Proposition 3.40} \), \( \mathcal{E} \) arbitrariness of \( x \) we have made use of Proposition 3.3. Hence, \( \mathcal{E} \) \( \mathcal{E} \) is nowhere dense, that is \( \mathcal{E} \) is not nowhere dense. Then, \( O \) \( O \) \( O \)'s are not all nowhere dense. Then, \( O \) \( O \)'s are closed.

Definition 3.42 A topological space \( X \) is said to be connected if there do not exist nonempty open sets \( O_1, O_2 \) such that \( X = O_1 \cup O_2 \) and \( O_1 \cap O_2 = \emptyset \). Such a pair of \( O_1 \) and \( O_2 \) is called a separation of \( X \) if it exists.
**Proposition 3.43** Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $f : \mathcal{X} \to \mathcal{Y}$ be a surjective continuous function. If $\mathcal{X}$ is connected then $\mathcal{Y}$ is connected.

**Proof** Suppose $\mathcal{Y}$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}_Y$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = \mathcal{Y}$ and $O_1 \cap O_2 = \emptyset$. Since $f$ is surjective, then $f^{-1}(O_1) \neq \emptyset$ and $f^{-1}(O_2) \neq \emptyset$. Since $f$ is continuous, then $f^{-1}(O_1), f^{-1}(O_2) \in \mathcal{O}_X$. By Proposition 2.5, $f^{-1}(O_1) \cup f^{-1}(O_2) = \mathcal{X}$ and $f^{-1}(O_1) \cap f^{-1}(O_2) = f^{-1}(O_1 \cap O_2) = \emptyset$. This contradicts with the assumption that $\mathcal{X}$ is connected. Therefore, $\mathcal{Y}$ is connected. This completes the proof of the proposition.

**Theorem 3.44 (Mean Value Theorem)** Let $\mathcal{X}$ be a topological space and $f : \mathcal{X} \to \mathbb{R}$ be a continuous function. Assume that $\mathcal{X}$ is connected and $\exists x, y \in \mathcal{X}$ such that $f(x) < c < f(y)$ for some $c \in \mathbb{R}$. Then, $\exists z \in \mathcal{X}$ such that $f(z) = c$.

**Proof** Suppose the result is false. Then, $f^{-1}([c]) = \emptyset$. Let $O_1 := f^{-1}((c, \infty))$ and $O_2 := f^{-1}((\infty, c))$. Then, $x \in O_2 \in \mathcal{O}_X$ and $y \in O_1 \in \mathcal{O}_X$, by assumptions of the theorem. By Proposition 2.5, $O_1 \cap O_2 = \emptyset$ and $O_1 \cup O_2 = \mathcal{X}$. This contradicts with the assumption that $\mathcal{X}$ is connected. Hence, the result is true. This completes the proof of the theorem.

**Proposition 3.45** Let $\mathcal{X}$ be a topological space, $U \subseteq V \subseteq \overline{U} \subseteq \mathcal{X}$, and $U$ be connected in the subset topology $\mathcal{O}_U$. Then, $V$ is connected in the subset topology $\mathcal{O}_V$.

**Proof** Suppose $V$ is not connected in its subset topology. Then, $\exists O_{V_1}, O_{V_2} \in \mathcal{O}_V$ with $O_{V_1} \neq \emptyset$ and $O_{V_2} \neq \emptyset$, such that $O_{V_1} \cup O_{V_2} = V$ and $O_{V_1} \cap O_{V_2} = \emptyset$. By Proposition 3.4, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_{V_1} = O_1 \cap V$ and $O_{V_2} = O_2 \cap V$. Let $x_1 \in O_{V_1}$ and $x_2 \in O_{V_2}$. Then, $x_i \in U \cap O_i$, $i = 1, 2$. By Proposition 3.4, $O_{U_1}, O_{U_2} \in \mathcal{O}_U$. Note that $O_{U_1} \cap O_{U_2} = U \cap O_1 \cap O_2 = U \cap V \cap O_1 \cap O_2 = U \cap (O_1 \cap V) \cap (O_2 \cap V) = U \cap O_{V_1} \cap O_{V_2} = \emptyset$ and $O_{U_1} \cup O_{U_2} = U \cap (O_1 \cup O_2) = U \cap V \cap (O_1 \cup O_2) = U \cap (V \cap O_1) \cup (V \cap O_2) = U \cap (O_1 \cup O_2) = U \cap V = U$. Hence, the pair $(O_{U_1}, O_{U_2})$ is a separation of $U$. This implies that $U$ is not connected in its subset topology. This contradicts with the assumption. Hence, $V$ is connected.

**Definition 3.46** Let $\mathcal{X}$ be a topological space, $x_0 \in \mathcal{X}$. Let

$$\mathcal{M} := \{ M \subseteq \mathcal{X} \mid x_0 \in M, M \text{ is connected in the subset topology.} \}$$

The component containing $x_0$ is defined by $A := \bigcup_{M \in \mathcal{M}} M$.

Clearly, $\mathcal{X}$ is the union of its components.

**Proposition 3.47** Let $\mathcal{X}$ be a topological space and $x_0 \in \mathcal{X}$. Then, the component $A$ of $\mathcal{X}$ containing $x_0$ is connected and closed.
Definition 3.49 Let $\mathcal{M}$ be as defined in Definition 3.46. Suppose $A$ is not connected. Let $\mathcal{O}_A$ be the subset topology on $A$. Then, $\exists O_1, O_2 \in \mathcal{O}_A$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = A$ and $O_1 \cap O_2 = \emptyset$. Without loss of generality, assume that $x_0 \in O_1$. By the definition of $A$, $\exists A_0 \in \mathcal{M}$ such that $A_0 \cap O_2 \neq \emptyset$. Note that $A_0 = (O_1 \cap A_0) \cup (O_2 \cap A_0)$ and $O_1 \cap A_0 \ni x_0$ and $O_2 \cap A_0$ are nonempty and open in the subset topology on $A_0$. Furthermore, $(O_1 \cap A_0) \cap (O_2 \cap A_0) = \emptyset$. This shows that $A_0$ is not connected, which contradicts with the fact that $A_0 \in \mathcal{M}$. Hence, $A$ is connected.

By Proposition 3.45, $\overline{A}$ is connected. Then, we have $\overline{A} \in \mathcal{M}$ and $\overline{A} = A$. By Proposition 3.3, $A$ is closed in $\mathcal{O}$.

This completes the proof of the proposition.

\begin{proof}
Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $A_\alpha \subseteq X$ be connected (in subset topology), $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set. Assume that $A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$, $\forall \alpha_1, \alpha_2 \in \Lambda$. Then, $A := \bigcup_{\alpha \in \Lambda} A_\alpha$ is connected (in subset topology).

Suppose $A$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_1 \cap A \neq \emptyset, O_2 \cap A \neq \emptyset$, $(O_1 \cap A) \cap (O_2 \cap A) = \emptyset$, and $(O_1 \cup A) \cap (O_2 \cup A) = (O_1 \cup O_2) \cap A = A$. Let $x_i \in O_i \cap A$, $i = 1, 2$. Then, $\exists \alpha_i \in \Lambda$ such that $x_i \in A_{\alpha_i}$, $i = 1, 2$. By the assumption, let $x_0 \in A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset$. Without loss of generality, assume $x_0 \in O_1$. Then, $x_0 \in O_1 \cap A_{\alpha_2} \neq \emptyset$, $x_2 \in O_2 \cap A_{\alpha_2} \neq \emptyset$, $(O_1 \cap A_{\alpha_2}) \cap (O_2 \cap A_{\alpha_2}) \subseteq (O_1 \cap A) \cap (O_2 \cap A) = \emptyset$, and $(O_1 \cap A_{\alpha_2}) \cup (O_2 \cap A_{\alpha_2}) = (O_1 \cup O_2) \cap A_{\alpha_2} = A_{\alpha_2}$. Hence, $O_1 \cap A_{\alpha_2}$ and $O_2 \cap A_{\alpha_2}$ form a separation of $A_{\alpha_2}$. This implies that $A_{\alpha_2}$ is not connected. This is a contradiction. Therefore, $A$ is connected. This completes the proof of the proposition.
\end{proof}

Definition 3.49 Let $\mathcal{X}$ be a topological space. It is said to be locally connected if there exists a basis $\mathcal{B}$ such that $B$ is connected (in the subset topology), $\forall B \in \mathcal{B}$.

Proposition 3.50 Any component of a locally connected topological space is open.

Suppose that $X$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ with $O_1 \neq \emptyset$ and $O_2 \neq \emptyset$ such that $O_1 \cup O_2 = X$ and $O_1 \cap O_2 = \emptyset$. Let $\mathcal{B}$ be
the basis defined in Proposition 3.25. Then, there exists nonempty disjoint index sets $A_1$ and $A_2$ such that $O_1 = \bigcup_{\lambda \in A_1} \Gamma_\lambda$ and $O_2 = \bigcup_{\lambda \in A_2} \Gamma_\lambda$, where $B_\lambda \in \mathcal{B}$, $\forall \lambda \in A_1$, and $C_\lambda \in \mathcal{B}$, $\forall \lambda \in A_2$. Let $B_\lambda = \bigoplus_{\alpha \in \Gamma} \Gamma_\alpha$, where $B_\lambda \subset \Gamma_\alpha$, $\forall \alpha \in \Gamma$, and $B_\lambda = X_\alpha$ for all except finitely many $\alpha$, say $\alpha \in \Gamma_\lambda$. Note that $\Gamma_\lambda \neq \emptyset$, since $B_\lambda \subset O_1 \subset X$. $\forall \lambda \in \Gamma_2$, $C_\lambda = \bigoplus_{\alpha \in \Gamma} \Gamma_\alpha$, where $C_\lambda \subset \Gamma_\alpha$, $\forall \alpha \in \Gamma$, and $C_\lambda = X_\alpha$ for all except finitely many $\alpha$, say $\alpha \in \Gamma_\lambda$. Note that $\Gamma_\lambda \neq \emptyset$, since $C_\lambda \subset O_2 \subset X$. Note that

$$\emptyset = O_1 \cap O_2 = \bigcup_{\lambda \in A_1} \bigcup_{\gamma \in A_2} (B_\lambda \cap C_\gamma)$$

Therefore, $B_\lambda \cap C_\gamma = \emptyset$, $\forall \lambda \in A_1$, $\forall \gamma \in A_2$, which implies that $\exists \alpha, \gamma \in \Gamma$ such that $B_\lambda \cap C_\gamma = 0$.

Fix $x_1 \in O_1$, and $x_2 \in O_2$. Then, $x_1 \neq x_2$, $\exists \lambda \in A_1$ such that $x_1 \in B_\lambda$. Let $\Gamma_1 := \Gamma_\lambda$, $\exists \lambda \in A_2$ such that $x_2 \in C_\lambda$. Let $\Gamma_2 := \Gamma_\lambda$. Note that $\forall \gamma \in X$ with $\pi_\alpha(x) = \pi_\alpha(x_1)$, $\forall \alpha \in \Gamma_1$, we have $x \in B_\lambda \subset O_1$. Similarly, $\forall \gamma \in X$ with $\pi_\alpha(x) = \pi_\alpha(x_2)$, $\forall \alpha \in \Gamma_2$, we have $x \in C_\lambda \subset O_2$. Therefore, starting with $x_1 \in O_1$ and switching, one by one, its coordinate $\pi_\alpha(x_1)$ to $\pi_\alpha(x_2)$, for all $\alpha \in \Gamma_2$, we will end up with a point $x_2 \in O_2$. Therefore, there must exist a step in this process such that switching one coordinate $\pi_\alpha(x_1)$ to $\pi_\alpha(x_2)$, for some $\alpha_0 \in \Gamma_2$, leads to the change of set membership from $x_1 \in O_1$ before the switch to $x_2 \in O_2$ after the switch. In summary, there exist $x_1 \in O_1$, $x_2 \in O_2$, and $\alpha_0 \in \Gamma$ such that $\pi_\alpha(x_1) = \pi_\alpha(x_2)$, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$. Since $O_1 \cap O_2 = \emptyset$, we must have $\pi_{\alpha_0}(x_1) \neq \pi_{\alpha_0}(x_2)$.

Define

$$A_{10} := \{ \lambda \in A_1 \mid B_\lambda = \bigoplus_{\alpha \in \Gamma} \Gamma_\alpha, \pi_\alpha(x_1) \in B_\lambda, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

and

$$A_{20} := \{ \lambda \in A_2 \mid C_\lambda = \bigoplus_{\alpha \in \Gamma} \Gamma_\alpha, \pi_\alpha(x_2) \in C_\lambda, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

Let $M := \bigoplus_{\alpha \in \Gamma} \Gamma_\alpha \subset X$ where $M_\alpha = \{ \pi_\alpha(x_1) \}$, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$, and $M_{\alpha_0} = X_{\alpha_0}$. Note that $M \subset X = O_1 \cup O_2 = \left( \bigcup_{\lambda \in A_1} B_\lambda \right) \cup \left( \bigcup_{\gamma \in A_2} C_\gamma \right)$ implies that

$$M \subset \left( \bigcup_{\lambda \in A_{10}} B_\lambda \right) \cup \left( \bigcup_{\gamma \in A_{20}} C_\gamma \right) = \bigcup_{\lambda \in A_{10}} \bigcup_{\gamma \in A_{20}} (B_\lambda \cup C_\gamma)$$

Therefore, we have $X_{\alpha_0} = \pi_{\alpha_0}(M) \subset \bigcup_{\lambda \in A_{10}} \bigcup_{\gamma \in A_{20}} (\pi_{\alpha_0}(B_\lambda) \cup \pi_{\alpha_0}(C_\gamma)) = \bigcup_{\lambda \in A_{10}} \bigcup_{\gamma \in A_{20}} (B_{\alpha_0} \cup C_{\alpha_0}) \subset X_{\alpha_0}$, by Proposition 2.5. Then,

$$X_{\alpha_0} = \bigcup_{\lambda \in A_{10}} \bigcup_{\gamma \in A_{20}} (B_{\alpha_0} \cup C_{\gamma \alpha_0}) = \bigcup_{\lambda \in A_{10}} B_{\alpha_0} \cup \left( \bigcup_{\gamma \in A_{20}} C_{\gamma \alpha_0} \right) = D_1 \cup D_2$$
By derivations in the first paragraph of this proof, we have \( \forall \lambda \in \Lambda_{10}, \forall \gamma \in \Lambda_{20}, B_\lambda \cap C_\gamma = \prod_{\alpha \in \Gamma} (B_\lambda \cap C_{\gamma_\alpha}) = \emptyset \). Note that, \( \forall \alpha \in \Gamma \setminus \{\alpha_0\} \), \( \pi_\alpha(\vec{x}_1) = \pi_\alpha(\vec{x}_2) \in B_\lambda \cap C_{\gamma_\alpha} \neq \emptyset \). Then, we must have \( B_{\lambda_0} \cap C_{\gamma_0} = \emptyset \).

Therefore, we have

\[
D_1 \cap D_2 = \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_{\lambda_0} \cap C_{\gamma_0}) = \emptyset
\]

Clearly, \( D_1, D_2 \in O_{\alpha_0}, \pi_{\alpha_0}(\vec{x}_1) \in D_1 \neq \emptyset \), and \( \pi_{\alpha_0}(\vec{x}_2) \in D_2 \neq \emptyset \) since \( \vec{x}_1 \in O_1 \) and \( \vec{x}_2 \in O_2 \). This shows that \( D_1 \) and \( D_2 \) form a separation of \( X_{\alpha_0} \). This contradicts with the assumption that \( (X_{\alpha_0}, O_{\alpha_0}) \) is connected, \( \forall \alpha \in \Gamma \). Therefore, \((X, O)\) is connected.

This completes the proof of the proposition. \( \square \)

**Definition 3.52** Let \( X \) be a topological space and \( I := [0, 1] \subset \mathbb{R} \) be endowed with the subset topology of \( \mathbb{R} \). A curve in \( X \) is a continuous mapping \( \gamma : I \rightarrow X \). \( \gamma(0) \) is called the beginning point and \( \gamma(1) \) is called the end point. A closed curve is such that \( \gamma(0) = \gamma(1) \). Two closed curves \( \gamma_1 \) and \( \gamma_2 \) are said to be homotopic to each other if there exists a continuous function \( \phi : I \times I \rightarrow X \) such that \( \phi(t, 0) = \gamma_1(t) \), \( \phi(t, 1) = \gamma_2(t) \), and \( \phi(0, t) = \phi(1, t), \forall t \in I \).

**Definition 3.53** Let \( X \) be a topological space. It is said to be arcwise connected if \( \forall x_1, x_2 \in X \), there exists a curve \( \gamma \) in \( X \) such that \( \gamma(0) = x_1 \) and \( \gamma(1) = x_2 \). \( X \) is said to be simply connected if it is arcwise connected and any closed curve is homotopic to a single point (that is a degenerate curve \( \gamma \) with \( \gamma(t) = x \in X \), \( \forall t \in [0, 1] \subset \mathbb{R} \)).

**Proposition 3.54** Let \((X, O)\) be a topological space. Then, it is connected if it is arcwise connected.

**Proof** Suppose \((X, O)\) is not connected. Then, \( \exists O_1, O_2 \in O \) with \( O_1 \neq \emptyset \) and \( O_2 \neq \emptyset \) such that \( O_1 \cup O_2 = X \) and \( O_1 \cap O_2 = \emptyset \). Fix \( x_1 \in O_1 \) and \( x_2 \in O_2 \). Then, \( x_1 \neq x_2 \). Since \((X, O)\) is arcwise connected, then there exists a curve \( \gamma \) such that \( \gamma(0) = x_1 \) and \( \gamma(1) = x_2 \). Define \( t := \sup \{ s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1 \} \). Then, \( t \in [0, 1] \subset \mathbb{R} \).

**Claim 3.54.1** \( \gamma(t) \in O_1 \).

**Proof of claim:** Suppose \( \gamma(t) \notin O_1 \), then \( \gamma(t) \in O_2 \). Then, \( t > 0 \), since \( \gamma(0) = x_1 \in O_1 \). Since \( \gamma \) is continuous, then \( \exists t_1 \in [0, t) \subset \mathbb{R} \) such that \( \gamma([t_1, t]) \subseteq O_2 \). \( \forall s \in (t_1, 1] \subset \mathbb{R}, \gamma([0, s]) \cap O_2 \neq \emptyset \). Then, \( s \notin \{ s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1 \} \). Hence, \( s \geq t \), which implies that \( t_1 \geq t \). This contradicts \( t_1 < t \). Hence, \( \gamma(t) \in O_1 \). This completes the proof of the claim. \( \square \)

Since \( \gamma(1) = x_2 \in O_2 \), then \( t < 1 \). By the continuity of \( \gamma \), \( \exists \gamma \in (t, 1] \subset \mathbb{R} \) such that \( \gamma([t, t_2]) \subseteq O_1 \). \( \forall s \in [0, t), \exists s_1 \in (s, t] \) such that
3.8. **Continuous Real-Valued Functions**

Theorem 3.55 (Urysohn’s Lemma) Let \((X, \mathcal{O})\) be a normal topological space, \(A, B \subseteq X\) be closed subsets, and \(A \cap B = \emptyset\). Then, there exists a continuous real-valued function \(f : X \to [0,1] \subset \mathbb{R}\) such that \(f(x) = 0, \forall x \in A\), and \(f(x) = 1, \forall x \in B\).

**Proof**  Since the set \(Q := Q \cap [0,1]\) is countable, then, by recursively applying Proposition 3.35, we may find \((O_r)_{r \in Q} \subseteq \mathcal{O}\) such that the following two properties are satisfied:

1. \(\forall r \in Q, A \subseteq O_r \subseteq \overline{O_r} \subseteq B\);
2. \(\forall r, s \in Q\) with \(r < s\), \(\overline{O_r} \subseteq O_s\).

Define the real-valued function \(f : X \to \mathbb{R}\) by

\[
f(x) = \inf\{r \in Q \mid x \in O_r\} \cup \{1\}
\]

Clearly, \(f : X \to [0,1]\), \(f(x) = 0\), \(\forall x \in O_0\), and \(f(x) = 1\), \(\forall x \in \overline{O_1}\). By 1, we have \(A \subseteq O_0\) and \(O_1 \subseteq B\). Hence, all we need to show is that \(f\) is continuous. \(\forall x \in X\), we will show that \(f\) is continuous at \(x_0\). Let \(a_0 = f(x_0) \in [0,1]\). \(\forall U \subseteq \mathbb{R}\) with \(U\) being open and \(a_0 \in U\), \(\exists a_1, a_2, a_3, a_4 \in Q\) such that \(a_1 < a_2 < a_0 < a_3 < a_4\) and \((a_1, a_4) \subseteq U\). Let \(\tilde{a}_2 = \max\{a_2, 0\}\) and \(\tilde{a}_3 = \min\{a_3, 1\}\). Then, we must have \(a_1 < \tilde{a}_2 < a_0 < \tilde{a}_3 < a_4\) and \(\tilde{a}_2, \tilde{a}_3 \in Q\). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \(a_0 \in (0,1)\); Case 2: \(a_0 = 0\); Case 3: \(a_0 = 1\).

Case 1: \(a_0 \in (0,1)\). Then, we must have \(a_1 < \tilde{a}_2 < a_0 < \tilde{a}_3 < a_4\). Let \(V = \overline{O_{a_2}} \cap O_{a_3} \subseteq \mathcal{O}\). \(\forall x \in V\), we have \(x \in O_{a_3}\) and \(f(x) \leq \tilde{a}_2\). Also, \(x \in \overline{O_{a_2}}\) implies that \(f(x) \geq \tilde{a}_2\). Hence, \(f(V) \subseteq [\tilde{a}_2, \tilde{a}_3] \subseteq (a_1, a_4) \subseteq U\). \(f(x_0) = a_0 < \tilde{a}_3\) implies that \(x_0 \in O_{a_3}\). \(f(x_0) = a_0 > \tilde{a}_2\) implies that \(\exists \tilde{a}_2 \in (\tilde{a}_2, a_0) \cap Q\) such that \(x_0 \in \overline{O_{\tilde{a}_2}} \subseteq \overline{O_{a_2}}\). Therefore \(x_0 \in V\). This shows that \(\exists V \in \mathcal{O}\) with \(x_0 \in V\) such that \(f(V) \subseteq U\).

Case 2: \(a_0 = 0\). Then, we must have \(a_1 < 0 = a_0 < \tilde{a}_3 < a_4\). Take \(V = O_{a_3} \subseteq \mathcal{O}\). We must have \(x_0 \in V\). \(\forall x \in V\), \(0 \leq f(x) \leq \tilde{a}_3\). Hence, \(f(V) \subseteq [0, \tilde{a}_3] \subseteq (a_1, a_4) \subseteq U\). Hence, \(\exists V \in \mathcal{O}\) with \(x_0 \in V\) such that \(f(V) \subseteq U\).

Case 3: \(a_0 = 1\). Then, we must have \(a_1 < \tilde{a}_2 < a_0 = 1 < a_4\). Take \(V = \overline{O_{a_2}} \subseteq \mathcal{O}\). Since \(f(x_0) = a_0 = 1\), then \(x_0 \in O_{a_0 + \epsilon} \subseteq \overline{O_{a_2}} = V\). \(\forall x \in V\),
$f(x) \geq a_2$. Hence, $f(V) \subseteq [a_2, 1) \subset (a_1, a_4) \subseteq U$. Hence, $\exists \, V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$.

Therefore, in all cases, $\exists \, V \in \mathcal{O}$ with $x_0 \in V$ such that $f(V) \subseteq U$. Hence, $f$ is continuous at $x_0$. By the arbitrariness of $x_0$ and Proposition 3.9, $f$ is continuous. This completes the proof of the theorem. \hfill \Box

**Proposition 3.56** Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $\mathcal{Y}$ be Hausdorff. $f_1 : \mathcal{X} \to \mathcal{Y}$ and $f_2 : \mathcal{X} \to \mathcal{Y}$ are continuous. Let $D \subseteq \mathcal{X}$ be dense. Assume that $f_1|_D = f_2|_D$. Then, $f_1 = f_2$.

**Proof** Suppose $f_1 \neq f_2$. Then, $\exists x \in \mathcal{X}$ such that $f_1(x) \neq f_2(x)$. Since $\mathcal{Y}$ is Hausdorff, then $\exists O_1, O_2 \in \mathcal{O}_Y$ such that $f_1(x) \in O_1$, $f_2(x) \in O_2$, and $O_1 \cap O_2 = \emptyset$. Since $f_1$ and $f_2$ are continuous, we have $U_1 := f_{1\text{inv}}(O_1) \in \mathcal{O}_X$ and $U_2 := f_{2\text{inv}}(O_2) \in \mathcal{O}_X$. Note that $x \in U_1 \cap U_2 \in \mathcal{O}_X$ and $x \in D$, then, by Proposition 3.3, $\exists \bar{x} \in D \cap U_1 \cap U_2$. Then, $f_1(\bar{x}) \in O_1$ and $f_2(\bar{x}) \in O_2$, which implies that $f_1|_D(\bar{x}) \neq f_2|_D(\bar{x})$. This is a contradiction. Hence, we must have $f_1 = f_2$.

This completes the proof of the proposition. \hfill \Box

**Theorem 3.57** (Tietze’s Extension Theorem) Let $(X, \mathcal{O})$ be a normal topological space, $A \subseteq X$ be closed, and $h : A \to \mathbb{R}$. Let $A$ be endowed with the subset topology $\mathcal{O}_A$. Assume that $h$ is continuous. Then, there exists a continuous function $k : X \to \mathbb{R}$ such that $k|_A = h$.

**Proof** Let $f := \frac{h}{1 + |h|}$. Then, $|f(x)| < 1$, $\forall x \in A$, and by Proposition 3.12, $f$ is continuous.

**Claim 3.57.1** Let $l : A \to \mathbb{R}$ be a continuous function such that $|l(x)| \leq c_1$ $\forall x \in A$, where $c_1 > 0$. Then, there exists a continuous function $g : X \to \mathbb{R}$ such that $|g(x)| \leq c_1/3$, $\forall x \in X$, and $|l(x) - g(x)| \leq 2c_1/3$, $\forall x \in A$.

**Proof of claim:** Let $B := \{ x \in A \mid l(x) \leq -c_1/3 \}$ and $C := \{ x \in A \mid l(x) \geq c_1/3 \}$. Then, $B$ and $C$ are closed sets in $\mathcal{O}_A$, by the continuity of $l$ and Proposition 3.10. Since $A$ is closed, then $B$ and $C$ are closed in $\mathcal{O}$, by Proposition 3.5. Clearly, $B \cap C = \emptyset$. By Urysohn’s Lemma, there exists a continuous function $g : X \to \mathbb{R}$ such that $|g(x)| \leq c_1/3$, $\forall x \in X$, $g(x) = -c_1/3$, $\forall x \in B$, and $g(x) = c_1/3$, $\forall x \in C$. Hence, $|l(x) - g(x)| \leq 2c_1/3$, $\forall x \in A$. This completes the proof of the claim. \hfill \Box

By repeated application of Claim 3.57.1, we may define $f_i : X \to \mathbb{R}$, $\forall i \in \mathbb{N}$, such that $f_i$ is continuous, $|f_i(x)| \leq \frac{2^{i-1}}{3^i}$, $\forall x \in X$, and $|f(x) - \sum_{k=1}^i f_k(x)| \leq \frac{2^i}{3^i}$, $\forall x \in A$.

Define $g : X \to \mathbb{R}$ by $g(x) = \lim_{i \in \mathbb{N}} \sum_{k=1}^i f_k(x)$, $\forall x \in X$. Clearly, $g$ is well-defined, $g|_A = f$, and $|g(x)| \leq \sum_{i=1}^\infty \frac{2^{i-1}}{3^i} = 1$, $\forall x \in X$. $\forall x_0 \in X$. 

3.8. CONTINUOUS REAL-VALUED FUNCTIONS

\[ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N} \text{ such that } |\sum_{i=N+1}^{\infty} f_i(x)| < \epsilon/3, \forall x \in X. \]

By the continuity of \( f_1, \ldots, f_N \) and Proposition 3.9, \( \exists U \in \mathcal{O} \) with \( x_0 \in U \) such that \( |\sum_{i=1}^{N} f_i(x) - \sum_{i=1}^{\infty} f_i(x_0)| < \epsilon/3, \forall x \in U \). Then, we have, \( \forall x \in U, \)

\[
|g(x) - g(x_0)| \leq \left| g(x) - \sum_{i=1}^{N} f_i(x) \right| + \left| \sum_{i=1}^{N} f_i(x) - \sum_{i=1}^{N} f_i(x_0) \right| + \left| \sum_{i=1}^{\infty} f_i(x_0) - g(x_0) \right| < \epsilon
\]

Therefore, \( g \) is continuous at \( x_0 \). Then, \( g \) is continuous, by the arbitrariness of \( x_0 \) and Proposition 3.9.

Let \( D := \{ x \in X \mid |g(x)| = 1 \} \). Clearly, \( D \) is a closed set, by Proposition 3.10. Note that \( A \cap D = \emptyset \), since \( g|_A = f \) and \( |f(x)| < 1 \), \( \forall x \in A \). Then, by Urysohn’s Lemma, there exists a continuous function \( \bar{g} : X \to [0, 1] \) such that \( \bar{g}|_A = 1 \) and \( \bar{g}|_D = 0 \). Define \( k : X \to \mathbb{R} \) by

\[
k(x) = \frac{\bar{g}(x) g(x)}{1 - (\bar{g}(x)) |g(x)|}, \quad \forall x \in X. \]

By Propositions 3.12 and 3.32 and the fact that \( 1 - \bar{g}(x) |g(x)| \neq 0, \forall x \in X \), we have \( k \) is continuous. \( \forall x \in A \),

\[
k(x) = \frac{g(x)}{1 - |g(x)|} = h(x). \]

Hence, \( k|_A = h \).

This completes the proof of the theorem. \( \square \)

**Definition 3.58** Let \( X \) be a set and \( \mathcal{F} \) be a collection of real-valued functions of \( X \). Then, there is the weakest topology on \( X \) such that all functions in \( \mathcal{F} \) are continuous. This topology is called the weak topology generated by \( \mathcal{F} \).

Let \( X \) be a set, \( I := [0, 1] \subset \mathbb{R} \), and \( \mathcal{F} \) be a collection of functions of \( X \) to \( I \) such that \( \forall x, y \in X \) with \( x \neq y \), \( \exists f \in \mathcal{F} \), we have \( f(x) \neq f(y) \). Each \( f \in \mathcal{F} \) is a point in \( I^X \) and \( \mathcal{F} \) can be identified with a subset of \( I^X \). The topology that \( \mathcal{F} \) inherits as a subspace of \( I^X \) is called the topology of pointwise convergence. Now, \( X \) can be identified with a subset of \( I^\mathcal{F} \) by, \( \forall x \in X, \pi_f(x) = f(x), \forall f \in \mathcal{F} \). Then, the topology of \( X \) as a subset of \( I^\mathcal{F} \) is the weak topology generated by \( \mathcal{F} \).

**Proposition 3.59** Let \( X \) be a topological space, \( I := [0, 1] \subset \mathbb{R} \), and \( \mathcal{F} \) be a collection of continuous functions of \( X \) to \( I \) such that \( \forall x, y \in X \) with \( x \neq y \), \( \exists f \in \mathcal{F} \), we have \( f(x) \neq f(y) \). Let \( E : X \to I^\mathcal{F} \) be the equivalence map given by, \( \forall x \in X, \pi_f(E(x)) = f(x), \forall f \in \mathcal{F} \). Then, \( E \) is continuous. Furthermore, if \( \forall \) closed set \( F \subseteq X \) and \( \forall x \in X \) with \( x \notin F \), \( \exists f \in \mathcal{F} \) with \( f(x) = 1 \) and \( f|_F = 0 \), then \( E : X \to E(X) \) is a homeomorphism.

**Proof** Fix a basis open set \( O \) in \( I^\mathcal{F} \) with \( E(x_0) \in O \). By Proposition 3.25, \( O = \prod_{f \in \mathcal{F}} O_f \), where \( O_f \in \mathcal{O}_I \) with \( O_f \) being the subset topology on \( I \), \( \forall f \in \mathcal{F} \), and \( O_f = I \) for all \( f \)'s except finitely many \( f \)'s, say...
Hence, \( f \in \mathcal{F}_N \). Let \( U = \bigcap_{f \in \mathcal{F}_N} f_{\text{inv}}(O_f) \in \mathcal{O}_X \). By \( E(x_0) \in O \), we have \( x_0 \in U \). 

\[ \forall x \in U, \text{ we have } \pi_f(E(x)) \in O_f = I, \forall f \in \mathcal{F} \setminus \mathcal{F}_N, \text{ and } \pi_f(E(x)) = f(x) \in O_f, \forall f \in \mathcal{F}_N. \text{ Hence, } E(x) \in O. \text{ Then, } E(U) \subseteq O. \text{ Therefore, } E \text{ is continuous at } x_0. \text{ By the arbitrariness of } x_0 \text{ and Proposition 3.9, } E \text{ is continuous.}

Under the additional assumption on \( \mathcal{F} \), we need to show that \( E \) is a homeomorphism between \( X \) and \( E(X) \). \( \forall x, y \in X \) with \( x \neq y, \exists f \in \mathcal{F} \) such that \( \pi_f(E(x)) = f(x) \neq f(y) = \pi_f(E(y)) \). Then, \( E(x) \neq E(y) \). Hence, \( E : X \to E(X) \) is injective. Clearly, \( E : X \to E(X) \) is surjective. Then, \( E : X \to E(X) \) is bijective and admits inverse \( E_{\text{inv}} : E(X) \to X \). 

\[ \forall x_0 \in X, \text{ we will show that } E_{\text{inv}} \text{ is continuous at } E(x_0). \forall O \in \mathcal{O}_X \text{ with } x_0 \in O. \widetilde{O} \text{ is closed and } x_0 \notin \widetilde{O}. \text{ Then, } \exists f_0 \in \mathcal{F} \text{ such that } f_0(x_0) = 1 \text{ and } f_0|\widetilde{O} = 0. \text{ Define } U = \bigcap_{f \in \mathcal{F}} U_f \subseteq I^F \text{ by } U_f = I, \forall f \in \mathcal{F} \setminus \{f_0\} \text{ and } U_{f_0} = (1/2, 1] \in O_f. \text{ Clearly, } U \text{ is open in } I^F. \text{ Clearly, } E(x_0) \in U. \forall x \in X \text{ with } E(x) \in U, \text{ we have } \pi_{f_0}(E(x)) = f_0(x) > 1/2. \text{ Then, } x \notin \widetilde{O} \text{ and } x \in O. \text{ This shows that } E_{\text{inv}}(E(X) \cap U) \subseteq O. \text{ Hence, } E_{\text{inv}} \text{ is continuous at } E(x_0). \text{ By the arbitrariness of } x_0 \text{ and Proposition 3.9, } E_{\text{inv}} : E(X) \to X \text{ is continuous. This implies that } E : X \to E(X) \text{ is a homeomorphism.}

This completes the proof of the proposition.

\[ \square \]

**Definition 3.60** A topological space \( X \) is said to be completely regular (or \( T_{3\frac{1}{2}} \)) if it is Tychonoff and \( \forall x_0 \in X \) and \( \forall \text{ closed set } F \subseteq X \) with \( x_0 \notin F \), there exists a continuous real-valued function \( f : X \to [0, 1] \) such that \( f(x_0) = 1 \) and \( f|_F = 0 \).

**Proposition 3.61** A normal topological space is completely regular. A completely regular topological space is regular.

**Proof** Let \( X \) be a normal topological space. Then, \( X \) is Tychonoff. 

\[ \forall x_0 \in X \text{ and } \forall \text{ closed set } F \subseteq X \text{ with } x_0 \notin F, \text{ we have } \{x_0\} \text{ is closed, by Proposition 3.34. By Urysohn’s Lemma, there exists a continuous real-valued function } f : X \to [0, 1] \text{ such that } f(x_0) = 1 \text{ and } f|_F = 0. \text{ Hence, } X \text{ is completely regular.}

Let \( X \) be a completely regular topological space. Then, \( X \) is Tychonoff. 

\[ \forall x_0 \in X \text{ and } \forall \text{ closed set } F \subseteq X \text{ with } x_0 \notin F, \text{ there exists a continuous real-valued function } f : X \to [0, 1] \text{ such that } f(x_0) = 1 \text{ and } f|_F = 0. \text{ Hence, } X \text{ is completely regular.}

Let \( X \) be a completely regular topological space. Then, \( X \) is Tychonoff. 

\[ \forall x_0 \in X \text{ and } \forall \text{ closed set } F \subseteq X \text{ with } x_0 \notin F, \text{ there exists a continuous real-valued function } f : X \to [0, 1] \text{ such that } f(x_0) = 1 \text{ and } f|_F = 0. \text{ Hence, } X \text{ is completely regular.}

Let \( O_1 := \{ x \in X \mid f(x) > 1/2 \} \text{ and } O_2 := \{ x \in X \mid f(x) < 1/2 \}. \text{ Then, } O_1, O_2 \in \mathcal{O} \text{ by the continuity of } f. \text{ Clearly, } x_0 \in O_1, F \subseteq O_2, \text{ and } O_1 \cap O_2 = \emptyset. \text{ Hence, } X \text{ is regular.}

This completes the proof of the proposition.

\[ \square \]

**Corollary 3.62** Let \( X \) be a completely regular topological space, \( I = [0, 1] \subseteq \mathbb{R} \), and \( \mathcal{F} := \{ f : X \to I \mid f \text{ is continuous} \}. \) Then, the equivalence map: \( E : X \to I^F \) defined by \( \pi_f(E(x)) = f(x), \forall x \in X, \forall f \in \mathcal{F}, \) is a homeomorphism between \( X \) and \( E(X) \subseteq I^F. \)
Proof Since \( X \) is completely regular, then \( X \) is Tychonoff and all singleton subset of \( X \) is closed. Then, it is easy to check that all assumptions in Proposition 3.59 are satisfied. Then, the result follows. This completes the proof of the corollary. \( \Box \)

3.9 Nets and Convergence

Definition 3.63 A directed system is a nonempty set \( A \) and a relation on \( A \), \( \prec \), such that

(i) \( \prec \) is transitive;

(ii) \( \forall \alpha, \beta \in A \), \( \exists \gamma \in A \) such that \( \alpha \prec \gamma \) and \( \beta \prec \gamma \).

A net is a mapping of a directed system \( A := (A, \prec) \) to a topological space \( X \). \( \forall \alpha \in A \), the image is \( x_\alpha \). The net is denoted by \( (x_\alpha)_{\alpha \in A} \), where we have abuse the notation to say \( \alpha \in A \) when \( \alpha \in A \). It is understood that the relation for \( A \) is \( \prec \), where we will ignore the subscript \( A \) if no confusion arises.

A point \( x \in X \) is a limit of the net \( (x_\alpha)_{\alpha \in A} \) if \( \forall O \in O \) with \( x \in O \), \( \exists \alpha_0 \in A \exists \cdot \forall \alpha \in A \) with \( \alpha_0 \prec \alpha \), we have \( x_\alpha \in O \). We also say that \( (x_\alpha)_{\alpha \in A} \) converges to \( x \).

A point \( x \in X \) is a cluster point of \( (x_\alpha)_{\alpha \in A} \) if \( \forall O \in O \) with \( x \in O \), \( \forall \alpha \in A \), \( \exists \beta \in A \) with \( \alpha \prec \beta \exists \cdot x_\beta \in O \).

Clearly, a limit point of a net is a cluster point of the net.

In Definition 3.63, we may restrict \( O \) to be a basis open set without changing the meaning of the definition.

Example 3.64 \((\mathbb{N}, \leq)\) is a directed system. A net over \((\mathbb{N}, \leq)\) corresponds to a sequence.

\( \diamond \)

Proposition 3.65 Let \( X \) be a topological space. Then, the following statements holds.

(i) \( X \) is Hausdorff if, and only if, for all net \((x_\alpha)_{\alpha \in A} \subseteq X\), there exists at most one limit point for the net. We then write \( x = \lim_{\alpha \in A} x_\alpha \) when the limit exists.

(ii) If \( X \) is Hausdorff, any convergent net \((x_\alpha)_{\alpha \in A} \subseteq X\) with limit \( x \in X \) has exactly one cluster point, which is \( x \).

Proof (i) “Only if” Suppose there exists a net \((x_\alpha)_{\alpha \in A} \subseteq X\) such that \( \exists x_A, x_B \in X \) with \( x_A \neq x_B \) and \( x_A \) and \( x_B \) are limit points of the net. Since \( X \) is Hausdorff, then \( \exists O_1, O_2 \in O \) such that \( x_A \in O_1, x_B \in O_2 \), and \( O_1 \cap O_2 = \emptyset \). Since \( x_A \) is the limit of the net, then \( \exists \alpha_1 \in A \), \( \forall \alpha \in A \) with \( \alpha_1 \prec \alpha \), we have \( x_\alpha \in O_1 \). Similarly, since \( x_B \) is the limit of the net, then
Proposition 3.66 Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $D \subseteq \mathcal{X}$ with subset topology $\mathcal{O}_D$, and $f : D \to \mathcal{Y}$. Then, the following are equivalent.

(i) $f$ is continuous at $x_0 \in D$;

(ii) $\forall$ net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with $x_0$ as a limit point, we have that the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$ has a limit point $f(x_0)$.

(iii) $\forall$ net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with $x_0$ as a cluster point, we have that the net $(f(x_\alpha))_{\alpha \in \mathcal{A}} \subseteq \mathcal{Y}$ has a cluster point $f(x_0)$.

Proof

(i) $\Rightarrow$ (ii). Fix a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with $x_0 \in D$ as a limit point. $\forall \mathcal{O} \in \mathcal{O}_Y$ with $f(x_0) \in \mathcal{O}_Y$. By the continuity of $f$ at $x_0$, $\exists \mathcal{O} \in \mathcal{O}_X$ with $x_0 \in \mathcal{O}_X$ such that $f(\mathcal{O}_X) \subseteq \mathcal{O}_Y$. Since $x_0$ is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$, this contradicts with the assumption that every net has at most one limit point. Therefore, $\mathcal{X}$ is Hausdorff.
then $\exists \alpha_0 \in \mathcal{A}$ such that, $\forall \alpha \in \mathcal{A}$ with $\alpha_0 \prec \alpha$, we have $x_\alpha \in O_\alpha$. Then, $f(x_\alpha) \in O_Y$. Hence, we have $f(x_0)$ is a limit point of $(f(x_\alpha))_{\alpha \in \mathcal{A}}$.

(ii) $\Rightarrow$ (i). Suppose $f$ is not continuous at $x_0 \in D$. Then, $\exists O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$ such that, $\forall O_X \in \mathcal{O}_X$ with $x_0 \in O_X$, we have $f(O_X) \not\subseteq O_Y$. Let $\mathcal{M} := \{O \in \mathcal{O}_X \mid x_0 \in O\}$. Clearly, $\mathcal{X} \in \mathcal{M}$ and $\mathcal{M} \neq \emptyset$. Define a relation $\prec$ on $\mathcal{M}$ by, $\forall O_1, O_2 \in \mathcal{M}$, we say $O_1 \prec O_2$ if $O_1 \supseteq O_2$. Clearly, $\prec$ is transitive on $\mathcal{M}$. $\forall O_1, O_2 \in \mathcal{M}$, let $O_3 = O_1 \cap O_2 \in \mathcal{O}_X$ and $x_0 \in O_3$. Then, $O_3 \in \mathcal{M}, O_1 \prec O_3$, and $O_2 \prec O_3$. Hence, $\mathcal{A} := (\mathcal{M}, \prec)$ is a directed system. $\forall O \in \mathcal{M}$, $f(O) \setminus O_Y \neq \emptyset$. By Axiom of Choice, we may define a net $(x_\alpha)_{\alpha \in \mathcal{A}}$ by $x_0 \in O \cap D$ with $f(x_0) \notin O_Y$. Clearly, $x_0$ is a limit point of $(x_\alpha)_{\alpha \in \mathcal{A}}$. Yet, $f(x_0) \in O_Y \in \mathcal{O}_Y$ and $f(x_\alpha) \notin O_Y$, $\forall O \in \mathcal{M}$. Then, $(x_0)$ is not a limit point of the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$. This contradicts with the assumption. Therefore, $f$ is continuous at $x_0$.

(i) $\Rightarrow$ (iii). Fix a net $(x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D$ with $x_0$ as a cluster point. $\forall O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$, by the continuity of $f$ at $x_0$, $\exists U \in O_X$ with $x_0 \in U$ such that $f(U) \subseteq O_Y$. By Definition 3.63, $\forall \alpha \in \mathcal{A}, \exists \alpha_0 \in \mathcal{A}$ with $\alpha \prec \alpha_0$, $x_{\alpha_0} \in U$. Then, $f(x_{\alpha_0}) \in O_Y$. Hence, $f(x_0)$ is a cluster point of the net $(f(x_\alpha))_{\alpha \in \mathcal{A}}$.

(iii) $\Rightarrow$ (i). Suppose $f$ is not continuous at $x_0$. Let $\mathcal{M} := \{O \in \mathcal{O}_X \mid x_0 \in O\}$. Clearly, $\mathcal{A} := (\mathcal{M}, \supseteq)$ is a directed system. $\exists O_Y \in \mathcal{O}_Y$ with $f(x_0) \in O_Y$ such that $\forall U \in \mathcal{M}$, we have $f(U) \not\subseteq O_Y$. By Axiom of Choice, we may assign to each $U \in \mathcal{M}$ an $x_U \in U \cap D$ such that $f(x_U) \in O_Y$. Consider the net $(x_U)_{U \in \mathcal{A}} \subseteq D$. Clearly, $x_0$ is a limit point of the net, and therefore is a cluster point of the net. Consider the net $(f(x_U))_{U \in \mathcal{A}}$. For the open set $O_Y \ni f(x_0), \forall U \in \mathcal{A}, f(x_U) \notin O_Y$. Then, $(x_0)$ is not a cluster point of $(f(x_U))_{U \in \mathcal{A}}$. This contradicts with the assumption. Therefore, $f$ must be continuous at $x_0$.

This completes the proof of the proposition. $\square$

**Proposition 3.67** Let $(X_\alpha, O_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set. Let $(X, \mathcal{O})$ be the product space $\prod_{\alpha \in \Lambda}(X_\alpha, O_\alpha)$. Let $(x_\beta)_{\beta \in \mathcal{A}} \subseteq X$ be a net. Then, $x_0 \in X$ is a limit point of $(x_\beta)_{\beta \in \mathcal{A}}$ if, and only if, $\forall \alpha \in \Lambda$, $\pi_\alpha(x_\alpha) \in X_\alpha$ is a limit point of $(\pi_\alpha(x_\beta))_{\beta \in \mathcal{A}}$.

**Proof** "Only if" $\forall \alpha \in \Lambda$, by Proposition 3.27, $\pi_\alpha$ is continuous. Then, $\pi_\alpha$ is continuous at $x_0 \in X$, by Proposition 3.9. By Proposition 3.66, $\pi_\alpha(x_\alpha)$ is a limit point of the net $(\pi_\alpha(x_\beta))_{\beta \in \mathcal{A}}$.

"If" Suppose that $x_0 \in X$ is not a limit point of the net $(x_\beta)_{\beta \in \mathcal{A}}$. Then, $\exists$ a basis open set $B \in \mathcal{O}$ with $x_0 \in B$ such that, $\forall \beta \in \mathcal{A}$, $\exists \beta_0 \prec \beta$, we have $x_\beta \notin B$. Then, $B = \prod_{\alpha \in \Lambda} O_\alpha, O_\alpha \in O_\alpha, \forall \alpha \in \Lambda$, and $O_\alpha = X_\alpha$ for all $\alpha$’s except finitely many $\alpha$’s, say $\alpha \in \Lambda_N$. Then, $\forall \beta_0 \in \mathcal{A}, \exists \beta \in \mathcal{A}$ with $\beta_0 \prec \beta$, we have $x_\beta \notin B$. This implies that $\pi_\alpha(x_\beta) \notin O_\alpha$, for some $\alpha \in \Lambda_N$. Then, by an argument of contradiction, we may show that $\exists \alpha_0$ such that, $\forall \beta_0 \in \mathcal{A}, \exists \beta \in \mathcal{A}$ with $\beta_0 \prec \beta$, we have $\pi_{\alpha_0}(x_\beta) \notin O_{\alpha_0}$. Hence, $\pi_{\alpha_0}(x_0) \in O_{\alpha_0}$ is not the limit of the net.
\[(\pi_{\alpha}(x_\beta))_{\beta \in \mathcal{A}}\]. This contradicts with the assumption. Hence, we have \(x_0\) is a limit point of \((x_\beta)_{\beta \in \mathcal{A}}\).

This completes the proof of the proposition. \(\square\)

**Proposition 3.68** Let \((X, \mathcal{O})\) be a topological space, \(E \subseteq X\), and \(x \in X\). \(x \in \overline{E}\) if, and only if, \(\exists\) a net \((x_\alpha)_{\alpha \in \mathcal{A}} \subseteq E\) such that \(x\) is a limit point of the net.

**Proof** "Only if" Let \(\mathcal{M} := \{O \in \mathcal{O} \mid x \in O\}\). Clearly, \(X \in \mathcal{M}\), then \(\mathcal{M} \neq \emptyset\). Clearly, \(\mathcal{A} := (\mathcal{M}, \supseteq)\) is a directed system. Since \(x \in \overline{E}\), then, by Proposition 3.3, \(\forall O \in \mathcal{A}\), \(O \cap E \neq \emptyset\). By Axiom of Choice, \(\exists\) a net \((x_\alpha)_{O \in \mathcal{A}} \subseteq E\) such that \(x_\alpha \in O \cap E\), \(\forall O \in \mathcal{A}\) with \(x \in O\), then \(O \in \mathcal{A}\). \(\forall O_1, O \in \mathcal{A}\) with \(O \supseteq O_1\), we have \(x_\alpha \in O_1 \cap E \subseteq O\). Hence, \(x\) is a limit point of \((x_\alpha)_{O \in \mathcal{A}}\).

"If" Let \((x_\alpha)_{\alpha \in \mathcal{A}} \subseteq E\) be the net such that \(x\) is a limit point of the net. \(\forall O \in \mathcal{O}\) with \(x \in O\), \(\exists\alpha_0 \in \mathcal{A}\), \(\forall\alpha \in \mathcal{A}\) with \(\alpha_0 \prec \alpha\), we have \(x_\alpha \in E \cap O\). Since \((x_\alpha)_{\alpha \in \mathcal{A}}\) is a net, then \(\exists\alpha_1 \in \mathcal{A}\) with \(\alpha_0 \prec \alpha_1\). Then, \(x_{\alpha_1} \in O \cap E \neq \emptyset\). By Proposition 3.3, \(x \in \overline{E}\).

This completes the proof of the proposition. \(\square\)

**Definition 3.69** Let \((X, \mathcal{O})\) be a topological space, \(\mathcal{A} := (\mathcal{A}, \prec)\) be a directed system, and \((x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X\) be a net. Let \(A_s \subseteq A\) be a subset with the same relation \(\prec\) as \(A\) such that \(\forall\alpha \in \mathcal{A}\), \(\exists\alpha_s \in \mathcal{A}\), such that \(\alpha \prec \alpha_s\). Then, \(\mathcal{A}_s := (A_s, \prec)\) is a directed system and \((x_\alpha)_{\alpha \in \mathcal{A}_s}\) is a net, which is called a subnet of \((x_\alpha)_{\alpha \in \mathcal{A}}\).

**Proposition 3.70** Let \((X, \mathcal{O})\) be a topological space and \((x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X\) be a net. Then, \(x_0 \in X\) is a limit point of \((x_\alpha)_{\alpha \in \mathcal{A}}\) if, and only if, any subnet \((x_\alpha)_{\alpha \in \mathcal{A}_s}\) has a limit point \(x_0\).

**Proof** "Only if" Since \(x_0 \in X\) is a limit point of \((x_\alpha)_{\alpha \in \mathcal{A}}\), then \(\forall O \in \mathcal{O}\) with \(x_0 \in O\), \(\exists\alpha_0 \in \mathcal{A}\) such that, \(\forall\alpha \in \mathcal{A}\) with \(\alpha_0 \prec \alpha\), we have \(x_\alpha \in O\). Let \((x_\alpha)_{\alpha \in \mathcal{A}_s}\) be a subnet. Then, \(\exists\alpha_{s_0} \in \mathcal{A}_s\) such that \(\alpha_0 \prec \alpha_{s_0}\). \(\forall\alpha_s \in \mathcal{A}_s\) with \(\alpha_0 \prec \alpha_s\), we have \(\alpha_0 \prec \alpha_s\) and \(x_{s_0} \in O\). Hence, \(x_0\) is a limit point of the subnet.

"If" Since \((x_\alpha)_{\alpha \in \mathcal{A}}\) is a subnet of itself, then it has limit \(x_0\).

This completes the proof of the proposition. \(\square\)

A cluster point of a subnet is clearly a cluster point of the net.

**Proposition 3.71** Let \((X, \mathcal{O})\) be a topological space and \((x_\alpha)_{\alpha \in \mathcal{A}} \subseteq X\) be a net. Then, \(x_0 \in X\) is a limit point of \((x_\alpha)_{\alpha \in \mathcal{A}}\) if, and only if, for every subnet \((x_\alpha)_{\alpha \in \mathcal{A}_s}\) of \((x_\alpha)_{\alpha \in \mathcal{A}}\), there exists a subnet \((x_\alpha)_{\alpha \in \mathcal{A}_{s_0}}\) that has a limit point \(x_0\).

**Proof** "Sufficiency" We assume that every subnet \((x_\alpha)_{\alpha \in \mathcal{A}_s}\) of \((x_\alpha)_{\alpha \in \mathcal{A}}\), there exists a subnet \((x_\alpha)_{\alpha \in \mathcal{A}_{s_0}}\) that has a limit point \(x_0\). We
will prove the result using an argument of contradiction. Suppose \( x_0 \) is not a limit point of \( (x_\alpha)_{\alpha \in \mathcal{A}} \). Then, \( \exists O_0 \in \mathcal{O} \) with \( x_0 \in O_0, \forall \alpha \in \mathcal{A}, \exists x \in \mathcal{A} \) with \( \alpha_0 < \alpha \) such that \( x_\alpha \in \overline{O}_0 \). Define \( \mathcal{A}_x := \{ \alpha \in \mathcal{A} \mid x_\alpha \in \overline{O}_0 \} \), \( \prec \).

Clearly, \( (x_\alpha)_{\alpha \in \mathcal{A}_x} \) is a subnet of \( (x_\alpha)_{\alpha \in \mathcal{A}} \). Any subnet \( (x_\alpha)_{\alpha \in \mathcal{A}_{ss}} \) of \( (x_\alpha)_{\alpha \in \mathcal{A}_x} \), \( \forall \alpha \in \mathcal{A}_{ss} \), we have \( x_{\alpha_{ss}} \in O_0 \). Then, \( x_0 \) is not a limit of \( (x_\alpha)_{\alpha \in \mathcal{A}_{ss}} \). This contradicts the assumption. Therefore, \( x_0 \) is a limit point of \( (x_\alpha)_{\alpha \in \mathcal{A}} \).

"Necessity" Let \( x_0 \) be a limit point of \( (x_\alpha)_{\alpha \in \mathcal{A}} \) and \( (x_\alpha)_{\alpha \in \mathcal{A}_x} \) be a subnet. By Proposition 3.70, \( x_0 \) is a limit point of \( (x_\alpha)_{\alpha \in \mathcal{A}_x} \), which is a subnet of itself. Then, the result holds.

This completes the proof of the proposition. \( \square \)

Definition 3.72 Let \( \mathcal{X} := (X, \mathcal{O}_X) \) and \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq \mathcal{X}, f : D \rightarrow \mathcal{Y}, \) and \( x_0 \in \mathcal{X} \) be an accumulation point of \( D \). \( y_0 \in \mathcal{Y} \) is said to be a limit point of \( f(x) \) as \( x \rightarrow x_0 \) if \( \forall O_Y \in \mathcal{O}_Y \) with \( y_0 \in O_Y \), \( \exists U \in \mathcal{O}_X \) with \( x_0 \in U \) such that \( f(U \setminus \{x_0\}) = f((D \cap U) \setminus \{x_0\}) \subseteq O_Y \).

We will also say that \( f(x) \) converges to \( y_0 \) as \( x \rightarrow x_0 \).

When basis are available on topological spaces \( \mathcal{X} \) and \( \mathcal{Y} \), in Definition 3.72, we may restrict the open sets \( O_Y \) and \( U \) to be basis open sets without changing the meaning of the definition.

Proposition 3.73 Let \( \mathcal{X} := (X, \mathcal{O}_X) \) and \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq \mathcal{X}, f : D \rightarrow \mathcal{Y}, \) and \( x_0 \in \mathcal{X} \) be an accumulation point of \( D \). If \( \mathcal{Y} \) is Hausdorff, then there is at most one limit point of \( f(x) \) as \( x \rightarrow x_0 \). In this case, we will write \( \lim_{x \rightarrow x_0} f(x) = y_0 \in \mathcal{Y} \) when the limit exists.

Proof Suppose \( f(x) \) admits limit points \( y_A, y_B \in \mathcal{Y} \) as \( x \rightarrow x_0 \) with \( y_A \neq y_B \). Since \( \mathcal{Y} \) is Hausdorff, then \( \exists U_A, U_B \in \mathcal{O}_Y \) such that \( y_A \in U_A, y_B \in U_B, \) and \( U_A \cap U_B = \emptyset \). Since \( y_A \) is a limit point of \( f(x) \) as \( x \rightarrow x_0 \), then \( \exists V_A \in \mathcal{O}_X \) with \( x_0 \in V_A \) such that \( f(V_A \setminus \{x_0\}) \subseteq U_A \). Since \( y_B \) is a limit point of \( f(x) \) as \( x \rightarrow x_0 \), then \( \exists V_B \in \mathcal{O}_X \) with \( x_0 \in V_B \) such that \( f(V_B \setminus \{x_0\}) \subseteq U_B \). Then, \( x_0 \in V := V_A \cap V_B \in \mathcal{O}_X \). Since \( x_0 \) is an accumulation point of \( D \), then \( \exists x \in (D \cap V) \setminus \{x_0\} \). Then, we have \( f(x) \in U_A \) since \( x \in (D \cap V_A) \setminus \{x_0\} \) and \( f(x) \in U_B \) since \( x \in (D \cap V_B) \setminus \{x_0\} \). Then, \( f(x) \in U_A \cap U_B \neq \emptyset \). This contradicts with \( U_A \cap U_B = \emptyset \). Hence, the result holds. This completes the proof of the proposition. \( \square \)

Proposition 3.74 Let \( \mathcal{X} := (X, \mathcal{O}_X) \) and \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq X \) with subset topology \( \mathcal{O}_D, f : D \rightarrow \mathcal{Y}, \) and \( x_0 \in D \). Then, the following statements are equivalent.

(i) \( f \) is continuous at \( x_0 \).

(ii) If \( x_0 \) is an accumulation point of \( D \), then \( f(x_0) \) is a limit point of \( f(x) \) as \( x \rightarrow x_0 \).
Proof

(i) $\Rightarrow$ (ii). This is straightforward.

(ii) $\Rightarrow$ (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0$ is not an accumulation point of $D$; Case 2: $x_0$ is an accumulation point of $D$.

Assume that $\lim_{x \to x_0} g(f(x)) = y_0$. We have $f(V) = \{f(x_0)\} \subseteq U$. Hence, $f$ is continuous at $x_0$.

Case 1: $x_0$ is an accumulation point of $D$. $\forall U \in O_Y$ with $f(x_0) \in U$, $\exists V \in O_X$ with $x_0 \in V$ such that $V \cap D = \{x_0\}$. $\forall U \in O_Y$ with $f(x_0) \in U$, we have $f(V) = \{f(x_0)\} \subseteq U$. Hence, $f$ is continuous at $x_0$.

Case 2: $x_0$ is an accumulation point of $D$. $\forall U \in O_Y$ with $f(x_0) \in U$, $\exists V \in O_X$ with $x_0 \in V$ such that $f(V \setminus \{x_0\}) \subseteq U$. Then, we have $f(V) \subseteq U$. Hence, $f$ is continuous at $x_0$.

In both cases, $f$ is continuous at $x_0$.

This completes the proof of the proposition. \hfill $\Box$

Proposition 3.75 Let $\mathcal{X} := (X, O_X)$, $\mathcal{Y} := (Y, O_Y)$, and $\mathcal{Z} := (Z, O_Z)$ be topological spaces, $D \subseteq X$, $f : D \to Y$, $x_0 \in X$ be an accumulation point of $D$, $y_0 \in Y$ be a limit point of $f(x)$ as $x \to x_0$, and $g : Y \to Z$ be continuous at $y_0$. Then, $g(y_0) \in Z$ is a limit point of $g(f(x))$ as $x \to x_0$. When $\mathcal{Y}$ and $\mathcal{Z}$ are Hausdorff, then we may write $\lim_{x \to x_0} g(f(x)) = g(\lim_{x \to x_0} f(x))$.

Proof

$\forall O_Z \in O_Z$ with $g(y_0) \in O_Z$, by the continuity of $g$ at $y_0$, $\exists O_Y \in O_Y$ with $y_0 \in O_Y$ such that $g(O_Y) \subseteq O_Z$. Since $y_0$ is the limit of $f(x)$ as $x \to x_0$, then $\exists O_X \in O_X$ with $x_0 \in O_X$ such that $f(O_X \setminus \{x_0\}) \subseteq O_Y$. Then, $g(f(O_X \setminus \{x_0\})) \subseteq O_Z$. Hence, $g(f(x))$ converges to $g(y_0)$ as $x \to x_0$. This completes the proof of the proposition. \hfill $\Box$

Proposition 3.76 Let $\mathcal{X}$ be a topological space, $\bar{D} \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of $D$, $\mathcal{Y}$ and $\mathcal{Z}$ be Hausdorff topological spaces, $\bar{D} \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of $D$, $f : \bar{D} \to D$, and $g : D \to \mathcal{Z}$.

Assume that

(i) $\exists O_0 \in O_X$ with $x_0 \in O_0$ such that $f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\}$;

(ii) $\lim_{x \to x_0} f(x) = y_0$ and $\lim_{y \to y_0} g(y) = z_0 \in \mathcal{Z}$.

Then, $\lim_{x \to x_0} g(f(x)) = z_0$.

Proof

$\forall O_Z \in O_Z$ with $z_0 \in O_Z$, by $\lim_{y \to y_0} g(y) = z_0$, $\exists O_Y \in O_Y$ with $y_0 \in O_Y$ such that $g(O_Y \setminus \{y_0\}) = g((O_Y \cap D) \setminus \{y_0\}) \subseteq O_Z$. By $\lim_{x \to x_0} f(x) = y_0$, $\exists O_X \in O_X$ with $x_0 \in O_X$ such that $f(O_X \setminus \{x_0\}) = f((O_X \cap \bar{D}) \setminus \{x_0\}) \subseteq O_Y$. Let $O_1 := O_0 \cap O_X \subseteq O_X$. Clearly, $x_0 \in O_1$.

Then, $\forall x \in (O_1 \cap D)$, we have $f(x) \in O_Y \cap (D \setminus \{y_0\}) = (O_Y \cap D) \setminus \{y_0\}$ and $g(f(x)) \in O_Z$. Hence, $g(f((O_1 \cap \bar{D}) \setminus \{x_0\})) \subseteq O_Z$. Hence, $\lim_{x \to x_0} g(f(x)) = z_0$. This completes the proof of the proposition. \hfill $\Box$

Proposition 3.77 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be Hausdorff topological spaces, $\bar{D} \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of $D$, $\bar{D} \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of $D$, $f : \bar{D} \to D$ be bijective, and $g : D \to \mathcal{Z}$. Assume that $\lim_{x \to x_0} f(x) = y_0$ and $\lim_{y \to y_0} f_{inv}(y) = x_0$. Then, $\lim_{x \to x_0} g(f(x)) = \lim_{y \to y_0} g(y)$ whenever one of the limits exists in $\mathcal{Z}$.
Proof. We will prove the result by distinguishing two exhaustive cases: Case 1: \( \lim_{y \to y_0} g(y) = z_0 \in Z \); Case 2: \( \lim_{x \to x_0} g(f(x)) = z_0 \in Z \).

Case 1: \( \lim_{y \to y_0} g(y) = z_0 \in Z \). We will further distinguish three exhaustive and mutually exclusive subcases: Case 1a: \( y_0 \notin D \); Case 1b: \( y_0 \in D \) and \( x_0 = f_{inv}(y_0) \); Case 1c: \( y_0 \in D \) and \( x_0 \neq f_{inv}(y_0) \). Case 1a: \( y_0 \notin D \). Then, \( (i) \) of Proposition 3.76 is satisfied with \( O_0 := X \). By Proposition 3.76, we have \( \lim_{x \to x_0} g(f(x)) = z_0 \in Z \). Case 1b: \( y_0 \in D \) and \( x_0 = f_{inv}(y_0) \). Let \( O_0 := X \). \( \forall x \in (O_0 \cap D) \setminus \{x_0\} \), we have \( D \ni f(x) \neq f(x_0) = y_0 \) by \( f \) being bijective. This implies that \( f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\} \).

Then, \( \lim_{x \to x_0} g(f(x)) = z_0 \in Z \) by Proposition 3.76. Case 1c: \( y_0 \in D \) and \( x_0 \neq f_{inv}(y_0) \). By \( X \) being Hausdorff, \( \exists O_0 \in \mathcal{O}_X \) such that \( x_0 \in O_0 \) and \( \bar{x}_0 \notin O_0 \). This leads to \( \forall x \in (O_0 \cap D) \setminus \{x_0\} \), we have \( D \ni f(x) \neq f(\bar{x}_0) = y_0 \) by \( f \) being bijective. This implies that \( f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\} \). Then, \( \lim_{x \to x_0} g(f(x)) = z_0 \in Z \) by Proposition 3.76. Hence, \( \lim_{x \to x_0} g(f(x)) = \lim_{y \to y_0} g(y) = z_0 \in Z \) in all three subcases. Hence, the result holds in this case.

Case 2: \( \lim_{x \to x_0} g(f(x)) = z_0 \in Z \). Define \( h : \tilde{D} \to Z \) by \( h(x) = g(f(x)) \), \( \forall x \in \tilde{D} \). Then, \( \lim_{x \to x_0} h(x) = z_0 \). By Case 1, we have \( \lim_{y \to y_0} h(f_{inv}(y)) = z_0 \in Z \). Then, \( \lim_{y \to y_0} g(f(f_{inv}(y))) = \lim_{y \to y_0} g(y) = z_0 \). Hence, the result holds in this case.

This completes the proof of the proposition.

Example 3.78 Let \( g : \mathbb{R} \to \mathbb{R} \). It is desired to calculate \( \lim_{y \to +\infty} g(y) \).

We will apply Proposition 3.77 to this calculation. Take \( X = \mathbb{R}, \mathcal{Y} = \mathbb{R}_c \), and \( Z = \mathbb{R}_c \). Let \( D := (-\infty, -1] \cup (0, +\infty) \subseteq \mathbb{R} = X \), \( x_0 = 0 \), \( D := \mathbb{R} \subseteq \mathbb{R}_c = \mathcal{Y} \), and \( y_0 = +\infty \). Then, \( g : D \to Z \). Define \( f : D \to D \) by \( f(x) = \begin{cases} 1/x & x > 0 \\ x + 1 & x \leq -1 \end{cases} \), \( \forall x \in \tilde{D} \). Clearly, \( f \) is bijective with \( f_{inv} : D \to \tilde{D} \) given by \( f_{inv}(y) = \begin{cases} 1/y & y > 0 \\ y - 1 & y \leq 0 \end{cases} \), \( \forall y \in D \). Clearly, \( \lim_{x \to 0} f(x) = +\infty \) and \( \lim_{y \to +\infty} f_{inv}(y) = 0 \). Then, by Proposition 3.77, \( \lim_{y \to +\infty} g(y) = \lim_{x \to 0} g(f(x)) \) whenever one of the limits exists in \( Z \).

Proposition 3.79 Let \( X := (X, \mathcal{O}_X) \) and \( Y := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq X \), \( x_0 \in X \) be an accumulation point of \( D \), \( y_0 \in Y \), and \( f : D \to Y \). Then, the following statements are equivalent.

\((i)\) \( y_0 \) is a limit point of \( f(x) \) as \( x \to x_0 \).

\((ii)\) \( \forall \) net \( (x_\alpha)_{\alpha \in A} \subseteq D \setminus \{x_0\} \) with \( x_0 \) as a limit, we have \( y_0 \) is a limit point of the net \( (f(x_\alpha))_{\alpha \in A} \).

Proof. \((i) \Rightarrow (ii)\). Fix any net \( (x_\alpha)_{\alpha \in A} \subseteq D \setminus \{x_0\} \) with \( x_0 \) as a limit. \( \forall U \in \mathcal{O}_Y \) with \( y_0 \in U \), by \((i)\), \( \exists V \in \mathcal{O}_X \) with \( x_0 \in V \) such that \( f(V \setminus \{x_0\}) \subseteq U \). \( \exists \alpha_0 \in A \) such that \( \forall \alpha \in A \) with \( \alpha \prec \alpha_0 \), we have \( x_\alpha \in V \). Then, \( x_\alpha \in (D \cap V) \setminus \{x_0\} \) and \( f(x_\alpha) \in U \). Hence, \( (f(x_\alpha))_{\alpha \in A} \) has a limit \( y_0 \).
(ii) \( \Rightarrow \) (i). Suppose \( y_0 \) is not a limit point of \( f(x) \) as \( x \to x_0 \). Then, \( \exists U_0 \in \mathcal{O}_X \) with \( y_0 \in U_0 \), \( \forall V \in \mathcal{O}_X \) with \( x_0 \in V \), we have \( f(V \setminus \{x_0\}) \not\subseteq U_0 \). Then, \( \exists x_V \in (D \cap V) \setminus \{x_0\} \) such that \( f(x_V) \not\in U_0 \). Let \( M := \{ V \in \mathcal{O}_X \mid x_0 \in V \} \) and \( \mathcal{A} := (M, \supseteq) \). Clearly, \( \mathcal{A} \) is a directed system. By Axiom of Choice, we may construct a net \( (x_V)_{V \in \mathcal{A}} \subseteq D \setminus \{x_0\} \). Clearly, \( x_0 \) is a limit point of this net. But, \( \forall V \in \mathcal{A} \), \( f(x_V) \not\in U_0 \). Hence, \( y_0 \) is not a limit point of the net \( (f(x_V))_{V \in \mathcal{A}} \). This contradicts with (ii). Therefore, \( y_0 \) is a limit point of \( f(x) \) as \( x \to x_0 \).

This completes the proof of the proposition. \( \square \)

**Example 3.80** The extended real line is \( \mathbb{R}_e := \mathbb{R} \cup \{ \infty \} \cup \{-\infty \} \). We assume that \( \forall x \in \mathbb{R}, \, -\infty < x < \infty \). We define the usual operations:

\[
x + \infty = \infty \quad -\infty + x = -\infty \\
x \cdot \infty = \infty; \quad \text{if } x > 0 \\
x \cdot (-\infty) = -\infty; \quad \text{if } x > 0 \\
\infty \cdot \infty = -\infty = -\infty \\
\infty \cdot (-\infty) = -\infty
\]

The operations \( \infty - \infty \) and \( 0 \cdot \infty \) are undefined.

On \( \mathbb{R}_e \), we introduce the countable collection of subsets of \( \mathbb{R}_e \), \( \mathcal{B}_{\mathbb{R}_e} \), as follows. \( \emptyset, \mathbb{R}, \mathbb{R}_e \in \mathcal{B}_{\mathbb{R}_e} \). \( \forall r_1, r_2 \in \mathbb{Q} \) with \( r_1 < r_2, \, [-\infty, r_1), (r_1, r_2), \, (r_2, +\infty] \in \mathcal{B}_{\mathbb{R}_e} \).

By Proposition 3.18, it is easy to show that \( \mathcal{B}_{\mathbb{R}_e} \) is a basis for a topology on \( \mathbb{R}_e \). This topology is denoted \( \mathcal{O}_{\mathbb{R}_e} \), which is the usual topology on \( \mathbb{R}_e \). It is easy to show that \( (\mathbb{R}_e, \mathcal{O}_{\mathbb{R}_e}) \) is an arcwise connected second countable Hausdorff topological space.

An important property of \( \mathbb{R}_e \) is that \( \forall E \subseteq \mathbb{R}_e, \, \sup_{x \in E} x \in \mathbb{R}_e \) and \( \inf_{x \in E} x \in \mathbb{R}_e \).

It is easy to see that \( \mathbb{R} \) as a subset of \( \mathbb{R}_e \) admits the subset topology \( \mathcal{O} \) that equals to the usual topology \( \mathcal{O}_\mathbb{R} \) on \( \mathbb{R} \).

\( \diamond \)

**Proposition 3.81** Let \( X \) be a set, \( f_1 : X \to \mathbb{R}, \, f_2 : X \to \mathbb{R}, \, g_1 : X \to \mathbb{R}_e, \) and \( g_2 : X \to \mathbb{R}_e \). Assume that \( g_1(x) + g_2(x) \in \mathbb{R}_e \) is well defined, \( \forall x \in X \). Then,

(i) \( \sup_{x \in X} g_1(x) \leq M \in \mathbb{R}_e \) if, and only if, \( \forall x \in X, \, g_1(x) \leq M \);
(ii) \( \inf_{x \in X} g_1(x) \geq m \in \mathbb{R}_e \) if, and only if, \( \forall x \in X, \, g_1(x) \geq m \);
(iii) \( \sup_{x \in X} g_1(x) > M \in \mathbb{R}_e \) if, and only if, \( \exists x \in X, \, g_1(x) > M \);
(iv) \( \inf_{x \in X} g_1(x) < m \in \mathbb{R}_e \) if, and only if, \( \exists x \in X, \, g_1(x) < m \);
(v) \( \sup_{x \in X} (-g_1)(x) = -\inf_{x \in X} g_1(x) \);
(vi) \( \sup_{x \in X} (f_1 + f_2)(x) \leq \sup_{x \in X} f_1(x) + \sup_{x \in X} f_2(x) \);
(viii) \( \forall \alpha \in (0, \infty) \subset \mathbb{R}, \sup_{x \in X} (\alpha g_1(x)) = \alpha \sup_{x \in X} g_1(x); \forall \alpha \in [0, \infty) \subset \mathbb{R}, \sup_{x \in X} (\alpha f_1(x)) = \alpha \sup_{x \in X} f_1(x) \) when \( \sup_{x \in X} f_1(x) \in \mathbb{R} \);

Proof This is straightforward, and is therefore omitted. \( \square \)

Definition 3.82 Let \((x_\alpha)_{\alpha \in A} \subseteq \mathbb{R}_e\) be a net. The limit superior and limit inferior of the net are defined by

\[
\limsup_{\alpha \in A} x_\alpha = \inf_{\alpha \in A} \sup_{\beta \in A, \alpha < \beta} x_\beta \quad \text{in} \quad \mathbb{R}_e.
\]

\[
\liminf_{\alpha \in A} x_\alpha = \sup_{\alpha \in A} \inf_{\beta \in A, \alpha < \beta} x_\beta \quad \text{in} \quad \mathbb{R}_e.
\]

Proposition 3.83 Let \((x_\alpha)_{\alpha \in A} \subseteq \mathbb{R}_e, (y_\alpha)_{\alpha \in A} \subseteq \mathbb{R}, \) and \((z_\alpha)_{\alpha \in A} \subseteq \mathbb{R}\) be nets over the same directed system. Then, we have

(i) \( \liminf_{\alpha \in A} x_\alpha \leq \limsup_{\alpha \in A} x_\alpha \);

(ii) \( -\liminf_{\alpha \in A} x_\alpha = \limsup_{\alpha \in A} (-x_\alpha) \);

(iii) \( \liminf_{\alpha \in A} x_\alpha = \limsup_{\alpha \in A} x_\alpha = L \in \mathbb{R}_e \) if, and only if, \( \lim_{\alpha \in A} x_\alpha = L \);

(iv) \( \limsup_{\alpha \in A} (y_\alpha + z_\alpha) \leq \limsup_{\alpha \in A} y_\alpha + \limsup_{\alpha \in A} z_\alpha, \) when the right-hand-side makes sense;

(v) \( \text{if} \lim_{\alpha \in A} y_\alpha = y \in \mathbb{R}, \text{then} \limsup_{\alpha \in A} (y_\alpha + z_\alpha) = y + \limsup_{\alpha \in A} z_\alpha. \)

Proof Let \( V_\alpha := \{ \beta \in A \mid \alpha < \beta \}, \forall \alpha \in A. \) Then, \( V_\alpha \neq \emptyset, \forall \alpha \in A, \) since \( A \) is a directed system, and \( V_\alpha \supseteq V_\beta, \forall \alpha, \beta \in A \) with \( \alpha < \beta. \)

(i) Let \( \ell := \liminf_{\alpha \in A} x_\alpha \in \mathbb{R}_e \) and \( L := \limsup_{\alpha \in A} x_\alpha \in \mathbb{R}_e. \) \( \forall m \in \mathbb{R} \) with \( m < \ell, \sup_{\alpha \in A} \inf_{\beta \in A, \alpha < \beta} x_\beta > m \) implies that, by Proposition 3.81, \( \exists \alpha_0 \in A \) such that \( \inf_{\beta \in V_{\alpha_0}} x_\beta > m. \) Then, \( \forall \alpha \in A, \exists \alpha_1 \in A \) such that \( \alpha_0 < \alpha_1 \) and \( \alpha < \alpha_1. \) Then, \( m < \inf_{\beta \in V_{\alpha_0}} x_\beta \leq \inf_{\beta \in V_{\alpha_1}} x_\beta \leq \sup_{\beta \in V_{\alpha_1}} x_\beta \leq \sup_{\beta \in V_{\alpha}} x_\beta. \) Hence, \( L \geq m. \) By the arbitrariness of \( m, \) we have \( L \geq \ell. \)

(ii) Note that, by Proposition 3.81,

\[
\limsup_{\alpha \in A} (-x_\alpha) = \inf_{\alpha \in A} \sup_{\beta \in V_{\alpha}} (-x_\beta) = \inf_{\alpha \in A} \left( -\inf_{\beta \in V_{\alpha}} x_\beta \right) = -\sup_{\alpha \in A} \inf_{\beta \in V_{\alpha}} x_\beta = -\liminf_{\alpha \in A} x_\alpha.
\]

(iii) "If" \( \forall m \in \mathbb{R} \) with \( m > L, \exists \alpha_0 \in A \) such that \( x_\alpha \in (-\infty, m), \) \( \forall \alpha \in A \) with \( \alpha < \alpha_0. \) Then, \( \sup_{\beta \in V_{\alpha_0}} x_\beta \leq m \) and \( \limsup_{\alpha \in A} x_\alpha \leq m. \) By the arbitrariness of \( m, \) we have \( \limsup_{\alpha \in A} x_\alpha \leq L. \) By (ii), we have
Proposition 3.85
Let \( x_0 \in X \) be an accumulation point of \( D \subseteq \mathcal{X} \), where \( f : D \to \mathbb{R} \) and \( g : D \to \mathbb{R} \). Then, we have

\[
\limsup_{x \to x_0} f(x) = \inf_{O \in \mathcal{O}} \sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \in \mathbb{R}_e
\]

\[
\liminf_{x \to x_0} f(x) = \sup_{O \in \mathcal{O}} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \in \mathbb{R}_e
\]
3.9. NETS AND CONVERGENCE

(i) \( \liminf_{x \to x_0} f(x) \leq \limsup_{x \to x_0} f(x) \);
(ii) \(- \liminf_{x \to x_0} f(x) = \limsup_{x \to x_0} (-f(x)) \);
(iii) \( \liminf_{x \to x_0} f(x) = \limsup_{x \to x_0} f(x) = L \in \mathbb{R} \) if, and only if, \( \lim_{x \to x_0} f(x) = L \);
(iv) \( \limsup_{x \to x_0} (f + g)(x) \leq \limsup_{x \to x_0} f(x) + \limsup_{x \to x_0} g(x) \), when the right-hand-side makes sense;
(v) if \( \lim_{x \to x_0} f(x) = y \in \mathbb{R} \), then \( \limsup_{x \to x_0} (f + g)(x) = y + \limsup_{x \to x_0} g(x) \).

Proof

(i) Let \( l := \liminf_{x \to x_0} f(x) \in \mathbb{R} \) and \( L := \limsup_{x \to x_0} f(x) \in \mathbb{R} \). \( \forall m \in \mathbb{R} \) with \( m < l \), \( \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) > m \)
implies that \( \exists U \in \mathcal{O} \) with \( x_0 \in U \) such that \( \inf_{x \in (D \cup U) \setminus \{x_0\}} f(x) > m \). \( \forall O \in \mathcal{O} \) with \( x_0 \in O \), we have \( x_0 \in V := O \cap U \in \mathcal{O} \) and \( (D \cap V) \setminus \{x_0\} \neq \emptyset \), since \( x_0 \) is an accumulation point of \( D \). Then, \( m < \inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq \liminf_{x \to x_0} f(x) \leq \inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \). Hence, \( L \geq m \). By the arbitrariness of \( m \), we have \( L \geq l \).

(ii) Note that, by Proposition 3.81,

\[
\limsup_{x \to x_0} (-f)(x) = \inf_{O \in \mathcal{O} \text{ with } x_0 \in O} \sup_{x \in (D \cap O) \setminus \{x_0\}} (-f(x)) = \inf_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) = \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) = \liminf_{x \to x_0} f(x)
\]

(iii) “If” \( \forall m \in \mathbb{R} \) with \( m > L \), \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( f(V \setminus \{x_0\}) \subseteq (-\infty, m) \). Then, \( \sup_{x \in (D \cup V) \setminus \{x_0\}} f(x) \leq m \) and \( \limsup_{x \to x_0} f(x) \leq m \).

By the arbitrariness of \( m \), we have \( \limsup_{x \to x_0} f(x) \leq L \). By (ii), we have

\[ -L = \lim_{x \to x_0} (-f)(x) \geq \limsup_{x \to x_0} (-f)(x) = -\liminf_{x \to x_0} f(x) \]

Then, by (i), we have \( L \leq \liminf_{x \to x_0} f(x) \leq \limsup_{x \to x_0} f(x) \leq L \).

“Only if” We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( L = -\infty \); Case 2: \( L \in \mathbb{R} \); Case 3: \( L = +\infty \). Case 1: \( L = -\infty \). \( \forall m \in \mathbb{R} \), \( \sup_{x \to x_0} f(x) < m \) implies that \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( \sup_{x \in (D \cup V) \setminus \{x_0\}} f(x) < m \). Then, \( f(V \setminus \{x_0\}) \subseteq (-\infty, m) \).

Hence, we have \( \lim_{x \to x_0} f(x) = -\infty = L \). Case 2: \( L \in \mathbb{R} \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( L = \liminf_{x \to x_0} f(x) \) implies that \( \exists V_1 \in \mathcal{O} \) with \( x_0 \in V_1 \) such that \( \inf_{x \in (D \cup V_1) \setminus \{x_0\}} f(x) > L - \epsilon \). Then, \( f(V_1 \setminus \{x_0\}) \subseteq (L - \epsilon, +\infty) \subseteq \mathbb{R} \).

\( L = \limsup_{x \to x_0} f(x) \) implies that \( \exists V_2 \in \mathcal{O} \) with \( x_0 \in V_2 \) such that \( \sup_{x \in (D \cup V_2) \setminus \{x_0\}} f(x) < L + \epsilon \). Then, \( f(V_2 \setminus \{x_0\}) \subseteq (-\infty, L + \epsilon) \).
Let \( V := V_1 \cap V_2 \in \mathcal{O} \). Clearly, \( x_0 \in V \) and, by Proposition 2.5, \( f(V \setminus \{x_0\}) \subseteq f(V_1 \setminus \{x_0\}) \cap f(V_2 \setminus \{x_0\}) \subseteq (L - \epsilon, L + \epsilon) \). Therefore, \( \lim_{x \to x_0} f(x) = L \). Case 3: \( L = +\infty \). \( \forall M \in \mathbb{R} \), \( \liminf_{x \to x_0} f(x) > M \) implies that \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( \inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) > M \). Then, \( f(V \setminus \{x_0\}) \subseteq (M, +\infty) \). Hence, we have \( \lim_{x \to x_0} f(x) = +\infty = L \).

(iv) Note that, \( \forall O \in \mathcal{O} \) with \( x_0 \in O \), by Proposition 3.81,

\[
\sup_{x \in (D \cap O) \setminus \{x_0\}} (f + g)(x) \leq \left( \sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) + \left( \sup_{x \in (D \cap O) \setminus \{x_0\}} g(x) \right)
\]

Then, by Proposition 3.81, we have, \( \forall O \in \mathcal{O} \) with \( x_0 \in O \),

\[
\limsup_{x \to x_0} (f + g)(x) \leq \left( \sup_{x \in (D \cap O) \setminus \{x_0\}} f(x) \right) + \left( \sup_{x \in (D \cap O) \setminus \{x_0\}} g(x) \right)
\]

\( \forall U \in \mathcal{O} \) with \( x_0 \in U \), we have \( x_0 \in V := U \cap O \in \mathcal{O} \). Then,

\[
\limsup_{x \to x_0} (f + g)(x) \leq \left( \sup_{x \in (D \cap (V \cap U)) \setminus \{x_0\}} f(x) \right) + \left( \sup_{x \in (D \cap (V \cap U)) \setminus \{x_0\}} g(x) \right)
\]

Hence, we have \( \limsup_{x \to x_0} (f + g)(x) \leq \limsup_{x \to x_0} f(x) + \limsup_{x \to x_0} g(x) \), when the right-hand-side makes sense.

(v) By (iv) and (iii), \( \limsup_{x \to x_0} (f + g)(x) \leq y + \limsup_{x \to x_0} g(x) \).

Note that \( \limsup_{x \to x_0} g(x) = \limsup_{x \to x_0} (f + g - f)(x) \leq \limsup_{x \to x_0} (f + g)(x) + \limsup_{x \to x_0} (-f)(x) \). Then, we have \( \limsup_{x \to x_0} (f + g)(x) \geq y + \limsup_{x \to x_0} g(x) \). Hence, we have \( \limsup_{x \to x_0} (f + g)(x) = y + \limsup_{x \to x_0} g(x) \).

This completes the proof of the proposition. \( \square \)

**Proposition 3.86** Let \( \mathcal{X} := (X, \mathcal{O}_X) \) be a topological space, \( D \subseteq X \) with the subset topology \( \mathcal{O}_D \), \( f : D \to \mathbb{R} \), and \( x_0 \in D \). Then, the following statements are equivalent.

(i) \( f \) is upper semicontinuous at \( x_0 \).

(ii) If \( x_0 \) is an accumulation point of \( D \), then \( \limsup_{x \to x_0} f(x) \leq f(x_0) \).

**Proof**

(i) \( \Rightarrow \) (ii). Let \( x_0 \) be an accumulation point of \( D \). By the upper semicontinuity of \( f \) at \( x_0 \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists O \in \mathcal{O}_X \) with \( x_0 \in O \) such that \( f(x) < f(x_0) + \epsilon, \forall x \in O \cap D \). Then, \( \sup_{x \in (O \cap D) \setminus \{x_0\}} f(x) \leq f(x_0) + \epsilon \), and \( \limsup_{x \to x_0} f(x) \leq f(x_0) + \epsilon \). By the arbitrariness of \( \epsilon \), we have \( \limsup_{x \to x_0} f(x) \leq f(x_0) \).

(ii) \( \Rightarrow \) (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_0 \) is not an accumulation point of \( D \); Case 2: \( x_0 \) is an accumulation point of \( D \). Case 1: \( x_0 \) is not an accumulation point of \( D \). \( \forall V \in \mathcal{O}_X \) with \( x_0 \in V \) such that \( V \cap D = \{x_0\} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), we have...
f(x) = f(x_0) < f(x_0) + \epsilon, \forall x \in V \cap D. Hence, f is upper semicontinuous at x_0.

Case 2: x_0 is an accumulation point of D. \forall \epsilon \in (0, \infty) \subset \mathbb{R},
\limsup_{x \to x_0} f(x) < f(x_0) + \epsilon implies that \exists O \in \mathcal{O} with x_0 \in O such that
sup_{x \in (O \cap D) \setminus \{x_0\}} f(x) < f(x_0) + \epsilon. Then, f(x) < f(x_0) + \epsilon, \forall x \in O \cap D.
Hence, f is upper semicontinuous at x_0.

In both cases, f is upper semicontinuous at x_0.

This completes the proof of the proposition.  

**Proposition 3.87** Let \mathcal{X} be a topological space, D \subseteq \mathcal{X}, x_0 \in \mathcal{X} be an accumulation point of D, and f : D \to \mathbb{R}. Assume there exists a net \( (x_\alpha)_{\alpha \in A} \subseteq D \setminus \{x_0\} \) such that \( \lim_{\alpha \in A} x_\alpha = x_0 \) and \( \liminf_{\alpha \in A} f(x_\alpha) = c \in \mathbb{R} \). Then, \( \liminf_{x \to x_0} f(x) \leq c \).

**Proof** By Definitions 3.84 and 3.82, we have
\[
\liminf_{x \to x_0} f(x) = \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x)
\]
\[
\liminf_{\alpha \in A} f(x_\alpha) = \sup_{\alpha \in A} \inf_{\beta \in A \text{ with } \alpha \prec \beta} f(x_\beta) = c
\]
\forall O \in \mathcal{O} \text{ with } x_0 \in O, \text{ since } \lim_{\alpha \in A} x_\alpha = x_0, \text{ then } \exists \alpha_0 \in A \text{ such that } \forall \alpha \in A \text{ with } \alpha_0 \prec \alpha, \text{ we have } x_\alpha \in O. \text{ Then, } x_\alpha \in (O \cap D) \setminus \{x_0\}. \text{ This leads to}
\[
\inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) \leq \inf_{\beta \in A \text{ with } \alpha_0 \prec \beta} f(x_\beta) \leq c
\]
Hence, we have \( \liminf_{x \to x_0} f(x) \leq c \). This completes the proof of the proposition.  

**Definition 3.88** Let \( A_i := (A_i, \prec_i) \) be a directed system, \( i = 1, 2 \), and \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space. Define a relation \( \prec \) on \( A_1 \times A_2 \) by, \( \forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in A_1 \times A_2 \), we say \( (\alpha_1, \beta_1) \prec (\alpha_2, \beta_2) \) if \( \alpha_1 \prec_1 \alpha_2 \) and \( \beta_1 \prec_2 \beta_2 \). It is easy to verify that \( A := (A_1 \times A_2, \prec) := (A_1, \prec_1) \times (A_2, \prec_2) \) is a directed system. A joint net is a mapping of the directed system \( A_1 \times A_2 \) to \( \mathcal{X} \), denoted by \( (x_{\alpha, \beta})_{(\alpha, \beta) \in A_1 \times A_2} \). The joint net is said to admit a (joint) limit point \( \hat{x} \in \mathcal{X} \) if it admits a limit point \( \hat{x} \) when viewed as a net over the directed system \( A \). When \( \mathcal{X} \) is Hausdorff, by Proposition 3.65, the joint net admits at most one joint limit point, which will be denoted by \( \lim_{(\alpha, \beta) \in A_1 \times A_2} x_{\alpha, \beta} \in \mathcal{X} \) if it exists.

**Proposition 3.89** Let \( X := [0, \infty) \subset \mathbb{R} \) with subset topology \( \mathcal{O} \), \( \mathcal{X} := (X, \mathcal{O}) \), and \( c \in (0, \infty) \subset \mathbb{R} \). Then, \( + : \mathcal{X} \times \mathcal{X} \to \mathcal{X} \) and \( c : \mathcal{X} \to \mathcal{X} \) are continuous.

**Proof** We will prove this using Proposition 3.9. \( \forall (x_1, x_2) \in X \times X \). We will distinguish four exhaustive and mutually exclusive cases. Case 1: \( x_1 < \infty \) and \( x_2 < \infty \). Then, \( x_1 + x_2 < \infty \). \forall basis open set \( U := (r_1, r_2) \cap X \in \mathcal{O} \)
with $x_1+x_2 \in U$, take $V_1 := (x_1-(x_1+x_2-r_1)/2, x_1+(r_2-x_1-x_2)/2) \cap X \in \mathcal{O}$ and $V_2 := (x_2-(x_1+x_2-r_1)/2, x_2+(r_2-x_1-x_2)/2) \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall(x_1, x_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous at $(x_1, x_2)$.

Case 2: $x_1 < \infty$ and $x_2 = \infty$. Then, $x_1 + x_2 = \infty$. $\forall$ basis open set $U := (r_1, \infty) \cap X \in \mathcal{O}$ with $x_1 + x_2 \in U$, take $V_1 := (x_1-1, x_1+1) \cap X \in \mathcal{O}$ and $V_2 := (r_1-x_1+1, \infty) \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall(x_1, x_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous at $(x_1, x_2)$.

Case 3: $x_1 = \infty$ and $x_2 < \infty$. By Case 2 and symmetry, $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous at $(x_1, x_2)$.

Case 4: $x_1 = \infty$ and $x_2 = \infty$. Then, $x_1 + x_2 = \infty$. $\forall$ basis open set $U := (r_1, \infty) \cap X \in \mathcal{O}$ with $x_1 + x_2 \in U$, take $V_1 := (r_1/2, \infty) \cap X \in \mathcal{O}$ and $V_2 := (r_1/2, \infty) \cap X \in \mathcal{O}$. Then, $V_1 \times V_2$ is an open set in $\mathcal{X} \times \mathcal{X}$ with $(x_1, x_2) \in V_1 \times V_2$ and $\forall(x_1, x_2) \in V_1 \times V_2$, we have $\bar{x}_1 + \bar{x}_2 \in U$. Hence, $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous at $(x_1, x_2)$.

In all cases, we have $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous at $(x_1, x_2)$. By Proposition 3.9, $+ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ is continuous.

$\forall x \in X$. We will distinguish two exhaustive and mutually exclusive cases. Case 1: $x < \infty$. Then, $cx < \infty$. $\forall$ basis open set $U := (r_1, r_2) \cap X \in \mathcal{O}$ with $cx \in U$, take $V := (r_1/c, r_2/c) \cap X \in \mathcal{O}$. Then, $x \in V$ and $\forall \bar{x} \in V$, we have $c\bar{x} \in U$. Hence, $c : \mathcal{X} \to \mathcal{X}$ is continuous at $x$.

Case 2: $x = \infty$. Then, $cx = \infty$. $\forall$ basis open set $U := (r_1, \infty) \cap X \in \mathcal{O}$ with $cx \in U$, take $V := (r_1/c, \infty) \cap X \in \mathcal{O}$. Then, $x \in V$ and $\forall \bar{x} \in V$, we have $c\bar{x} \in U$. Hence, $c : \mathcal{X} \to \mathcal{X}$ is continuous at $x$.

In both cases, we have $c : \mathcal{X} \to \mathcal{X}$ is continuous at $x$. By Proposition 3.9, $c : \mathcal{X} \to \mathcal{X}$ is continuous. This completes the proof of the proposition.

$\Box$
Chapter 4

Metric Spaces

4.1 Fundamental Notions

Definition 4.1 A metric space \((X, \rho)\) is a set \(X\) together with a metric \(\rho : X \times X \to \mathbb{R}\) such that, \(\forall x, y, z \in X\),

(i) \(\rho(x, y) \geq 0\);
(ii) \(\rho(x, y) = 0 \iff x = y\);
(iii) \(\rho(x, y) = \rho(y, x)\);
(iv) \(\rho(x, y) \leq \rho(x, z) + \rho(z, y)\).

Let \(S \subseteq X\). Then, \((S, \rho|_{S \times S})\) is also a metric space.

Example 4.2 \((\mathbb{R}, \rho)\), with \(\rho(x, y) = |x - y|, \forall x, y \in \mathbb{R}\), is a metric space. \((\mathbb{R}^n, \rho)\), with \(\rho(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^2\right)^{1/2}, \forall x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n\), is a metric space, where \(n \in \mathbb{N}\). \((\mathbb{R}^n, \rho)\), with \(\rho(x, y) = \sum_{i=1}^n |x_i - y_i|, \forall x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n\), is a metric space, where \(n \in \mathbb{N}\).

Proposition 4.3 Let \(X := (X, \rho)\) be a metric space, an open ball centered at \(x_0 \in X\) with radius \(r \in (0, \infty) \subseteq \mathbb{R}\) is defined by \(B_X (x_0, r) := \{x \in X \mid \rho(x, x_0) < r\}\).

The metric space generates a natural topology \(O\) on \(X\) with the basis collection \(\mathcal{B}\) given by \(\mathcal{B} := \{B_X (x_0, r) \mid x_0 \in X, r \in (0, \infty) \subseteq \mathbb{R}\}\).

The closed ball centered at \(x_0 \in X\) with radius \(r \in [0, \infty) \subseteq \mathbb{R}\) is defined by \(\overline{B}_X (x_0, r) := \{x \in X \mid \rho(x, x_0) \leq r\}\), which is closed.

Proof We will show that \(\mathcal{B}\) is the basis for its generated topology by Proposition 3.18. (i) \(\forall x \in X, x \in B_X (x, 1)\). (ii) \(\forall B_X (x_1, r_1), B_X (x_2, r_2) \in \mathcal{B}\), let \(x \in B_X (x_1, r_1) \cap B_X (x_2, r_2)\). Then, we have \(\rho(x, x_1) < r_1\) and \(\rho(x, x_2) < r_2\). Let \(r := \min\{r_1 - \rho(x, x_1), r_2 - \rho(x, x_2)\} \in (0, \infty) \subseteq \mathbb{R}\).
Then, \( x \in \mathcal{B}_X(x,r) \in \mathcal{B} \). \( \forall x_3 \in \mathcal{B}_X(x,r) \), we have \( \rho(x_3,x_1) \leq \rho(x_3,x) + \rho(x_1,x) < r_1 \) and \( \rho(x_3,x_2) \leq \rho(x_3,x) + \rho(x_2,x) < r_2 \). Hence, we have \( x_3 \in \mathcal{B}_X(x_1,r_1) \cap \mathcal{B}_X(x_2,r_2) \). Therefore, we have \( \mathcal{B}_X(x,r) \subseteq \mathcal{B}_X(x_1,r_1) \cap \mathcal{B}_X(x_2,r_2) \). Hence, the assumptions of Proposition 3.18 are satisfied, and then \( \mathcal{B} \) is a basis for \( \mathcal{O} \).

Next, we show that \( \overline{\mathcal{B}_X(x_0,r)} \) is a closed set. \( \forall x \in (\overline{\mathcal{B}_X(x_0,r)}) \), we have \( \rho(x,x_0) > r \). Let \( r_1 := \rho(x,x_0) - r \in (0, \infty) \subseteq \mathbb{R} \). \( \forall x_1 \in \mathcal{B}_X(x,r_1) \), we have \( \rho(x_1,x_0) \geq \rho(x,x_0) - \rho(x,x_1) > r \). Hence, we have \( x \in \mathcal{B}_X(x_1,r_1) \subseteq (\overline{\mathcal{B}_X(x_0,r)}) \). Therefore, \( (\overline{\mathcal{B}_X(x_0,r)}) \) is open and \( \overline{\mathcal{B}_X(x_0,r)} \) is closed.

This completes the proof of the proposition. \( \square \)

We will sometimes talk about a metric space \( \mathcal{X} := (X, \rho) \) without referring the the components of \( \mathcal{X} \), where the metric is understood to be \( \rho \), the natural topology is understood to be \( \mathcal{O}_X \). When it is clear from the context, we will also neglect the subscript \( \mathcal{X} \). We will abuse the notation and say \( x \in \mathcal{X} \) and \( A \subseteq \mathcal{X} \) when \( x \in X \) and \( A \subseteq X \).

On a metric space, we can talk about open and closed sets, and all those concepts defined in Chapter 3, all with respect to the natural topology.

A metric space is clearly first countable, where a countable basis at \( x_0 \in \mathcal{X} \) is \( \{ \mathcal{B}(x_0,r) \mid r \in \mathbb{Q}, r > 0 \} \).

**Proposition 4.4** A metric space \( (X, \rho) \) is separable if, and only if, it is second countable.

**Proof**  “Only if” Let \( D \subseteq X \) be a countable dense set. Let \( \mathcal{M} := \{ \mathcal{B}(x,r) \mid x \in D, r \in \mathbb{Q}, \text{ and } r > 0 \} \). Clearly, \( \mathcal{M} \) is countable and \( \mathcal{O} \subseteq \mathcal{M} \). \( \forall \mathcal{O} \in \mathcal{O}, \forall x \in \mathcal{O}, \exists r \in (0, \infty) \cap \mathbb{Q} \) such that \( \mathcal{B}(x,r) \subseteq \mathcal{O} \). Since \( D \) is dense, then \( \exists r_1 \in \mathcal{B}(x,r)/2 \cap D \). Let \( r_1 := r/2 \in (0, \infty) \cap \mathbb{Q} \). Then, we have \( x \in \mathcal{B}(x_0,r_1) \subseteq \mathcal{B}(x,r) \subseteq \mathcal{O} \) and \( \mathcal{B}(x_1,r_1) \in \mathcal{M} \). Hence, \( \mathcal{M} \) is a basis for \( \mathcal{O} \). Hence, \( (X, \mathcal{O}) \) is second countable.

“If” Let \( \mathcal{O} \) has a countable basis \( \mathcal{B} \). By Axiom of Choice, we may assign \( a_{x_B} \in B, VB \in \mathcal{B} \) with \( B \neq \emptyset \). Let \( D = \{ x_B \in X \mid B \in \mathcal{B}, B \neq \emptyset \} \). Then, \( D \) is countable. \( \forall x \in X, \forall \mathcal{O} \in \mathcal{O} \) with \( x \in \mathcal{O} \), \( \exists B \in \mathcal{B} \) such that \( x \in B \subseteq \mathcal{O} \). Then, \( x_B \in D \cap B \subseteq D \cap \mathcal{O} \neq \emptyset \). Hence, by Proposition 3.3, we have \( x \in \overline{\mathcal{B}} \). Therefore, by the arbitrariness of \( x \), we have \( D \) is dense. Hence, \( (X, \mathcal{O}) \) is separable.

This completes the proof of the proposition. \( \square \)

**Proposition 4.5** Let \( \mathcal{X} \) be a topological space and \( \mathcal{Y} \) and \( \mathcal{Z} \) be metric spaces. Let \( f : \mathcal{X} \rightarrow \mathcal{Y}, g : \mathcal{Y} \rightarrow \mathcal{Z}, h : \mathcal{Y} \rightarrow \mathcal{X}, x_0 \in \mathcal{X} \), and \( y_0 \in \mathcal{Y} \). Then, the following statements hold.

1. \( f \) is continuous at \( x_0 \) if, and only if, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists U \in \mathcal{O}_X \) with \( x_0 \in U \) such that \( \rho_x(f(x), f(x_0)) < \epsilon, \forall x \in U \).

2. \( g \) is continuous at \( y_0 \) if, and only if, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \rho_z(g(y), g(y_0)) < \epsilon, \forall y \in \mathcal{B}_Y (y_0, \delta) \).
3. $h$ is continuous at $y_0$ if, and only if, $\forall U \in \mathcal{O}_X$ with $h(y_0) \in U$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $h(y) \in U$, $\forall y \in B_Y(y_0, \delta)$.

**Proof** The proof is straightforward, and is therefore omitted. □

**Definition 4.6** Let $X$ and $Y$ be metric spaces and $f : X \to Y$ be a homeomorphism. $f$ is said to be an isometry between $X$ and $Y$ if $\rho_Y(f(x_1), f(x_2)) = \rho_X(x_1, x_2)$, $\forall x_1, x_2 \in X$. Then, the two metric spaces are said to be isometric.

**Definition 4.7** Let $X$ be a set and $\rho_1$ and $\rho_2$ be two metric on $X$. $\rho_1$ and $\rho_2$ are said to be equivalent if the identity map from $(X, \rho_1)$ to $(X, \rho_2)$ is a homeomorphism.

When two metrics are equivalent, then the natural topologies generated by them are equal to each other.

Clearly, a metric space is Hausdorff.

### 4.2 Convergence and Completeness

**Proposition 4.8** Let $X$ be a metric space and $(x_\alpha)_{\alpha \in A} \subseteq X$ be a net. Then, $\lim_{\alpha \in A} x_\alpha = x \in X$ if, and only if, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \alpha_0 \in A$, $\forall \alpha \in A$ with $\alpha_0 \prec \alpha$, we have $\rho(x_\alpha, x) < \epsilon$.

**Proof** This is straightforward, and is omitted. □

Since metric spaces are Hausdorff, then the limit is unique if it exists.

**Definition 4.9** Let $X$ be a metric space, $x_0 \in X$, and $S \subseteq X$. The distance from $x_0$ to $S$ is $\text{dist}(x_0, S) := \inf_{s \in S} \rho(x_0, s) \in [0, \infty) \subset \mathbb{R}_e$.

$\text{dist}(x_0, S) = \infty$ if, and only if, $S = \emptyset$.

**Proposition 4.10** Let $X$ be a metric space, $x_0 \in X$, $S \subseteq X$, and $S$ is closed. Then, $x_0 \in S$ if, and only if, $\text{dist}(x_0, S) = 0$.

**Proof** “Only if” This is obvious.

“If” By the fact that $\text{dist}(x_0, S) = 0$, $\forall n \in \mathbb{N}$, $\exists x_n \in S$ such that $\rho(x_0, x_n) < 1/n$. Then, $\lim_{n \in \mathbb{N}} x_n = x_0$. By Proposition 3.68, $x_0 \in \overline{S}$. Since $S$ is closed, then, by Proposition 3.3, $S = \overline{S}$. Hence, $x_0 \in S$.

This completes the proof of the proposition. □

**Proposition 4.11** A metric space with its natural topology is normal.

**Proof** Let $X$ be the metric space. Clearly, $X$ is Hausdorff. $\forall$ closed sets $F_1, F_2 \subseteq X$ with $F_1 \cap F_2 = \emptyset$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $F_1 = \emptyset$ or $F_2 = \emptyset$; Case 2: $F_1 \neq \emptyset$ and $F_2 \neq \emptyset$. Case 1: $F_1 = \emptyset$ or $F_2 = \emptyset$. Without loss
of generality, assume $F_1 = \emptyset$. Take $O_1 = \emptyset \in O$ and $O_2 = X \in O$, then $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, $O_1 \cap O_2 = \emptyset$. Case 2: $F_1 \neq \emptyset$ and $F_2 \neq \emptyset$. 

\( \forall x \in F_1 \), \( \text{dist}(x, F_2) \in (0, \infty) \subseteq \mathbb{R} \), by Proposition 4.10. Define $O_1 \in O$ by $O_1 := \bigcup_{x \in F_1} \mathcal{B}(x, \text{dist}(x, F_2)/3)$. \( \forall x \in F_2 \), \( \text{dist}(x, F_1) \in (0, \infty) \subseteq \mathbb{R} \) by Proposition 4.10. Define $O_2 \in O$ by $O_2 := \bigcup_{x \in F_2} \mathcal{B}(x, \text{dist}(x, F_1)/3)$. 

Clearly, $F_1 \subseteq O_1$ and $F_2 \subseteq O_2$. Note that $O_1 \cap O_2 = \emptyset$, since otherwise, \( \exists x_0 \in O_1 \cap O_2 \), \( \exists x_1 \in F_1 \) such that \( x_0 \in \mathcal{B}(x_1, \text{dist}(x_1, F_2)/3) \), \( \exists x_2 \in F_2 \) such that \( x_0 \in \mathcal{B}(x_2, \text{dist}(x_2, F_1)/3) \), without loss of generality, assume \( \text{dist}(x_1, F_2) \leq \text{dist}(x_2, F_1) \), then \( \text{dist}(x_2, F_1) \leq \rho(x_2, x_1) \leq \rho(x_2, x_0) + \rho(x_0, x_1) < \text{dist}(x_2, F_1)/3 + \text{dist}(x_1, F_2)/3 \leq 2\text{dist}(x_2, F_1)/3 \), which is a contradiction.

Hence, in both cases, \( \exists O_1, O_2 \in O \) such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, $O_1 \cap O_2 = \emptyset$. Hence, $X$ is normal. This completes the proof of the proposition. \( \square \)

**Definition 4.12** Let $\mathcal{X}$ be a metric space and \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \). The sequence is said to be a Cauchy sequence if \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists N \in \mathbb{N} \), \( \forall n, m \geq N \), \( \rho(x_n, x_m) < \epsilon \).

Clearly, every convergent sequence in a metric space is a Cauchy sequence.

**Proposition 4.13** Let $\mathcal{X}$ be a metric space, $E \subseteq \mathcal{X}$, and $x_0 \in \mathcal{X}$. Then, $x_0 \in \overline{E}$ if, and only if, \( \exists (x_n)_{n=1}^{\infty} \subseteq E \) such that $\lim_{n \in \mathbb{N}} x_n = x_0$.

**Proof** \( "\text{Only if}" \) \( \forall n \in \mathbb{N} \), since $x_0 \in \overline{E}$, then, by Proposition 3.3, \( \exists x_n \in E \cap \mathcal{B}(x_0, 1/n) \). Clearly, \( (x_n)_{n=1}^{\infty} \subseteq E \) and $\lim_{n \in \mathbb{N}} x_n = x_0$.

\( "\text{If}" \) This is immediate by Proposition 3.68. This completes the proof of the proposition. \( \square \)

**Definition 4.14** A metric space is said to be complete if every Cauchy sequence in the metric space converges to a point in the space.

**Proposition 4.15** Let $\mathcal{X} := (X, \rho)$ be a metric space, $\mathcal{Y} := (Y, O)$ be a topological space, $f : \mathcal{X} \to \mathcal{Y}$, and $x_0 \in \mathcal{X}$. Then, the following statements are equivalent.

(i) $f$ is continuous at $x_0$;

(ii) if $x_0$ is an accumulation point of $\mathcal{X}$, then $f(x)$ converges to $f(x_0)$ as $x \to x_0$;

(iii) \( \forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) with $\lim_{n \in \mathbb{N}} x_n = x_0$, we have \( (f(x_n))_{n=1}^{\infty} \subseteq \mathcal{Y} \) converges to $f(x_0)$;

(iv) \( \forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) with $x_0$ as a cluster point, we have that \( (f(x_n))_{n=1}^{\infty} \subseteq \mathcal{Y} \) admits a cluster point $f(x_0)$. 


4.2. CONVERGENCE AND COMPLETENESS

Proof: (i) \iff (ii). This follows from Proposition 3.74.

(i) \Rightarrow (iii). This follows from Proposition 3.66.

(iii) \Rightarrow (iv). \forall (x_n)_{n=1}^\infty \subseteq X with \( x_0 \) as a cluster point. Since \( X \) is a metric space and therefore first countable, then \( \exists \) a subsequence \((x_{n_i})_{i=1}^\infty\) of \((x_n)_{n=1}^\infty\) such that \( \lim_{i \to \infty} x_{n_i} = x_0 \). Then, by (iii), \((f(x_{n_i}))_{i=1}^\infty\) converges to \( f(x_0) \). Then, \( f(x_0) \) is a cluster point of \((f(x_n))_{n=1}^\infty\) and therefore a cluster point of \((f(x_n))_{n=1}^\infty\).

(iv) \Rightarrow (i). Suppose \( f \) is not continuous at \( x_0 \). \( \exists O_{y_0} \subseteq \mathcal{O}_y \) with \( f(x_0) \in \mathcal{O}_{y_0} \) such that \( \forall n \in \mathbb{N} \), we have \( f(B(x_0, 1/n)) \not\subseteq O_{y_0} \). Then, \( \exists n \in B(x_0, 1/n) \) such that \( f(x_n) \in \mathcal{O}_{y_0} \). Consider the sequence \((x_n)_{n=1}^\infty\).

Clearly, \( x_0 = \lim_{n \to \infty} x_n \) and therefore is a cluster point of the sequence. Consider the sequence \((f(x_n))_{n=1}^\infty\). For the open set \( O_{y_0} = f(x_0) \), \( \forall n \in \mathbb{N} \), \( f(x_n) \in \mathcal{O}_{y_0} \). Then, \( f(x_0) \) is not a cluster point of \((f(x_n))_{n=1}^\infty\). This contradicts with the assumption. Therefore, \( f \) must be continuous at \( x_0 \).

This completes the proof of the proposition.

\[ \square \]

Proposition 4.16 Let \( X \) be a metric space, \( \mathcal{Y} := (Y, \mathcal{O}) \) be a topological space, \( D \subseteq X \), \( f : D \to \mathcal{Y} \), \( x_0 \in X \) be an accumulation point of \( D \), and \( y_0 \in \mathcal{Y} \). Then, the following statements are equivalent.

(i) \( f(x) \) converges to \( y_0 \) as \( x \to x_0 \).

(ii) \( \forall (x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\} \) with \( \lim_{n \to \infty} x_n = x_0 \), we have that \( y_0 \) is a limit point of \((f(x_n))_{n=1}^\infty\).

Proof. (i) \Rightarrow (ii). By (i), \( \forall O \in \mathcal{O} \) with \( y_0 \in O \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \forall \epsilon \in (D \cap B_X(x_0, \delta)) \setminus \{x_0\}, f(x) \in O \). \( \forall (x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\} \) with \( \lim_{n \to \infty} x_n = x_0 \), \( \exists N \in \mathbb{N} \) such that \( \forall n \geq N \), \( x_n \in B_X(x_0, \delta) \). Then, \( x_n \in (D \cap B_X(x_0, \delta)) \setminus \{x_0\} \) and \( f(x_n) \in O \). Hence, \( y_0 \) is a limit point of \((f(x_n))_{n=1}^\infty\).

(ii) \Rightarrow (i). We will show this by an argument of contradiction. Suppose (i) does not hold. Then, \( \exists O_0 \in \mathcal{O} \) with \( y_0 \in O_0 \), \( \forall n \in \mathbb{N} \), \( \exists x_n \in (D \cap B_X(x_0, 1/n)) \setminus \{x_0\} \) such that \( f(x_n) \in X \setminus O_0 \). Clearly, the sequence \((x_n)_{n=1}^\infty \subseteq D \setminus \{x_0\} \) and \( \lim_{n \to \infty} x_n = x_0 \). But, \( f(x_n) \not\in O_0 \), \( \forall n \in \mathbb{N} \). Then, \( y_0 \) is not a limit point of \((f(x_n))_{n=1}^\infty\). This contradicts (ii). Hence, (i) must hold.

This completes the proof of the proposition.

\[ \square \]

Proposition 4.17 Let \( X \) be a metric space, \( D \subseteq X \), \( f : D \to \mathbb{R} \), and \( x_0 \in X \) be an accumulation point of \( D \). Then, we have

\[
\limsup_{x \to x_0} f(x) = \inf_{\epsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap B(x_0, \epsilon)) \setminus \{x_0\}} f(x)
\]

\[
\liminf_{x \to x_0} f(x) = \sup_{\epsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap B(x_0, \epsilon)) \setminus \{x_0\}} f(x)
\]
Proof Let \( L = \limsup_{x \to x_0} f(x) \in \mathbb{R} \) and \( \bar{L} := \inf_{x \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap B(x_0, \epsilon) \backslash \{x_0\})} f(x) \in \mathbb{R}^e \). \( \forall m \in \mathbb{R} \) with \( m < L \), we have \( m < \sup_{x \in (D \cap B(x_0, \epsilon) \backslash \{x_0\})} f(x) \), \( \forall V \in \mathcal{O}_X \) with \( x_0 \in V \). Then, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \sup_{x \in (D \cap B(x_0, \epsilon) \backslash \{x_0\})} f(x) > m \). Hence, \( m \leq \bar{L} \). By the arbitrariness of \( m \), we have \( \bar{L} \leq L \). By the arbitrariness of \( \bar{L} \), we have \( L \leq \bar{L} \). Then, the two metrics are said to be uniformly equivalent.

\[
\liminf_{x \to x_0} f(x) = -\limsup_{x \to x_0} (-f(x)) = -\inf_{\epsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap B(x_0, \epsilon) \backslash \{x_0\})} (-f(x)) = \sup_{\epsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap B(x_0, \epsilon) \backslash \{x_0\})} f(x)
\]

This completes the proof of the proposition. \( \square \)

4.3 Uniform Continuity and Uniformity

Definition 4.18 Let \( X \) and \( Y \) be metric spaces and \( f : X \to Y \). \( f \) is said to be uniformly continuous if \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_Y(f(x_1), f(x_2)) < \epsilon \), \( \forall x_1, x_2 \in X \) with \( \rho_X(x_1, x_2) < \delta \).

Clearly, a function \( f \) is uniformly continuous implies that it is continuous.

Definition 4.19 Let \( X \) and \( Y \) be metric spaces and \( f : X \to Y \). \( f \) is said to be a uniform homeomorphism if \( f \) is bijective and both \( f \) and \( f^{-1} \) are uniformly continuous.

Properties preserved under uniform homeomorphisms are called uniform properties. These includes Cauchy sequences, completeness, uniform continuity, and total boundedness.

Definition 4.20 Let \( X \) be a set and \( \rho_1 \) and \( \rho_2 \) be two metrics defined on \( X \). Then, the two metrics are said to be uniformly equivalent if the identity map from \( (X, \rho_1) \) to \( (X, \rho_2) \) is a uniform homeomorphism.

Proposition 4.21 Let \( (X, \rho) \) be a metric space. Define \( \sigma : X \times X \to \mathbb{R} \) by \( \sigma(x_1, x_2) = \frac{\rho(x_1, x_2)}{1 + \rho(x_1, x_2)} \), \( \forall x_1, x_2 \in X \). Then, \( \sigma \) is a metric on \( X \) and \( \rho \) and \( \sigma \) are uniformly equivalent.

Proof \( \forall x_1, x_2, x_3 \in X \), \( \sigma(x_1, x_2) \in [0, 1) \); \( \sigma(x_1, x_2) = 0 \iff \rho(x_1, x_2) = 0 \iff x_1 = x_2 \); \( \sigma(x_1, x_2) = \sigma(x_2, x_1) \); by the monotonicity of \( \frac{\rho}{1 + \rho} \) on \( s > -1 \),
we have

\[ \sigma(x_1, x_2) \leq \frac{\rho(x_1, x_3) + \rho(x_3, x_2)}{1 + \rho(x_1, x_3) + \rho(x_3, x_2)} \leq \frac{\rho(x_1, x_3)}{1 + \rho(x_1, x_3)} + \frac{\rho(x_3, x_2)}{1 + \rho(x_3, x_2)} = \sigma(x_1, x_3) + \sigma(x_3, x_2) \]

Hence, \( \sigma \) defines a metric on \( X \).

\[ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x_1, x_2 \in X \text{ with } \rho(x_1, x_2) < \epsilon, \text{ we have } \sigma(x_1, x_2) \leq \rho(x_1, x_2) < \epsilon. \] Hence, \( \text{id}_X : (X, \rho) \to (X, \sigma) \) is uniformly continuous.

On the other hand, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x_1, x_2 \in X \) with \( \sigma(x_1, x_2) < \frac{\epsilon}{1 + \epsilon} \), we have \( \rho(x_1, x_2) < \epsilon. \) Hence, \( \text{id}_X : (X, \sigma) \to (X, \rho) \) is uniformly continuous.

Therefore, \( \rho \) and \( \sigma \) are uniformly equivalent. \( \square \)

**Definition 4.22** A metric space \( \mathcal{X} \) is said to be totally bounded if \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) there exists finitely many open balls with radius \( \epsilon \) that cover \( \mathcal{X} \).

**Proposition 4.23** Let \( \mathcal{X}, \mathcal{Y}, \mathcal{W}, \) and \( \mathcal{Z} \) be metric spaces, \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) be a Cauchy sequence, \( f : \mathcal{X} \to \mathcal{Y} \) and \( g : \mathcal{Y} \to \mathcal{Z} \) be uniformly continuous functions, and \( h : \mathcal{X} \to \mathcal{W} \) be a uniform homeomorphism. Then, the following statements hold.

(i) \( (f(x_n))_{n=1}^{\infty} \) is a Cauchy sequence.

(ii) \( g \circ f \) is uniformly continuous.

(iii) If \( \mathcal{X} \) is complete, then \( \mathcal{W} \) is complete.

(iv) If \( \mathcal{X} \) is totally bounded and \( f \) is surjective, then \( \mathcal{Y} \) is totally bounded.

**Proof**

(i) \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by the uniform continuity of \( f, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_{\mathcal{Y}}(f(x_a), f(x_b)) < \epsilon, \forall x_a, x_b \in \mathcal{X} \) with \( \rho_{\mathcal{X}}(x_a, x_b) < \delta. \) Since \( (x_n)_{n=1}^{\infty} \) is Cauchy, then \( \exists N \in \mathbb{N} \) such that \( \rho_{\mathcal{X}}(x_n, x_m) < \delta, \forall n, m \geq N. \) Then, \( \rho_{\mathcal{Y}}(f(x_n), f(x_m)) < \epsilon. \) Hence, \( (f(x_n))_{n=1}^{\infty} \) is a Cauchy sequence.

(ii) \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by the uniform continuity of \( g, \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_{\mathcal{Z}}(g(y_1), g(y_2)) < \epsilon, \forall y_1, y_2 \in \mathcal{Y} \) with \( \rho_{\mathcal{Y}}(y_1, y_2) < \delta_1. \) By the uniform continuity of \( f, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_{\mathcal{X}}(f(x_a), f(x_b)) < \delta_1, \forall x_a, x_b \in \mathcal{X} \) with \( \rho_{\mathcal{X}}(x_a, x_b) < \delta. \) Then, we have \( \rho_{\mathcal{Z}}(g(f(x_n)), g(f(x_m))) < \epsilon. \) Hence, \( g \circ f \) is uniformly continuous.

(iii) \( \forall \text{ Cauchy sequence } (w_i)_{i=1}^{\infty} \subseteq \mathcal{W}. \) By (i), \( (h_{\text{inv}}(w_i))_{i=1}^{\infty} \subseteq \mathcal{X} \) is a Cauchy sequence. Since \( \mathcal{X} \) is complete, then \( \lim_{i \in \mathbb{N}} h_{\text{inv}}(w_i) = x_0 \in \mathcal{X}. \) By Proposition 3.66, we have \( \lim_{i \in \mathbb{N}} w_i = \lim_{i \in \mathbb{N}} h(h_{\text{inv}}(w_i)) = h(x_0) \in \mathcal{W}. \) Hence, \( \mathcal{W} \) is complete.

(iv) \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by the uniform continuity of \( f, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_{\mathcal{Y}}(f(x_a), f(x_b)) < \epsilon, \forall x_a, x_b \in \mathcal{X} \) with \( \rho_{\mathcal{X}}(x_a, x_b) < \delta. \) By the total boundedness of \( \mathcal{X}, \) there exists a finite set \( X_N \subseteq \mathcal{X}, \) such that
Definition 4.27 Let \( X \) be a metric space. Let \( B_X(x, \delta) = \{ y \in X \mid d(x, y) < \delta \} \) be a family of open balls in \( X \) for each \( x \in X \) and \( \delta > 0 \). Then, the union of all balls is \( \bigcup_{x \in X} B_X(x, \delta) = X \). Then, by the surjectiveness of \( f \) and Proposition 2.5, we have \( \bigcup_{x \in X} f(B_X(x, \delta)) = Y \). Note that \( f(B_X(x, \delta)) \subseteq B_Y(f(x), \epsilon) \), \( \forall x \in X \). Then, we have \( \bigcup_{x \in X} B_Y(f(x), \epsilon) = Y \). Hence, \( Y \) is totally bounded.

This completes the proof of the proposition.

Definition 4.24 Let \( X \) be a set and \( Y := (Y, \rho) \) be a metric space. Let \((f_\alpha)_{\alpha \in A}\) be a net of functions of \( X \) to \( Y \). Then, the net is said to converge uniformly to a function \( f : X \to Y \) if \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists \alpha_0 \in A \), \( \forall \alpha \in A \) with \( \alpha_0 < \alpha \), we have \( \rho(f_\alpha(x), f(x)) < \epsilon \), \( \forall x \in X \).

Definition 4.25 Let \( X \) be a set and \( Y := (Y, \rho) \) be a metric space. Let \((f_n)_{n=1}^\infty\) be a sequence of functions of \( X \) to \( Y \). Then, the sequence is said to be a uniform Cauchy sequence if \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists N \in \mathbb{N} \), \( \forall n, m \in \mathbb{N} \) with \( n, m \geq N \), we have \( \rho(f_n(x), f_m(x)) < \epsilon \), \( \forall x \in X \).

A uniformly convergent sequence \((f_n)_{n=1}^\infty\) is a uniform Cauchy sequence. A uniform Cauchy sequence in a complete metric space is uniformly convergent.

Proposition 4.26 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space and \( \mathcal{Y} := (Y, \rho) \) be a metric space. Let \((f_n)_{n=1}^\infty\) be a uniformly convergent sequence of functions of \( \mathcal{X} \) to \( \mathcal{Y} \) whose limit is \( f : \mathcal{X} \to \mathcal{Y} \). Assume that, \( \forall n \in \mathbb{N} \), \( f_n \) is continuous at \( x_0 \in \mathcal{X} \). Then, \( f \) is continuous at \( x_0 \).

Proof \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists N \in \mathbb{N} \), we have \( \rho(f_N(x), f(x)) < \epsilon/3 \), \( \forall x \in \mathcal{X} \). Since \( f_N \) is continuous at \( x_0 \), then \( \exists U \in \mathcal{O} \) with \( x_0 \in U \), \( \forall x \in U \), we have \( \rho(f_N(x), f_N(x_0)) < \epsilon/3 \). Then, \( \forall x \in U \), we have

\[
\rho(f(x), f(x_0)) \leq \rho(f(x), f_N(x)) + \rho(f_N(x), f_N(x_0)) + \rho(f_N(x_0), f(x_0)) < \epsilon
\]

Therefore, \( f \) is continuous at \( x_0 \).

This completes the proof of the proposition.

Definition 4.27 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space and \( \mathcal{Y} := (Y, \rho) \) be a metric space. Let \( \mathcal{F} \) be a family of continuous functions of \( \mathcal{X} \) to \( \mathcal{Y} \). \( \mathcal{F} \) is said to be equicontinuous at \( x_0 \in \mathcal{X} \) if \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists U \in \mathcal{O} \) with \( x_0 \in U \), \( \forall x \in U \), we have \( \rho(f(x), f(x_0)) < \epsilon \), \( \forall f \in \mathcal{F} \). \( \mathcal{F} \) is said to be equicontinuous if it is equicontinuous at \( x_0 \), \( \forall x_0 \in \mathcal{X} \).

4.4 Product Metric Spaces

Definition 4.28 Let \( (X, \rho_X) \) and \( (Y, \rho_Y) \) be metric spaces. Then, the product metric space \( (X \times Y, \rho) := (X, \rho_X) \times (Y, \rho_Y) \) is defined by

\[
\rho((x_1, y_1), (x_2, y_2)) := ((\rho_X(x_1, x_2))^2 + (\rho_Y(y_1, y_2))^2)^{1/2}
\]

\( \forall x_1, x_2 \in X \), \( \forall y_1, y_2 \in Y \). \( \rho \) is called the Cartesian metric.
In the above definition, it is straightforward to show that $\rho$ defines a metric on $X \times Y$.

**Proposition 4.29** Let $\mathcal{X} := (X, \rho_X)$ be a metric space with natural topology $\mathcal{O}_X$. Let $\mathcal{Y} := (Y, \rho_Y)$ be a metric space with natural topology $\mathcal{O}_Y$. Let $\mathcal{X} \times \mathcal{Y} := (X \times Y, \rho)$ be the product metric space with the natural topology $\mathcal{O}_m$. Let $(X \times Y, \mathcal{O})$ be the product topological space $(X, \mathcal{O}_X) \times (Y, \mathcal{O}_Y)$. Then, $\mathcal{O} = \mathcal{O}_m$.

**Proof** A basis for $\mathcal{O}$ is $\mathcal{M} := \{O_X \times O_Y \mid O_X \in \mathcal{O}_X, O_Y \in \mathcal{O}_Y\}$. A basis for $\mathcal{O}_m$ is $\mathcal{M}_m := \{B_{X \times Y}((x, y), r) \mid x \in X, y \in Y, r \in (0, \infty) \subset \mathbb{R}\}$.

$\forall B_{X \times Y}((x_0, y_0), r) \in \mathcal{M}_m$. $\forall (x, y) \in B_{X \times Y}((x_0, y_0), r)$, let $d := r - \rho((x, y), (x_0, y_0)) \in (0, \infty) \subset \mathbb{R}$. $\forall (\bar{x}, \bar{y}) \in B_X(x, d/\sqrt{2}) \times B_Y(y, d/\sqrt{2}) \in \mathcal{M}$, we have $\rho((\bar{x}, \bar{y}),(x_0, y_0)) \leq \rho((\bar{x}, \bar{y}),(x, y)) + \rho((x, y),(x_0, y_0)) = ((\rho_X(\bar{x}, x))^2 + (\rho_Y(\bar{y}, y))^2)^{1/2} + r - d < r$. Then, $(\bar{x}, \bar{y}) \in B_{X \times Y}((x_0, y_0), r)$. Then, $B_X(x, d/\sqrt{2}) \times B_Y(y, d/\sqrt{2}) \subseteq B_{X \times Y}((x_0, y_0), r)$. This implies that $B_{X \times Y}((x_0, y_0), r) \in \mathcal{O}$. Then, we have $\mathcal{O}_m \subseteq \mathcal{O}$.

On the other hand, $\forall O_X \times O_Y \in \mathcal{M}$. $\forall (x, y) \in O_X \times O_Y$, we have $x \in O_X$ and $y \in O_Y$. Then, $\exists d_1 \in (0, \infty) \subset \mathbb{R}$ such that $B_X(x, d_1) \subseteq O_X$ and $\exists d_2 \in (0, \infty) \subset \mathbb{R}$ such that $B_Y(y, d_2) \subseteq O_Y$. Let $d = \min\{d_1, d_2\} \in (0, \infty) \subset \mathbb{R}$. $\forall (\bar{x}, \bar{y}) \in B_{X \times Y}((x, y), d) \in \mathcal{M}_m$, we have $\rho_X(\bar{x}, x) < d$ and $\rho_Y(\bar{y}, y) < d$, which implies that $\bar{x} \in B_X(x, d_1)$ and $\bar{y} \in B_Y(y, d_2)$. Then, $B_{X \times Y}((x, y), d) \subseteq O_X \times O_Y$. Hence, we have $O_X \times O_Y \in \mathcal{O}_m$. Then, $\mathcal{O} \subseteq \mathcal{O}_m$.

Hence, $\mathcal{O} = \mathcal{O}_m$. This completes the proof of the proposition. □

The above proposition shows that the natural topology induced by the product metric defined in Definition 4.28 is the product topology on $X \times Y$.

**Proposition 4.30** Let $\mathcal{X}$ be a metric space. Then the metric $\rho_X$ is a uniformly continuous function on $\mathcal{X} \times \mathcal{X}$.

**Proof** $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$. Let $\delta = \epsilon/\sqrt{2} \in (0, \infty) \subset \mathbb{R}$. $\forall (x_1, x_2), (y_1, y_2) \in \mathcal{X} \times \mathcal{X}$ with $\rho_{X \times X}((x_1, x_2), (y_1, y_2)) < \delta$. Then, we have $\delta > ((\rho_X(x_1, y_1))^2 + (\rho_X(x_2, y_2))^2)^{1/2} = (\rho_X(x_1, y_1) + \rho_X(x_2, y_2))/\sqrt{2}$. Note that

\[
\rho_X(x_1, x_2) \leq \rho_X(x_1, y_1) + \rho_X(y_1, x_2)
\]

\[
\rho_X(y_1, x_2) \leq \rho_X(x_1, x_2) + \rho_X(x_1, y_1)
\]

\[
\rho_X(y_1, x_2) \leq \rho_X(y_1, y_2) + \rho_X(x_2, y_2)
\]

\[
\rho_X(y_1, y_2) \leq \rho_X(x_2, y_1) + \rho_X(x_2, x_2)
\]

This implies that

\[-\epsilon = -\sqrt{2} \delta < -\rho_X(x_1, y_1) - \rho_X(x_2, y_2) \leq \rho_X(x_1, x_2) - \rho_X(x_2, y_1) + \rho_X(x_2, y_1) - \rho_X(y_1, y_2) = \rho_X(x_1, x_2) - \rho_X(y_1, y_2) \leq \rho_X(x_1, y_1)\]
Hence, we have \( |\rho_X(x_1, x_2) - \rho_X(y_1, y_2)| < \epsilon \). Hence, \( \rho_X \) is uniformly continuous on \( X \times X \). This completes the proof of the proposition. \( \square \)

**Proposition 4.31** Let \( X \) and \( Y \) be complete metric spaces and \( Z = X \times Y \) be the product metric space with the cartesian metric \( \rho \). Then, \( Z \) is complete.

**Proof** Fix any Cauchy sequence \( (x_n, y_n) \) in \( Z \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N} \) such that \( \rho((x_n, y_n), (x_m, y_m)) < \epsilon, \forall n, m \geq N \). Then, we have \( \rho_X(x_n, x_m) < \epsilon \) and \( \rho_Y(y_n, y_m) < \epsilon \). Hence, \( (x_n)_{n=1}^\infty \subseteq X \) and \( (y_n)_{n=1}^\infty \subseteq Y \) are Cauchy sequences. By the completeness of \( X \) and \( Y \), \( \exists \lim x_n = x \) and \( \exists \lim y_n = y \) such that \( \lim_n x_n = x_0 \) and \( \lim_n y_n = y_0 \). By Proposition 3.67, we have \( \lim_n (x_n, y_n) = (x_0, y_0) \in Z \). Hence, \( Z \) is complete. This completes the proof of the proposition. \( \square \)

Clearly, Definition 4.28 and Propositions 4.29 and 4.31 may be easily generalized to the case of \( X_1 \times \cdots \times X_n \), where \( n \in \mathbb{N} \) and \( X_i \) are metric spaces, \( i = 1, \ldots, n \). When \( n = 0 \), it should be noted that \( \prod_{\beta \in \emptyset} \beta = (\{\emptyset\}, \rho) \), where \( \rho(\emptyset, \emptyset) = 0 \).

**Proposition 4.32** Let \( X_\alpha \) be a metric space, \( \alpha \in \Lambda \), where \( \Lambda \) is a finite set. Let \( \Lambda = \bigcup_{\beta \in \Gamma} \Lambda_\beta \), where \( \Lambda_\beta \)'s are pairwise disjoint and finite and \( \Gamma \) is also finite. \( \forall \beta \in \Gamma \), let \( X^{(\beta)} := \prod_{\alpha \in \Lambda_\beta} X_\alpha \) be the product metric space. Let \( X^{(\Gamma)} := \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \) be the product metric space of product metric spaces, and \( X := \prod_{\alpha \in \Lambda} X_\alpha \) be the product metric space. Then, \( X \) and \( X^{(\Gamma)} \) are isometric.

**Proof** Define \( E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \to \prod_{\alpha \in \Lambda} X_\alpha \) by, \( \forall x \in X^{(\Gamma)} \), \( \forall \alpha \in \Lambda, \exists \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha} \), \( \pi_\alpha(E(x)) = \pi^{(\beta_\alpha)}_\alpha(\pi^{(\Gamma)}_\beta(x)) \). By Proposition 3.30, \( E \) is a homeomorphism. \( \forall x, y \in X^{(\Gamma)} \), we have

\[
\rho(E(x), E(y)) = \left( \sum_{\alpha \in \Lambda} (\rho_\alpha(\pi_\alpha(E(x)), \pi_\alpha(E(y))))^2 \right)^{1/2}
\]

\[
= \left( \sum_{\beta \in \Gamma} \sum_{\alpha \in \Lambda_{\beta}} (\rho_\alpha(\pi_\alpha(E(x)), \pi_\alpha(E(y))))^2 \right)^{1/2}
\]

\[
= \left( \sum_{\beta \in \Gamma} \sum_{\alpha \in \Lambda_{\beta}} (\rho_\alpha(\pi^{(\beta)}_\alpha(\pi^{(\Gamma)}_\beta(x)), \pi^{(\beta)}_\alpha(\pi^{(\Gamma)}_\beta(y))))^2 \right)^{1/2}
\]

\[
= \left( \sum_{\beta \in \Gamma} \left( \sum_{\alpha \in \Lambda_{\beta}} (\rho_\alpha(\pi^{(\beta)}_\alpha(\pi^{(\Gamma)}_\beta(x)), \pi^{(\beta)}_\alpha(\pi^{(\Gamma)}_\beta(y))))^2 \right)^{1/2} \right)^{1/2}
\]

\[
= \left( \sum_{\beta \in \Gamma} (\rho^{(\beta)}(\pi^{(\Gamma)}_\beta(x), \pi^{(\Gamma)}_\beta(y)))^2 \right)^{1/2} = \rho^{(\Gamma)}(x, y)
\]

Hence, \( E \) is an isometry. This completes the proof of the proposition. \( \square \)
Proposition 4.33 Let $\mathcal{X}_\alpha := (X_\alpha, \rho_{X_\alpha})$ and $\mathcal{Y}_\alpha := (Y_\alpha, \rho_{Y_\alpha})$ be uniformly homeomorphic metric spaces, $\forall \alpha \in \Lambda$, where $\Lambda$ is a finite index set. Define the product metric spaces $\mathcal{X} := (X, \rho_X) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha$ and $\mathcal{Y} := (Y, \rho_Y) := \prod_{\alpha \in \Lambda} \mathcal{Y}_\alpha$, where $\rho_X$ and $\rho_Y$ are the Cartesian metric. Then, $\mathcal{X}$ and $\mathcal{Y}$ are uniformly homeomorphic.

**Proof** Let $F_{\alpha} : \mathcal{X}_\alpha \to \mathcal{Y}_\alpha$ be a uniform homeomorphism, $\forall \alpha \in \Lambda$. Define $F : \mathcal{X} \to \mathcal{Y}$ by, $\forall x \in \mathcal{X}$, $\pi_{\alpha}^{(X)}(F(x)) = F_{\alpha}(\pi_{\alpha}^{(X)}(x))$, $\forall \alpha \in \Lambda$. By Propositions 4.29 and 3.31, $F$ is a homeomorphism between $\mathcal{X}$ and $\mathcal{Y}$. We need only to show that $F$ and $F_{\text{inv}}$ are uniformly continuous.

Let $m \in \mathbb{Z}_+$ be the number of elements in $\Lambda$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_\alpha \in (0, \infty) \subset \mathbb{R}$ such that, $\forall x_{\alpha_1}, x_{\alpha_2} \in \mathcal{X}_\alpha$ with $\rho_{X_\alpha}(x_{\alpha_1}, x_{\alpha_2}) < \delta_\alpha$, we have $\rho_{Y_\alpha}(F_{\alpha}(x_{\alpha_1}), F_{\alpha}(x_{\alpha_2})) < \epsilon/\sqrt{1 + m}$, by the uniform continuity of $F_{\alpha}$. Let $\delta = \min\{\inf_{\alpha \in \Lambda} \delta_\alpha, 1\} \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\rho_X(x_1, x_2) < \delta$, we have, $\forall \alpha \in \Lambda$,

$$\rho_{X_\alpha}(\pi_{\alpha}^{(X)}(x_1), \pi_{\alpha}^{(X)}(x_2)) < \delta \leq \delta_\alpha$$

This implies that $\rho_{Y_\alpha}(F_{\alpha}(\pi_{\alpha}^{(X)}(x_1)), F_{\alpha}(\pi_{\alpha}^{(X)}(x_2))) < \epsilon/\sqrt{1 + m}$. Hence, we have

$$\rho_Y(F(x_1), F(x_2)) = \left( \sum_{\alpha \in \Lambda} (\rho_{Y_\alpha}(F_{\alpha}(\pi_{\alpha}^{(X)}(x_1)), F_{\alpha}(\pi_{\alpha}^{(X)}(x_2))))^2 \right)^{1/2}$$

$$= \left( \sum_{\alpha \in \Lambda} (\rho_{Y_\alpha}(F_{\alpha}(\pi_{\alpha}^{(X)}(x_1)), F_{\alpha}(\pi_{\alpha}^{(X)}(x_2))))^2 \right)^{1/2} < \epsilon$$

Hence, $F$ is uniformly continuous.

It is easy to see that $F_{\text{inv}} : \mathcal{Y} \to \mathcal{X}$ is given by, $\forall y \in \mathcal{Y}$, $\pi_{\alpha}^{(X)}(F_{\text{inv}}(y)) = F_{\text{inv}}(\pi_{\alpha}^{(Y)}(y))$, $\forall \alpha \in \Lambda$. Then, it is straightforward to apply a similar argument as above to show that $F_{\text{inv}}$ is uniformly continuous.

This completes the proof of the proposition. \qed

Proposition 4.34 Let $X_i := (X_i, \rho_i)$ be a metric space with natural topology $\mathcal{O}_i$, $\forall i \in \mathbb{N}$. On $X := \prod_{i=1}^{\infty} X_i$ define the metric, $\forall x, y \in X$,

$$\rho(x, y) = \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(y))}$$

Then, the natural topology, $\mathcal{O}_m$, on $\mathcal{X} := (X, \rho)$ induced by the metric $\rho$ equals to the product topology $\mathcal{O}$, where $(X, \mathcal{O}) = \prod_{i=1}^{\infty} (X_i, \mathcal{O}_i)$.

**Proof** We will first show that $\rho$ defines a metric on $X$. $\forall x, y, z \in X$. Clearly, we have $0 \leq \rho(x, y) < 1$ and $\rho(x, y) = \rho(y, x)$. When $x = y$, we have $\rho(x, y) = 0$. If $\rho(x, y) = 0$, then $\rho_i(\pi_i(x), \pi_i(y)) = 0$, $\forall i \in \mathbb{N}$,
which implies that \( \pi_i(x) = \pi_i(y), \forall i \in \mathbb{N} \), and hence \( x = y \). Note that the function \( \frac{1}{1+s} = 1 - \frac{1}{1+s} \) is strictly increasing on \( s > -1 \). Then, we have

\[
\rho(x, y) \leq \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(z)) + \rho_i(\pi_i(z), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(z)) + \rho_i(\pi_i(z), \pi_i(y))}
\]

\[
\leq \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(z))}{1 + \rho_i(\pi_i(x), \pi_i(z))} + \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(z), \pi_i(y))}{1 + \rho_i(\pi_i(z), \pi_i(y))}
\]

\[
= \rho(x, z) + \rho(z, y)
\]

Hence, \( \rho \) is a metric on \( X \).

Fix any basis open set \( B = B_X(x, r) \in \mathcal{O}_m \), where \( x \in X \) and \( r \in (0, \infty) \subseteq \mathbb{R} \). \( \forall y \in B \). Let \( \delta := r - \rho(x, y) > 0 \) and \( N \in \mathbb{N} \) be such that \( 2^{-N} < \delta/2 \). Consider the set \( C_y = \prod_{i=1}^{\infty} C_{y_i} \subseteq X \) given by \( C_{y_i} = B_X(\pi_i(y), \delta/2), i = 1, \ldots, N \), and \( C_{y_i} = X_i, i = N + 1, N + 2, \ldots \). Clearly, \( y \in C_y \subseteq \mathcal{O} \). \( \forall z \in C_y \), we have \( \rho(y, z) < \sum_{i=1}^{N} 2^{-i-1} \delta + \sum_{i=N+1}^{\infty} 2^{-i} \delta/2 + 2^{-N} < \delta \). Then, \( \rho(x, z) \leq \rho(x, y) + \rho(y, z) < \rho(x, y) < r \). Hence, \( z \in B \). Then, \( C_y \subseteq B \). This shows that \( B \in \mathcal{O} \). Hence, \( \mathcal{O}_m \subseteq \mathcal{O} \).

Fix any basis open set \( C \in \mathcal{O} \). Then, \( C = \prod_{i=1}^{\infty} C_i \), where \( C_i \in \mathcal{O}_i, i = 1, \ldots, N \), and \( C_i = X_i, i = N + 1, N + 2, \ldots \), for some \( N \in \mathbb{N} \). \( \forall x \in C \), we have \( \pi_i(x) \in C_i, \forall i \in \mathbb{N} \). Then, \( \forall i = 1, \ldots, N \), \( \exists \delta_i \in (0, \infty) \subseteq \mathbb{R} \) such that \( B_X(\pi_i(x), \delta_i) \subseteq C_i \). Let \( \delta_x := \min_{1 \leq i \leq N} 2^{-i-1} \delta_i > 0 \). Let \( B = B_X(x, \delta_x) \in \mathcal{O}_m \). \( \forall y \in B \), \( \rho(x, y) < \delta_x \) implies that \( 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(y))} < 2^{-i} \frac{\delta_i}{1 + \delta_i} \) and \( \rho_i(\pi_i(x), \pi_i(y)) < \delta_i, i = 1, \ldots, N \). Hence, \( y \in C \). Then, we have \( x \in B \subseteq C \). This shows that \( C \in \mathcal{O}_m \). Hence, \( \mathcal{O} \subseteq \mathcal{O}_m \).

Hence, \( \mathcal{O} = \mathcal{O}_m \). This completes the proof of the proposition. \( \square \)

**Proposition 4.35** Let \( \mathcal{X}_i := (X_i, \rho_i) \) be a complete metric space, \( \forall i \in \mathbb{N} \). On \( X := \prod_{i=1}^{\infty} X_i \), define the metric \( \rho \) as in Proposition 4.34. Then, the product metric space \( \mathcal{X} := (X, \rho) \) is complete.

**Proof** \( \forall \) Cauchy sequence \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \), clearly, \( (\pi_i(x_n))_{n=1}^{\infty} \subseteq \mathcal{X}_i \) is a Cauchy sequence, \( \forall i \in \mathbb{N} \). By the completeness of \( \mathcal{X}_i, \exists y_i \in \mathcal{X}_i \) such that \( \lim_{n \in \mathbb{N}} \pi_i(x_n) = y_i, \forall i \in \mathbb{N} \). Define \( y \in \mathcal{X} \) by \( \pi_i(y) = y_i, \forall i \in \mathbb{N} \). By Propositions 3.67 and 4.34, we have \( y \) is the limit of \( (x_n)_{n=1}^{\infty} \). Hence, \( \mathcal{X} \) is complete. This completes the proof of the proposition. \( \square \)

**Proposition 4.36** Let \( \mathcal{X}_i := (X_i, \rho_{X_i}) \) and \( \mathcal{Y}_i := (Y_i, \rho_{Y_i}) \) be uniformly homeomorphic metric spaces, \( \forall i \in \mathbb{N} \). Define the infinite product metric spaces \( \mathcal{X} := (X, \rho_{X}) := (\prod_{i=1}^{\infty} X_i, \rho_{X}) \) and \( \mathcal{Y} := (Y, \rho_{Y}) := (\prod_{i=1}^{\infty} Y_i, \rho_{Y}) \), where \( \rho_{X} \) and \( \rho_{Y} \) are defined as in Proposition 4.34. Then, \( \mathcal{X} \) and \( \mathcal{Y} \) are uniformly homeomorphic.

**Proof** Let \( F_i : \mathcal{X}_i \rightarrow \mathcal{Y}_i \) be a uniform homeomorphism, \( \forall i \in \mathbb{N} \). Define \( F : \mathcal{X} \rightarrow \mathcal{Y} \) by, \( \forall x \in \mathcal{X}, \pi_i^{(Y)}(F(x)) = F_i(\pi_i^{(X)}(x)), \forall i \in \mathbb{N} \). By
4.5. Subspaces

Propositions 4.34 and 3.31, F is a homeomorphism between X and Y. We need only to show that F and F_inv are uniformly continuous.

∀ε ∈ (0, ∞) ⊆ R, ∃N ∈ N such that 2^{-N} < ε/2. ∀i = 1, ..., N, ∃δ_i ∈ (0, ∞) ⊆ R such that, ∀x_{i1}, x_{i2} ∈ X, with ρ_X(x_{i1}, x_{i2}) < δ_i, we have ρ_Y(F_i(x_{i1}), F_i(x_{i2})) < ε/2, by the uniform continuity of F_i. Let

δ = \min_{1 \leq i \leq N} 2^{-i} \frac{\delta_i}{1 + \delta_i} > 0. ∀x_1, x_2 ∈ X with ρ_X(x_1, x_2) < δ, we have, ∀i = 1, ..., N,

\frac{\rho_X(\pi_i^X(x_1), \pi_i^X(x_2))}{1 + \rho_X(\pi_i^X(x_1), \pi_i^X(x_2))} < 2^i δ \leq \frac{\delta_i}{1 + \delta_i}

This implies that ρ_X(\pi_i^X(x_1), \pi_i^X(x_2)) < δ_i, which further implies that ρ_Y(F_i(\pi_i^X(x_1)), F_i(\pi_i^X(x_2))) < ε/2. Hence, we have

\rho_Y(F(x_1), F(x_2)) = \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_Y(\pi_i^Y(F(x_1)), \pi_i^Y(F(x_2)))}{1 + \rho_Y(\pi_i^Y(F(x_1)), \pi_i^Y(F(x_2)))}

= \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_Y(F_i(\pi_i^X(x_1)), F_i(\pi_i^X(x_2)))}{1 + \rho_Y(F_i(\pi_i^X(x_1)), F_i(\pi_i^X(x_2)))}

< \sum_{i=1}^{N} 2^{-i} \frac{\epsilon}{2} + 2^{-N} < \epsilon

Hence, F is uniformly continuous.

It is easy to see that F_inv : Y → X is given by, ∀y ∈ Y, π_Y(y)(F_inv(y)) = F_inv(π_i^Y(y)), ∀i ∈ N. Then, it is straightforward to apply a similar argument as above to show that F_inv is uniformly continuous.

This completes the proof of the proposition.

4.5 Subspaces

Proposition 4.37 Let X := (X, ρ) be a metric space with natural topology O and A ⊆ X be a subset. A := (A, ρ) is a metric space with natural topology A. A may also be endowed with the subset topology A with respect to (X, O). Then, O_A = O_s.

Proof A basis for the natural topology A is

B := \{ B_A(x, r) \mid x ∈ A, r ∈ (0, ∞) \subseteq R \}

= \{ B_X(x, r) \cap A \mid x ∈ A, r ∈ (0, ∞) \subseteq R \}

O_s = \{ O \cap A \mid O \subseteq O \}. Clearly, we have B ⊆ O_s. Then, O_A ⊆ O_s. On the other hand, ∀O_s ∈ O_s, O_s = O \cap A with O ∈ O. ∀x ∈ O_s, x ∈ A and \exists r ∈ (0, ∞) ⊆ R such that B_X(x, r) ⊆ O. Then, B_A(x, r) ⊆ O_s. Then, O_s ∈ O_A. Then, O_s ⊆ O_A. Hence, O_A = O_s.

This completes the proof of the proposition.
Proposition 4.38 Let $\mathcal{X}$ be a metric space and $S \subseteq \mathcal{X}$. If $\mathcal{X}$ is separable then $S := (S, \rho)$ is separable. If $S$ is separable, then $\overline{S}$ is separable.

Proof Since $\mathcal{X}$ is separable, by Proposition 4.4, there exists a countable basis $\mathcal{B}$ for $\mathcal{X}$. Let $\mathcal{B}_S := \{B \cap S \mid B \in \mathcal{B}\}$, which is a countable basis for $S$. By Proposition 4.4, $S$ is separable.

Let $S$ be separable. Then, there exists a countable dense subset $D \subseteq S$. The closure of $D$ in $\mathcal{X}$ is $\overline{D}$. By Proposition 3.3, $S \subseteq \overline{D}$. By Proposition 3.3, $\overline{\mathcal{B}} \subseteq \overline{S} \subseteq \overline{\mathcal{D}}$. Hence, $\overline{\mathcal{D}} = \overline{S}$ and $D$ is dense in $\overline{S}$. Hence, $\overline{S}$ is separable. This completes the proof of the proposition. \qed

Proposition 4.39 Let $\mathcal{X} := (X, \rho)$ be a metric space and $S \subseteq X$.

(i) If $(S, \rho)$ is complete then $S$ is closed in $\mathcal{X}$.

(ii) If $\mathcal{X}$ is complete and $S$ is closed in $\mathcal{X}$, then $(S, \rho)$ is complete.

Proof (i) Any convergent sequence is Cauchy. Then, by Proposition 4.13, $(\overline{S}, \rho)$ is complete implies that any point of closure of $S$ is in $S$. Then, $S = \overline{S}$ and $S$ is closed. (ii) Any Cauchy sequence in $S$ will converge in $\mathcal{X}$ since $\mathcal{X}$ is complete. Then, by Proposition 4.13, the limit lies in $S$ since $S$ is closed. Hence, $(S, \rho)$ is complete. This completes the proof of the proposition. \qed

4.6 Baire Category

Theorem 4.40 (Baire) Let $\mathcal{X}$ be a complete metric space and $(O_k)_{k=1}^{\infty}$ be a sequence of open dense sets in $\mathcal{X}$. Then, $\bigcap_{k=1}^{\infty} O_k$ is dense in $\mathcal{X}$.

Proof $\forall U \in \mathcal{O}$ with $U \neq \emptyset$. Since $O_1$ is dense, then $\exists x_1 \in O_1 \cap U$. By the openness of $U$ and $O_1$, $\exists r_1 \in (0, \infty) \subset \mathbb{R}$ such that $B(x_1, r_1) \subseteq O_1 \cap U$. $\forall n \in \mathbb{N}$ with $n > 1$, since $O_n$ is dense, then $\exists x_n \in O_n \cap B(x_{n-1}, r_{n-1}/2)$. Since $O_n, B(x_{n-1}, r_{n-1}) \in \mathcal{O}$, then $\exists r_n \in (0, r_{n-1}/2] \subseteq \mathbb{R}$ such that $B(x_n, r_n) \subseteq O_n \cap B(x_{n-1}, r_{n-1})$.

Note that $\forall n \in \mathbb{N}$, $r_n \leq 2^{1-n}r_1$ and, $\forall m > n$,

$$\rho(x_n, x_m) \leq \rho(x_n, x_{n+1}) + \cdots + \rho(x_{m-1}, x_m) < \frac{1}{2}(r_n + \cdots + r_{m-1}) \leq 2^{-n}r_1 + \cdots + 2^{1-m}r_1 < 2^{1-n}r_1$$

Hence, $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence, which converges to $x_0 \in \mathcal{X}$, by the completeness of $\mathcal{X}$. By Propositions 3.66, 3.67, and 4.30, we have, $\forall n \in \mathbb{N}$,

$$\rho(x_n, x_0) = \lim_{m \to \infty} \rho(x_n, x_m) < \frac{1}{2}(r_n + 2^{-1}r_n + 2^{-2}r_n + \cdots) = r_n$$

Then, $x_0 \in B(x_n, r_n) \subseteq O_n$. Hence, $x_0 \in (\bigcap_{k=1}^{\infty} O_k) \cap U \neq \emptyset$. Hence, $\bigcap_{k=1}^{\infty} O_k$ is dense in $\mathcal{X}$. This completes the proof of the theorem. \qed
Theorem 4.41 (Baire Category Theorem) Let \( \mathcal{X} \) be a complete metric space, then it is second category everywhere, that is no nonempty open subset of \( \mathcal{X} \) is of first category.

Proof. We will prove this theorem by using Proposition 3.38. Let \( (O_\alpha)_{\alpha \in \Lambda} \) be a countable collection of open dense sets in \( \mathcal{X} \). When \( \Lambda \) is finite, we may add the open dense set \( \mathcal{X} \) into the collection to make it infinite and countable. By Baire’s Theorem, we then have \( \bigcap_{\alpha \in \Lambda} O_\alpha \) is dense in \( \mathcal{X} \). Then, by Proposition 3.38, \( \mathcal{X} \) is second category everywhere. This completes the proof of the theorem. \( \square \)

4.7 Completion of Metric Spaces

Proposition 4.42 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space and \( \mathcal{Y} := (Y, \rho) \) be a complete metric space. Let \( (f_n)_{n=1}^\infty \) be a uniform Cauchy sequence of continuous functions of \( \mathcal{X} \) to \( \mathcal{Y} \). Then, there exists a continuous function \( f : \mathcal{X} \to \mathcal{Y} \) such that \( (f_n)_{n=1}^\infty \) converges uniformly to \( f \).

Proof. \( \forall x \in \mathcal{X}, \ (f_n(x))_{n=1}^\infty \) is a Cauchy sequence in \( \mathcal{Y} \). By the completeness of \( \mathcal{Y} \) and Proposition 3.65, \( \exists y \in \mathcal{Y} \) such that \( \lim_{n \to \infty} f_n(x) = y \). This defines a function \( f : \mathcal{X} \to \mathcal{Y} \).

Since \( (f_n)_{n=1}^\infty \) is a uniform Cauchy sequence, then, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \forall n, m \geq N, \text{ we have } \rho(f_n(x), f_m(x)) < \epsilon, \forall x \in \mathcal{X} \). By Propositions 4.30, 3.66, and 3.67, we have \( \rho(f_n(x), f(x)) = \lim_{m \to \infty} \rho(f_n(x), f_m(x)) \leq \epsilon, \forall x \in \mathcal{X} \). Hence, \( (f_n)_{n=1}^\infty \) converges uniformly to \( f \). By Proposition 4.26, \( f \) is continuous.

This completes the proof of the proposition. \( \square \)

Definition 4.43 Let \( \mathcal{X} \) be a metric space. A net \( (x_\alpha)_{\alpha \in A} \subseteq \mathcal{X} \) is said to be Cauchy if \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \alpha_0 \in A, \forall \alpha_1, \alpha_2 \in A \) with \( \alpha_0 < \alpha_1 \) and \( \alpha_0 < \alpha_2 \), we have \( \rho(x_{\alpha_1}, x_{\alpha_2}) < \epsilon \).

Any convergent net in a metric space is Cauchy.

Proposition 4.44 Let \( \mathcal{X} \) be a metric space. Then, \( \mathcal{X} \) is complete if, and only if, every Cauchy net has a limit in \( \mathcal{X} \).

Proof. “If” A Cauchy sequence is a Cauchy net and admits a limit in \( \mathcal{X} \). Hence, \( \mathcal{X} \) is complete.

“Only if” Let \( (x_\alpha)_{\alpha \in A} \subseteq \mathcal{X} \) be a Cauchy net. \( \exists \alpha_1 \in A \ni \forall \alpha_1, \alpha_2 \in A \) with \( \alpha_1 < \alpha_2 \), we have \( \rho(x_{\alpha_1}, x_{\alpha_2}) < 1 \). \( \forall n \in \mathbb{N} \) with \( n > 1 \), \( \exists \alpha_\alpha \in A \) with \( \alpha_{n-1} < \alpha_n \), \( \forall \alpha_1, \alpha_2 \in A \) with \( \alpha_n < \alpha_1 \) and \( \alpha_n < \alpha_2 \), we have \( \rho(x_{\alpha_1}, x_{\alpha_2}) < 1/n \). \( \forall n \in \mathbb{N}, \forall m_1, m_2 \in \mathbb{N} \) with \( m_1, m_2 > n \), we have \( \alpha_n < \alpha_{n+1} < \cdots < \alpha_m \) and \( \alpha_n < \alpha_{n+1} < \cdots < \alpha_m \). Then, \( \rho(x_{\alpha_m}, x_{\alpha_m}) < 1/n \). Hence, \( (x_{\alpha_n})_{n=1}^\infty \subseteq \mathcal{X} \) is a Cauchy sequence. By the completeness of \( \mathcal{X} \), \( \exists x_0 \in \mathcal{X} \) such that \( \lim_{n \to \infty} x_{\alpha_n} = x_0 \). \( \forall n \in \mathbb{N}, \forall \alpha \in A \)

Going back to the proof, we proceed as follows. We aim to show that \( f : \mathcal{X} \to \mathcal{Y} \) is continuous. Fix \( x \in \mathcal{X} \) and \( \epsilon > 0 \). Since \( f_n \to f \) uniformly, there exists a \( N \) such that for all \( n \geq N \), we have \( \rho(f_n(x), f(x)) < \epsilon/2 \). By completeness of \( \mathcal{Y} \), there exists a \( y \) such that \( \lim_{n \to \infty} f_n(x) = y \). Thus, \( \rho(y, f(x)) \leq \rho(f_n(x), f(x)) + \rho(f_n(x), y) < \epsilon/2 + \epsilon/2 = \epsilon \), which shows that \( f \) is continuous at \( x \). Since \( x \) was arbitrary, \( f \) is continuous on \( \mathcal{X} \). Therefore, \( \mathcal{X} \) is complete.
with \( \alpha_n < \alpha, \forall m \in \mathbb{N} \) with \( m > n \), we have \( \alpha_n < \alpha_{n+1} < \cdots < \alpha_m \) and \( \rho(x_n, x_m) < 1/n \). Then, by Propositions 4.30, 3.66, and 3.67, we have \( \rho(x_n, x_0) = \lim_{m \to \infty} \rho(x_n, x_m) \leq 1/n \). Hence, we have \( \lim_{x \to x_0} f(x) = x_0 \).

This completes the proof of the proposition. \( \square \)

**Proposition 4.45** Let \( X := (X, O) \) be a topological space, \( Y := (Y, \rho) \) be a complete metric space, \( D \subseteq X \), \( x_0 \in X \) be an accumulation point of \( D \), and \( f : D \to Y \). Then, \( \lim_{x \to x_0} f(x) \in Y \) if, and only if, \( \forall \epsilon > 0, \exists \delta > 0, \forall x \in D \cap O \) \( \{ x \} \subseteq \{ x_0 \} \), \( f(x) \in Y \). Then, \( \lim_{x \to x_0} f(x) = y_0 \) in \( Y \).

**Proof** “Sufficiency” Assume that \( \forall \epsilon > 0, \exists \delta > 0, \forall x \in D \cap O \) \( \{ x \} \subseteq \{ x_0 \} \), \( f(x) \in Y \). Define \( M := \{ O \in \mathcal{O} \mid x_0 \in O \} \). Clearly, \( X \in M \) and \( M \neq \emptyset \). It is easy to see that \( A := (M, \supseteq) \) is a directed system. Since \( x_0 \) is an accumulation point of \( D \), then \( \forall O \in A, (D \cap O) \not\subseteq \{ x_0 \} \). By Axiom of Choice, \( \exists \) a net \( \{ x_0 \}_{0 \in A} \subseteq X \) such that \( x_0 \in (D \cap O) \) \( \{ x_0 \} \subseteq (D \cap O) \) \( \{ x_0 \} \). This also defines a net \( \{ f(x_0) \}_{0 \in A} \subseteq Y \) by Axiom of Replacement. \( \forall \epsilon > 0, \exists \delta > 0, \forall x \in D \cap O \) \( \{ x \} \subseteq \{ x_0 \} \), \( f(x) \in Y \). Then, \( \lim_{x \to x_0} f(x) \in Y \).

“Necessity” Let \( \lim_{x \to x_0} f(x) = y_0 \in Y \). Then, \( \forall \epsilon > 0, \exists \delta > 0, \forall x \in D \cap O \) \( \{ x \} \subseteq \{ x_0 \} \), \( f(x) \in Y \). Hence, \( \lim_{x \to x_0} f(x) = y_0 \in Y \).

This completes the proof of the proposition. \( \square \)

**Proposition 4.46** Let \( (X, \rho_X) \) and \( (Y, \rho_Y) \) be metric spaces, \( (Y, \rho_Y) \) be complete, \( E \subseteq X \), and \( f : E \to Y \) be uniformly continuous. Then, there is a unique continuous extension \( g : \overline{E} \to Y \). Furthermore, \( g \) is uniformly continuous.

**Proof** \( \forall (x_i)_{i=1}^\infty \subseteq E \) with \( \lim_{i \to \infty} x_i = x_0 \in X \), by Proposition 4.23 and the uniform continuity of \( f \), \( \{ f(x_i) \}_{i=1}^\infty \) is a Cauchy sequence in \( Y \). Since \( (Y, \rho_Y) \) is complete, then \( \exists y_0 \in Y \) such that \( \lim_{i \to \infty} f(x_i) = y_0 \). Let \( (\tilde{x}_i)_{i=1}^\infty \subseteq E \) be any other sequence with \( \lim_{i \to \infty} \tilde{x}_i = x_0 \). Then, the sequence \( (x_1, \tilde{x}_1, x_2, \tilde{x}_2, \ldots) \) converges to \( x_0 \). By Proposition 4.23 and the uniform continuity of \( f \), \( \{ f(x_1), f(\tilde{x}_1), f(x_2), f(\tilde{x}_2), \ldots \} \) is a Cauchy sequence in \( Y \), which converges since \( (Y, \rho_Y) \) is complete. The limit for this Cauchy sequence must be \( y_0 \) by Proposition 3.70. Hence, we have \( \lim_{i \to \infty} f(\tilde{x}_i) = y_0 \). Hence, \( y_0 \) is dependent only on \( x_0 \) but not on the sequence \( (x_i)_{i=1}^\infty \).
By Proposition 4.13, we may define a function \( g : \overline{E} \to Y \) by \( g(x_0) = y_0 \), \( \forall x_0 \in \overline{E} \). \( \forall x_0 \in E \), choose a sequence \((x_0, x_0, \ldots)\), which converges to \( x_0 \), then \( y_0 = f(x_0) \). Hence, we have \( g|_E = f \).

Next, we show that \( g \) is uniformly continuous. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the uniform continuity of \( f \) on \( E \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_Y(f(x_1), f(x_2)) < \epsilon \), \( \forall x_1, x_2 \in E \) with \( \rho_X(x_1, x_2) < \delta \). \( \forall x_1, x_2 \in \overline{E} \) with \( \rho_X(x_1, x_2) < \delta \). By Proposition 4.13, \( \exists \left( x_i^{(1)} \right)_{i=1}^{\infty}, \left( x_i^{(2)} \right)_{i=1}^{\infty} \subset E \) such that \( \lim_{i \in \mathbb{N}} x_i^{(j)} = x_j, \ j = 1, 2 \). By the definition of \( g \), we have \( \lim_{i \in \mathbb{N}} f(x_i^{(j)}) = g(x_j), \ j = 1, 2 \). By Proposition 3.67, we have \( \lim_{i \in \mathbb{N}} (f(x_i^{(1)}), f(x_i^{(2)})) = (g(x_1), g(x_2)) \). By Propositions 3.66 and 4.30, we have \( \lim_{i \in \mathbb{N}} \rho_Y(f(x_i^{(1)}), f(x_i^{(2)})) = \rho_Y(g(x_1), g(x_2)) \). Hence, we have \( \rho_Y(g(x_1), g(x_2)) \leq \epsilon \). Hence, \( g \) is uniformly continuous.

Let \( h : \overline{E} \to Y \) be any continuous mapping such that \( h|_E = f \). \( \forall \bar{x} \in \overline{E} \), by Proposition 4.13, \( \exists (x_i)_{i=1}^{\infty} \subset E \) such that \( \lim_{i \in \mathbb{N}} x_i = \bar{x} \). By continuity of \( h \) and \( g \) and Proposition 3.66, we have \( h(\bar{x}) = \lim_{i \in \mathbb{N}} h(x_i) = \lim_{i \in \mathbb{N}} f(x_i) = \lim_{i \in \mathbb{N}} g(x_i) = g(\bar{x}) \). Then, \( h = g \). This shows that \( g \) is unique in the class of continuous functions that extends \( f \).

This completes the proof of the proposition.

**Definition 4.47** A pseudo-metric \( \rho \) on a set \( X \) satisfies (i), (iii), and (iv) in Definition 4.1 and \( \rho(x, x) = 0 \), \( \forall x \in X \), but not necessarily (ii).

**Lemma 4.48** Let \( X \) be a set and \( \rho \) be a pseudo-metric on \( X \). Define an equivalence relation \( \equiv \) on \( X \) by \( x_1 \equiv x_2 \) if \( \rho(x_1, x_2) = 0 \), \( \forall x_1, x_2 \in X \). Let \( \mathcal{Y} \) be the quotient set \( X/\equiv \), that is \( \mathcal{Y} := \{ F \subseteq X \mid F \neq \emptyset, \forall x_1, x_2 \in F, x_1 \equiv x_2, \forall x_3 \in X \setminus F, x_1 \not\equiv x_3 \} \). Define a mapping \( \rho_Y : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R} \) by \( \rho_Y(Y_1, Y_2) = \rho(x_1, x_2) \), \( \forall Y_1, Y_2 \in \mathcal{Y} \) and \( x_1 \in Y_1 \) and \( x_2 \in Y_2 \). Then, \( (\mathcal{Y}, \rho_Y) \) is a metric space and said to be the quotient space of \( (X, \rho) \) modulo \( \rho \).

**Proof** We first show that \( \equiv \) is an equivalence relation. Clearly, \( \equiv \) is reflexive and symmetric.

**Claim 4.48.1** \( \forall x, y, z \in X \), if \( \rho(y, z) = 0 \), then \( \rho(x, y) = \rho(x, z) \).

**Proof of claim:** \( \rho(x, z) \leq \rho(x, y) + \rho(y, z) = \rho(x, y) \leq \rho(x, z) \). This completes the proof of the claim. \( \square \)

By Claim 4.48.1, \( \equiv \) is transitive. Hence, \( \equiv \) is a equivalence relation.

Again by Claim 4.48.1, the mapping \( \rho_Y \) is well-defined.

To show that \( \rho_Y \) is a metric on \( \mathcal{Y} \), we note that, \( \forall Y_1, Y_2, Y_3 \in \mathcal{Y} \), (i) \( 0 \leq \rho_Y(Y_1, Y_2) < \infty \); (ii) \( \rho_Y(Y_1, Y_2) = 0 \iff \forall x_1 \in Y_1, \forall x_2 \in Y_2, \rho(x_1, x_2) = 0 \) and \( x_1 \equiv x_2 \iff Y_1 = Y_2 \); (iii) \( \rho_Y(Y_1, Y_2) = \rho(x_1, x_2) = \rho(x_2, x_1) = \rho_Y(Y_2, Y_1) \), for some \( x_1 \in Y_1 \) and \( x_2 \in Y_2 \); (iv) \( \rho_Y(Y_1, Y_2) = \rho(x_1, x_2) \leq \rho(x_1, x_3) + \rho(x_3, x_2) = \rho_Y(Y_1, Y_3) + \rho_Y(Y_3, Y_2) \), for some \( x_1 \in Y_1 \), \( x_2 \in Y_2 \), and \( x_3 \in Y_3 \). Hence, \( \rho_Y \) is a metric on \( \mathcal{Y} \).

This completes the proof of the lemma. \( \square \)
Theorem 4.49 Let \((X, \rho)\) be an metric space. Then,

(i) there exists a complete metric space \((\tilde{X}, \tilde{\rho})\) such that \(X\) is isometrically embedded as a dense subset in \(\tilde{X}\);

(ii) let \((Y, \rho_Y)\) be any complete metric space with \(X \subseteq Y\) and \(\rho_Y|_{X \times X} = \rho\). Then, \((\tilde{X}, \tilde{\rho})\) is isometric with \(\overline{X}\) in \(Y\).

Proof  (i) Let \(\mathcal{X} := \{ (x_n)_{n=1}^{\infty} \subseteq X \mid (x_n)_{n=1}^{\infty} \) is a Cauchy sequence \}. Define \(\sigma : \mathcal{X} \times \mathcal{X} \to \mathbb{R}\) by \(\sigma \left( \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(2)} \right)_{n=1}^{\infty} \right) \right) = \lim_{n \in \mathbb{N}} \rho(\langle x_n^{(1)} \rangle_{n=1}^{\infty}, \langle x_n^{(2)} \rangle_{n=1}^{\infty})\),

\(\forall \left( x_n^{(1)} \right)_{n=1}^{\infty}, \left( x_n^{(2)} \right)_{n=1}^{\infty} \in \mathcal{X}\). The mapping \(\sigma\) is well defined by Propositions 4.30 and 4.23 and the fact that \(\left( (x_n^{(1)}), (x_n^{(2)}) \right)_{n=1}^{\infty} \) is a Cauchy sequence.

\(\sigma \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(2)} \right)_{n=1}^{\infty} \right) \in [0, \infty) \subset \mathbb{R}; \sigma \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(2)} \right)_{n=1}^{\infty} \right) = \lim_{n \in \mathbb{N}} \rho(\langle x_n^{(1)} \rangle_{n=1}^{\infty}, \langle x_n^{(2)} \rangle_{n=1}^{\infty}) = \lim_{n \in \mathbb{N}} \rho(\langle x_n^{(1)} \rangle_{n=1}^{\infty}, \langle x_n^{(2)} \rangle_{n=1}^{\infty}) = \sigma \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(2)} \right)_{n=1}^{\infty} \right) \) and also \(0 = \lim_{n \in \mathbb{N}} \rho(\langle x_n^{(1)} \rangle_{n=1}^{\infty}, \langle x_n^{(1)} \rangle_{n=1}^{\infty}) = \sigma \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(1)} \right)_{n=1}^{\infty} \right) \). Hence, \(\sigma\) defines a pseudo-metric on \(\mathcal{X}\).

By Lemma 4.48, we can define the equivalence relation \(\equiv\) on \(\mathcal{X}\) by \(\langle x_n^{(1)} \rangle_{n=1}^{\infty} \equiv \langle x_n^{(2)} \rangle_{n=1}^{\infty} \) if \(\lim_{n \in \mathbb{N}} \rho(\langle x_n^{(1)} \rangle_{n=1}^{\infty}, \langle x_n^{(2)} \rangle_{n=1}^{\infty}) = 0\), \(\forall \left( x_n^{(1)} \right)_{n=1}^{\infty}, \left( x_n^{(2)} \right)_{n=1}^{\infty} \in \mathcal{X}\). Then, define the set \(\tilde{X} := \mathcal{X}/\equiv := \left\{ F \subseteq \mathcal{X} \mid F \neq \emptyset, \forall \left( x_n^{(1)} \right)_{n=1}^{\infty}, \left( x_n^{(2)} \right)_{n=1}^{\infty} \in F, \left( x_n^{(1)} \right)_{n=1}^{\infty} \equiv \left( x_n^{(2)} \right)_{n=1}^{\infty}, \forall \left( x_n^{(3)} \right)_{n=1}^{\infty} \in \mathcal{X} \backslash F, \left( x_n^{(1)} \right)_{n=1}^{\infty} \neq \left( x_n^{(3)} \right)_{n=1}^{\infty} \right\}\). A metric on \(\tilde{X}\) is \(\tilde{\rho}\), defined by \(\tilde{\rho}(F_1, F_2) = \sigma \left( \left( x_n^{(1)} \right)_{n=1}^{\infty} , \left( x_n^{(2)} \right)_{n=1}^{\infty} \right), \forall F_1, F_2 \in \tilde{X}\), for some \(\left( x_n^{(1)} \right)_{n=1}^{\infty} \in F_1\) and \(\left( x_n^{(2)} \right)_{n=1}^{\infty} \in F_2\).

Define mapping \(T : X \to \tilde{X}\) by, \(\forall x \in X, T(x) = F \in \tilde{X}\) such that the Cauchy sequence \(x, x, \ldots \in F\). Let \(X_I\) be the image of \(X\) under \(T\). Then, \(T : X \to X_I\) is surjective. \(\forall x_1, x_2 \in X\) with \(T(x_1) = T(x_2)\), we have \(\langle x_1, x_1, \ldots \rangle \equiv \langle x_2, x_2, \ldots \rangle\), which implies that \(\sigma(\langle x_1, x_1, \ldots \rangle, \langle x_2, x_2, \ldots \rangle) = \rho(x_1, x_2) = 0\), and hence \(x_1 = x_2\). Therefore, \(T\) is injective. Hence, \(T : X \to X_I\) is bijective and admits an inverse \(T_{inv} : X_I \to X\). \(\forall x_1, x_2 \in X, \tilde{\rho}(T(x_1), T(x_2)) = \sigma(\langle x_1, x_1, \ldots \rangle, \langle x_2, x_2, \ldots \rangle) = \rho(x_1, x_2)\). Hence, \(T\) is metric preserving. Then, both \(T\) and \(T_{inv}\) are uniformly continuous. Hence, \(T\) is an isometry between \(X\) and \(X_I\).
4.7. COMPLETION OF METRIC SPACES

Therefore, \((X, \rho)\) is isometrically embeded in \((\bar{X}, \bar{\rho})\).

Next, we show \((\bar{X}, \bar{\rho})\) is complete. Fix a Cauchy sequence \((F_n)_{n=1}^{\infty} \subseteq \bar{X}\). \(\forall n \in \mathbb{N}\), let \(\left( x_i^{(n)} \right)_{i=1}^{\infty} \in F_n\). Note that \(\left( x_i^{(n)} \right)_{i=1}^{\infty} \subseteq X\) is a Cauchy sequence. \(\forall n, m \in \mathbb{N}\), \(\bar{\rho}(F_n, F_m) = \sigma(\left( x_i^{(n)} \right)_{i=1}^{\infty}, \left( x_i^{(m)} \right)_{i=1}^{\infty}) = \lim_{i \to \infty} \rho(x_i^{(n)}, x_i^{(m)})\). Define a sequence \(\left( x_i^{(0)} \right)_{i=1}^{\infty} \subseteq X\) as following. Set \(n_0 = 0\). \(\forall i \in \mathbb{N}\), choose \(n_i \in \mathbb{N}\) with \(n_i > n_{i-1}\) such that \(\forall m_1, m_2 \geq n_i\), we have \(\rho(x_{m_1}^{(i)}, x_{m_2}^{(i)}) < 1/i\) and set \(x_i^{(0)} = x_{n_i}^{(i)}\).

Claim 4.49.1 \(\left( x_i^{(0)} \right)_{i=1}^{\infty} \in \mathcal{X}\).

Proof of claim: \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists N_1 \in \mathbb{N}\) such that \(\forall n, m \geq N_1\), we have \(\bar{\rho}(F_n, F_m) < \epsilon\). This is valid since \((F_n)_{n=1}^{\infty}\) is Cauchy in \(\bar{X}\). Let \(N_2 \in \mathbb{N}\) be such that \(1/N_2 < \epsilon\). Let \(N = \max\{N_1, N_2\} \in \mathbb{N}\). \(\forall i, j \geq N\), \(\bar{\rho}(x_i^{(0)}, x_j^{(0)}) = \rho(x_i^{(0)}, x_j^{(0)})\). Since \(\bar{\rho}(F_i, F_j) = \lim_{i \to \infty} \rho(x_i^{(0)}, x_j^{(0)}) < \epsilon\), then \(\exists i_0 \in \mathbb{N}\) with \(i_0 \geq \max\{n_i, n_j\}\) such that \(0 \leq \rho(x_{i_0}^{(i)}, x_{i_0}^{(j)}) < \bar{\rho}(F_i, F_j) + \epsilon < 2\epsilon\). Then, we have

\[
\rho(x_i^{(0)}, x_j^{(0)}) \leq \rho(x_{i_0}^{(i)}, x_{i_0}^{(i)}) + \rho(x_{i_0}^{(i)}, x_{i_0}^{(j)}) + \rho(x_{i_0}^{(j)}, x_{i_0}^{(j)}) < 1/i + 2\epsilon + 1/j < 4\epsilon
\]

Hence, \(\left( x_i^{(0)} \right)_{i=1}^{\infty} \in \mathcal{X}\). This completes the proof of the claim. \(\square\)

Let \(F_0 \in \bar{X}\) be such that \(\left( x_i^{(0)} \right)_{i=1}^{\infty} \in F_0\).

Claim 4.49.2 \(\lim_{n \to \infty} F_n = F_0\).

Proof of claim: \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists N_1 \in \mathbb{N}\) such that \(\forall n, m \geq N_1\), we have \(\bar{\rho}(F_n, F_m) < \epsilon\). Let \(N_2 \in \mathbb{N}\) be such that \(1/N_2 < \epsilon\). Let \(N := \max\{N_1, N_2\} \in \mathbb{N}\). \(\forall m \geq N\), fix \(i_0 := n_m \geq m \geq N\). \(\forall i \geq i_0\), \(\exists i_1 > n_i \geq i \geq n_m\) such that \(0 \leq \rho(x_i^{(m)}, x_i^{(i)}) < \bar{\rho}(F_m, F_i) + \epsilon < 2\epsilon\). Then,

\[
0 \leq \rho(x_i^{(m)}, x_i^{(0)}) = \rho(x_i^{(m)}, x_i^{(i)}) \\
\leq \rho(x_i^{(m)}, x_i^{(m)}) + \rho(x_i^{(m)}, x_i^{(i)}) + \rho(x_i^{(i)}, x_i^{(0)}) \\
< 1/m + 2\epsilon + 1/i < 4\epsilon
\]

Hence, \(0 \leq \bar{\rho}(F_m, F_0) = \lim_{i \to \infty} \rho(x_i^{(m)}, x_i^{(0)}) \leq 4\epsilon\). Therefore, we have \(\lim_{n \to \infty} \bar{\rho}(F_n, F_0) = 0\). This completes the proof of the claim. \(\square\)

Hence, we have shown that any Cauchy sequence \((F_n)_{n=1}^{\infty} \subseteq \bar{X}\) admits a limit \(F_0 \in \bar{X}\). Hence, \((\bar{X}, \bar{\rho})\) is complete.

To complete the proof of (i), we need only to show that \(X_f\) is dense in \(\bar{X}\). \(\forall F \in \bar{X}\), let \(\left( x_n \right)_{n=1}^{\infty} \in F\). \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists N \in \mathbb{N}\) such that \(\forall n \in \mathbb{N}\)
with \( n \geq N \), we have \( \rho(x_N, x_n) < \epsilon/2 \). Let \( F_N := T(x_N) \in X_I \). Then, by Propositions 4.30, 3.66, and 3.67, \( 0 \leq \rho(F_N, F) = \lim_{n \in \mathbb{N}} \rho(x_N, x_n) \leq \epsilon/2 < \epsilon \). Hence, \( F_N \in B(\bar{X}, \rho)(F, \epsilon) \cap X_I \neq \emptyset \). Hence, \( X_I \) is dense in \( \bar{X} \), by Proposition 3.3.

(ii) Let \( X_c \) be the closure of \( X \) in \( (Y, \rho_Y) \). Since \( T \) is an isometry between \( X \) and \( X_I \), then \( T : X \to \bar{X} \) is uniformly continuous. By Proposition 4.46, \( \exists \mathcal{G} : X_c \to \bar{X} \) such that \( \mathcal{G}|_X = T \) and \( \mathcal{G} \) is uniformly continuous. Note that \( T_{\text{inv}} : X_I \to X \subseteq Y \) is also uniformly continuous, since \( T_{\text{inv}} \) is metric preserving. Then, by Propositions 4.39 and 4.46, \( \exists \mathcal{H} : \bar{X} \to X_c \) such that \( H|_{X_I} = T_{\text{inv}} \) and \( H \) is uniformly continuous. \( G \circ H : \bar{X} \to \bar{X} \) is uniformly continuous, by Proposition 4.23. Note that \( (G \circ H)|_{X_I} = \mathcal{G} \). By Proposition 3.56, we have \( G \circ H = \mathcal{H} \). \( H \circ G : X_c \to X_c \) is uniformly continuous, by Proposition 4.23. Note that \( (H \circ G)|_X = \mathcal{H}|_X \). By Proposition 3.56, we have \( H \circ G = \mathcal{H} \). Therefore, \( G : X_c \to \bar{X} \) is bijective with inverse \( H : \bar{X} \to X_c \), by Proposition 2.4. Hence, \( X_c \) and \( \bar{X} \) are uniformly homeomorphic.

\[ \forall y_1, y_2 \in X_c, \exists \left( x_i^{(1)} \right)_{i=1}^{\infty}, \left( x_i^{(2)} \right)_{i=1}^{\infty} \in \mathcal{X} \text{ such that } \lim_{i \in \mathbb{N}} x_i^{(1)} = y_j, j = 1, 2, \text{ by Proposition 4.13.} \]

Then, by Proposition 3.66, we have \( G(y_j) = \lim_{i \in \mathbb{N}} G(x_i^{(j)}) = \lim_{i \in \mathbb{N}} T(x_i^{(j)}) = y_j, j = 1, 2. \) By Propositions 3.66, 3.67, and 4.30 and the fact that \( T \) is metric preserving on \( X \), we have \( \rho(G(y_1), G(y_2)) = \lim_{i \in \mathbb{N}} \rho(T(x_i^{(1)}), T(x_i^{(2)})) = \lim_{i \in \mathbb{N}} \rho(x_i^{(1)}, x_i^{(2)}) = \lim_{i \in \mathbb{N}} \rho(y_1, y_2). \) Hence, \( G : X_c \to \bar{X} \) is also metric preserving. Hence, \( G \) is an isometry between \( X_c \) and \( \bar{X} \).

This completes the proof of the theorem. \( \square \)

### 4.8 Metrization of Topological Spaces

**Definition 4.50** A topological space \( (X, \mathcal{O}) \) is said to be metrizable if there is a metric \( \rho \) on \( X \) whose natural topology is exactly the same as \( \mathcal{O} \).

**Proposition 4.51** Let \( \mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha) \) be metrizable topological spaces, \( \forall \alpha \in \Lambda \), where \( \Lambda \) is a countable set. Then, the product topological space \( \mathcal{X} = (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} X_\alpha \) is metrizable.

**Proof** \( \forall \alpha \in \Lambda \), let \( \rho_\alpha \) be a metric on \( X_\alpha \) whose natural topology is exactly \( \mathcal{O}_\alpha \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( \Lambda = \emptyset \); Case 2: \( \Lambda \neq \emptyset \) and is finite; Case 3: \( \Lambda \) is countably infinite.

Case 1: \( \Lambda = \emptyset \). Then, \( (X, \mathcal{O}) = (\{\emptyset\}, \emptyset) \). It is clearly metrizable, where the metric \( \rho \) may be defined by \( \rho(\emptyset, \emptyset) = 0 \).

Case 2: \( \Lambda \neq \emptyset \) and is finite. We may define a metric \( \rho \) on \( X \) be to the Cartesian metric as in Definition 4.28. By Proposition 4.29, the natural topology induced by \( \rho \) is exactly the same as the product topology \( \mathcal{O} \). Hence, \( \mathcal{X} \) is metrizable.
Case 3: $\Lambda$ is countable infinite. The result follows by Proposition 4.34. This completes the proof of the proposition.

**Definition 4.52** Let $I := [0,1] \subset \mathbb{R}$ and $A$ be a set. $I^A$ is called a cube. Associate with $I$ the subset topology of $\mathbb{R}$. Denote $X^\infty$ the countably infinite product space of $X$. The cube $I^\infty$ is metrizable, by Proposition 4.51, and is called a Hilbert cube.

**Theorem 4.53 (Urysohn Metrization Theorem)** Every normal topological space satisfying the second axiom of countability is metrizable.

**Proof** Let $\mathcal{X} := (X, \mathcal{O})$ be a second countable normal topological space. Let $(B_\alpha)_{\alpha \in \Lambda}$ be a countable basis for the topology $\mathcal{O}$. $\forall \alpha_1, \alpha_2 \in \Lambda$, if $B_{\alpha_1} \cap \overline{B_{\alpha_2}} = \emptyset$, then, by Urysohn’s Lemma, $\exists$ a continuous function $f_{\alpha_1,\alpha_2} : \mathcal{X} \rightarrow [0,1] := I \subset \mathbb{R}$ such that $f_{\alpha_1,\alpha_2}|_{\overline{B_{\alpha_1}}} = 0$ and $f_{\alpha_1,\alpha_2}|_{\overline{B_{\alpha_2}}} = 1$.

Define $\mathcal{F} := \{ f_{\alpha_1,\alpha_2} : \mathcal{X} \rightarrow I \mid \alpha_1, \alpha_2 \in \Lambda, \overline{B_{\alpha_1}} \cap \overline{B_{\alpha_2}} = \emptyset \}$. Clearly, $\mathcal{F}$ is a countable set. $\forall x_0 \in \mathcal{X}$, $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, we have $\exists \alpha_1 \in \Lambda$ such that $x_0 \in \overline{B_{\alpha_1}} \subseteq \overline{F}$. By Definition 3.33 and Propositions 3.34 and 3.35, $\exists U \in \mathcal{O}$ such that $\{x_0\} \subseteq U \subseteq \overline{U} \subseteq B_{\alpha_1}$. Then, $\exists \alpha_2 \in \Lambda$ such that $x_0 \in B_{\alpha_2} \subseteq U$. Then, we have $\{x_0\} \subseteq B_{\alpha_2} \subseteq \overline{B_{\alpha_2}} \subseteq \overline{B_{\alpha_1}}$, and hence $\overline{B_{\alpha_1}} \cap \overline{B_{\alpha_2}} = \emptyset$. Then, $\exists f_{\alpha_1,\alpha_2} \in \mathcal{F}$ such that $f_{\alpha_1,\alpha_2}|_{\overline{B_{\alpha_1}}} = 0$ and $f_{\alpha_1,\alpha_2}|_{\overline{B_{\alpha_2}}} = 1$, which implies that $f_{\alpha_1,\alpha_2}|_{F} = 0$ and $f_{\alpha_1,\alpha_2}(x_0) = 1$. Hence, $\mathcal{F}$ is a countable collection of continuous $[0,1]$-valued functions and satisfies $\forall x_0 \in \mathcal{X}$, $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, $\exists f \in \mathcal{F}$ such that $f(x_0) = 1$ and $f|_{F} = 0$. Since $\mathcal{X}$ is normal, then it is Tychonoff and any singleton subset is closed. $\forall x, y \in \mathcal{X}$ with $x \neq y$, we have $\{y\}$ is closed and therefore, $\exists f \in \mathcal{F}$ such that $f(x) = 1 \neq 0 = f(y)$.

By Proposition 3.59, the equivalence map $E : \mathcal{X} \rightarrow I^\mathcal{X}$ is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X})$. Since $I$ is metrizable and $\mathcal{F}$ is countable, then, by Proposition 4.51, $I^\mathcal{X}$ is metrizable with a metric $\rho_{I^\mathcal{X}}$. By Proposition 4.37, the metric $\rho_{I^\mathcal{X}}$ generates the subset topology on $E(\mathcal{X})$. Since $E$ is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X})$, then $\mathcal{X}$ is metrizable with a metric $\rho_\mathcal{X} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ given by $\rho_\mathcal{X}(x,y) := \rho_{I^\mathcal{X}}(E(x),E(y))$, $\forall x,y \in \mathcal{X}$.

This completes the proof of the theorem.

### 4.9 Interchange Limits

**Theorem 4.54 (Joint Limit Theorem)** Let $\mathcal{Y} := (Y, \rho)$ be a metric space and $(y_{\alpha,\beta})_{(\alpha,\beta) \in A_1 \times A_2} \subseteq \mathcal{Y}$ be a joint net. Assume that $\lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \bar{y} \in \mathcal{Y}$ and, $\forall \alpha \in A_1$, $\lim_{\beta \in A_2} y_{\alpha,\beta} = \bar{y}_\alpha \in \mathcal{Y}$. Then, $\lim_{\alpha \in A_1} \bar{y}_\alpha = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta}$.

**Proof** $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, by Definition 3.88 and Proposition 4.8, $\exists (\alpha_0, \beta_0) \in A_1 \times A_2$, $\forall (\alpha, \beta) \in A_1 \times A_2$ with $(\alpha, \beta) \prec (\alpha_0, \beta_0)$, we have
\[ \rho(y_{\alpha,\beta}, \bar{y}) < \epsilon/2. \] Fix any \( \alpha \in A_1 \) with \( \alpha_0 \prec_1 \alpha \). By the fact that \( \lim_{\beta \in A_2} y_{\alpha,\beta} = \bar{y}_{\alpha} \in \mathcal{Y} \), \( \exists \beta_{\alpha} \in A_2 \), \( \forall \beta \in A_2 \) with \( \beta_{\alpha} \prec_2 \beta \), we have \( \rho(y_{\alpha,\beta}, y_{\alpha}) < \epsilon/2 \). Since \( A_2 \) is a directed system, then \( \exists \beta_1 \in A_2 \) such that \( \beta_0 \prec_2 \beta_1 \) and \( \beta_{\alpha} \prec_2 \beta_1 \). Then, we have \( \rho(y_{\alpha,\beta}, \bar{y}) \leq \rho(y_{\alpha,\beta}, y_{\alpha}) + \rho(y_{\alpha,\beta}, \bar{y}) < \epsilon \). This shows that \( \lim_{\alpha \in A_1} \bar{y}_{\alpha} = \bar{y} \). Hence, we have \( \lim_{\alpha \in A_1} \bar{y}_{\alpha} = \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} \). This completes the proof of the theorem.

**Corollary 4.55** Let \( \mathcal{Y} := (Y, \rho) \) be a metric space and \( (y_{\alpha,\beta})_{(\alpha,\beta) \in A_1 \times A_2} \subseteq \mathcal{Y} \) be a joint net. Assume that \( \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \bar{y} \in \mathcal{Y} \), \( \forall \alpha \in A_1 \), \( \lim_{\beta \in A_2} y_{\alpha,\beta} = \bar{y}_{\alpha} \in \mathcal{Y} \), and, \( \forall \beta \in A_2 \), \( \lim_{\alpha \in A_1} y_{\alpha,\beta} = \bar{y}_{\beta} \in \mathcal{Y} \). Then, \( \lim_{\alpha \in A_1} \bar{y}_{\alpha} = \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \lim_{\beta \in A_2} \bar{y}_{\beta} = \lim_{\beta \in A_2} \lim_{\alpha \in A_1} y_{\alpha,\beta} \).

**Proof** This is straightforward by Joint Limit Theorem 4.54.

**Theorem 4.56 (Iterated Limit Theorem)** Let \( \mathcal{Y} := (Y, \rho) \) be a complete metric space and \( (y_{\alpha,\beta})_{(\alpha,\beta) \in A_1 \times A_2} \subseteq \mathcal{Y} \) be a joint net. Assume that

1. \( \forall \alpha \in A_1 \), \( \lim_{\beta \in A_2} y_{\alpha,\beta} = \bar{y}_{\alpha} \in \mathcal{Y} \);
2. the nets \( (y_{\alpha,\beta})_{\alpha \in A_1} \) converge uniformly to \( \bar{y}_{\beta} \in \mathcal{Y} \), \( \forall \beta \in A_2 \).

Then, \( \lim_{\alpha \in A_1} \bar{y}_{\alpha} = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \lim_{\beta \in A_2} \lim_{\alpha \in A_1} y_{\alpha,\beta} \).

**Proof** By (ii), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \alpha_0 \in A_1 \), \( \forall \beta \in A_2 \), \( \forall \alpha \in A_1 \) with \( \alpha_0 \prec_1 \alpha \) and \( \beta \prec_2 \beta \), we have \( \rho(y_{\alpha,\beta}, \bar{y}) < \epsilon/2 \). Fix an \( \alpha_1 \in A_1 \) with \( \alpha_0 \prec_1 \alpha_1 \). Since \( \lim_{\beta \in A_2} y_{\alpha_1,\beta} = \bar{y}_{\alpha_1} \), \( \exists \beta_{\alpha_1} \in A_2 \), \( \forall \beta \in A_2 \) with \( \beta_{\alpha_1} \prec_2 \beta \), we have \( \rho(y_{\alpha_1,\beta}, \bar{y}_{\alpha_1}) < \epsilon/4 \). Then, \( \forall \beta_1, \beta_2 \in A_2 \) with \( \beta_1 \prec_2 \beta_2 \) and \( \beta_1 \prec_2 \beta_2 \), we have \( \rho(y_{\alpha_1,\beta_1}, y_{\alpha_1,\beta_2}) \leq \rho(y_{\alpha_1,\beta_1}, \bar{y}_{\alpha_1}) + \rho(y_{\alpha_1,\beta_1}, \bar{y}_{\alpha_1}) + \rho(y_{\alpha_1,\beta_2}, \bar{y}_{\alpha_1}) < \epsilon \). Hence, the net \( (\bar{y}_{\beta})_{\beta \in A_2} \) is a Cauchy net. By Proposition 4.44, \( \lim_{\beta \in A_2} \bar{y}_{\beta} = \bar{y} \in \mathcal{Y} \).

Then, \( \exists \beta \in A_2 \), \( \forall \beta \in A_2 \) with \( \beta \prec_2 \beta \), we have \( \rho(\bar{y}_{\beta} - \bar{y}, \bar{y}) < \epsilon/2 \). Note that \( \forall \alpha, \beta \in A_1 \times A_2 \) with \( (\alpha_0, \beta) \prec (\alpha, \beta) \), we have \( \rho(y_{\alpha,\beta}, \bar{y}) \leq \rho(y_{\alpha,\beta}, \bar{y}) + \rho(\bar{y}_{\beta} - \bar{y}, \bar{y}) < \epsilon \). Hence, \( \lim_{\alpha,\beta \in A_1 \times A_2} y_{\alpha,\beta} = \bar{y} \). Finally, by Joint Limit Theorem 4.54, we have \( \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{\alpha \in A_1} \bar{y}_{\alpha} = \bar{y} \).

This completes the proof of the theorem.

**Proposition 4.57** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( \mathcal{Y} := (Y, \rho) \) be a complete metric space, \( D \subseteq \mathcal{X} \), \( x_0 \in \mathcal{X} \setminus D \) be an accumulation point of \( D \), and \( (f_\alpha)_{\alpha \in A} \) be a net of functions of \( D \) to \( \mathcal{Y} \). Assume that \( (f_\alpha)_{\alpha \in A} \) converges uniformly to function \( f : D \to \mathcal{Y} \) and \( \lim_{x \to x_0} f_\alpha(x) = y_\alpha \in \mathcal{Y} \), \( \forall \alpha \in A \).

Then, \( \lim_{x \to x_0} f_\alpha(x) = \lim_{x \to x_0} f_\alpha(x) = \lim_{x \to x_0} f_\alpha(x) = \lim_{x \to x_0} f_\alpha(x) = \lim_{x \to x_0} f(x) \in \mathcal{Y} \).
**4.9. INTERCHANGE LIMITS**

**Proof**  Fix any net \((x_\beta)_{\beta \in \hat{A}} \subseteq D\) with \(x_0\) as a limit. Such net exists by Proposition 3.68. By the assumption that \((f_\alpha)_{\alpha \in A}\) converges uniformly to \(f\), then, for the joint net \((f_\alpha(x_\beta))_{(\alpha,\beta) \in A \times \hat{A}}\), we have the nets \((f_\alpha(x_\beta))_{\alpha \in A}\) converge uniformly to \(f(x_\beta) \in \mathcal{Y}\), \(\forall \beta \in \hat{A}\). By Proposition 3.79, we have \(\lim_{\beta \in \hat{A}} f_\alpha(x_\beta) = y_\alpha\), \(\forall \alpha \in A\). By Iterated Limit Theorem 4.56, we have \(\lim_{\alpha \in A} y_\alpha = \lim_{\beta \in \hat{A}} f_\alpha(x_\beta) = \lim_{(\alpha,\beta) \in A \times \hat{A}} f_\alpha(x_\beta) = \lim_{\beta \in \hat{A}} \lim_{\alpha \in A} f_\alpha(x_\beta) = \lim_{\beta \in \hat{A}} f(x_\beta) = \hat{y} \in \mathcal{Y}\). By the arbitrariness of \((x_\beta)_{\beta \in \hat{A}}\) and Proposition 3.79, we have \(\hat{y} = \lim_{x \to x_0} f(x)\). Hence, we have \(\lim_{\alpha \in A} y_\alpha = \lim_{\alpha \in A} \lim_{x \to x_0} f_\alpha(x) = \hat{y} = \lim_{x \to x_0} f(x) = \lim_{x \to x_0} \lim_{\alpha \in A} f_\alpha(x) \in \mathcal{Y}\). This completes the proof of the proposition. \(\square\)
Chapter 5

Compact and Locally Compact Spaces

5.1 Compact Spaces

Definition 5.1 A topological space \((X, \mathcal{O})\) is called compact if every open covering has a finite subcovering, that is \(\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}\) with \(X \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha\), where \(\Lambda\) is an index set, \(\exists\) a finite set \(\Lambda_N \subseteq \Lambda\) such that \(X \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha\).

A subset \(K \subseteq X\) is called compact if it is compact in its subset topology. This is equivalent to any open covering of \(K\) in \(X\) has a finite subcovering of \(K\).

Definition 5.2 Let \(X\) be a set and \(\mathcal{F}\) be a collection of subsets in \(X\). \(\mathcal{F}\) is said to admit the finite intersection property if any finite subcollection of \(\mathcal{F}\) has a nonempty intersection.

Proposition 5.3 A topological space \(\mathcal{X}\) is compact if, and only if, any collection of closed sets \(\mathcal{F}\) with the finite intersection property has a nonempty intersection.

Proof “If” Suppose \(\mathcal{X}\) is not compact. Then, there exists an open covering \((O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}\) of \(\mathcal{X}\), which does not have any finite subcovering. Then, \((\widetilde{O}_\alpha)_{\alpha \in \Lambda}\) is a collection of closed sets with the finite intersection property. Yet, \(\bigcap_{\alpha \in \Lambda} \widetilde{O}_\alpha = (\bigcup_{\alpha \in \Lambda} O_\alpha)^\sim = \emptyset\). This contradicts the assumption. Hence, \(\mathcal{X}\) is compact.

“Only if” Suppose that the result does not hold. Then, there exists a collection of closed sets \(\mathcal{F}\) with finite intersection property such that \(\bigcap_{F \in \mathcal{F}} F = \emptyset\). Then, we have \(\bigcup_{F \in \mathcal{F}} \widetilde{F} = (\bigcap_{F \in \mathcal{F}} F)^\sim = \mathcal{X}\). By the compactness of \(\mathcal{X}\), there exists a finite collection \(\mathcal{F}_N \subseteq \mathcal{F}\) such that
\[ \bigcup_{F \in \mathcal{F}_N} \overline{F} = \mathcal{X}. \] Then, \( \bigcap_{F \in \mathcal{F}_N} F = \emptyset. \) This contradicts with the assumption that \( \mathcal{F} \) has finite intersection property. Hence, the result holds.

This completes the proof of the proposition. \( \square \)

**Proposition 5.4** A topological space \( \mathcal{X} \) is compact if, and only if, \( \forall \) net \( (x_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{X} \) has a cluster point in \( \mathcal{X}. \)

**Proof** "Only if" Fix any net \( (x_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{X}, \forall \alpha \in \Lambda, \) let \( F_\alpha := \{ x_\beta \mid \beta \in \Lambda, \alpha \prec \beta \}. \) Let \( \overline{F_\alpha} \) be the closure of \( F_\alpha. \) For any finite set \( A \subseteq \Lambda, \exists \alpha_0 \in \Lambda \) such that \( \alpha \prec \alpha_0, \forall \alpha \in A, \) since \( \Lambda \) is a directed system. Then, \( x_{\alpha_0} \in F_\alpha, \forall \alpha \in \Lambda. \) Hence, we have \( x_{\alpha_0} \in \bigcap_{\alpha \in A} \overline{F_\alpha} \neq \emptyset. \) Hence, \( (\overline{F_\alpha})_{\alpha \in \Lambda} \) is a collection of closed sets in \( \mathcal{X} \) with the finite intersection property. By Proposition 5.3 and the compactness of \( \mathcal{X}, \) we have \( \bigcap_{\alpha \in \Lambda} \overline{F_\alpha} = \emptyset. \) Fix an \( x \in \bigcap_{\alpha \in \Lambda} \overline{F_\alpha}. \forall \alpha \in A \) with \( x \in O, \forall \alpha_0 \in \Lambda, \) we have \( x \in \overline{F_{\alpha_0}} \) and \( O \cap F_{\alpha_0} \neq \emptyset. \) Hence, \( x_{\alpha_0} \in A \) with \( \alpha_0 \prec \alpha \) such that \( x_{\alpha_0} \in O \cap F_{\alpha_0} \subseteq O. \) Hence, \( x \) is a cluster point of the net \( (x_\alpha)_{\alpha \in \Lambda}. \)

"If" Suppose \( \mathcal{X} \) is not compact. Then, there exists an open covering \( (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O} \) of \( \mathcal{X}, \) where \( \Lambda \) is an index set, such that \( \mathcal{X} \) is not covered by any finite subcollection of sets in \( (O_\alpha)_{\alpha \in \Lambda}. \) Let \( \Gamma := \{ B \subseteq \Lambda \mid B \) is a finite set \}. \) Define a relation \( \prec \) on \( \Gamma \) by \( \forall B_1, B_2 \in \Gamma, B_1 \prec B_2 \) if \( B_1 \subseteq B_2. \) Clearly, \( \Lambda := (\Gamma, \prec) \) is a directed system. \( \forall B \in \Gamma, \) let \( x_B \in (\bigcup_{\alpha \in B} O_\alpha)^\sim. \) This constructs a net \( (x_B)_{B \in \Lambda} \subseteq \mathcal{X}. \) By the assumption, \( \exists x_0 \in \mathcal{X} \) such that \( x_0 \) is a cluster point of the net. \( \forall \alpha_0 \in \Lambda, \{0_0\} \subseteq \Gamma. \forall O \in \mathcal{O} \) with \( x_0 \in O, \exists B \in \Gamma \) with \( \{0_0\} \prec B \) such that \( x_B \in O. \) Then, \( x_B \in O \cap (\bigcup_{\alpha \in B} O_\alpha)^\sim \subseteq O \cap \overline{O_{\alpha_0}} \neq \emptyset. \) Hence, by the arbitrariness of \( O \) and Proposition 3.3, \( x_0 \in \overline{O_{\alpha_0}} = \overline{O_{\alpha_0}} \). Then, we have \( x_0 \in \bigcap_{\alpha \in \Lambda} \overline{O_\alpha} = (\bigcup_{\alpha \in \Lambda} O_\alpha)^\sim = \emptyset. \) This is a contradiction. Hence, \( \mathcal{X} \) must be compact.

This completes the proof of the proposition. \( \square \)

**Proposition 5.5** A closed subset of a compact space is compact. A compact subset of a Hausdorff space is closed.

**Proof** Let \( (X, \mathcal{O}) \) be a compact topological space and \( K \subseteq X \) be closed. \( \forall \) open covering \( (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O} \) of \( K, \) where \( \Lambda \) is an index set.

We have \( X = (\bigcup_{\alpha \in \Lambda} O_\alpha) \cup \overline{K}, \) which forms an open covering of \( X. \) By the compactness of \( X, \) there exists a finite set \( \Lambda_N \subseteq \Lambda \) such that \( X = (\bigcup_{\alpha \in \Lambda_N} O_\alpha) \cup \overline{K}. \) Then, we have \( K \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha. \) Therefore, \( K \) is compact.

Let \( (X, \mathcal{O}) \) be a Hausdorff space and \( K \subseteq X \) be compact. Suppose that \( K \) is not closed. Then, by Proposition 3.3, \( \exists x_0 \in \overline{K} \setminus K. \forall x \in K, \) we have \( x \neq x_0. \) Since \( (X, \mathcal{O}) \) is Hausdorff, \( \exists O_1^{(x)}, O_2^{(x)} \subseteq \mathcal{O} \) such that \( x_0 \in O_1^{(x)}, x \in O_2^{(x)}, \) and \( O_1^{(x)} \cap O_2^{(x)} = \emptyset. \) Then, \( \left( O_2^{(x)} \right)_{x \in K} \) forms an open covering of \( K. \) By the compactness of \( K, \) \( \exists \) a finite set \( K_N \subseteq K \) such that \( \left( O_2^{(x)} \right)_{x \in K_N} \)
forms an open covering of \( K \). Then, \( x_0 \in O := \bigcap_{x \in K} O_1^{(x)} \in \mathcal{O} \), and 
\( O \cap K \subseteq O \cap \left( \bigcup_{x \in K} O_2^{(x)} \right) = \emptyset \). This contradicts with \( x_0 \in \overline{K} \), by Proposition 3.3. Hence, \( K \) is closed.

This completes the proof of the proposition. \( \square \)

**Theorem 5.6 (Heine-Borel)** Let \( A \) be a subset of \( \mathbb{R} \). \( A \) is compact if, and only if, \( A \) is closed and bounded.

**Proof**  
"If" Consider first the special case \( A = [a, b] \), where \( a, b \in \mathbb{R} \) and \( a < b \). Let \( (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_\mathbb{R} \) be an arbitrary open covering of \( A \), where \( \Lambda \) is an index set. Let \( B := \{ x \in A \mid \text{The interval } [a, x] \text{ can be covered by finitely many sets in } (O_\alpha)_{\alpha \in \Lambda} \} \). Clearly, \( a \in B \). Let \( c = \sup B \). Then, \( c \in A \). \exists \alpha_0 \in \Lambda \text{ such that } c \in O_{\alpha_0}. \exists \delta \in (0, \infty) \subset \mathbb{R} \text{ such that } [c-\delta, c+\delta] \subseteq O_{\alpha_0}. \) By the definition of \( c \), \( \exists d \in B \text{ such that } d \in [c-\delta, c]. \) By \( d \in B \), \( [a, d] \) can be covered by finitely many sets in \( (O_\alpha)_{\alpha \in \Lambda} \). Hence, \( c \in B \). Note that we must have \( c = b \) since otherwise \( c < b \) and \( \min\{c+\delta, b\} \in B \) which contradicts the definition of \( c \). Therefore, there exists a finite subcovering of \( [a, b] = A \). This shows that \( A \) is compact.

Now, let \( A \) be any closed and bounded subset of \( \mathbb{R} \). Let \( (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_\mathbb{R} \) be an arbitrary open covering of \( A \), where \( \Lambda \) is an index set. Since \( A \) is bounded, then \( A \subseteq [a, b] \), for some \( a, b \in \mathbb{R} \) with \( a < b \). Note that \( \left( \bigcup_{\alpha \in \Lambda} O_\alpha \right) \cup A \supseteq [a, b] \). By the compactness of \( [a, b] \), there exists a finite set \( \Lambda_N \subseteq \Lambda \text{ such that } [a, b] \subseteq \left( \bigcup_{\alpha \in \Lambda_N} O_\alpha \right) \cup A. \) Then, \( A \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha \). Hence, \( A \) is compact.

"Only if" Since \( \mathbb{R} \) is Hausdorff, by Proposition 5.5, \( A \) is closed. Let \( I_n := (a-n, a) \subset \mathbb{R}, \forall n \in \mathbb{N} \). Then, \( \bigcup_{n \in \mathbb{N}} I_n = A \). By the compactness of \( A \), \( \exists N \in \mathbb{N} \text{ such that } A \subseteq \bigcup_{1 \leq n \leq N} I_n = I_N. \) Hence, \( A \) is bounded.

This completes the proof of the theorem. \( \square \)

**Proposition 5.7** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be topological spaces, \( \mathcal{X} \) be compact, and \( f : \mathcal{X} \to \mathcal{Y} \) be continuous. Then, \( f(\mathcal{X}) \subseteq \mathcal{Y} \) is compact.

**Proof**  
Let \( (O_{Y_\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}_Y \) be an arbitrary open covering of \( f(\mathcal{X}) \), where \( \Lambda \) is an index set. Then, \( (f_{\text{inv}}(O_{Y_\alpha}))_{\alpha \in \Lambda} \subseteq \mathcal{O}_X \) is an open covering of \( \mathcal{X} \), by the continuity of \( f \). By the compactness of \( \mathcal{X} \), there exists a finite set \( \Lambda_N \subseteq \Lambda \text{ such that } (f_{\text{inv}}(O_{Y_\alpha}))_{\alpha \in \Lambda_N} \text{ is a subcovering of } \mathcal{X}. \) Then, by Proposition 2.5, \( (O_{Y_\alpha})_{\alpha \in \Lambda_N} \) is a finite subcovering of \( f(\mathcal{X}) \). Hence, \( f(\mathcal{X}) \) is compact. This completes the proof of the proposition. \( \square \)

**Proposition 5.8** Let \( \mathcal{X} \) be a compact space, \( \mathcal{Y} \) be a Hausdorff topological space, and \( f : \mathcal{X} \to \mathcal{Y} \) be a bijective continuous function. Then, \( f \) is a homeomorphism.

**Proof**  
By Proposition 5.7, \( \mathcal{Y} \) is compact. By the assumption of the proposition, \( f \) is invertible with inverse \( f_{\text{inv}} : \mathcal{Y} \to \mathcal{X}. \) \( \forall \) closed set \( F \subseteq \mathcal{X}. \)
By Proposition 5.5, \( F \) is compact. By Proposition 5.7, \( f(F) \) is compact. By Proposition 5.5, \( f(F) \) is closed. By Proposition 3.10, \( f_{\text{inv}} \) is continuous. Hence, \( f \) is a homeomorphism. This completes the proof of the proposition.

\[ \square \]

**Definition 5.9** Let \((X, \mathcal{O})\) be a topological space and \( \mathcal{U} \subseteq \mathcal{O} \) be an open covering of \( X \). \( \mathcal{V} \subseteq \mathcal{O} \) is said to be an open refinement of \( \mathcal{U} \) (or to refine \( \mathcal{U} \)) if \( \mathcal{V} \) is an open covering of \( X \) and \( \forall V \in \mathcal{V}, \exists U \in \mathcal{U} \) such that \( V \subseteq U \).

**Proposition 5.10** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space. \( \mathcal{X} \) is compact if, and only if, any open covering \( \mathcal{U} \subseteq \mathcal{O} \) has a finite open refinement \( \mathcal{V} \).

**Proof** “Only if” Let \( \mathcal{U} \subseteq \mathcal{O} \) be an open covering of \( X \). By the compactness of \( \mathcal{X} \), there exists a finite set \( \mathcal{V} \subseteq \mathcal{U} \) that covers \( X \). Then, \( \mathcal{V} \) is a finite open refinement of \( \mathcal{U} \).

“If” Let \( \mathcal{U} \subseteq \mathcal{O} \) be an arbitrary open covering of \( X \). Then, there exists a finite open refinement \( \mathcal{V} \) of \( \mathcal{U} \). \( \forall V \in \mathcal{V}, \exists U \in \mathcal{U} \) such that \( V \subseteq U \). By Axiom of Choice and Axiom of Replacement, we may define a set \( \bar{U} := \{ U \in \mathcal{U} \mid V \in \mathcal{V} \} \). Clearly, \( \bar{U} \subseteq \mathcal{U} \) is a finite subcovering of \( X \). Hence, \( \mathcal{X} \) is compact.

This completes the proof of the proposition.

\[ \square \]

**Proposition 5.11** Let \((X, \mathcal{O})\) be a Hausdorff topological space and \((K_n)_{n=1}^{\infty}\) be a sequence of compact subsets of \( X \) with \( K_{n+1} \subseteq K_n, \forall n \in \mathbb{N} \). Let \( O \in \mathcal{O} \) and \( \bigcap_{n \in \mathbb{N}} K_n \subseteq O \). Then, \( \exists n_0 \in \mathbb{N} \) such that \( K_n \subseteq O \).

**Proof** By Proposition 5.5, \( K_n \) is closed, \( \forall n \in \mathbb{N} \). Let \( O_n := O \cup \tilde{K}_n \), \( \forall n \in \mathbb{N} \). Then, we have

\[
\bigcup_{n=1}^{\infty} O_n = O \cup \left( \bigcap_{n=1}^{\infty} \tilde{K}_n \right) = O \cup \left( \bigcap_{n=1}^{\infty} K_n \right) = X \supseteq K_1
\]

By the compactness of \( K_1 \), \( \exists N \in \mathbb{N} \) such that \( K_1 \subseteq \bigcup_{n=1}^{N} O_n = O \cup \left( \bigcap_{n=1}^{N} K_n \right) \). Then, \( K_N \subseteq O \). This completes the proof of the proposition.

\[ \square \]

**Proposition 5.12** Let \( \mathcal{X} \) be a Hausdorff topological space and \( K_\alpha \subseteq \mathcal{X} \) be a compact subset, \( \forall \alpha \in \Lambda \), where \( \Lambda \) is an index set. Assume that \( \bigcap_{\alpha \in \Lambda} K_\alpha = \emptyset \). Then, there exists a finite set \( \Lambda_N \subseteq \Lambda \) such that \( \bigcap_{\alpha \in \Lambda_N} K_\alpha = \emptyset \).

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \Lambda \) is finite; Case 2: \( \Lambda \) is infinite. Case 1: \( \Lambda \) is finite. Choose \( \Lambda_N := \Lambda \) and the result follows.

Case 2: \( \Lambda \) is infinite. \( \forall \alpha \in \Lambda \), \( K_\alpha \) is closed in \( \mathcal{X} \) by Proposition 5.5. Note that \( \bigcup_{\alpha \in \Lambda} \hat{K}_\alpha = \left( \bigcap_{\alpha \in \Lambda} K_\alpha \right) = \mathcal{X} \supseteq K_{\alpha_0} \), for some \( \alpha_0 \in \Lambda \). By the compactness of \( K_{\alpha_0} \), there exists a finite set \( \hat{\Lambda}_N \subseteq \Lambda \) such that \( K_{\alpha_0} \subseteq \bigcup_{\alpha \in \hat{\Lambda}_N} \hat{K}_\alpha \). Let \( \Lambda_N := \hat{\Lambda}_N \cup \{ \alpha_0 \} \), which is a finite set. Then, \( \bigcap_{\alpha \in \Lambda_N} K_\alpha = K_{\alpha_0} \cap \left( \bigcap_{\alpha \in \hat{\Lambda}_N} K_\alpha \right) = \emptyset \). This completes the proof of the proposition.

\[ \square \]
Proposition 5.13 Let $X := (X, \mathcal{O})$ be a Hausdorff topological space, $K_1, K_2 \subseteq X$ be compact subsets, and $K_1 \cap K_2 = \emptyset$. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $K_1 \subseteq O_1$, $K_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$.

**Proof** Since $(X, \mathcal{O})$ is Hausdorff, then, $\forall x_1 \in K_1$ and $\forall x_2 \in K_2$, we must have $x_1 \neq x_2$ and $\exists O^{(1)}_{x_1, x_2}, O^{(2)}_{x_1, x_2} \in \mathcal{O}$ such that $x_j \in O^{(j)}_{x_1, x_2}$, $j = 1, 2$, and $O^{(1)}_{x_1, x_2} \cap O^{(2)}_{x_1, x_2} = \emptyset$. Then, $\forall x_1 \in K_1$, $K_2 \subseteq \bigcup_{x_2 \in K_2} O^{(2)}_{x_1, x_2}$. By the compactness of $K_2$, there exists a finite set $K_{2, x_1} \subseteq K_2$ such that $K_2 \subseteq \bigcup_{x_2 \in K_{2, x_1}} O^{(2)}_{x_1, x_2}$. It is clear that $x_1 \in \bigcap_{x_2 \in K_{2, x_1}} O^{(1)}_{x_1, x_2} =: O^{(1)}_{x_1}$ and $O^{(1)}_{x_1} \cap O^{(2)}_{x_1} = \emptyset$. Then, $K_1 \subseteq \bigcup_{x_1 \in K_1} O^{(1)}_{x_1}$. By the compactness of $K_1$, there exists a finite set $K_{1, N} \subseteq K_1$ such that $K_1 \subseteq \bigcup_{x_1 \in K_{1, N}} O^{(1)}_{x_1} := O_1 \in \mathcal{O}$. Note that $K_2 \subseteq \bigcap_{x_1 \in K_{1, N}} O^{(2)}_{x_1} =: O_2 \in \mathcal{O}$. Clearly, $O_1 \cap O_2 = \emptyset$. This completes the proof of the proposition. \qed

Proposition 5.14 A compact Hausdorff topological space is normal.

**Proof** Let $(X, \mathcal{O})$ be a compact Hausdorff topological space. Fix any closed sets $F_1, F_2 \subseteq X$ with $F_1 \cap F_2 = \emptyset$. By Proposition 5.5, $F_1$ and $F_2$ are compact. By Proposition 5.13, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Hence, $(X, \mathcal{O})$ is normal. This completes the proof of the proposition. \qed

Proposition 5.15 Let $X$ be a set and $\mathcal{O}$ and $\mathcal{O}_1$ be topologies on $X$. Assume that $\mathcal{O}_1$ is weaker than $\mathcal{O}$, that is $\mathcal{O}_1 \subseteq \mathcal{O}$, and $(X, \mathcal{O})$ is compact. Then, $(X, \mathcal{O}_1)$ is compact.

**Proof** Let $U_1 \subseteq \mathcal{O}_1$ be any open covering of $(X, \mathcal{O}_1)$. Then, it is an open covering of $(X, \mathcal{O})$. By the compactness of $(X, \mathcal{O})$, there exists finite subcovering $U_N \subseteq U_1$ of $(X, \mathcal{O})$. Clearly $U_N \subseteq \mathcal{O}_1$ is a finite subcovering of $(X, \mathcal{O}_1)$. Therefore, $(X, \mathcal{O}_1)$ is compact. This completes the proof of the proposition. \qed

Proposition 5.16 Let $X$ be a set and $\mathcal{O}$ and $\mathcal{O}_1$ be topologies on $X$. Assume that $\mathcal{O}_1$ is stronger than $\mathcal{O}$, that is $\mathcal{O} \subseteq \mathcal{O}_1$, and $(X, \mathcal{O})$ is Hausdorff. Then, $(X, \mathcal{O}_1)$ is Hausdorff.

**Proof** $\forall x, y \in X$ with $x \neq y$, $\exists O_1, O_2 \in \mathcal{O}$ such that $x \in O_1$, $y \in O_2$, and $O_1 \cap O_2 = \emptyset$, since $(X, \mathcal{O})$ is Hausdorff. Note that $O_1, O_2 \in \mathcal{O}_1$ since $\mathcal{O} \subseteq \mathcal{O}_1$. Hence, $(X, \mathcal{O}_1)$ is Hausdorff. This completes the proof of the proposition. \qed

Proposition 5.17 Let $(X, \mathcal{O})$ be a compact Hausdorff space. Then, any weaker topology $\mathcal{O}_1 \subseteq \mathcal{O}$ is not Hausdorff, and any stronger topology $\mathcal{O}_2 \supset \mathcal{O}$ is not compact.
Proof Since \( \mathcal{O}_1 \subseteq \mathcal{O} \), then \( \exists O_0 \in \mathcal{O} \setminus \mathcal{O}_1 \). Then, \( \tilde{O}_0 \) is closed in \((X, \mathcal{O})\). By Proposition 5.5, \( \tilde{O}_0 \) is compact in \((X, \mathcal{O})\). By Proposition 5.15, \( \tilde{O}_0 \) is compact in \((X, \mathcal{O}_1)\). Suppose \((X, \mathcal{O}_1)\) is Hausdorff. Then, by Proposition 5.5, \( \tilde{O}_0 \) is closed in \((X, \mathcal{O}_1)\). This implies that \( O_0 \in \mathcal{O}_1 \), which contradicts with \( O_0 \in \mathcal{O} \setminus \mathcal{O}_1 \). Hence, \((X, \mathcal{O}_1)\) is not Hausdorff.

Since \( \mathcal{O}_2 \supset \mathcal{O} \), then \( \exists O_0 \in \mathcal{O}_2 \setminus \mathcal{O} \). Suppose \((X, \mathcal{O}_2)\) is compact. Then, by Proposition 5.5, \( \tilde{O}_0 \) is compact in \((X, \mathcal{O}_2)\). By Proposition 5.15, \( \tilde{O}_0 \) is compact in \((X, \mathcal{O})\). By Proposition 5.5, \( \tilde{O}_0 \) is closed in \((X, \mathcal{O})\). Then, \( O_0 \in \mathcal{O} \), which contradicts with \( O_0 \in \mathcal{O}_2 \setminus \mathcal{O} \). Hence, \((X, \mathcal{O}_2)\) is not compact.

This completes the proof of the proposition.

Proposition 5.18 Let \( \mathcal{X} \) be a compact topological space, \( \mathcal{Y} \) be a Hausdorff topological space, \( \mathcal{Z} \) be a topological space, \( f : \mathcal{X} \to \mathcal{Y} \) be surjective and continuous, and \( g : \mathcal{Y} \to \mathcal{Z} \) be such that \( T := g \circ f : \mathcal{X} \to \mathcal{Z} \) is continuous. Then, \( g \) is continuous.

Proof By Proposition 5.7 and the compactness of \( \mathcal{X} \), \( \mathcal{Y} = f(\mathcal{X}) \) is compact. Then, \( \mathcal{Y} \) is compact and Hausdorff.

Claim 5.18.1 \( \forall O_Z \subseteq \mathcal{Z}, g_{\text{inv}}(O_Z) = f(T_{\text{inv}}(O_Z)) \).

Proof of claim: \( \forall y_0 \in g_{\text{inv}}(O_Z) \subseteq \mathcal{Y}, \exists x_0 \in \mathcal{X} \) such that \( y_0 = f(x_0) \) since \( f \) is surjective. Then, \( g(y_0) = g(f(x_0)) = T(x_0) \in O_Z \). This implies that \( x_0 \in T_{\text{inv}}(O_Z) \). Then, \( y_0 \in f(T_{\text{inv}}(O_Z)) \). Hence, \( g_{\text{inv}}(O_Z) \subseteq f(T_{\text{inv}}(O_Z)) \).

On the other hand, \( \forall y_0 \in f(T_{\text{inv}}(O_Z)), \exists x_0 \in T_{\text{inv}}(O_Z) \) such that \( y_0 = f(x_0) \). Then, \( O_Z \ni T(x_0) = g(f(x_0)) = g(y_0) \). This implies that \( y_0 \in g_{\text{inv}}(O_Z) \). Hence, \( f(T_{\text{inv}}(O_Z)) \subseteq g_{\text{inv}}(O_Z) \).

Therefore, we have \( g_{\text{inv}}(O_Z) = f(T_{\text{inv}}(O_Z)) \). This completes the proof of the claim.

\( \forall \) closed set \( F \subseteq \mathcal{Z}, T_{\text{inv}}(F) \) is closed in \( \mathcal{X} \) by Proposition 3.10 and the continuity of \( T \). Then, \( T_{\text{inv}}(F) \) is compact by Proposition 5.5. This implies that, by Proposition 5.7, \( f(T_{\text{inv}}(F)) \) is compact. By Proposition 5.5, \( g_{\text{inv}}(F) = f(T_{\text{inv}}(F)) \) is closed. Then, by Proposition 3.10 and the arbitrariness of \( F \), \( g \) is continuous.

This completes the proof of the proposition.

Definition 5.19 A compact and connected Hausdorff topological space containing more than one point is called a continuum.

Proposition 5.20 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a Hausdorff topological space, \( (K_n)_{n=1}^{\infty} \) be a sequence of compact and connected subsets of \( \mathcal{X} \), and \( K_{n+1} \subseteq K_n \), \( \forall n \in \mathbb{N} \). Then, \( \bigcap_{n=1}^{\infty} K_n := K \) is compact and connected.
5.2 Countable and Sequential Compactness

Definition 5.21 Let \((X, \mathcal{O})\) be a topological space. A subset \(K \subseteq X\) is said to be countably compact if any countable open covering of \(K\) admits a finite subcovering.

Proposition 5.22 A topological space is compact if, and only if, it is Lindelöf and countably compact.

Proof This is straightforward and therefore omitted. \(\Box\)

Proposition 5.23 A second countable topological space is compact if, and only if, it is countably compact.

Proof A second countable topological space is Lindelöf, by Proposition 3.24. Then, the result follows from Proposition 5.22. \(\Box\)

Proposition 5.24 Let \(\mathcal{X}\) be a countably compact topological space, \(\mathcal{Y}\) be a topological space, and \(f : \mathcal{X} \to \mathcal{Y}\) be continuous. Then, \(f(\mathcal{X})\) is countably compact.

Proof Fix any countable open covering \((O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}_Y\) of \(f(\mathcal{X})\), where \(\Lambda\) is a countable index set. By Proposition 2.5 and the continuity of \(f\), \((f_{\text{inv}}(O_\alpha))_{\alpha \in \Lambda}\) is a countable open covering of \(\mathcal{X}\). By the countable compactness of \(\mathcal{X}\), there exists a finite set \(\Lambda_N \subseteq \Lambda\) such that \(\bigcup_{\alpha \in \Lambda_N} f_{\text{inv}}(O_\alpha) = \mathcal{X}\). Then, we have \(\bigcup_{\alpha \in \Lambda_N} O_\alpha \supseteq \bigcup_{\alpha \in \Lambda_N} f(f_{\text{inv}}(O_\alpha)) = f \left( \bigcup_{\alpha \in \Lambda_N} f_{\text{inv}}(O_\alpha) \right) = f(\mathcal{X})\), by Proposition 2.5. Hence, \(f(\mathcal{X})\) is countably compact. This completes the proof of the proposition. \(\Box\)
Definition 5.25 Let \( \mathcal{X} \) be a topological space. It is said to have the Bolzano-Weierstrass property if \( \forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \), there exists a cluster point \( x \in \mathcal{X} \) of the sequence.

Proposition 5.26 A topological space \( \mathcal{X} \) is countably compact if, and only if, it has the Bolzano-Weierstrass property.

Proof "If" Suppose \( \mathcal{X} \) is not countably compact. Then, there exists an open covering \( (O_n)_{n=1}^{\infty} \subseteq \mathcal{O} \) of \( \mathcal{X} \) which does not have any finite sub-covering. \( \forall n \in \mathbb{N}, \exists x_n \in \bigcup_{i=1}^{n} O_i = \bigcap_{i=1}^{n} \overline{O_i} \). Then, \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \). By the Bolzano-Weierstrass property, \( \exists x_0 \in \mathcal{X} \) such that \( x_0 \) is a cluster point of \( (x_n)_{n=1}^{\infty} \). By Proposition 3.3, \( x_0 \in (x_n)_{n=1}^{\infty}, \forall n \in \mathbb{N}, x_0 \in (x_i)_{i=n}, (x_i)_{i=n}^{\infty} \subseteq \bigcap_{i=1}^{n} \overline{O_i} \), and \( \bigcap_{i=1}^{n} \overline{O_i} \) is closed, then \( x_0 \in \bigcap_{i=1}^{n} \overline{O_i} \). Then, \( x_0 \in \bigcap_{i=1}^{\infty} \overline{O_i} \), which contradicts with the fact that \( (O_n)_{n=1}^{\infty} \) is an open covering of \( \mathcal{X} \). Therefore, \( \mathcal{X} \) is countably compact.

"Only if" Fix any sequence \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \). Let \( B_n := (x_i)_{i=n}, \forall n \in \mathbb{N} \). Suppose \( \bigcap_{n=1}^{\infty} B_n = \emptyset \). Then, \( (B_n)_{n \in \mathbb{N}} \) form a countable open covering of \( \mathcal{X} \). By the countable compactness of \( \mathcal{X} \), \( \exists N \in \mathbb{N} \) such that \( \bigcup_{n=1}^{N} \overline{B_n} = \mathcal{X} \). Hence, we have \( \bigcap_{n=1}^{N} \overline{B_n} = \emptyset \). This contradicts with the fact that \( x_{N+1} \in \bigcap_{n=1}^{N} \overline{B_n} \). Therefore, \( \bigcap_{n=1}^{\infty} \overline{B_n} \neq \emptyset \). Let \( x_0 \in \bigcap_{n=1}^{\infty} \overline{B_n} \subseteq \mathcal{X} \). \( \forall O \in \mathcal{O} \) with \( x_0 \in O, \forall n \in \mathbb{N}, x_0 \in \overline{B_n} \) implies that \( O \cap B_n \neq \emptyset \) by Proposition 3.3. Then, \( \exists m \in \mathbb{N} \) with \( m \geq n \) such that \( x_m \in O \). This shows that \( x_0 \) is a cluster point of \( (x_n)_{n=1}^{\infty} \). Hence, \( \mathcal{X} \) has the Bolzano-Weierstrass property.

This completes the proof of the proposition. \( \square \)

Definition 5.27 A topological space \( \mathcal{X} \) is said to be sequentially compact if \( \forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \), there exists a subsequence \( (x_n)_{n=1}^{\infty} \) that converges to some \( x_0 \in \mathcal{X} \).

Proposition 5.28 Let \( \mathcal{X} \) be a topological space. Then, the following statements hold.

(i) If it is sequentially compact, then \( \mathcal{X} \) is countably compact.

(ii) If it is first countable and countably compact, then \( \mathcal{X} \) is sequentially compact.

Proof (i) \( \mathcal{X} \) is sequentially compact implies that \( \mathcal{X} \) has the Bolzano-Weierstrass property since the limit of a convergent subsequence is a cluster point for the subsequence and hence a cluster point for the original sequence. Then, \( \mathcal{X} \) is countably compact by Proposition 5.26.

(ii) Let \( \mathcal{X} \) be countably compact and first countable. Fix any \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \). By Proposition 5.26, \( \mathcal{X} \) has the Bolzano-Weierstrass property. Then, \( \exists x_0 \in \mathcal{X} \) such that \( x_0 \) is a cluster point of \( (x_n)_{n=1}^{\infty} \). Let \( B_{x_0} := (B_n)_{n=1}^{\infty} \subseteq \mathcal{O} \) be a countable basis at \( x_0 \). (In case \( B_{x_0} \) is finite, we will pad \( \mathcal{X} \) as
5.3. REAL-VALUED FUNCTIONS AND COMPACTNESS

additional basis sets to make it countably infinite.) Let \( n_0 = 0 \), \( \forall i \in \mathbb{N}, \exists n_i \in \mathbb{N} \) with \( n_i > n_{i-1} \) such that \( x_{n_i} \in \bigcap_{j=1}^{i} B_j \in \mathcal{O} \). Clearly, \( (x_{n_i})_{i=1}^{\infty} \) is a subsequence of \( (x_n)_{n=1}^{\infty} \) and admits a limit point \( x_0 \). Hence, \( \mathcal{X} \) is sequentially compact.

This completes the proof of the proposition. \( \square \)

The relationships between different compactness concepts are: compactness implies countable compactness, which is equivalent to the Bolzano-Weierstrass property; sequential compactness implies countable compactness.

5.3 Real-Valued Functions and Compactness

**Proposition 5.29** Let \( \mathcal{X} \) be a countably compact topological space and \( f : \mathcal{X} \to \mathbb{R} \) be continuous. Then, \( f \) is bounded, that is, \( \exists a, b \in \mathbb{R} \) such that \( a \leq f(x) \leq b, \forall x \in \mathcal{X} \). Furthermore, if \( \mathcal{X} \neq \emptyset \), \( \exists x_m, x_M \in \mathcal{X} \) such that \( f(x_m) = \min_{x \in \mathcal{X}} f(\bar{x}) \leq f(x) \leq \max_{x \in \mathcal{X}} f(\bar{x}) = f(x_M) \), \( \forall x \in \mathcal{X} \), i.e., \( f \) achieves its minimum and maximum on \( \mathcal{X} \).

**Proof** Note that \( \mathbb{R} = \bigcup_{n \in \mathbb{N}} (-n, n) \). Then, \( \mathcal{X} = \bigcup_{n \in \mathbb{N}} f_{\text{inv}}((-n, n)) \), by Proposition 2.5. By the continuity of \( f \), \( f_{\text{inv}}((-n, n)) \) is open, \( \forall n \in \mathbb{N} \). By the countable compactness of \( \mathcal{X} \), \( \exists N \in \mathbb{N} \) such that \( \mathcal{X} = \bigcup_{n=1}^{N} f_{\text{inv}}((-n, n)) = f_{\text{inv}}((-N, N)) \). By Proposition 2.5, \( f(\mathcal{X}) \subseteq (-N, N) \). Hence, \( f \) is bounded.

Let \( \mathcal{X} \neq \emptyset \). Define \( M := \sup_{x \in \mathcal{X}} f(x) \) and \( m := \inf_{x \in \mathcal{X}} f(x) \). Then, \( -N \leq m \leq M \leq N \), \( \forall n \in \mathbb{N} \), \( \exists x_n \in \mathcal{X} \) such that \( M - 1/n < f(x_n) \leq M \). We thus obtain a sequence \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \). By Proposition 5.26 and the countable compactness of \( \mathcal{X} \), \( (x_n)_{n=1}^{\infty} \) admits a cluster point \( x_M \in \mathcal{X} \). By Proposition 3.66, the sequence \( (f(x_n))_{n=1}^{\infty} \) admits a cluster point \( f(x_M) \). By construction, \( \lim_{n \in \mathbb{N}} f(x_n) = M \), which is the only cluster point of \( (f(x_n))_{n=1}^{\infty} \). Then, we have \( f(x_M) = M \). By an argument that is similar to the above, \( \exists x_m \in \mathcal{X} \) such that \( f(x_m) = m \). This completes the proof of the proposition. \( \square \)

**Proposition 5.30** Let \( \mathcal{X} \) be a countably compact topological space and \( f : \mathcal{X} \to \mathbb{R} \) be upper semicontinuous. Then, \( f \) is bounded from above, that is, \( \exists b \in \mathbb{R} \) such that \( f(x) \leq b, \forall x \in \mathcal{X} \). Furthermore, if \( \mathcal{X} \neq \emptyset \), then \( \exists x_M \in \mathcal{X} \) such that \( f(x) \leq \max_{x \in \mathcal{X}} f(\bar{x}) = f(x_M) \), \( \forall x \in \mathcal{X} \), i.e., \( f \) achieves its maximum on \( \mathcal{X} \).

**Proof** Note that \( \mathbb{R} = \bigcup_{n \in \mathbb{N}} (-\infty, n) \). Then, \( \mathcal{X} = \bigcup_{n \in \mathbb{N}} f_{\text{inv}}((-\infty, n)) \), by Proposition 2.5. By the upper semicontinuity of \( f \), \( f_{\text{inv}}((-\infty, n)) \) is open, \( \forall n \in \mathbb{N} \). By the countable compactness of \( \mathcal{X} \), \( \exists N \in \mathbb{N} \) such that \( \mathcal{X} = \bigcup_{n=1}^{N} f_{\text{inv}}((-\infty, n)) = f_{\text{inv}}((-\infty, N)) \). By Proposition 2.5, \( f(\mathcal{X}) \subseteq (-\infty, N) \).
Let $X \neq \emptyset$. Define $M := \sup_{x \in X} f(x)$. Then, $-\infty < M \leq N$. For all $n \in \mathbb{N}$, there exists $x_n \in X$ such that $M - 1/n < f(x_n) \leq M$. We thus obtain a sequence $(x_n)_{n=1}^{\infty}$ in $X$. By Proposition 5.26 and the countable compactness of $X$, $(x_n)_{n=1}^{\infty}$ admits a cluster point $x_M \in X$. By Proposition 3.66, the sequence $(f(x_n))_{n=1}^{\infty}$ is nonincreasing and converges to $0$ pointwise. By Dini’s Lemma, $(f(x_n))_{n=1}^{\infty}$ converges uniformly to $f(x_M)$. By construction, $\lim_{n \in \mathbb{N}} f(x_n) = M$, which is the only cluster point of $(f(x_n))_{n=1}^{\infty}$. Therefore, $f(x_M) = M$. This completes the proof of the proposition.

**Lemma 5.31 (Dini’s Lemma)** Let $X := (X, \mathcal{O})$ be a countably compact topological space and $(f_n)_{n=1}^{\infty}$ be a sequence of upper semicontinuous real-valued functions on $X$. Assume that $\forall x \in X$, $(f_n(x))_{n=1}^{\infty} \subseteq \mathbb{R}$ is nonincreasing and converges to $0$. Then, $(f_n)_{n=1}^{\infty}$ converges to the zero function uniformly on $X$.

**Proof** $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, let $O_n := \{x \in X \mid f_n(x) < \epsilon\}$, $\forall n \in \mathbb{N}$. By the upper semicontinuity of $f_n$, $O_n \in \mathcal{O}$, $\forall n \in \mathbb{N}$. Since $\lim_{n \in \mathbb{N}} f_n(x) = 0$, $\forall x \in X$, then $X = \bigcup_{n \in \mathbb{N}} O_n$. By the countable compactness of $X$, $\exists N \in \mathbb{N}$ such that $X = \bigcup_{n=1}^{N} O_n$. $\forall n \in \mathbb{N}$ with $n \geq N$, $\forall x \in X$, $\exists i \in \{1, \ldots, N\}$ such that $x \in O_i$, then $0 \leq f_n(x) \leq f_i(x) < \epsilon$. Hence, the result is established. This completes the proof of the lemma.

**Proposition 5.32** Let $X := (X, \mathcal{O})$ be a topological space and $(f_n)_{n=1}^{\infty}$ be a sequence of upper semicontinuous real-valued functions on $X$. Assume that $\forall x \in X$, $(f_n(x))_{n=1}^{\infty} \subseteq \mathbb{R}$ is nonincreasing and converges to $f_0(x)$. Then,

(i) $f_0$ is upper semicontinuous;
(ii) if, in addition, $X$ is countably compact and $f_0$ is lower semicontinuous, then, $(f_n)_{n=1}^{\infty}$ converges to $f_0$ uniformly on $X$.

**Proof**
(i) $\forall a \in \mathbb{R}$, since, $\forall x \in X$, $f_0(x) < a \iff f_n(x) < a$ for all $n \geq m$, for some $m \in \mathbb{N}$; and $f_{inv}((\infty, a)) \subseteq f_{inv}((\infty, a))$, $\forall i \leq j$, then $f_{inv}((\infty, a)) = \bigcup_{n \in \mathbb{N}} f_{inv}((\infty, a))$. By the upper semicontinuity of $f_n$, $f_{inv}((\infty, a)) \in \mathcal{O}$, $\forall n \in \mathbb{N}$. Hence, $f_{inv}((\infty, a)) \in \mathcal{O}$. Therefore, $f_0$ is upper semicontinuous.

(ii) $f_0$ is continuous by Proposition 3.16. Then, by Proposition 3.16, $(f_n - f_0)_{n=1}^{\infty}$ is a sequence of upper semicontinuous functions, which is nonincreasing and converges to $0$ pointwise. By Dini’s Lemma, $(f_n - f_0)_{n=1}^{\infty}$ converges to the zero function uniformly. Then, $(f_n)_{n=1}^{\infty}$ converges to $f_0$ uniformly.

This completes the proof of the proposition.

**Proposition 5.33** Let $X$ be a topological space and $(f_n)_{n=1}^{\infty}$ be a sequence of upper semicontinuous real-valued functions on $X$ that converges uniformly to $f : X \to \mathbb{R}$. Then, $f$ is upper semicontinuous.
5.4 Compactness in Metric Spaces

Lemma 5.34 A sequentially compact metric space is totally bounded.

Proof Let $X$ be a sequentially compact metric space. Suppose $X$ is not totally bounded. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}$, $X$ cannot be covered by finitely many open balls with radius $\epsilon_0$. Clearly, $X \neq \emptyset$. Fix an $x_1 \in X$. $\forall n \in \mathbb{N}$ with $n \geq 2$, $\exists x_n \in X \setminus \left( \bigcup_{i=1}^{n-1} B(x_i, \epsilon_0) \right)$. The sequence $(x_n)_{n=1}^\infty$ is such that $\rho(x_n, x_m) \geq \epsilon_0$, $\forall n, m \in \mathbb{N}$ with $n \neq m$. Clearly, this sequence does not have any convergent subsequence. This contradicts with the assumption that $X$ is sequentially compact. Therefore, $X$ must be totally bounded. This completes the proof of the lemma.

Definition 5.35 Let $X$ be a sequentially compact metric space and $(O_\alpha)_{\alpha \in \Lambda}$ be an open covering of $X$, where $\Lambda$ is an index set. The Lebesgue number of $(O_\alpha)_{\alpha \in \Lambda}$ is defined by

$$\epsilon := \sup \{ \delta \in (0, \infty) \subset \mathbb{R} \mid \forall x \in X, \exists \bar{x} \in X \ni \rho(x, \bar{x}) \geq \delta \}
\text{and } B(x, \delta) \subseteq O_\alpha \text{ for some } \alpha \in \Lambda\}$$

Lemma 5.36 Let $X$ be a sequentially compact metric space containing at least two distinct points and $(O_\alpha)_{\alpha \in \Lambda}$ be an open covering of $X$, where $\Lambda$ is an index set. Then, the Lebesgue number $\epsilon$ of $(O_\alpha)_{\alpha \in \Lambda}$ is positive and belongs to $\mathbb{R}$.

Proof Define a function $\phi : X \rightarrow \mathbb{R}$ by, $\forall x \in X$,

$$\phi(x) = \sup \{ c \in (0, \infty) \subset \mathbb{R} \mid B(x, c) \subseteq O_\alpha \text{ for some } \alpha \in \Lambda \text{ and} \exists \bar{x} \in X \ni \rho(\bar{x}, x) \geq c \}$$

By Proposition 5.34, $X$ is totally bounded. Then, $\forall x \in X$, $\exists c_x \in (0, \infty) \subset \mathbb{R}$, such that $\bar{x} \in X$ with $\rho(\bar{x}, x) \geq c_x$. Then, $\phi(x) \leq c_x$. Since $(O_\alpha)_{\alpha \in \Lambda}$ is an open cover of $X$, $\exists c_1 \in (0, \infty) \subset \mathbb{R}$ such that $B(x, c_1) \subseteq O_\alpha$ for some $\alpha \in \Lambda$. Since $X$ contains at least two distinct points, then $\exists \bar{x} \in X$.
with $\bar{x} \neq x$ such that $\rho(\bar{x}, x) =: c_2 > 0$. Then, $c := \min\{c_1, c_2\} > 0$ and $\phi(x) \geq c$. Hence, $0 < \phi(x) < +\infty, \forall x \in \mathcal{X}$.

Next, we will show that $\phi$ is continuous. \(\forall x, y \in \mathcal{X}\), we will distinguish two exhaustive and mutually exclusive cases: Case 1: $\phi(x) - \rho(x, y) \leq 0$; Case 2: $\phi(x) - \rho(x, y) > 0$. Then, we have $\phi(y) \geq \phi(x) - \rho(x, y)$. Case 2: $\phi(x) - \rho(x, y) =: l > 0$. Then, $\forall \delta \in (0, l)$, $\exists \alpha \in \Lambda$ and $\exists \bar{x} \in \mathcal{X}$ such that $\mathcal{B}(x, \phi(x) - \delta) \subseteq O_\alpha$ and $\rho(x, \bar{x}) \geq \phi(x) - \delta$. Then, $\rho(\bar{x}, y) \geq \rho(x, \bar{x}) - \rho(x, y) \geq \phi(x) - \phi(x) - \rho(x, y) = l - \delta$. Then, $\forall \bar{y} \in \mathcal{B}(y, l - \delta)$, we have $\rho(x, \bar{y}) \leq \rho(x, y) + \rho(y, \bar{y}) < l - \delta + \rho(x, y) = \phi(x) - \delta$. Then, $\bar{y} \in \mathcal{B}(x, \phi(x) - \delta) \subseteq O_\alpha$. This implies that $\mathcal{B}(y, l - \delta) \subseteq O_\alpha$. Hence, we have $\phi(y) \geq l - \delta$. By the arbitrariness of $\delta$, we have $\phi(y) \geq \phi(x) - \rho(x, y)$. Hence, in both cases, we have arrived at $\phi(y) \geq \phi(x) - \rho(x, y)$. Then, $\phi(x) - \phi(y) \leq \rho(x, y)$. This further implies that $\phi(y) - \phi(x) \leq \rho(x, y)$. Hence, we have $|\phi(x) - \phi(y)| \leq \rho(x, y)$. Hence, $\phi$ is continuous.

By the sequential compactness of $\mathcal{X}$, $\mathcal{X} \neq \emptyset$, and Propositions 5.28 and 5.29, $\exists x_m \in \mathcal{X}$ such that $\phi(x_m) = \min_{x \in \mathcal{X}} \phi(x) \in (0, \infty) \subset \mathbb{R}$. Then, the Lebesgue number $\epsilon = \phi(x_m)$ is positive and belongs to $\mathbb{R}$. This completes the proof of the lemma.

**Theorem 5.37 (Borel-Lebesgue Theorem)** Let $\mathcal{X}$ be a metric space. Then, the following are equivalent.

(i) $\mathcal{X}$ is compact.

(ii) $\mathcal{X}$ is countably compact.

(iii) $\mathcal{X}$ is sequentially compact.

(iv) $\mathcal{X}$ has the Bolzano-Weierstrass property.

**Proof**

(i) $\Rightarrow$ (ii). This follows directly from Definitions 5.1 and 5.21.

(ii) $\Rightarrow$ (iii). Clearly, $\mathcal{X}$ is first countable. By Proposition 5.28, $\mathcal{X}$ is sequentially compact.

(iii) $\Rightarrow$ (i). We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X}$ is a singleton set; Case 3: $\mathcal{X}$ contains at least two distinct points. Case 1: $\mathcal{X} = \emptyset$. Clearly, $\mathcal{X}$ is compact. Case 2: $\mathcal{X}$ is a singleton set. Clearly, $\mathcal{X}$ is compact. Case 3: $\mathcal{X}$ contains at least two distinct points. Fix any open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}$ of $\mathcal{X}$. Let $\epsilon$ be the Lebesgue number of the covering, which is a positive real number by Lemma 5.36. By Proposition 5.34, $\mathcal{X}$ is totally bounded. Then, there exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} = \bigcup_{x \in X_N} \mathcal{B}(x, \epsilon/2), \forall x \in X_N, \mathcal{B}(x, \epsilon/2) \subseteq O_\alpha_x$ for some $\alpha_x \in \Lambda$. Hence, $\mathcal{X} = \bigcup_{x \in X_N} O_{\alpha_x},$ which is a finite subcovering. Therefore, $\mathcal{X}$ is compact.

(ii) $\Leftrightarrow$ (iv). This is proved in Proposition 5.26.

This completes the proof of the theorem.

**Proposition 5.38** A metric space $\mathcal{X}$ is compact if, and only if, it is complete and totally bounded.
5.5. **THE ASCOLI-ARZELÁ THEOREM**

**Proof**  
"Only if" Fix any Cauchy sequence \((x_n)_{n=1}^\infty \subseteq X\). By the compactness of \(X\) and Borel-Lebesgue Theorem, there exists a subsequence \((x_{n_k})_{k=1}^\infty\) that converges to \(x_0 \in X\). Then, it is straightforward to show that \(\lim_{n \to \infty} x_n = x_0\). (Here, the notation \(\lim_{n \to \infty} x_n\) makes sense since metric spaces are Hausdorff.) Hence, \(X\) is complete. By Borel-Lebesgue Theorem and Proposition 5.34, \(X\) is totally bounded.

"If" Fix any sequence \((x_n)_{n=1}^\infty \subseteq X\). We will construct subsequences as following. Take \(\left( x^{(1)}_n \right)_{n=1}^\infty = (x_n)_{n=1}^\infty\). \(\forall m \in \mathbb{N}\) with \(m \geq 2\), since \(X\) is covered by finite many balls of radius \(1/m\), we may find a subsequence \(\left( x^{(m)}_{n_k} \right)_{k=1}^\infty\) of \(\left( x^{(m-1)}_{n_k} \right)_{k=1}^\infty\) such that \(\left( x^{(m)}_{n_k} \right)_{k=1}^\infty \subseteq B(x^{(m)}, 1/m)\) for some \(x^{(m)} \in X\). Now, consider the sequence of diagonal elements \(\left( x^{(m)}_{n_k} \right)_{k=1}^\infty \subseteq X\), which is a subsequence of the original sequence \((x_n)_{n=1}^\infty\). Clearly, this sequence is Cauchy by construction. By the completeness of \(X\), \(\lim_{n \to \infty} x_n = x_0 \in X\). Hence, we have shown that \(X\) is sequentially compact. By Borel-Lebesgue Theorem, \(X\) is compact.

This completes the proof of the proposition. \(\square\)

**Proposition 5.39** Let \(X := (X, \rho)\) be a compact metric space, \(Y := (Y, \sigma)\) be a metric space, and \(f : X \to Y\) be continuous. Then, \(f\) is uniformly continuous.

**Proof** \(\forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \forall x \in X\), by the continuity of \(f\), \(\exists \delta_x \in (0, \infty) \subset \mathbb{R}\) such that \(\sigma(f(x), f(\tilde{x})) < \varepsilon / 2, \forall \tilde{x} \in B(x, \delta_x)\). Then, \(X = \bigcup_{x \in X} B(x, \delta_x / 2)\). By the compactness of \(X\), there exists a finite set \(X_N \subseteq X\) such that \(X = \bigcup_{x \in X_N} B(x, \delta_x / 2)\). Let \(\delta = \min \{ \inf_{x \in X_N} \delta_x / 2, 1 \} \in (0, \infty) \subset \mathbb{R}\). \(\forall x_1, x_2 \in X\) with \(\rho(x_1, x_2) < \delta\), \(\exists x_0 \in X_N\) such that \(x_1 \in B(x_0, \delta_x / 2)\). Then, \(x_2 \in B(x_0, \delta_x)\). This implies that \(x_1, x_2 \in B(x_0, \delta_x)\). Then, \(\sigma(f(x_1), f(x_2)) \leq \sigma(f(x_1), f(x_0)) + \sigma(f(x_0), f(x_2)) < \varepsilon\). Hence, \(f\) is uniformly continuous. This completes the proof of the proposition. \(\square\)

**Proposition 5.40** Let \(n \in \mathbb{Z}_+\) and \(K \subseteq \mathbb{R}^n\). \(K\) is compact if, and only if, \(K\) is closed and bounded.

**Proof** When \(n \in \mathbb{N}\), note that \(\mathbb{R}^n\) is a complete metric space. Then, the result follows from Propositions 5.38 and 4.39. When \(n = 0\), note that \(\mathbb{R}^n\) is a singleton set and is compact. Then, the result follows immediately. \(\square\)

### 5.5 The Ascoli-Arzelá Theorem

**Lemma 5.41** Let \((f_n)_{n=1}^\infty\) be a sequence of functions of a countable set \(D\) to a metric space \(Y\) such that \(\forall x \in D, (f_n(x))_{n \in \mathbb{N}}\) is compact. Then, there
exists a subsequence \((f_{n_k})_{k=1}^\infty\) such that \(\forall x \in D\), the sequence \((f_{n_k}(x))_{k=1}^\infty\) converges to some point \(f_0(x) \in \mathcal{Y}\). 

**Proof** We will distinguish three exhaustive and mutually exclusive cases: Case 1: \(D = \emptyset\); Case 2: \(D \neq \emptyset\) and is finite; Case 3: \(D\) is countably infinite. Case 1: \(D = \emptyset\). Then, \((f_{n})_{n=1}^\infty\) is the subsequence we seek. Case 2: \(D \neq \emptyset\) and is finite. Take \(D = \{x_1, \ldots, x_m\}\) for some \(m \in \mathbb{N}\). Let 

\[
(f_n^{(0)})_{n=1}^\infty = (f_n)_{n=1}^\infty, \quad \forall k = 1, \ldots, m, \quad \left(\frac{\langle f_n^{(k)}(x_k) \rangle}{\rho} \right)_n \subseteq (f_n(x_k))_{n \in \mathbb{N}}
\]

is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence \((f_n^{(k)})_{n=1}^\infty\) of \((f_n^{(k-1)})_{n=1}^\infty\) such that \((f_n^{(k)}(x_k))_{n=1}^\infty\) converges. Hence, the sequence \((f_n^{(m)})_{n=1}^\infty\) is the subsequence we seek. Case 3: \(D\) is countably infinite. Take \(D = \{x_m\}_{m=1}^\infty\). Let 

\[
(f_n^{(0)})_{n=1}^\infty = (f_n)_{n=1}^\infty, \quad \forall k \in \mathbb{N}, \quad \left(\frac{\langle f_n^{(k)}(x_k) \rangle}{\rho} \right)_n \subseteq (f_n(x_k))_{n \in \mathbb{N}}
\]

is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence \((f_n^{(k)})_{n=1}^\infty\) of \((f_n^{(k-1)})_{n=1}^\infty\) such that \((f_n^{(k)}(x_k))_{n=1}^\infty\) converges. Now, the diagonal sequence \((f_n^{(k)})_{n=1}^\infty\) is the subsequence we seek. This completes the proof of the lemma.

**Lemma 5.42** Let \(\mathcal{X} := (X, \mathcal{O})\) be a compact topological space, \(\mathcal{Y} := (Y, \rho)\) be a metric space, \(\mathcal{F}\) be an equicontinuous family of functions of \(\mathcal{X}\) to \(\mathcal{Y}\), and \((f_n)_{n=1}^\infty \subseteq \mathcal{F}\) be such that \(\lim_{n \in \mathbb{N}} f_n(x) = f_0(x), \forall x \in \mathcal{X}\), for some \(f_0 : \mathcal{X} \to \mathcal{Y}\). Then, \(f_0\) is continuous and \((f_n)_{n=1}^\infty\) converges to \(f_0\) uniformly.

**Proof** \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \forall x_0 \in \mathcal{X}\), by the equicontinuity of \(\mathcal{F}\), \(\exists U_{x_0, \epsilon} \in \mathcal{O}\) with \(x_0 \in U_{x_0, \epsilon}\) such that \(\forall x \in U_{x_0, \epsilon}, \forall f \in \mathcal{F}, \rho(f(x), f_0(x)) < \epsilon\).

By \(\lim_{n \in \mathbb{N}} f_n(x_0) = f_0(x_0)\), \(\exists N_{x_0, \epsilon} \in \mathbb{N}\), \(\forall n \in \mathbb{N}\) with \(n \geq N_{x_0, \epsilon}\), we have \(\rho(f_n(x_0), f_0(x_0)) < \epsilon\).

\(\forall x_0 \in \mathcal{X}\), \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\forall x \in U_{x_0, \epsilon}\), we have, by Propositions 4.30, 3.66, and 3.67,

\[
\rho(f_0(x), f_0(x_0)) = \lim_{n \in \mathbb{N}} \rho(f_n(x), f_n(x_0)) \leq \epsilon
\]

Hence, \(f_0\) is continuous at \(x_0\). By the arbitrariness of \(x_0\) and Proposition 3.9, \(f_0\) is continuous.

\(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\mathcal{X} \subseteq \bigcup_{x \in \mathcal{X}} U_{x, \epsilon}\). By the compactness of \(\mathcal{X}\), there exists a finite set \(X_N \subseteq \mathcal{X}\) such that \(\mathcal{X} \subseteq \bigcup_{x \in X_N} U_{x, \epsilon}\). Let \(N := \max\{\sup_{x \in X_N} N_{x, \epsilon}, 1\} \in \mathbb{N}\), \(\forall n \in \mathbb{N}\) with \(n \geq N\), \(\forall x \in \mathcal{X}\), \(\exists x_0 \in X_N\) such that \(x \in U_{x_0, \epsilon}\). Then, we have

\[
\rho(f_n(x), f_0(x)) \leq \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f_0(x_0)) + \rho(f_0(x_0), f_0(x)) < 3\epsilon
\]

Hence, we have \((f_n)_{n=1}^\infty\) converges to \(f_0\) uniformly.

This completes the proof of the lemma. \(\square\)
Lemma 5.43 Let \( X := (X, \mathcal{O}) \) be a topological space, \( \mathcal{Y} := (Y, \rho) \) be a metric space, \( (f_n)_{n=1}^{\infty} \) be an equicontinuous sequence of functions of \( \mathcal{X} \) to \( \mathcal{Y} \), and \( (f_n(x))_{n=1}^{\infty} \) converge at every \( x \in D \), where \( D \subseteq \mathcal{X} \) is dense. Assume that \( \forall x \in \mathcal{X} \), \( (f_n(x))_{n=1}^{\infty} \subseteq \mathcal{Y} \) is complete. Then, \( \exists \) continuous function \( f_0 : \mathcal{X} \rightarrow \mathcal{Y} \) such that \( \lim_{n \in \mathbb{N}} f_n(x) = f_0(x) \), \( \forall x \in \mathcal{X} \).

Proof \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \forall x \in \mathcal{X} \), since \( (f_n)_{n=1}^{\infty} \) is equicontinuous, then \( \exists \mathcal{O} \in \mathcal{O} \) with \( x \in \mathcal{O} \) such that, \( \forall \bar{x} \in \mathcal{O} \), \( \rho(f_n(x), f_n(\bar{x})) < \epsilon \), \( \forall n \in \mathbb{N} \).

Since \( D \) is dense, then, \( \exists x_0 \in D \cap \mathcal{O} \). \( (f_n(x_0))_{n=1}^{\infty} \subseteq \mathcal{Y} \) is Cauchy since it is convergent. Then, \( \exists N \in \mathbb{N} \), \( \forall n, m \geq N \), we have \( \rho(f_n(x_0), f_m(x_0)) < \epsilon \).

Note that
\[
\rho(f_n(x), f_m(x)) \leq \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f_m(x_0)) + \rho(f_m(x_0), f_m(x)) < 3\epsilon
\]
Therefore, \( (f_n(x))_{n=1}^{\infty} \subseteq (f_n(x_0))_{n=1}^{\infty} \subseteq \mathcal{Y} \) is a Cauchy sequence, which converges since \( (f_n(x))_{n=1}^{\infty} \) is complete. Then, we may define \( f_0 : \mathcal{X} \rightarrow \mathcal{Y} \) by \( f_0(x) = \lim_{n \in \mathbb{N}} f_n(x) \).

By Proposition 4.30, 3.66, and 3.67, we have, \( \forall \bar{x} \in \mathcal{O} \),
\[
\rho(f_0(x), f_0(\bar{x})) = \lim_{n \in \mathbb{N}} \rho(f_n(x), f_n(\bar{x})) \leq \epsilon
\]
Hence, \( f_0 \) is continuous at \( x \). By the arbitrariness of \( x \) and Proposition 3.9, \( f_0 \) is continuous. This completes the proof of the lemma. \( \square \)

Theorem 5.44 (Ascoli-Arzelà Theorem) Let \( \mathcal{X} := (X, \mathcal{O}) \) be a separable topological space, \( \mathcal{Y} := (Y, \rho) \) be a metric space, \( \mathcal{F} \) be an equicontinuous family of functions of \( \mathcal{X} \) to \( \mathcal{Y} \), and \( (f_n)_{n=1}^{\infty} \subseteq \mathcal{F} \) be such that \( (f_n(x))_{n \in \mathbb{N}} \subseteq \mathcal{Y} \) is compact, \( \forall x \in \mathcal{X} \). Then, there exists a subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) that converges pointwise to a continuous function \( f_0 : \mathcal{X} \rightarrow \mathcal{Y} \) and the convergence is uniform on any compact subset of \( \mathcal{X} \).

Proof Let \( D \subseteq \mathcal{X} \) be a countable dense set. By Proposition 5.41, there exists a subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) that converges pointwise on \( D \) to a function \( \bar{f} : D \rightarrow \mathcal{Y} \). Note that \( \forall x \in \mathcal{X} \), \( (f_n(x))_{n \in \mathbb{N}} \subseteq \mathcal{Y} \) is compact implies that \( (f_n(x))_{n \in \mathbb{N}} \) is complete by Proposition 5.38. Then, by Proposition 5.43, there exists a continuous function \( f_0 : \mathcal{X} \rightarrow \mathcal{Y} \) to which \( (f_{n_k})_{k=1}^{\infty} \) converges pointwise. By Proposition 5.42, \( (f_{n_k})_{k=1}^{\infty} \) converges to \( f_0 \) uniformly on compact subsets of \( \mathcal{X} \). This completes the proof of the theorem. \( \square \)

5.6 Product Spaces

Lemma 5.45 Let \( \mathcal{A} \) be a collection of subsets of a set \( X \) with the finite intersection property. Then, there is a collection \( \mathcal{M} \) of subsets in \( X \) such
that $\mathcal{A} \subseteq \mathcal{M}$, $\mathcal{M}$ has the finite intersection property, and $\mathcal{M}$ is maximal with respect to this property, that is $\forall C \subseteq \mathcal{X}^2$ with the finite intersection property and $\mathcal{M} \subseteq C$, we have $C = \mathcal{M}$.

**Proof** Clearly, $X \neq \emptyset$ by the fact that $\mathcal{A}$ has the finite intersection property. Define $M := \{C \subseteq \mathcal{X}^2 \mid \mathcal{A} \subseteq C$ and $C$ has the finite intersection property.$\}$. Set containment $\subseteq$ forms an antisymmetric partial ordering on $M$ and $\mathcal{A} \subseteq M \neq \emptyset$. Let $N \subseteq M$ be any nonempty totally ordered subset. Let $B := \bigcup_{C \in N} C \subseteq \mathcal{X}^2$. Fix any finite subcollection $B_N \subseteq B$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $B_N = \emptyset$. Case 2: $B_N \neq \emptyset$. Take $B_N = \{B_1, \ldots, B_n\}$ for some $n \in \mathbb{N}$. $\forall B_i \in B_N$, $\exists C_i \in N$ such that $B_i \in C_i$. Since $N$ is totally ordered by $\subseteq$, without loss of generality, assume that $C_1 \subseteq C_2 \subseteq \cdots \subseteq C_n$. Then, $B_i \in C_n, \forall B_i \in B_N$. By the definition of the set $M$, we have $\bigcap_{i=1}^n B_i = \emptyset$. In both cases, we have arrived at $\bigcap_{B \in B_N} B \neq \emptyset$. Hence, $B$ has finite intersection property. Since $N \neq \emptyset$, then $\mathcal{A} \subseteq B$. Hence, $B \in M$ and is an upper bound of $N$. By Zorn’s Lemma, there exists a maximal element $\mathcal{M}$ of $M$. This completes the proof of the lemma. \qed

**Lemma 5.46** Let $\mathcal{B}$ be a collection of subsets of a set $X$ that is maximal with respect to the finite intersection property. Then, each intersection of finite number of sets in $\mathcal{B}$ is again in $\mathcal{B}$, and each set that meets every set in $\mathcal{B}$ is itself in $\mathcal{B}$.

**Proof** Fix any finite subcollection $B_N \subseteq B$. Let $\bar{B} := \bigcap_{B \in B_N} B$. Then, $\bar{B} \neq \emptyset$. Let $\bar{B} := B \cup \{\bar{B}\} \supseteq B$. It is easy to see that $\bar{B}$ has the finite intersection property. Then, by maximality of $\mathcal{B}$, we have $\bar{B} \in \mathcal{B}$. Hence, $\bar{B} \in \mathcal{B}$.

Let $C \subseteq X$ be such that $C \cap B \neq \emptyset$, $\forall B \in \mathcal{B}$. Let $\bar{B} := B \cup \{C\} \supseteq B$. Fix any finite subcollection $\mathcal{M} \subseteq \bar{B}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $C \notin \mathcal{M}$; Case 2: $C \in \mathcal{M}$. Case 1: $C \notin \mathcal{M}$. Then, $\mathcal{M} \subseteq B$. Then, $\bigcap_{B \in \mathcal{M}} B \neq \emptyset$ since $\mathcal{B}$ has the finite intersection property. Case 2: $C \in \mathcal{M}$. Then, $\mathcal{M} := \mathcal{M} \setminus \{C\} \subseteq \mathcal{B}$ and is a finite set. Then, $\bigcap_{B \in \mathcal{M}} B \in \mathcal{B}$. Then, $\bigcap_{B \in \mathcal{M}} B = C \cap \left(\bigcap_{B \in \mathcal{M}} B\right) \neq \emptyset$. Hence, in both cases, we have $\bigcap_{B \in \mathcal{M}} B \neq \emptyset$. This shows that $\bar{B}$ has the finite intersection property. By the maximality of $\mathcal{B}$, we have $\bar{B} = B$. Hence, $C \in \mathcal{B}$.

This completes the proof of the lemma. \qed

**Theorem 5.47 (Tychonoff Theorem)** Let $X_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be a compact topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set. Then, the product topological space $X := (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} X_\alpha$ is compact.

**Proof** (Bourbaki) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\Lambda = \emptyset$; Case 2: $\Lambda \neq \emptyset$. Case 1: $\Lambda = \emptyset$. Then, $X$ is a singleton set, which is clearly compact.
5.6. PRODUCT SPACES

117

Case 2: \( \Lambda \neq \emptyset \). We will prove this case by Proposition 5.3. Note that the a basis for \( \mathcal{X} \) is given by \( \mathcal{B} := \{ \prod_{\alpha \in \Lambda} O_\alpha \subseteq \chi \mid O_\alpha \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda, \) and \( O_\alpha = X_\alpha \) for all \( \alpha \)'s except finitely many \( \alpha \)'s. Fix a collection \( \mathcal{A} \) of closed sets in \( \mathcal{X} \) with the finite intersection property. By Lemma 5.45, let \( \mathcal{C} \) be a collection of subsets of \( \mathcal{X} \) such that \( \mathcal{A} \subseteq \mathcal{C} \) and \( \mathcal{C} \) is maximal with respect to the finite intersection property.

\[ \forall \alpha \in \Lambda, \text{ let } C_\alpha := \{ C_\alpha \subseteq X_\alpha \mid \exists C \in \mathcal{C} \ni \pi_\alpha(C) = C_\alpha \}. \forall \text{ finite subcollection } \mathcal{C}_{\alpha \in \mathcal{N}} \subseteq \mathcal{C}, \forall C_\alpha \in \mathcal{C}_{\alpha \in \mathcal{N}}, \exists \mathcal{C}_\alpha \in \mathcal{C} \text{ such that } C_\alpha = \pi_\alpha(\mathcal{C}_\alpha). \] Since \( \mathcal{C} \) has the finite intersection property, then \( \bigcap_{\alpha \in \mathcal{C}_{\alpha \in \mathcal{N}}} C_\alpha \neq \emptyset \). Then, by Proposition 2.5, \( \bigcap_{\alpha \in \mathcal{C}_{\alpha \in \mathcal{N}}} C_\alpha = \bigcap_{\alpha \in \mathcal{C}_{\alpha \in \mathcal{N}}} \pi_\alpha(\mathcal{C}_\alpha) \supseteq \pi_\alpha \left( \bigcap_{\alpha \in \mathcal{C}_{\alpha \in \mathcal{N}}} \mathcal{C}_\alpha \right) \neq \emptyset \). Hence, \( \mathcal{C} \) has finite intersection property. Let \( \mathcal{C}_\alpha := \{ \mathcal{C}_\alpha \mid C_\alpha \in \mathcal{C} \}. \) Then, \( \mathcal{C}_\alpha \) has finite intersection property. By the compactness of \( X_\alpha \) and Proposition 5.3, \( \exists x_\alpha \in \bigcap_{\alpha \in \mathcal{C}_\alpha} \mathcal{C}_\alpha. \)

Let \( x \in \mathcal{X} \) be given by \( \pi_\alpha(x) = x_\alpha, \forall \alpha \in \Lambda \). Consider a set \( S \) of the form \( S = \pi_\alpha^{-1}(O_\alpha) \) for some \( \alpha \in \Lambda \) and for some \( O_\alpha \in \mathcal{O}_\alpha \). Then, \( S \cap C \neq \emptyset, \forall C \in \mathcal{C} \), since \( x_\alpha \in \pi_\alpha(C) \). Hence, \( S \in \mathcal{C} \) by Lemma 5.46 and the maximality of \( \mathcal{C} \). \( \forall B \in \mathcal{B} \) with \( x \in B, \exists n \in \mathbb{N}, \exists \alpha_1, \ldots, \alpha_n \in \Lambda, \) and \( \exists O_\alpha_i \in \mathcal{O}_\alpha_i, i = 1, \ldots, n, \text{ such that } x_\alpha_i \in O_\alpha_i \) and \( B = \bigcap_{i=1}^n \pi_\alpha_i^{-1}(O_\alpha_i). \) Then, \( B \in \mathcal{C} \) by Lemma 5.46.

\( \forall F \in \mathcal{A}, F \) is closed. Then, \( F \cap B \neq \emptyset, \forall B \in \mathcal{B} \) with \( x \in B, \) by the finite intersection property of \( \mathcal{C} \). By Proposition 3.3, \( x \in \overline{F} = F. \) Hence, \( x \in \bigcap_{F \in \mathcal{A}} F \neq \emptyset. \) By Proposition 5.3, \( \mathcal{X} \) is compact.

This completes the proof of the theorem. \( \Box \)

**Proposition 5.48** Let \( \mathcal{X}_\alpha := (X_\alpha, \mathcal{O}_\alpha) \) be a sequentially compact topological space, \( \forall \alpha \in \Lambda, \) where \( \Lambda \) is a countable index set. Then, the product topological space \( \mathcal{X} := (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} \mathcal{X}_\alpha \) is sequentially compact.

**Proof** We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( \Lambda = \emptyset; \) Case 2: \( \Lambda \neq \emptyset \) and is finite; Case 3: \( \Lambda \) is countably infinite. Case 1: \( \Lambda = \emptyset. \) Then, \( X = \{ \emptyset \} \) is a singleton set. Clearly, \( \mathcal{X} \) is sequentially compact.

Case 2: \( \Lambda \neq \emptyset \) and is finite. Without loss of generality, assume that \( \Lambda = \{1, \ldots, m\} \) for some \( m \in \mathbb{N}. \) Fix a sequence \( (x_n^{(0)})_{n=1}^\infty \subseteq \mathcal{X}. \) \( \forall k \in \{1, \ldots, m\}, \left( \pi_k(x_n^{(k-1)}) \right)_{n=1}^\infty \subseteq \mathcal{X}_k. \) By the sequential compactness of \( \mathcal{X}_k, \) there exists a subsequence \( \left( x_n^{(k)} \right)_{n=1}^\infty \subseteq (x_n^{(k-1)})_{n=1}^\infty \) such that \( \left( \pi_k(x_n^{(k)}) \right)_{n=1}^\infty \) converges to \( x_k \in \mathcal{X}_k. \) Let \( x \in \mathcal{X} \) be given by \( \pi_k(x) = x_k, \) \( \forall k \in \{1, \ldots, m\}. \) The sequence \( (x_n^{(m)})_{n=1}^\infty \subseteq \mathcal{X} \) is a subsequence of \( (x_n^{(0)})_{n=1}^\infty. \) Clearly, \( \forall k \in \{1, \ldots, m\}, \left( \pi_k(x_n^{(m)}) \right)_{n=1}^\infty \) converges to \( x_k. \) By Proposition 3.67, \( (x_n^{(m)})_{n=1}^\infty \) converges to \( x. \) Hence, \( \mathcal{X} \) is sequentially compact.
Proposition 5.51 A topological space \( \mathcal{X} \) is locally compact if, and only if, the collection \( \mathcal{B}_1 := \{ O \in \mathcal{O} \mid \overline{O} \) is compact \( \} \) forms a basis for \( \mathcal{O} \).

**Proof**

"Only if" Let \( \mathcal{X} \) be locally compact. \( \forall x \in \mathcal{X}, \exists O \in \mathcal{O} \) such that \( x \in B \subseteq O \subseteq \overline{O} \) is compact. Let \( B_1 := \mathcal{O} \cap B \). Then, \( x \in B_1 \subseteq \mathcal{O} \) and \( B_1 \) is closed. By Proposition 5.51, \( \overline{B_1} \) is compact. Therefore, \( B_1 \) is a basis for \( \mathcal{O} \).

"If" This is straightforward.

This completes the proof of the proposition.

\( \qed \)

5.7 Locally Compact Spaces

5.7.1 Fundamental notion

**Definition 5.49** A topological space \( \mathcal{X} \) is locally compact if \( \forall x \in \mathcal{X}, \exists O \in \mathcal{O} \) such that \( \overline{O} \) is compact.

Clearly, a compact space is locally compact.

**Example 5.50** \( \mathbb{R}^n \) is a locally compact space but not a compact space, \( \forall n \in \mathbb{N} \).

**Proposition 5.51** A topological space \( \mathcal{X} \) is locally compact if, and only if, the collection \( \mathcal{B}_1 := \{ O \in \mathcal{O} \mid \overline{O} \) is compact \( \} \) forms a basis for \( \mathcal{O} \).

**Proof**

"Only if" Let \( \mathcal{X} \) be locally compact. \( \forall x \in \mathcal{X}, \exists B \in \mathcal{B}_1 \) such that \( x \in B \subseteq O \subseteq \overline{O} \) is compact. By the compactness of \( K \), there exists a finite set \( K_N \subseteq K \) such that \( K \subseteq \bigcup_{x \in K_N} O_x =: O \subseteq \mathcal{O} \). By Proposition 3.3, \( \overline{O} = \bigcup_{x \in K_N} \overline{O_x} = \bigcup_{x \in K_N} \overline{O_x} \), which is compact. This completes the proof of the proposition.

**Proposition 5.53** Let \( \mathcal{X} \) be a locally compact Hausdorff topological space and \( \mathcal{B}_1 \) be the basis of \( \mathcal{X} \) defined in Proposition 5.51. \( \forall O \in \mathcal{O}, \forall x \in O, \exists B \in \mathcal{B}_1 \) such that \( x \in B \subseteq \overline{B} \subseteq O \) and \( \overline{B} \) is compact.
5.7. **LOCALLY COMPACT SPACES**

**Proof**  \(\forall O \in \mathcal{O}, \forall x \in O,\) we will distinguish two exhaustive and mutually exclusive cases: Case 1: \(O = \emptyset\); Case 2: \(O \subset \mathcal{X}\). Case 1: \(O = \emptyset\). The result holds by Definition 5.49. Case 2: \(O \subset \mathcal{X}\). \(\forall y \in \emptyset,\) since \(\mathcal{X}\) is Hausdorff, then \(\exists O_y^{(1)}, O_y^{(2)} \in \mathcal{O}\) such that \(x \in O_y^{(1)}\), \(y \in O_y^{(2)}\), and \(O_y^{(1)} \cap O_y^{(2)} = \emptyset\). By Proposition 5.51 and Definition 3.17, \(\exists B_y \in \mathcal{B}_I\) such that \(x \in B_y \subseteq O_y^{(1)}\). Then, \(B_y \subseteq O_y^{(2)}\) and \(\overline{B_y} \subseteq O_y^{(2)}\), since \(O_y^{(2)}\) is closed in \(\mathcal{X}\). This implies that \(y \notin \overline{B_y}\). Note that \(\overline{O}\) is closed in \(\mathcal{X}\), then \(\overline{\overline{B_y}} = \emptyset\). By Proposition 3.3. This completes the proof of the proposition.

**Proposition 5.54** Let \(\mathcal{X}\) be a locally compact Hausdorff topological space, \(U \in \mathcal{O},\) and \(K \subseteq U\) be closed. Then, \(\exists V \in \mathcal{O}\) such that \(K \subseteq V \subseteq \overline{V} \subseteq U\) and \(\overline{V}\) is compact.

**Proof**  \(\forall x \in K,\) by Proposition 5.53, \(\exists V_x \in \mathcal{O}\) such that \(x \in V_x \subseteq \overline{V_x} \subseteq U\) and \(\overline{V_x}\) is compact. Then, \(K \subseteq \bigcup_{x \in K} V_x\). By the compactness of \(K\), there exists a finite set \(K_N \subseteq K\) such that \(K \subseteq \bigcup_{x \in K_N} V_x =: V \in \mathcal{O}\). By Proposition 3.3, \(V = \bigcup_{x \in K_N} V_x \subseteq U\) and is compact. This completes the proof of the proposition.

**Proposition 5.55** Let \(\mathcal{X}\) be a locally compact space and \(F \subseteq \mathcal{X}\). Then, \(F\) is closed if, and only if, for any closed and compact set \(K \subseteq \mathcal{X}\), we have \(F \cap K\) is closed.

**Proof**  “Only if” Let \(F\) be closed. For any closed and compact \(K \subseteq \mathcal{X},\) \(F \cap K\) is closed.

“\(\forall x \in F\)” By local compactness of \(\mathcal{X}\), \(\exists O \in \mathcal{O}\) such that \(x \in O \subseteq \overline{O}\) and \(\overline{O}\) is compact. By the assumption, \(\overline{O} \cap F\) is closed. \(\forall U \in \mathcal{O}\) with \(x \in U,\) we have \(x \in O \cap U \subseteq O\). By Proposition 3.3, \((O \cap U) \cap F \neq \emptyset,\) since \(x \in F\).

Then, we have \(U \cap (\overline{O} \cap F) \neq \emptyset\). Hence, we have \(x \in \overline{O} \cap F = \overline{O} \cap F \subseteq F,\) by Proposition 3.3. Hence, \(\overline{F} \subseteq F\) and \(F\) is closed.

This completes the proof of the proposition.

**Proposition 5.56** Let \(\mathcal{X}\) be a locally compact Hausdorff topological space and \((O_n)_{n=1}^{\infty} \subseteq \mathcal{O}\) be a sequence of open dense subsets in \(\mathcal{X}\). Then, \(\bigcap_{n=1}^{\infty} O_n\) is dense. Therefore, \(\mathcal{X}\) is second category everywhere.

**Proof**  \(\forall U \in \mathcal{O}\) with \(U \neq \emptyset,\) since \(O_1\) is dense, then \(\exists x_1 \in U \cap O_1.\) By Proposition 5.53, \(\exists V_1 \in \mathcal{O}\) such that \(x_1 \in V_1 \subseteq \overline{V_1} \subseteq U \cap O_1\) and \(\overline{V_1}\) is
compact. \( \forall n \in \mathbb{N} \) with \( n \geq 2 \), since \( O_n \) is dense, then \( \exists x_n \in V_{n-1} \cap O_n \). By Proposition 5.53, \( \exists V_n \in \mathcal{O} \) such that \( x_n \in V_n \subseteq \overline{V_n} \subseteq V_{n-1} \cap O_n \) and \( \overline{V_n} \) is compact. Hence, \( (\overline{V_n})_{n=1}^{\infty} \) is a sequence of nonempty closed sets which is nonincreasing (that is \( V_{n+1} \subseteq V_n \), \( \forall n \in \mathbb{N} \)). Clearly, this sequence have the finite intersection property. By the compactness of \( \overline{V_n} \) and Proposition 5.3, \( \bigcap_{n=1}^{\infty} \overline{V_n} \neq \emptyset \). Clearly, \( \bigcap_{n=1}^{\infty} V_n \subseteq U \cap \bigcap_{n=1}^{\infty} O_n \neq \emptyset \). Therefore, \( \bigcap_{n=1}^{\infty} O_n \) is dense in \( X \), by the arbitrariness of \( U \).

By Proposition 3.38, \( X \) is second category everywhere. This completes the proof of the proposition. \( \Box \)

**Proposition 5.57** Let \( X := (X, \mathcal{O}) \) be a Hausdorff space, \( Y \subseteq X \) be a dense subset, and \( \mathcal{O}_Y \) be the subset topology on \( Y \). Assume that \( Y := (Y, \mathcal{O}_Y) \) is locally compact. Then, \( Y \) is open in \( X \).

**Proof** \( \forall y \in Y \), by the local compactness of \( Y \), \( \exists U_y \in \mathcal{O}_Y \) such that \( y \in U_y \subseteq U_{yc} \subseteq Y \), where \( U_{yc} \) is the closure of \( U_y \) in \( Y \), and \( U_{yc} \) is compact. Then, \( U_{yc} \) is compact in \( X \) and is closed, by Proposition 5.5. Then, \( U_{yc} \supseteq \overline{U_y} \). On the other hand, by Proposition 3.5, \( U_{yc} = \overline{U_y} \cap Y \subseteq \overline{U_y} \). Hence, we have \( U_{yc} = \overline{U_y} \). Since \( U_y \in \mathcal{O}_Y \), \( \exists O \in \mathcal{O} \) such that \( U_y \subseteq O \cap Y \subseteq \overline{O} \). By Proposition 3.37, we have \( U_{yc} \subseteq \overline{O} \cap Y \subseteq \overline{O} \). Hence, we have \( y \in U_y \subseteq O \subseteq \mathcal{O} \subseteq U_{yc} \subseteq Y \). Hence, \( Y \) is open in \( X \). This completes the proof of the proposition. \( \Box \)

**Lemma 5.58** Let \( X := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \in \mathcal{O} \), and \( \mathcal{O}_Y \) be the subset topology on \( Y \). Then, \( Y := (Y, \mathcal{O}_Y) \) is locally compact.

**Proof** \( \forall x \in Y \), by Proposition 5.53, \( \exists U \in \mathcal{O} \) such that \( x \in U \subseteq \overline{U} \subseteq Y \) and \( \overline{U} \) is compact. Then, \( U \in \mathcal{O}_Y \) and the closure of \( U \) in \( Y \) is \( \overline{U} \), by Proposition 3.5. Then, \( \overline{U} \) is compact in \( Y \). Hence, \( Y \) is locally compact. This completes the proof of the lemma. \( \Box \)

**Lemma 5.59** Let \( X := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \subseteq X \) be closed, and \( \mathcal{O}_Y \) be the subset topology on \( Y \). Then, \( Y := (Y, \mathcal{O}_Y) \) is locally compact.

**Proof** \( \forall x \in Y \), by the local compactness of \( X \), \( \exists U \in \mathcal{O} \) such that \( x \in U \subseteq \overline{U} \) and \( \overline{U} \) is compact. Then, \( \overline{U} := U \cap Y \in \mathcal{O}_Y \). Let \( \overline{U} \) be the closure of \( U \) in \( Y \). Then, \( \overline{U} \subseteq U \cap Y \subseteq \overline{U} \). Note that \( \overline{U} \) is a closed set in \( Y \) and therefore a closed set in \( X \) since \( Y \) is closed in \( X \) and Proposition 3.5. Then, \( \overline{U} \) is compact by Proposition 5.5. Thus, we have \( x \in U \subseteq \overline{U} \subseteq Y \) and \( \overline{U} \) is compact. Hence, \( Y \) is locally compact. This completes the proof of the lemma. \( \Box \)

**Proposition 5.60** Let \( X := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \subseteq X \), and \( \mathcal{O}_Y \) be the subset topology on \( Y \). \( Y := (Y, \mathcal{O}_Y) \) is locally compact if, and only if, \( Y \) is relatively open in \( \overline{Y} \).
5.7. LOCALLY COMPACT SPACES

Proof  “Only if” Let \( \mathcal{O}_Y \) be the subset topology on \( \overline{Y} \). Then, \( (\overline{Y}, \mathcal{O}_Y) \) is Hausdorff. By Proposition 5.57, \( Y \) is relative open in \( \mathcal{O}_Y \).

“If” Since \( \overline{Y} \) is closed in \( X \), then, by Lemma 5.59, \( (\overline{Y}, \mathcal{O}_Y) \) is locally compact. Clearly, \( (\overline{Y}, \mathcal{O}_Y) \) is Hausdorff. In \( (\overline{Y}, \mathcal{O}_Y) \), by Lemma 5.58, \( Y \) is locally compact.

This completes the proof of the proposition. \( \square \)

5.7.2 Partition of unity

Definition 5.61 Let \( \mathcal{X} \) be a topological space. For a function \( f : \mathcal{X} \to \mathbb{R} \), the support of \( f \) is the set \( \text{supp}(f) := \{ x \in \mathcal{X} \mid f(x) \neq 0 \} \). A collection of real-valued functions \( (\phi_\alpha)_\alpha \in \Gamma \) on \( \mathcal{X} \) is said to be subordinate to a collection of open sets \( (O_\lambda)_\lambda \in \Lambda \) if \( \forall \alpha \in \Gamma \), \( \exists \lambda \in \Lambda \) such that \( \text{supp}(\phi_\alpha) \subseteq O_\lambda \).

Proposition 5.62 Let \( \mathcal{X} \) be a locally compact Hausdorff topological space, \( U \in \mathcal{O} \), and \( K \subseteq U \) be compact. Then, there exists a continuous function \( f : \mathcal{X} \to [0,1] \subseteq \mathbb{R} \) such that \( f|_K = 1 \) and \( f|_{\overline{K}} = 0 \). Furthermore, \( \text{supp}(f) \subseteq U \) is compact.

Proof  Since the set \( Q := \mathbb{Q} \cap [0,1] \) is countable, then, by recursively applying Proposition 5.54, we may find \( (O_r)_{r \in Q} \subseteq \mathcal{O} \) such that the following two properties are satisfied:

1. \( \forall r \in Q, K \subseteq O_r \subseteq \overline{O_r} \subseteq U \) and \( \overline{O_r} \) is compact;
2. \( \forall r, s \in Q \) with \( r < s \), \( \overline{O_r} \subseteq O_s \).

Define the real-valued function \( \tilde{f} : X \to \mathbb{R} \) by

\[ \tilde{f}(x) = \inf(\{ r \in Q \mid x \in O_r \} \cup \{1\}) \]

Clearly, \( \tilde{f} : X \to [0,1] \), \( \tilde{f}(x) = 0 \), \( \forall x \in O_0 \), and \( \tilde{f}(x) = 1 \), \( \forall x \in \overline{O_1} \). By 1, we have \( K \subseteq O_0 \) and \( O_1 \subseteq U \). Next, we will show that \( \tilde{f} \) is continuous. \( \forall x_0 \in X \), we will show that \( \tilde{f} \) is continuous at \( x_0 \). Let \( a_0 = \tilde{f}(x_0) \in [0,1] \).

\( \forall B \subseteq \mathbb{R} \) with \( B \) being open and \( a_0 \in B \), \( \exists a_1, a_2, a_3, a_4 \in \mathbb{Q} \) such that \( a_1 < a_2 < a_0 < a_3 < a_4 \) and \( (a_1, a_4) \subseteq B \). Let \( \bar{a}_2 = \max\{a_2, 0\} \) and \( \bar{a}_3 = \min\{a_3, 1\} \). Then, we must have \( a_1 < \bar{a}_2 \leq a_0 \leq \bar{a}_3 < a_4 \) and \( \bar{a}_2, \bar{a}_3 \in Q \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( a_0 \in (0,1) \); Case 2: \( a_0 = 0 \); Case 3: \( a_0 = 1 \).

Case 1: \( a_0 \in (0,1) \). Then, we must have \( a_1 < \bar{a}_2 < a_0 < \bar{a}_3 < a_4 \). Let \( V = \overline{O_{\bar{a}_2}} \cap O_{\bar{a}_3} \in \mathcal{O} \). \( \forall x \in V \), we have \( x \in O_{\bar{a}_3} \) and \( \tilde{f}(x) \leq \bar{a}_3 \). Also, \( x \in \overline{O_{\bar{a}_2}} \) implies that \( \tilde{f}(x) \geq \tilde{a}_2 \). Hence, \( \tilde{f}(V) \subseteq [\tilde{a}_2, \bar{a}_3] \subseteq (a_1, a_4) \subseteq B \).

\( \tilde{f}(x_0) = a_0 \leq \bar{a}_3 \) implies that \( x_0 \in O_{\bar{a}_3} \). \( \tilde{f}(x_0) = a_0 > \tilde{a}_2 \) implies that \( \exists \tilde{a}_2 \in (\tilde{a}_2, a_0) \cap Q \) such that \( x_0 \in \overline{O_{\tilde{a}_2}} \subseteq \overline{O_{\bar{a}_2}} \). Therefore \( x_0 \in V \). This shows that \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( \tilde{f}(V) \subseteq B \).

Case 2: \( a_0 = 0 \). Then, we must have \( a_1 < 0 = a_0 < \bar{a}_3 < a_4 \). Take \( V = O_{\bar{a}_3} \in \mathcal{O} \). We must have \( x_0 \in V \). \( \forall x \in V, 0 \leq \tilde{f}(x) \leq \bar{a}_3 \). Hence,
Theorem 5.63 (Partition of Unity) Let $X$ be a locally compact Hausdorff topological space, $K \subseteq X$ be compact, and $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}$ be an open covering of $K$, where $\Lambda$ is an index set. Let $\mathcal{F}$ be a collection of continuous real-valued functions on $X$ such that

(i) the constantly zero function is in $\mathcal{F}$;
(ii) $\forall f, g \in \mathcal{F}, f + g \in \mathcal{F}$;
(iii) $\forall f, g \in \mathcal{F}$ with $\text{supp}(f) \subseteq M := \{x \in X \mid g(x) \neq 0\}$, define $f/g : X \to \mathbb{R}$ by

$$
(f/g)(x) = \begin{cases} 
\frac{f(x)}{g(x)} & \forall x \in \text{supp}(f) \\
0 & \forall x \in \text{supp}(f)
\end{cases}
$$

then $f/g \in \mathcal{F}$;
(iv) $\forall O \in \mathcal{O}, \forall x_0 \in O, \exists f \in \mathcal{F}$ such that $f(x_0) = 1$, $f|_{\bar{O}} = 0$, and $f : X \to [0, 1]$.

Then, there exists a finite collection $\Phi \subseteq \mathcal{F}$ of continuous nonnegative real-valued functions on $X$, which is subordinate to $(O_\lambda)_{\lambda \in \Lambda}$, such that

$$
\left(\sum_{\phi \in \Phi} \phi\right)|_K = 1, \quad \sum_{\phi \in \Phi} \phi(x) \in [0, 1] \subseteq \mathbb{R}, \forall x \in X, \text{ and supp}(\phi) \text{ is compact, } \forall \phi \in \Phi.
$$

Remark 5.64 Note that in (iii) above $f/g$ is always continuous since $(f/g)|_{\text{supp}(f)} = 0$ is continuous, $(f/g)|_M$ is continuous, $\text{supp}(f)$ and $M$ are open, $\text{supp}(f) \cup M = X$, and Theorem 3.11 implies the result.

Proof Let $O := \bigcup_{\lambda \in \Lambda} O_\lambda \supseteq K$. By Proposition 5.54, there exists $U \in \mathcal{O}$ such that $K \subseteq U \subseteq \bar{U} \subseteq O$ and $\bar{U}$ is compact. $\forall \lambda \in \Lambda$, let $U_\lambda := U \cap O_\lambda \in \mathcal{O}$. Then, $K \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda = U$. 

\[ \bar{f}(V) \subseteq [0, \bar{a}_3] \subset (a_1, a_4) \subseteq B. \] Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Case 3: $a_0 = 1$. Then, we must have $a_1 < \bar{a}_2 < a_0 = 1 < a_4$. Take $V = \overline{O_{\bar{a}_2}} \in \mathcal{O}$. Since $\bar{f}(x_0) = a_0 = 1$, then $x_0 \in \overline{O_{\bar{a}_2}} \subseteq \overline{O_{a_2}} = V$. $\forall x \in V$, $\bar{f}(x) \geq \bar{a}_2$. Hence, $\bar{f}(V) \subseteq [\bar{a}_2, 1] \subset (a_1, a_4) \subseteq B$. Hence, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Therefore, in all cases, $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$. Then, $\bar{f}$ is continuous at $x_0$. By the arbitrariness of $x_0$ and Proposition 3.9, $\bar{f}$ is continuous. Define $f : X \to [0, 1] \subseteq \mathbb{R}$ by $f(x) = 1 - \bar{f}(x), \forall x \in X$. Clearly, $f$ is continuous by Proposition 3.12, $f|_K = 1$, and $f|_{\bar{O}} = 0$.

Note that $\text{supp}(f) = \{x \in X \mid f(x) < 1\} \subseteq \overline{O_1} \subseteq U$ and is compact by Proposition 5.5. This completes the proof of the proposition. \(\square\)
∀x₀ ∈ K, ∃λ₀ ∈ Λ such that x₀ ∈ Uₜₐₜ λ₀. By Proposition 5.53, ∃Uₜₐₜ λ₀ ∈ O such that x₀ ∈ Uₜₐₜ λ₀ ⊆ Uₜₐₜ λ₀ and Uₜₐₜ λ₀ is compact. By (iv), there exists a continuous function fₚₜ x₀ : X → [0, 1] in F such that fₚₜ x₀(x₀) = 1 and fₚₜ x₀(x) = 0. Let Vₚₜ x₀ := \{ x ∈ X | fₚₜ x₀(x) > 0 \}. Then, x₀ ∈ Vₚₜ x₀ ∈ O since fₚₜ x₀ is continuous. Note that Vₚₜ x₀ ⊆ Ûₜₐₜ λ₀, which implies that supp(fₚₜ x₀) = Vₚₜ x₀ ⊆ Ûₜₐₜ λ₀ ⊆ Uₜₐₜ λ₀ ⊆ Oₜₐₜ λ₀, which is compact by Proposition 5.5. Hence, fₚₜ x₀ has compact support.

∀x₀ ∈ U \ K, K ∈ O, by Proposition 5.5. By (iv), there exists a continuous function gₚₜ x₀ : X → [0, 1] in F such that gₚₜ x₀(x₀) = 1 and gₚₜ x₀(x) = 0. Let Wₚₜ x₀ := \{ x ∈ X | gₚₜ x₀(x) > 0 \}. Then, x₀ ∈ Wₚₜ x₀ ∈ O since gₚₜ x₀ is continuous. Clearly, Wₚₜ x₀ ⊆ K. Note that U ∈ (∪x∈K Vₓ) ∪ (∪x∈∪K \ K Wₓ). By the compactness of U, there exist finite sets K₆ ≺ K and U "₇ ≺ U \ K such that U ∈ (∪x∈K₆ Vₓ) ∪ (∪x∈∪₇ Vₓ). By the construction of Wₓ’s, K₆ ⊆ ∪x∈K₆ Vₓ. Let f := ∑x∈K₆ fₓ ∈ F and g := ∑x∈∪₇ gₓ ∈ F (here f, g ∈ F by (i) and (ii)). Then, f is nonnegative, f(x) > 0, ∀x ∈ K, g is nonnegative, f(x) + g(x) > 0, ∀x ∈ U.

Define Φ := \{ φ : X → [0, 1] | φ = fₓ/(f + g), x ∈ K₆ \}, which is a finite set. Note that ∀x ∈ K₆, supp(fₓ) ⊆ Uₜₐₜ λ₀ ⊆ U ⊆ U ≺ \{ x ∈ X | (f + g)(x) ≠ 0 \}, for some λ₀ ∈ Λ. Then, fₓ/(f + g) ∈ F. Hence, Φ ⊆ F. ∀φ ∈ Φ, ∃x ∈ K₆ such that φ = fₓ/(f + g). Then, φ is continuous by the fact that φ ∈ F. supp(φ) = supp(fₓ) is compact and supp(φ) ⊆ Oₜₐₜ λ₀, for some λ ∈ Λ. Hence, Φ is subordinate to (Oₜₐₜ λ₀)ₗ∈Λ. Clearly, φ is nonnegative, ∀φ ∈ Φ, ∀x ∈ K, we have

\[
\sum_{φ ∈ Φ} φ(x) = \frac{f(x)}{f(x) + g(x)} = \frac{f(x)}{f(x)} = 1
\]

∀x ∈ X, we have

\[
0 ≤ \sum_{φ ∈ Φ} φ(x) = (f/(f + g))(x) ≤ 1
\]

This completes the proof of the theorem.

**Corollary 5.65** Let X be a locally compact Hausdorff topological space, K ≺ X be compact, and (Oₜₐₜ λ₀)ₗ∈Λ ⊆ O be an open covering of K, where Λ is an index set. Then, there exists a finite collection Φ of continuous nonnegative real-valued functions on X, which is subordinate to (Oₜₐₜ λ₀)ₗ∈Λ, such that \( \left( \sum_{φ ∈ Φ} φ \right) \big|_{K} = 1 \), \( \sum_{φ ∈ Φ} φ(x) ∈ [0, 1] ⊆ R \), ∀x ∈ X, and supp(φ) is compact, ∀φ ∈ Φ.

**Proof** Let F be the collection of continuous real-valued functions on X. Clearly, F satisfies (i) – (iii) in Theorem 5.63. Note that \( \{ x₀ \} \) is compact, by Proposition 5.62, (iv) of Theorem 5.63 is also satisfied by F. Then, the result follows.
5.7.3 The Alexandroff one-point compactification

**Theorem 5.66 (Alexandroff One-Point Compactification)** Let $(X, \mathcal{O})$ be a locally compact Hausdorff topological space. The Alexandroff one-point compactification of $X$ is the set $X_\omega := X \cup \{\omega\}$ with the topology $\mathcal{O}_\omega := \{O \subseteq X_\omega \mid O \in \mathcal{O} \text{ or } X_\omega \setminus O \text{ is compact in } X\}$. Then, $X_\omega := (X_\omega, \mathcal{O}_\omega)$ is a compact Hausdorff space and the identity map id : $X \to X_\omega \setminus \{\omega\}$ is a homeomorphism. The element $\omega$ is called the point at infinity in $X_\omega$.

**Proof** We first show that $\mathcal{O}_\omega$ is a topology on $X_\omega$. (i) $\emptyset \in \mathcal{O}_\omega$, then $\emptyset \in \mathcal{O}_\omega$; $X_\omega \setminus X \subset X_\omega \setminus \{\omega\}$ is compact in $X$, then $X \in \mathcal{O}_\omega$. (ii) $\forall O_1, O_2 \in \mathcal{O}_\omega$, we will distinguish four exhaustive and mutually exclusive cases: Case 1: $O_1, O_2 \in \mathcal{O}$; Case 2: $X \setminus O_1$ and $X \setminus O_2$ are compact in $X$; Case 3: $X \setminus O_1$ is compact in $X$ and $O_2 \in \mathcal{O}$; Case 4: $O_1 \in \mathcal{O}$ and $X \setminus O_2$ is compact in $X$. Case 1: $O_1 \cap O_2 \in \mathcal{O}$. Then, $O_1 \cap O_2 \in \mathcal{O}$ and hence $O_1 \cap O_2 \in \mathcal{O}$. Case 2: $X \setminus O_1$ and $X \setminus O_2$ are compact in $X$. Then, $X \setminus (O_1 \cap O_2) = (X \setminus O_1) \cup (X \setminus O_2)$, which is compact in $X$. This implies that $O_1 \cap O_2 \in \mathcal{O}_\omega$. Case 3: $X \setminus O_1$ is compact in $X$ and $O_2 \in \mathcal{O}$. Let $O_3 := O_1 \setminus \{\omega\}$, then $O_1 = X \setminus (X \setminus O_3)$. Since $X \setminus O_1$ is compact in $X$ and $X$ is Hausdorff, then, by Proposition 5.5, $X \setminus O_1$ is closed in $X$. Then, $O_1 \subseteq O_\omega$. Note that $O_1 \cap O_2 = O_1 \cap O_2 \subseteq O_\omega$. Then, $O_1 \cap O_2 \in \mathcal{O}_\omega$. Case 4: $O_1 \in \mathcal{O}$ and $X \setminus O_2$ is compact in $X$. By an argument that is similar to Case 3, we have $O_1 \cap O_2 \in \mathcal{O}_\omega$. Hence, in all four cases, we have $O_1 \cap O_2 \in \mathcal{O}_\omega$.

(iii) $\forall (O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}_\omega$, where $\Lambda$ is an index set, we will distinguish two exhaustive and mutually exclusive cases: Case A: $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}$; Case B: $\exists \lambda_0 \in \Lambda$ such that $\omega \in O_{\lambda_0}$. Case A: $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}$ and hence $\bigcup_{\lambda \in \Lambda} O_\lambda \subseteq \mathcal{O}$. Case B: $\exists \lambda_0 \in \Lambda$ such that $\omega \in O_{\lambda_0}$. Then, $X \setminus O_{\lambda_0}$ is compact in $X$. We may partition $\Lambda$ into two disjoint set $\Lambda_1$ and $\Lambda_2$ such that $\Lambda = \Lambda_1 \cup \Lambda_2$, $\Lambda_1 \cap \Lambda_2 = \emptyset$, $\forall \lambda \in \Lambda_1$, $O_{\lambda} \in \mathcal{O}$, $\forall \lambda \in \Lambda_2$, $X \setminus O_{\lambda}$ is compact in $X$. Note that

$$X_\omega \setminus \left( \bigcup_{\lambda \in \Lambda} O_{\lambda} \right) = \left( \bigcap_{\lambda \in \Lambda_1} (X \setminus O_{\lambda}) \right) \cap \left( \bigcap_{\lambda \in \Lambda_2} (X \setminus O_{\lambda}) \right)$$

$$= \left( \bigcap_{\lambda \in \Lambda_1} (X \setminus O_{\lambda}) \right) \cap \left( \bigcap_{\lambda \in \Lambda_2} (X \setminus O_{\lambda}) \right) \cap (X \setminus O_{\lambda_0})$$

$\forall \lambda \in \Lambda_1$, $X \setminus O_{\lambda}$ is a closed set in $X$. $\forall \lambda \in \Lambda_2$, $X \setminus O_{\lambda}$ is compact in $X$ and therefore closed in $X$ by Proposition 5.5. Hence, $X \setminus \left( \bigcup_{\lambda \in \Lambda} O_{\lambda} \right)$ is a closed subset of $X \setminus O_{\lambda_0}$ and hence compact in $X$ by Proposition 5.5. Then, $\bigcup_{\lambda \in \Lambda} O_{\lambda} \subseteq \mathcal{O}_\omega$. Hence, in both cases, we have $\bigcup_{\lambda \in \Lambda} O_{\lambda} \subseteq \mathcal{O}_\omega$. Summarizing the above, $\mathcal{O}_\omega$ is a topology on $X_\omega$.

Next, we show that $X_\omega$ is compact. Fix an open covering $(O_\lambda)_{\lambda \in \Lambda} \subseteq \mathcal{O}_\omega$ of $X_\omega$. We may partition $\Lambda$ into two disjoint set $\Lambda_1$ and $\Lambda_2$ such that $\Lambda = \Lambda_1 \cup \Lambda_2$, $\Lambda_1 \cap \Lambda_2 = \emptyset$, $\forall \lambda \in \Lambda_1$, $O_{\lambda} \in \mathcal{O}$, $\forall \lambda \in \Lambda_2$, $X \setminus O_{\lambda}$ is compact
in \( X \). Since \( \omega \in X_c \), then \( \exists \lambda_0 \in \Lambda_2 \) such that \( \omega \in O_{c\lambda_0} \). Then, \( X_c \setminus O_{c\lambda_0} \) is compact in \( X \). \( \forall \lambda \in \Lambda_2 \), let \( \bar{O}_{c\lambda} := O_{c\lambda} \setminus \{ \omega \} \), then \( O_{c\lambda} = X \setminus (X_c \setminus O_{c\lambda}) \).

By Proposition 5.5 and the compactness of \( X_c \setminus O_{c\lambda} \), \( \bar{O}_{c\lambda} \in \mathcal{O} \). Note that

\[
X_c = \left( \bigcup_{\lambda \in \Lambda_1} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_2} O_{c\lambda} \right) \cup O_{c\lambda_0} = O_{c\lambda_0} \cup (X_c \setminus O_{c\lambda_0})
\]

Then, \( X_c \setminus O_{c\lambda_0} \subseteq \left( \bigcup_{\lambda \in \Lambda_1} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_2} \bar{O}_{c\lambda} \right) \). By the compactness of \( X_c \setminus O_{c\lambda_0} \), there exist finite sets \( \Lambda_{N1} \subseteq \Lambda_1 \) and \( \Lambda_{N2} \subseteq \Lambda_2 \) such that \( X_c \setminus O_{c\lambda_0} \subseteq \left( \bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_{N2}} \bar{O}_{c\lambda} \right) \). Then, \( X_c = \left( \bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_{N2}} \bar{O}_{c\lambda} \right) \). Hence, \( X_c \) is compact.

Next, we show that \( X_c \) is Hausdorff. \( \forall x_1, x_2 \in X_c \) with \( x_1 \neq x_2 \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_1, x_2 \in X_c \); Case 2: \( x_1 = \omega \) or \( x_2 = \omega \). Case 1: \( x_1, x_2 \in X_c \). Since \( X \) is Hausdorff, then \( \exists O_1, O_2 \in \mathcal{O} \) such that \( x_1 \in O_1 \), \( x_2 \in O_2 \), and \( O_1 \cap O_2 = \emptyset \). Clearly, \( O_1, O_2 \in \mathcal{O}_c \). Case 2: \( x_1 = \omega \) or \( x_2 = \omega \). Without loss of generality, assume \( x_2 = \omega \). Then, \( x_1 \in X \). By local compactness of \( X \), \( \exists O \in \mathcal{O} \) such that \( x_1 \in O \subseteq \bar{O} \subseteq X \) and \( \bar{O} \) is compact. Then, \( O_1 \in \mathcal{O}_c \) and \( O_2 := X_c \setminus \bar{O} \in \mathcal{O}_c \). Clearly, \( x_2 \in O_2 \) and \( O_1 \cap O_2 = \emptyset \). Hence, in both cases, we have obtained \( O_1, O_2 \in \mathcal{O}_c \) such that \( x_1 \in O_1 \), \( x_2 \in O_2 \), and \( O_1 \cap O_2 = \emptyset \). Hence, \( X_c \) is Hausdorff.

Finally, we show that \( \text{id} : X \rightarrow X_c \setminus \{ \omega \} \) is a homeomorphism. Clearly, \( \text{id} \) is bijective. \( \forall O_c \in \mathcal{O}_c \), we either have \( O_c \in \mathcal{O} \), which implies that \( O_c \cap (X_c \setminus \{ \omega \}) = O_c \in \mathcal{O}_c \), or we have \( X_c \setminus O_c \) is compact, which implies that \( \bar{O}_c := O_c \setminus \{ \omega \} = X \setminus (X_c \setminus O_c) \in \mathcal{O} \), by Proposition 5.5, and hence \( O_c \cap (X_c \setminus \{ \omega \}) = \bar{O}_c \in \mathcal{O}_c \). Hence, the subset topology \( \mathcal{O}_c \) on \( X_c \setminus \{ \omega \} \) with respect to \( X_c \) is contained in \( \mathcal{O} \). It is easy to see that \( \mathcal{O} \subseteq \mathcal{O}_c \). Then, \( \mathcal{O} = \mathcal{O}_c \). Hence, \( \text{id} : X \rightarrow X_c \setminus \{ \omega \} \) is a homeomorphism.

This completes the proof of the theorem. \( \square \)

5.7.4 Proper functions

**Definition 5.67** Let \( X \) and \( Y \) be topological spaces and \( f : X \rightarrow Y \) be continuous. \( f \) is said to be proper if \( \forall \) compact set \( K \subseteq Y \), we have \( f^{-1}(K) \subseteq X \) is compact. \( f \) is said to be countably proper if \( \forall \) compact set \( K \subseteq Y \), we have \( f^{-1}(K) \subseteq X \) is countably compact.

**Proposition 5.68** Let \( X := (X, \mathcal{O}_X) \) and \( Y := (Y, \mathcal{O}_Y) \) be locally compact Hausdorff topological spaces and \( f : X \rightarrow Y \) be continuous. Let \( X_c := (X_c, \mathcal{O}_{X_c}) \) and \( Y_c := (Y_c, \mathcal{O}_{Y_c}) \) be the Alexandroff one-point compactifications of \( X \) and \( Y \), respectively, where \( X_c = X \cup \{ \omega_x \} \) and \( Y_c = Y \cup \{ \omega_y \} \).

Define a function \( f_c : X_c \rightarrow Y_c \) by \( f_c(x) = f(x), \forall x \in X \), and \( f_c(\omega_x) = \omega_y \).

Then, \( f \) is proper if, and only if, \( f_c \) is continuous.

**Proof** “Only if” Let \( f \) be proper. \( \forall \mathcal{O}_{Y_c} \in \mathcal{O}_{Y_c} \), we will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{O}_{Y_c} \in \mathcal{O}_{Y_c} \); Case
2: \( Y_c \cap O_Y \) is compact in \( Y \). Case 1: \( O_Y \in O_Y \). Then, \( f_{\text{inv}}(O_Y \cup \{x\}) = f_{\text{inv}}(O_Y \cup \{x\}) \in O_X \). This implies that \( f_{\text{inv}}(Y_c \setminus K) \in O_X \). Hence, in both cases, we have \( f_{\text{inv}}(Y_c \setminus K) \in O_X \). Hence, \( f_{\text{inv}}(Y_c \setminus K) \) is compact by Proposition 5.5.

“Proof” Let \( f_{\text{inv}}(Y_c \setminus K) \) be continuous. Fix a compact set \( K \subseteq Y \). Then, \( \omega = \omega - \omega \in Y_c \setminus K \). By the continuity of \( f_{\text{inv}}(Y_c \setminus K) \), we have \( \omega = \omega \). This implies that \( K \subseteq K \). By Proposition 2.5, \( f^{-1}(Y_c \setminus K) \) is compact in \( X \). By Proposition 5.4, the net \( \left( f^{-1}(Y_c \setminus K) \right) \) is closed. This completes the proof of the proposition.

**Proposition 5.69** Let \( X \) be a topological space, \( Y \) be a metric space, \( F \subseteq X \) be closed, and \( f : X \to Y \) be continuous and proper. Then, \( f(F) \subseteq Y \) is closed.

**Proof** \( \forall y \in f(F) \), by Proposition 4.13, \( \exists (y_n)_{n=1}^{\infty} \subseteq f(F) \) such that \( \lim_{n \to \infty} y_n = y \). Then, \( \forall n \in \mathbb{N}, \exists x_n \in F \) such that \( f(x_n) = y_n \). Let \( K = \{y\} \cup \{y_n\}_{n=1}^{\infty} \). Then, \( K \) is compact. By the properness of \( f \), \( f_{\text{inv}}(K) \) is compact. By Propositions 3.5 and 5.5, \( F \cap f_{\text{inv}}(K) \) is compact. Note that \( (x_n)_{n=1}^{\infty} \subseteq F \cap f_{\text{inv}}(K) \). Then, by Proposition 5.4, \( (x_n)_{n=1}^{\infty} \) admits a cluster point \( x \in F \cap f_{\text{inv}}(K) \). By the continuity of \( f \) and Propositions 3.66, we have the sequence \( (f(x_n))_{n=1}^{\infty} = (y_n)_{n=1}^{\infty} \) admits a cluster point \( f(x) \). Since \( Y \) is Hausdorff and \( \lim_{n \to \infty} y_n = y \), then \( y = f(x) \in f(F) \). Therefore, \( f(F) \subseteq f(F) \) and \( f(F) \) is closed. This completes the proof of the proposition.

In the above proposition, the assumption on \( Y \) may be relaxed to first countable Hausdorff space.

**Proposition 5.70** Let \( X \) be a topological space, \( Y \) be a locally compact Hausdorff topological space, \( F \subseteq X \) be closed, and \( f : X \to Y \) be continuous and proper. Then, \( f(F) \subseteq Y \) is closed.

**Proof** \( \forall y \in f(F) \), by the local compactness of \( Y \), \( \exists O \in O_Y \) such that \( y \in \overline{f(F) \cap O} \). \( \forall U \in O_Y \) with \( y \in U \), by Proposition 3.3, we have \( \emptyset \neq (U \cap f(F)) = U \cap f(F) \). Hence, \( y \in f(F) \). By Proposition 3.68, there exists a net \( (y_b)_{b \in A} \subseteq f(F) \) such that \( \lim_{b \in A} y_b = y \). \( \forall \alpha \in A, \exists x_\alpha \in F \) such that \( y_\alpha = f(x_\alpha) \). Then, \( x_\alpha \in O \cap f_{\text{inv}}(O) \). Then, the net \( (x_\alpha)_{\alpha \in A} \subseteq f(F) \). By the properness of \( f \), we have \( f_{\text{inv}}(O) \subseteq X \). By Proposition 3.5, \( F \cap f_{\text{inv}}(O) \) is closed relative to \( f_{\text{inv}}(O) \), which further implies that \( F \cap f_{\text{inv}}(O) \) is compact by Proposition 5.5. By Proposition 5.4, the net \( (x_\alpha)_{\alpha \in A} \) admits a cluster point \( x \in F \cap f_{\text{inv}}(O) \). By the continuity of \( f \) and Proposition 3.66, \( f(x) \) is a cluster point of the net \( (f(x_\alpha))_{\alpha \in A} = (y_\alpha)_{\alpha \in A} \). Since \( Y \) is Hausdorff and \( \lim_{\alpha \in A} y_\alpha = y \), then \( y = f(x) \in f(F) \). Hence, \( f(F) \subseteq f(F) \) and \( f(F) \) is closed. This completes the proof of the proposition.
5.8 σ-Compact Spaces

**Definition 5.71** A topological space is said to be σ-compact if it is the union of countably infinitely many compact sets.

**Proposition 5.72** Let \( \mathcal{X} \) be a locally compact topological space. Then, the following statements are equivalent.

(i) \( \mathcal{X} \) is Lindelöf.

(ii) \( \mathcal{X} \) is σ-compact.

(iii) \( \exists (O_n)_{n=1}^{\infty} \subseteq O \) such that \( \forall n \in \mathbb{N}, \overline{O_n} \subseteq O_{n+1} \) is compact and \( \mathcal{X} = \bigcup_{n=1}^{\infty} O_n \). This sequence is called an exhaustion of \( \mathcal{X} \).

Furthermore, if \( \mathcal{X} \) is Hausdorff, then the above is equivalent to

(iv) \( \exists \phi : \mathcal{X} \to [0, \infty) \subseteq \mathbb{R} \) that is proper and continuous.

**Proof**

(i) \( \Rightarrow \) (ii). ∀\( x \in \mathcal{X} \), by local compactness of \( \mathcal{X} \), \( \exists O_x \subseteq O \) such that \( x \in O_x \subseteq \overline{O_x} \) and \( \overline{O_x} \) is compact. \( \mathcal{X} = \bigcup_{x \in \mathcal{X}} O_x \). Since \( \mathcal{X} \) is Lindelöf, then \( \exists \) a countable set \( X_C \subseteq \mathcal{X} \) such that \( \mathcal{X} = \bigcup_{x \in X_C} O_x \).

We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( X_C = \emptyset \); Case 2: \( X_C \neq \emptyset \) is finite; Case 3: \( X_C \) is countably infinite.

Case 1: \( X_C = \emptyset \). Let \( K_n = \emptyset, \forall n \in \mathbb{N} \), which are clearly compact. Then, \( \mathcal{X} = \emptyset = \bigcup_{n=1}^{\infty} K_n \). Hence, \( \mathcal{X} \) is σ-compact. Case 2: \( X_C \neq \emptyset \) is finite. Without loss of generality, assume that \( X_C = \{x_1, \ldots, x_n\} \) for some \( n \in \mathbb{N} \). Let \( K_i = \overline{O_{x_i}}, i = 1, \ldots, n \), and \( K_i = \emptyset, i = n + 1, n + 2, \ldots \). Clearly, \( K_i \)'s are compact and \( \mathcal{X} = \bigcup_{i=1}^{\infty} K_i \). Hence, \( \mathcal{X} \) is σ-compact.

Case 3: \( X_C \) is countably infinite. Without loss of generality, assume that \( X_C = \{x_1, x_2, \ldots\} \). Let \( K_i = \overline{O_{x_i}}, \forall i \in \mathbb{N} \). Clearly, \( K_i \)'s are compact and \( \mathcal{X} = \bigcup_{i=1}^{\infty} K_i \). Hence, \( \mathcal{X} \) is σ-compact. In all cases, \( \mathcal{X} \) is σ-compact.

(ii) \( \Rightarrow \) (iii). Let \( \mathcal{X} = \bigcup_{n=1}^{\infty} K_n \), where \( K_n \) is compact, \( \forall n \in \mathbb{N} \). Without loss of generality, we assume that \( K_n \subseteq K_{n+1}, \forall n \in \mathbb{N} \), since otherwise, we may let \( K_n = \bigcup_{i=1}^{n} K_n \) and consider \( \mathcal{X} = \bigcup_{n=1}^{\infty} K_n \) instead. Let \( O_0 := \emptyset \in O \). Then, \( \overline{O_0} = \emptyset \) is compact. \( \forall n \in \mathbb{N}, K_n \subseteq \overline{O_n} \subseteq \overline{O_{n-1}} \) is compact, by Proposition 5.52, \( \exists O_n \subseteq O \) such that \( K_n \subseteq \overline{O_{n-1}} \subseteq O_n \subseteq \overline{O_n} \), and \( \overline{O_n} \) is compact. Then, \( (O_n)_{n=1}^{\infty} \) is an exhaustion of \( \mathcal{X} \) that we seek.

(iii) \( \Rightarrow \) (iv). Let \( (U_n)_{n=1}^{\infty} \) be an exhaustion of \( \mathcal{X} \). Fix any open covering \( (O_\alpha)_{\alpha \in \Lambda} \subseteq O \) of \( \mathcal{X} \), where \( \Lambda \) is an index set. \( \forall n \in \mathbb{N}, U_n \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha \) by the compactness of \( \overline{U_n} \), there exists a finite set \( \Lambda_N \subseteq \Lambda \) such that \( \overline{U_n} \subseteq \bigcup_{\alpha \in \Lambda_N} O_\alpha \). Then, \( \mathcal{X} = \bigcup_{n=1}^{\infty} U_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{\alpha \in \Lambda_N} O_\alpha \), which is a countable subcovering. Hence, \( \mathcal{X} \) is Lindelöf.

Now, assume that \( \mathcal{X} \) is Hausdorff.

(iii) \( \Rightarrow \) (iv). Let \( (O_n)_{n=1}^{\infty} \) be an exhaustion of \( \mathcal{X} \). \( \forall n \in \mathbb{N}, \overline{O_n} \subseteq O_{n+1} \) is compact. Then, by Proposition 5.52, there exists a continuous function \( \phi_n : \mathcal{X} \to [0, 1] \) such that \( \phi_n|_{\overline{O_n}} = 1 \) and \( \phi_n|_{\overline{O_{n+1}}} = 0 \). Let \( \phi := \sum_{n=1}^{\infty} (1 - \phi_n) \).
Clearly, \( \phi : \mathcal{X} \to [0, \infty) \subset \mathbb{R} \). \( \forall x_0 \in \mathcal{X}, \exists n_0 \in \mathbb{N} \) such that \( x_0 \in O_{n_0} \). Then, \( \phi|_{O_{n_0}} = \sum_{i=1}^{n_0-1} (1 - \phi_i) |_{O_{n_0}} \), which is clearly continuous. \( \forall U \in \mathcal{O}_\mathbb{R} \) with \( \phi(x_0) \in U \), \( \exists V \in \mathcal{O} \) with \( x_0 \in V \subseteq O_{n_0} \) such that \( \phi|_{O_{n_0}}(V) \subseteq U \). Then, \( \phi(V) \subseteq U \). Hence, \( \phi \) is continuous at \( x_0 \). By the arbitrariness of \( x_0 \) and Proposition 3.9, \( \phi \) is continuous. Fix any compact subset \( K \subseteq \mathbb{R} \). Then, \( K \) is closed and bounded, by Heine-Borel Theorem. \( \exists N \in \mathbb{N} \) such that \( K \subseteq [-N, N] \subseteq \mathbb{R} \). \( \forall x \in O_n \), \( n = 1, \ldots, N + 2 \), and hence \( \phi_n(x) = 0 \), \( n = 1, \ldots, N + 1 \), and \( \phi(x) \geq N + 1 > N \). Then, \( \phi_{inv}(K) \subseteq O_{N+2} \subseteq \overline{O_{N+2}} \). By Proposition 3.10, \( \phi_{inv}(K) \) is closed in \( \mathbb{R} \). By the compactness of \( O_{N+2} \) and Proposition 5.5, \( \phi_{inv}(K) \) is compact. Hence, \( \phi \) is proper.

(iv) \( \Rightarrow \) (ii). Let \( K_n = [-n, n] \subseteq \mathbb{R} \), \( \forall n \in \mathbb{N} \). Then, \( K_n \)'s are compact in \( \mathbb{R} \) and \( \mathbb{R} \subseteq \bigcup_{n=1}^{\infty} K_n \). By the properness of \( \phi \), \( \phi_{inv}(K_n) \) is compact in \( \mathcal{X} \), \( \forall n \in \mathbb{N} \). By Proposition 2.5, \( \mathcal{X} = \bigcup_{n=1}^{\infty} \phi_{inv}(K_n) \). Hence, \( \mathcal{X} \) is \( \sigma \)-compact.

This completes the proof of the proposition.

5.9 Paracompact Spaces

**Definition 5.73** Let \( \mathcal{X} \) be a topological space and \( \mathcal{A} \) be a collection of subsets in \( \mathcal{X} \). \( \mathcal{A} \) is said to be locally finite if \( \forall x \in \mathcal{X}, \exists U \in \mathcal{O} \) with \( x \in U \) such that \( U \) meets only finitely many members of \( \mathcal{A} \), that is \( U \cap A \neq \emptyset \) for finitely many \( A \in \mathcal{A} \).

**Proposition 5.74** Let \( \mathcal{X} \) be a topological space and \( (E_\lambda)_{\lambda \in \Lambda} \) be a locally finite collection of subsets of \( \mathcal{X} \), where \( \Lambda \) is an index set.

(i) Let \( E = \bigcup_{\lambda \in \Lambda} E_\lambda \). Then, \( \overline{E} = \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \).

(ii) Let \( K \subseteq \mathcal{X} \) be compact. Then, \( K \) meets only finitely many members of \( (E_\lambda)_{\lambda \in \Lambda} \).

**Proof**

(i). \( \forall x \in \overline{E} \), by local finiteness of \( (E_\lambda)_{\lambda \in \Lambda} \), \( \exists U \in \mathcal{O} \) with \( x \in U \), then \( U \) meets only finitely many members of \( (E_\lambda)_{\lambda \in \Lambda} \). Let \( \Lambda_N \subseteq \Lambda \) be the finite set such that \( U \cap E_\lambda \neq \emptyset \), \( \forall \lambda \in \Lambda_N \). Suppose \( x \notin \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \). Then, \( \forall \lambda \in \Lambda_N, x \notin \overline{E_\lambda} \). \( \exists U_\lambda \in \mathcal{O} \) with \( x \in U_\lambda \) such that \( U_\lambda \cap E_\lambda = \emptyset \), by Proposition 3.3. Let \( O := (\bigcap_{\lambda \in \Lambda_N} U_\lambda) \cap U \in \mathcal{O} \). Then, \( x \in O \) and \( O \cap E_\lambda = \emptyset \), \( \forall \lambda \in \Lambda_N \). Then, \( O \cap E = \emptyset \). This contradicts with \( x \in \overline{E} \), by Proposition 3.3. Hence, \( x \in \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \). Then, \( \overline{E} \subseteq \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \).

On the other hand, \( \forall x \in \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \), \( \exists \lambda_0 \in \Lambda \) such that \( x \in \overline{E_{\lambda_0}} \). \( \forall O \in \mathcal{O} \) with \( x \in O \), \( O \cap E_{\lambda_0} \neq \emptyset \). Then, \( O \cap \overline{E} \neq \emptyset \) and hence \( x \in \overline{E} \). Therefore, we have \( \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \subseteq \overline{E} \).

Hence, \( \overline{E} = \bigcup_{\lambda \in \Lambda} \overline{E_\lambda} \).

(ii). Let \( K \subseteq \mathcal{X} \) be compact. \( \forall x \in K \), by the local finiteness of \( (E_\lambda)_{\lambda \in \Lambda} \), \( \exists U_x \in \mathcal{O} \) with \( x \in U_x \) such that \( U_x \) meets only finitely many members of \( (E_\lambda)_{\lambda \in \Lambda} \), for finitely many \( \lambda \in \Lambda \) such that \( x \in \overline{E_\lambda} \). Then,
(E_\lambda)_{\lambda \in \Lambda}. Then, K \subseteq \bigcup_{x \in K} U_x. By the compactness of K, there exists a finite set K_N \subseteq K such that K \subseteq \bigcup_{x \in K_N} U_x. Clearly, \bigcup_{x \in K_N} U_x meets only finitely many members in (E_\lambda)_{\lambda \in \Lambda}. Hence, K meets only finitely many members in (E_\lambda)_{\lambda \in \Lambda}.

This completes the proof of the proposition.

\[ \square \]

**Definition 5.75** A topological space \( X \) is said to be paracompact if every open covering of \( X \) has a locally finite open refinement.

**Proposition 5.76** A closed subset of a paracompact space is paracompact in the subset topology.

**Proof** Let \( X \) be a paracompact space, \( F \subseteq X \) be closed, and \( \mathcal{O}_F \) be the subset topology on \( F \). Let \((O_{\lambda})_{\lambda \in \Lambda} \subseteq \mathcal{O}_F \) be any open covering of \( F \), where \( \Lambda \) is an index set (open in the subset topology \( \mathcal{O}_F \)). By Proposition 3.4, \( \forall \alpha \in \Lambda, \exists \beta_{\alpha} \in \mathcal{O} \) such that \( O_{\alpha} = O_{\alpha} \cap F \). Then, \( F \subseteq \bigcup_{\alpha \in \Lambda} O_{\alpha} \) and \( X \subseteq \bigcup_{\alpha \in \Lambda} O_{\alpha} \cup \bar{F} \). By the paracompactness of \( X \), there exists a locally finite open refinement \( \mathcal{V} \) of \((O_{\alpha})_{\alpha \in \Lambda} \cup \{\bar{F}\} \). Let \( \mathcal{V} := \{V \in \mathcal{V} \mid V \subseteq O_{\alpha}, \text{ for some } \alpha \in \Lambda\} \). Then, \( F \subseteq \bigcup_{V \in \mathcal{V}} V \), since other \( V \)'s in \( \mathcal{V} \) are subset of \( \bar{F} \). Clearly, \( \mathcal{V} \) is locally finite. Then, \( \{V \cap F \mid V \in \mathcal{V}\} \subseteq \mathcal{O}_F \) is a locally finite open refinement of \((O_{\lambda})_{\lambda \in \Lambda} \) (open in the subset topology \( \mathcal{O}_F \)). Hence, \((F, \mathcal{O}_F)\) is paracompact. This completes the proof of the proposition.

\[ \square \]

**Theorem 5.77** Metric spaces are paracompact.

**Proof** Let \( X \) be a metric space. Let \((O_{\alpha})_{\alpha \in \Lambda}\) be any open covering of \( X \), where \( \Lambda \) is an index set. By Well-Ordering Principle, \( \Lambda \) may be well-ordered by \( \leq \). We will construct an open refinement of \((O_{\alpha})_{\alpha \in \Lambda}\) in the following steps. Let \( X_0 = X \).

Step \( n \) (\( n \in \mathbb{N}\)): \( \forall \alpha \in \Lambda, \text{ let } P_{\alpha n} := \{x \in O_{\alpha} \cap X_{n-1} \mid \alpha \text{ is the least element of } \{ \beta \in \Lambda \mid x \in O_{\beta}\} \}. \text{ Let } Q_{\alpha n} := \{x \in P_{\alpha n} \mid B(x, 1/2^n) \subseteq O_{\alpha}\} \text{ and } D_{\alpha n} := \bigcup_{x \in Q_{\alpha n}} B(x, 1/2^n) \in \mathcal{O}. \text{ Let } X_n := X_{n-1} \setminus (\bigcup_{\alpha \in \Lambda} D_{\alpha n}). \)

We will now show that \( \mathcal{V} := \{D_{\alpha n} \mid \alpha \in \Lambda, n \in \mathbb{N}\} \) is a locally finite open refinement of \((O_{\alpha})_{\alpha \in \Lambda}\). \( \forall x_0 \in X \). Then, \( \{\beta \in \Lambda \mid x_0 \in O_{\beta}\} \neq \emptyset \).

Let \( \alpha_0 \) be the least element of the set, which exists since \( \Lambda \) is well-ordered. Then, \( x_0 \in O_{\alpha_0} \). Then, \( \exists n_0 \in \mathbb{N} \) such that \( B(x_0, 3/2^{n_0}) \subseteq O_{\alpha_0} \), since \( O_{\alpha_0} \) is open. Then, \( x_0 \in D_{\alpha_0 n_0} \) or \( x_0 \in D_{\beta n} \) for some \( \beta \in \Lambda \) and for some \( n \in \mathbb{N} \) with \( n < n_0 \). Hence, \( \mathcal{V} \) covers \( X \). Clearly, \( D_{\alpha n} \subseteq O_{\alpha} \) and is open by construction, \( \forall \alpha \in \Lambda \) and \( \forall n \in \mathbb{N} \). Then, \( \mathcal{V} \) is an open refinement of \((O_{\alpha})_{\alpha \in \Lambda}\).

Fix any \( x_0 \in X \). Consider the set \( \{\beta \in \Lambda \mid \exists n \in \mathbb{N}, x_0 \in D_{\beta n}\} \neq \emptyset \).

Let \( \alpha_1 \) be the least element of this set. Then, \( x_0 \in D_{\alpha_1 n_0} \) for some \( n_0 \in \mathbb{N} \). Furthermore, \( \exists j_0 \in \mathbb{N} \) such that \( B(x_0, 2^{-j_0}) \subseteq D_{\alpha_1 n_0} \). Consider the open set \( B(x_0, 2^{-j_0 - n_0}) \ni x_0 \).
∀n ≥ n₀ + j₀, ∀α ∈ Λ, since B (x₀, 2⁻j₀) ⊆ D₁αn₀ ⊆ O₁α, then ∀x ∈ B (x₀, 2⁻j₀−n₀), we have B (x, 2⁻j₀−2⁻j₀−n₀) ⊆ B (x₀, 2⁻j₀) ⊆ D₁αn₀, ∀y ∈ Dₐn, since n > n₀, then 3x ∈ Qₐn with 3x /∈ D₁αn₀ such that y ∈ B (3x, 2⁻n). Then, ∃x ∈ B (y, 2⁻n) such that 3x ∈ Dₐn \ D₁αn₀. Note that 2⁻n ≤ 2⁻j₀−n₀ ≤ 2⁻j₀−2⁻j₀−n₀. Therefore, x /∈ Dₐn. Hence, B (x₀, 2⁻j₀−n₀) ∩ Dₐn = ∅. Thus, B (x₀, 2⁻j₀−n₀) does not intersect any Dₐn, ∀α ∈ Λ, ∀n ≥ n₀ + j₀.

Claim 5.77.1 ∀n < n₀ + j₀, B (x₀, 2⁻j₀−n₀) intersects at most one of the set Dₐn’s, α ∈ Λ.

Proof of claim: Suppose the result is not true. Then, ∃α, β ∈ Λ such that B (x₀, 2⁻j₀−n₀) ∩ Dₐn ≠ ∅, B (x₀, 2⁻j₀−n₀) ∩ Dₐβn ≠ ∅, and α ≠ β.

Without loss of generality, assume that α ≤ β. Let p ∈ B (x₀, 2⁻j₀−n₀) ∩ Dₐn and q ∈ B (x₀, 2⁻j₀−n₀) ∩ Dₐβn. By the definition of Dₐn, ∃p ∈ Qₐn ⊆ Dₐn such that p ∈ B (p, 2⁻n) ⊆ Dₐn. Similarly, ∃q ∈ Qₐβn ⊆ Dₐβn such that q ∈ B (q, 2⁻n) ⊆ Dₐβn. Then, by the definition of Qₐn, we have B (p, 3/2ⁿ) ⊆ O₂α. Similarly, B (q, 3/2ⁿ) ⊆ O₂β. Note that

\[ ρ (\bar{p}, \bar{q}) ≤ ρ (\bar{p}, p) + ρ (p, x₀) + ρ (x₀, q) + ρ (q, \bar{q}) ≤ 2⁻n + 2⁻j₀−n₀ + 2⁻j₀−n₀ + 2⁻n ≤ 3/2ⁿ \]

Then, q ∈ O₂α. But, q ∈ O₂β ⊆ P₂β, α ≤ β, and α ≠ β implies that q /∈ O₂α. This is a contradiction. Therefore, the result must hold. This completes the proof of the claim.

Hence, B (x₀, 2⁻j₀−n₀) can meet at most n₀ + j₀ − 1 sets in V. Hence, V is locally finite.

Therefore, we have obtained a locally finite open refinement of (O₂α)α ∈ Λ. Hence, X is paracompact. This completes the proof of the theorem.

Definition 5.78 Let X be a topological space and (Eₐ)α ∈ Λ be a collection of subsets of X, where Λ is an index set. (Eₐ)α ∈ Λ is said to be star-finite if ∀α₀ ∈ Λ, Eₐ₀ meets only finitely many members in the collection.

It is easy to see that a star-finite open covering of X is locally finite. But the converse need not hold.

Proposition 5.79 A σ-compact locally compact space is paracompact.

Proof Let X be a σ-compact locally compact space. Fix any open covering (Oₐ)α ∈ Λ of X, where Λ is an index set. By Proposition 5.72, there exists an exhaustion (Uₙ)∞ₙ=1 of X. Let U₋₁ = ∅ and U₀ = ∅. Let Vₐn := Oₐ ∩ (Uₙ₊₁ \ Uₙ₋₂) ∈ O, ∀α ∈ Λ, ∀n ∈ N. It is easy to see that V := {Vₐn | α ∈ Λ, n ∈ N} is an open refinement of (Oₐ)α ∈ Λ that covers X. ∀n ∈ N, by the compactness of Uₙ and Proposition 5.5, Uₙ \ Uₙ₋₁ ⊆ ∪ₐ ∈ Λ Vₐn is compact. Then, there exists a finite set Vₙ ⊆
5.9. PARACOMPACT SPACES

\{ \alpha \subseteq \Lambda \} \text{ such that } \overline{U_n} \setminus U_{n-1} \subseteq \bigcup_{\alpha \subseteq \Lambda} V_{\alpha}. \text{ Then, } \mathcal{V}_L := \bigcup_{n=1}^{\infty} V_n \text{ is an open refinement of } \{ \alpha \subseteq \Lambda \} \text{ that covers } \mathcal{X} \text{ since } \bigcup_{n=1}^{\infty} (U_n \setminus U_{n-1}) = \mathcal{X}.

Hence, \exists \text{ regular.}

Proof

"If" Let \mathcal{U} be any open covering of \mathcal{X}. \forall x \in \mathcal{X}, \exists U_x \in \mathcal{U} such that \bar{x} \subseteq U_x. By Proposition 5.53, \exists O_x \in \mathcal{O} \text{ such that } x \in O_x, O_x \subseteq U_x \text{ and } O_x \text{ is compact. Hence, } (O_x)_{x \in \mathcal{X}} \text{ is an open refinement of } \mathcal{U}. \text{ By the paracompactness of } \mathcal{X}, \text{ there exists a locally finite open refinement } \mathcal{V} \text{ of } (O_x)_{x \in \mathcal{X}} \text{ that covers } \mathcal{X}. \forall V \in \mathcal{V}, \exists x \in \mathcal{X} \text{ such that } V \subseteq O_x. \text{ Then, } \nabla \subseteq \bigcup_{x \in \mathcal{X}} O_x. \text{ By Propositions 5.5 and 3.5, we have } \nabla \text{ is compact. By Proposition 5.74, } \nabla \text{ meets only finitely many members of } \mathcal{V}. \text{ Hence, } \mathcal{V} \text{ is star-finite. Therefore, } \mathcal{V} \text{ is a star-finite open refinement of } \mathcal{U}.

"Only if" Let \mathcal{U} be any open covering of \mathcal{X}. Then, there is a star-finite open refinement \mathcal{V} \subseteq \mathcal{U}. \text{ Then, } \mathcal{V} \text{ is a locally finite open refinement of } \mathcal{U}. \text{ Hence, } \mathcal{X} \text{ is paracompact.}

This completes the proof of the proposition. \qed

Proposition 5.81

A paracompact Hausdorff topological space is normal.

Proof

Let \mathcal{X} be a paracompact Hausdorff topological space. Clearly, \mathcal{X} is Tychonoff. \forall x_0 \in \mathcal{X} \text{ and } \forall \text{ closed set } F \subseteq \mathcal{X} \text{ with } x_0 \notin F, \text{ we will show that } \exists O_1, O_2 \in \mathcal{O} \text{ such that } x_0 \in O_1, F \subseteq O_2, \text{ and } O_1 \cap O_2 = \emptyset. \forall x \in F, \text{ then } x \neq x_0. \text{ Since } \mathcal{X} \text{ is Hausdorff, } \exists O_1^{(1)}, O_2^{(2)} \in \mathcal{O} \text{ such that } x_0 \in O_1^{(1)}, x \in O_2^{(2)}, \text{ and } O_1^{(1)} \cap O_2^{(2)} = \emptyset. \text{ Then, } \mathcal{X} \subseteq \overline{F} \cup \bigcup_{x \in F} O_2^{(2)}. \text{ By the paracompactness of } \mathcal{X}, \text{ there exists a locally finite open refinement } \mathcal{V} \subseteq \mathcal{O} \text{ of } \{ \overline{F} \} \cup \Big\{ \bigcup_{x \in F} O_2^{(2)} \Big\}. \text{ Let } \mathcal{V} := \left\{ V \in \mathcal{V} \mid V \subseteq O_2^{(2)}, \text{ for some } x \in F \right\}. \text{ Then, } \mathcal{V} \text{ is an open covering of } F, \text{ since } \forall V \in \mathcal{V} \setminus \mathcal{V} \text{ we have } V \subseteq \overline{F}. \text{ Furthermore, } \mathcal{V} \text{ is locally finite. } \forall V \in \mathcal{V}, \exists x \in F \text{ such that } V \subseteq O_2^{(2)}, x_0 \in O_1^{(1)}, \text{ and } O_1^{(1)} \cap O_2^{(2)} = \emptyset. \text{ Then, } V \subseteq O_2^{(2)} \subseteq O_1^{(1)}. \text{ Hence, we have } \mathcal{V} \subseteq O_1^{(1)} \text{ and } x_0 \notin \mathcal{V}. \text{ By Proposition 5.74, } \bigcup_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V}} \overline{V} = : \overline{O_1}. \text{ Then, } O_1, O_2 \in \mathcal{O}, F \subseteq \bigcup_{V \in \mathcal{V}} V = : O_2 \subseteq O_1, \text{ and } x_0 \in O_1. \text{ Hence, } \exists O_1, O_2 \in \mathcal{O} \text{ such that } x_0 \in O_1, F \subseteq O_2, \text{ and } O_1 \cap O_2 = \emptyset. \text{ Then, } \mathcal{X} \text{ is regular.}

Next, we show that \mathcal{X} is normal. Fix any closed sets \mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{X} \text{ with } \mathcal{F}_1 \cap \mathcal{F}_2 = \emptyset. \forall x \in \mathcal{F}_2, \text{ then } x \notin \mathcal{F}_1. \text{ Since } \mathcal{X} \text{ is regular, } \exists O_1^{(1)}, O_2^{(2)} \in \mathcal{O} \text{ such that } \mathcal{F}_1 \subseteq O_1^{(1)}, x \in O_2^{(2)}, \text{ and } O_1^{(1)} \cap O_2^{(2)} = \emptyset. \text{ Then, } \mathcal{X} \subseteq \overline{\mathcal{F}_2} \cup
By the paracompactness of $X$, there exists a locally finite open refinement $V \subseteq \mathcal{O}$ of $\{F_2\} \cup \left\{O_x^{(2)} \mid x \in F_2\right\}$. Let $\mathcal{V} := \left\{V \in \mathcal{V} \mid V \subseteq O_x^{(2)} \text{, for some } x \in F_2\right\}$. Then, $\mathcal{V}$ is an open covering of $F_2$, since $\forall V \in \mathcal{V} \setminus \mathcal{V}$ we have $V \subseteq \tilde{F}_2$. Furthermore, $\mathcal{V}$ is locally finite. $\forall V \in \mathcal{V}$, $\exists x \in F_2$ such that $V \subseteq O_x^{(2)}$, $F_1 \subseteq O_x^{(1)}$, and $O_x^{(2)} \cap O_x^{(1)} = \emptyset$. Then, $V \subseteq O_x^{(2)} \subseteq O_x^{(1)}$. Hence, we have $V \subseteq O_x^{(1)}$ and $F_1 \cap V = \emptyset$. By Proposition 5.74, $\bigcup_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V} \setminus \mathcal{V}} V =: \bar{O}_1$. Then, $O_1 \in \mathcal{O}$, $F_2 \subseteq \bigcup_{V \in \mathcal{V}} V =: O_2 \in \mathcal{O}$, $O_2 \subseteq \bar{O}_1$, and $F_1 \cap \bar{O}_1 = \emptyset$. Hence, $\exists O_1, O_2 \in \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Then, $X$ is normal.

This completes the proof of the proposition. \hfill \Box

### 5.10 The Stone-Čech Compactification

**Definition 5.82** Let $X := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, and $\mathcal{F}$ be the family of continuous functions of $X$ to $I$. By Corollary 3.62, $X$ is homeomorphic to $E(X) \subseteq I^\mathcal{F}$, where $E$ is the equivalence map. Let $F = \bar{E}(X)$ and $\mathcal{O}_F$ be the subset topology of $F$. Then, $\beta(X) := (F, \mathcal{O}_F)$ is a compact Hausdorff topological space, by Tychonoff Theorem and Proposition 5.5. $\beta(X)$ is said to be the Stone-Čech compactification of $X$.

**Proposition 5.83** Let $X := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, $\mathcal{F}$ be the family of continuous functions of $X$ to $I$, and $E : X \to E(X) \subseteq I^\mathcal{F}$ be the equivalence map. Then, there exists a unique compact Hausdorff topological space $\beta(X)$ with the following properties:

(i) the space $E(X)$ is dense in $\beta(X)$;

(ii) each bounded continuous real-valued function on $E(X)$ extends to a bounded continuous real-valued function on $\beta(X)$;

(iii) if $X$ is a dense subset of a compact Hausdorff topological space $Y$, then, there exists a unique continuous mapping $\phi : \beta(X) \to Y$ such that $\phi$ is surjective and $\phi|_{E(X)} = E\circ inv$.

Furthermore, if $X$ is locally compact, then $E(X)$ is an open subset of $\beta(X)$.

**Proof** We will first show that $\beta(X)$ defined in Definition 5.82 satisfies properties (i), (ii), and (iii). (i). By Proposition 3.5, $E(X)$ is dense in $F = \bar{E}(X) = \beta(X)$. (ii). Fix any bounded continuous real-valued function $g : E(X) \to \mathbb{R}$. Then, $\exists N \in \mathbb{N}$ such that $g : E(X) \to [-N, N] \subset \mathbb{R}$. Then, $\bar{g} := ((g + N)/(2N)) \circ E \in \mathcal{F}$ by Proposition 3.12. Now, consider the function $h := \pi_{\bar{g}|_{\beta(X)}} : \beta(X) \to [0, 1] \subset \mathbb{R}$, which is continuous by
5.10. THE STONE-ČECH COMPACTIFICATION

Proposition 3.27. Note that $\tilde{h}|_{E(X)} = \bar{g} \circ E_{\text{inv}} = ((g + N)/(2N))$. Then, by Proposition 3.12, $h := 2N\tilde{h} - N$ is a continuous function of $\beta(X)$ to $[-N,N]$. Furthermore, $h|_{E(X)} = g$. Hence, $h$ is the desired extension that we seek. Clearly, $h : \beta(X) \to [-N,N] \subset \mathbb{R}$.

(iii). Let $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be a compact Hausdorff space such that $X$ is dense in $\mathcal{Y}$. Let $G$ be the family of continuous real-valued functions of $\mathcal{Y}$ to $\mathbb{R}$. Let $E^{(3)} : \mathcal{Y} \to I^G$ be the equivalence map and $\pi_g^{(3)} : I^G \to I$ be the projection function, $\forall g \in G$. By Propositions 3.14 and 3.61, $\mathcal{Y}$ is completely regular. By Corollary 3.62, $E^{(3)} : \mathcal{Y} \to E^{(3)}(\mathcal{Y}) \subseteq I^G$ is a homeomorphism. Define a mapping $\psi : G \to F$ by $\psi(g) = g|_{X}$, $\forall g \in G$. By Proposition 3.56, $\psi$ is injective. By Tychonoff Theorem, $I^G$ is a compact Hausdorff space. Define a mapping $\Psi : I^F \to I^G$ by $\pi^{(G)}_{\psi(i)}(\Psi(i)) = \pi^{(G)}_{\psi(g)}(i), \forall i \in I^F, \forall g \in G$. $\forall x \in X$, $E^{(3)}(x) \in I^G$ satisfies $\pi^{(G)}_{\psi(i)}(E^{(3)}(x)) = g(x) = g|_{X}(x) = \psi(g)(x) = \pi^{(G)}_{\psi(g)}(E(x))$, $\forall g \in G$. Then, we have $E^{(3)}(x) = \Psi(E(x)), \forall x \in X$. Hence, $\Psi \circ E = E^{(3)} \big|_{X}$.

Claim 5.83.1 $\Psi$ is continuous.

Proof of claim: Fix any basis open set $U \subseteq I^G$. Then, $U = \bigcap_{g \in G} U_g$, where $U_g \subseteq I$ is open, $\forall g \in G$, and $U_g = I$ for all $g$'s except finitely many $g$'s, say $g \in G \subseteq G$. Let $V := \bigcap_{f \in F} V_f \subseteq I^F$ be given by $V_f = I$, $\forall f \notin \text{range}(\psi)$, and $V_f = U_g, \forall f \in \text{range}(\psi)$ and $f = \psi(g)$. The set $V$ is well defined since $\psi$ is injective. $\forall i_f \in V, \forall g \in G$, $\pi^{(G)}_{\psi(g)}(\Psi(i_f)) = \pi^{(G)}_{\psi(g)}(i_f) \in V_{\psi(g)} = U_g$. Hence, we have $\Psi(i_f) \in U$. Then, $V \subseteq \Psi_{\text{inv}}(U)$. $\forall i_f \in I^F \setminus V$, $\exists ! f_0 \in \text{range}(\psi)$ such that $\pi^{(G)}_{\psi(g)}(i_f) \notin V_{\psi(g)}$. Then, $f_0 = \psi(g_0), \forall g_0 \in G$. Note that $\pi^{(G)}_{\psi(g_0)}(\Psi(i_f)) = \pi^{(G)}_{\psi(g_0)}(i_f) \notin V_{\psi(g_0)} = U_{g_0}$. Then, $\Psi(i_f) \notin U$. This shows that $I^F \setminus V \subseteq \Psi_{\text{inv}}(U) \setminus U$. Hence, we have $V = \Psi_{\text{inv}}(U)$, by Proposition 2.5. Clearly, $V$ is a basis open set in $I^F$. Hence, $\Psi$ is continuous. This completes the proof of the claim. □

Since $X$ is dense in $\mathcal{Y}$ and $E^{(3)}(\mathcal{Y})$ is a homeomorphism between $\mathcal{Y}$ and $E^{(3)}(\mathcal{Y})$, then $E^{(3)}(X)$ is dense in $E^{(3)}(\mathcal{Y})$. Since $\mathcal{Y}$ is compact and $E^{(3)}$ is continuous, then $E^{(3)}(\mathcal{Y})$ is compact in $I^G$, by Proposition 5.7. Furthermore, by Proposition 5.5, $E^{(3)}(\mathcal{Y})$ is closed in $I^G$. Then, $\overline{E^{(3)}(X)} = E^{(3)}(\mathcal{Y})$, where $\overline{E^{(3)}(X)}$ is the closure of $E^{(3)}(X)$ in $I^G$. By the compactness of $\beta(X)$, the continuity of $\Psi$, and Proposition 5.7, $\Psi(\beta(X)) \subseteq I^G$ is compact. Furthermore, by Proposition 5.5, $\Psi(\beta(X))$ is closed in $I^G$. Note that $E^{(3)}(X) = \Psi(E(X)) \subseteq \Psi(\beta(X))$. Then, $\Psi(\beta(X)) \supseteq E^{(3)}(\mathcal{Y})$.

Claim 5.83.2 $\Psi(\beta(X)) = E^{(3)}(\mathcal{Y})$.

Proof of claim: Suppose $E^{(3)}(\mathcal{Y}) \subset \Psi(\beta(X))$, then $\Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \subset \beta(X)$ and $\beta(X) \cap (I^F \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y}))) \neq \emptyset$. Note that $\Psi_{\text{inv}}(E^{(3)}(\mathcal{Y}))$ is closed in $I^F$ by the closedness of $E^{(3)}(\mathcal{Y})$, the continuity of $\Psi$, and Proposition 3.10, which further implies that $I^F \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \in \mathcal{O}_{I^F}$. By the denseness of $E(X)$ in $\beta(X)$, we have $E(X) \cap (I^F \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y}))) \neq \emptyset$ and
hence \( \exists x \in \mathcal{X} \) such that \( E(x) \in I^X \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y})) \). By Proposition 2.5, \( I^X \setminus \Psi_{\text{inv}}(E^{(\mathcal{Y})}(\mathcal{Y})) = \Psi_{\text{inv}}(I^Y \setminus E^{(\mathcal{Y})}(\mathcal{Y})) \). Then, \( \Psi(E(x)) \in I^Y \setminus E^{(\mathcal{Y})}(\mathcal{Y}) \). This contradicts with the fact that \( \Psi(E(x)) = E^{(\mathcal{Y})}(x) \in E^{(\mathcal{Y})}(\mathcal{Y}) \). Therefore, we must have \( \Psi(\beta(\mathcal{X})) = E^{(\mathcal{Y})}(\mathcal{Y}) \). This completes the proof of the claim.

Define \( \phi : \beta(\mathcal{X}) \to \mathcal{Y} \) by \( \phi = E^{(\mathcal{Y})} \circ \Psi|_{\beta(\mathcal{X})} \). Clearly, \( \phi \) is continuous and surjective by Proposition 3.12. Clearly, \( \phi \circ E = E^{(\mathcal{Y})} \circ \Psi|_{\beta(\mathcal{X})} |_{\mathcal{X}} = \text{id}|_{\mathcal{X}} \). Hence \( \phi|_{E(\mathcal{X})} = E_{\text{inv}} \).

Let \( \tilde{\phi} : \beta(\mathcal{X}) \to \mathcal{Y} \) be any continuous and surjective mapping such that \( \tilde{\phi}|_{E(\mathcal{X})} = E_{\text{inv}} \). By Proposition 3.56 and the denseness of \( E(\mathcal{X}) \) in \( \beta(\mathcal{X}) \), we have \( \phi = \tilde{\phi} \). Hence, \( \phi \) is unique.

Thus, we have shown that the Stone-Čech compactification \( \beta(\mathcal{X}) \) satisfies (i), (ii), and (iii). Next, we will show that \( \beta(\mathcal{X}) \) is unique. Let \( \mathcal{Y} \) be any compact Hausdorff space that satisfies (i), (ii), and (iii). Since \( \mathcal{X} \) and \( E(\mathcal{X}) \) are homeomorphic, we may identify \( \mathcal{X} \) as \( E(\mathcal{X}) \). Then, by (iii), \( \exists! \phi : \beta(\mathcal{X}) \to \mathcal{Y} \), which is continuous and surjective, such that \( \phi|_{\mathcal{X}} = \text{id}|_{\mathcal{X}} \). Since \( \mathcal{Y} \) satisfies (iii), then \( \exists! \lambda : \mathcal{Y} \to \beta(\mathcal{X}) \), which is continuous and surjective, such that \( \lambda|_{\mathcal{X}} = \text{id}|_{\mathcal{X}} \). Then, \( \phi \circ \lambda : \mathcal{Y} \to \mathcal{Y} \) is continuous and satisfies \( (\phi \circ \lambda)|_{\mathcal{X}} = \text{id}|_{\mathcal{X}} \). By Proposition 3.56, we have \( \phi \circ \lambda = \text{id}|_{\mathcal{Y}} \). On the other hand, \( \lambda \circ \phi : \beta(\mathcal{X}) \to \beta(\mathcal{X}) \) is continuous and satisfies \( (\lambda \circ \phi)|_{\mathcal{X}} = \text{id}|_{\mathcal{X}} \). Then, by Proposition 3.56, we have \( \lambda \circ \phi = \text{id}|_{\beta(\mathcal{X})} \). By Proposition 2.4, we have \( \lambda = \phi_{\text{inv}} \). Hence, \( \beta(\mathcal{X}) \) and \( \mathcal{Y} \) are homeomorphic. Hence, \( \beta(\mathcal{X}) \) is unique.

If, in addition, \( \mathcal{X} \) is locally compact, then, by Proposition 5.57, \( \mathcal{X} \) is open in \( \beta(\mathcal{X}) \). This completes the proof of the proposition.

\( \square \)

**Example 5.84** \( \mathbb{R}_e \) with the topology \( \mathcal{O}_{\mathbb{R}_e} \) introduced in Example 3.80 is Hausdorff and second countable. It is easy to show that \( \mathbb{R}_e \) is compact. By Proposition 5.14, \( \mathbb{R}_e \) is a normal topological space. By Urysohn Metrization Theorem 4.53, \( \mathbb{R}_e \) is metrizable. Hence, \( \mathbb{R}_e \) is a second countable metrizable compact Hausdorff topological space.
Chapter 6

Vector Spaces

6.1 Group

**Definition 6.1** A group is the triple $(G, +, e)$, which consists a nonempty set $G$, an operation $+: G \times G \to G$, and a unit element $e \in G$, satisfying, $\forall g_1, g_2, g_3 \in G$,

(i) $(g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)$; \hspace{1em} (associativity)
(ii) $e + g_1 = g_1 + e = g_1$; \hspace{1em} (unit element)
(iii) $\exists (-g_1) \in G$ such that $g_1 + (-g_1) = e = (-g_1) + g_1$. \hspace{1em} (existence of inverse, which is clearly unique)

We have the following result.

**Proposition 6.2** Let $(G, +, e)$ be a group and $H \subseteq G$ be nonempty. Then, $(H, +, e)$ is a group (which will be called a subgroup of $(G, +, e)$) if, and only if, $\forall g_1, g_2 \in H$, we have $g_1 + (-g_2) \in H$.

**Proof** “Necessity” This is obvious.

“Sufficiency” Since $H \neq \emptyset$, then $\exists g \in H$. Then, $e = g + (-g) \in H$. $\forall g_1, g_2 \in H$, $e + (-g_2) = (-g_2) \in H$. Then, $g_1 + g_2 = g_1 + (-(-g_2)) \in H$. Hence, $H$ is closed under $+$. Then, it is straightforward to check that $(H, +, e)$ satisfies all the properties of Definition 6.1. Hence, it is a group.

This completes the proof of the proposition. \hfill \square

**Proposition 6.3** Let $(G, +, e)$ be a group and $g_1, g_2, g_3 \in G$. If $g_1 + g_2 = g_1 + g_3$ then $g_2 = g_3$. On the other hand, if $g_1 + g_3 = g_2 + g_3$ then $g_1 = g_2$.

**Proof** If $g_1 + g_2 = g_1 + g_3$, then we have

$$g_2 = e + g_2 = ((-g_1) + g_1) + g_2 = (-g_1) + (g_1 + g_2)$$
$$= (-g_1) + (g_1 + g_3) = ((-g_1) + g_1) + g_3 = e + g_3 = g_3$$
If $g_1 + g_3 = g_2 + g_3$, then we have

$$g_1 = g_1 + e = g_1 + (g_3 + (-g_3)) = (g_1 + g_3) + (-g_3)$$

$$= (g_2 + g_3) + (-g_3) = g_2 + (g_3 + (-g_3)) = g_2 + e = g_2$$

This completes the proof of the proposition.

**Definition 6.4** Let $(G, +, e)$ be a group, the order of the group is the number of elements in $G$, if $G$ is finite. $\forall g \in G$, the order of $g$ is the integer $n > 0$ such that $g + \cdots + g = e$.

**Definition 6.5** Let $(G, +_G, e_G)$ and $(H, +_H, e_H)$ be two groups, and $T : G \to H$. $T$ is said to be a homomorphism if, $\forall g_1, g_2 \in G$, we have $T(g_1) +_H T(g_2) = T(g_1 +_G g_2)$. $T$ is said to be an isomorphism if it is bijective and a homomorphism, in this case, the two groups are said to be isomorphic.

Let $T : G \to H$ be a homomorphism, then $T(e_G) = e_H$.

**Definition 6.6** Let $(G, +, e)$ be a group. $g_1, g_2 \in G$ are said to be conjugate if $\exists g_3 \in G$ such that $g_2 = (-g_3) + g_1 + g_3$. Let $(H, +, e)$ be a subgroup of $(G, +, e)$. It is said to be normal (self-conjugate) if, $\forall h \in H, \forall g \in G$, we have $(-g) + h + g \in H$.

**Definition 6.7** Let $(G, +, 0)$ be a group. It is said to be abelian if, $\forall g_1, g_2 \in G$, we have $g_1 + g_2 = g_2 + g_1$ (commutativeness). Then, the unit element $0$ is also called the zero-element.

Sometimes, we are interested in an algebraic structure that is weaker than a group. For example, the structure of functions $f : X \to X$ with respect to the function composition operation. This leads us to the following definition.

**Definition 6.8** A semigroup is the triple $(G, \circ, e)$, which consists a nonempty set $G$, an operation $\circ : G \times G \to G$, and a unit element $e \in G$, satisfying, $\forall g_1, g_2, g_3 \in G$,

(i) $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$; (associativeness)

(ii) $e \circ g_1 = g_1 = g_1 \circ e$. (unit element)

Furthermore, a semigroup that satisfies $g_1 \circ g_2 = g_2 \circ g_1, \forall g_1, g_2 \in G$, is called an abelian semigroup.
6.2 Ring

Definition 6.9 A ring \((R, +, \times, 0)\) is an abelian group \((R, +, 0)\) with an operation \(\times: R \times R \to R\) such that, \(\forall r_1, r_2, r_3 \in R\),

(i) \((r_1 + r_2) \times r_3 = r_1 \times (r_2 + r_3)\); \ (associativeness)

(ii) \(r_1 \times (r_2 + r_3) = r_1 \times r_2 + r_1 \times r_3\); \ (right distributiveness)

(iii) \((r_1 + r_2) \times r_3 = r_1 \times r_3 + r_2 \times r_3\). \ (left distributiveness)

A ring is commutative if, \(\forall r_1, r_2 \in R\), we have \(r_1 \times r_2 = r_2 \times r_1\). A ring is with identity element if \(\exists 1 \in R\) such that, \(\forall r \in R\), \(1 \times r = r \times 1 = r\). (The identity element is unique if it exists.)

Proposition 6.10 Let \((R, +, \times, 0)\) be a ring. \(S \subseteq R\). Then, \((S, +, \times, 0)\) is a ring (which will be called a subring) if \((S, +, 0)\) is a subgroup of \((R, +, 0)\) and \(s_1 \times s_2 \in S\), \(\forall s_1, s_2 \in S\).

Proof Clearly, \((S, +, 0)\) is an abelian group. Then, it is straightforward to show that \((S, +, \times, 0)\) is a ring. \(\Box\)

Proposition 6.11 Let \((R, +, \times, 0)\) be a ring, and \(r_1, r_2 \in R\), then, we have

(i) \(r_1 \times 0 = 0 \times r_1 = 0\);

(ii) \(r_1 \times (−r_2) = (−r_1) \times r_2 = −(r_1 \times r_2)\);

(iii) \((−r_1) \times (−r_2) = r_1 \times r_2\).

Proof Note that \(0 + 0 \times r_1 = 0 \times r_1 = (0 + 0) \times r_1 = 0 \times r_1 + 0 \times r_1\). By Proposition 6.3, we have \(0 \times r_1 = 0\). Similarly, we can show that \(r_1 \times 0 = 0\).

Note that \(r_1 \times r_2 + r_1 \times (−r_2) = r_1 \times (r_2 + (−r_2)) = r_1 \times 0 = 0\). Then, we have \(r_1 \times (−r_2) = −(r_1 \times r_2)\). Similarly, we can show that \((−r_1) \times r_2 = −(r_1 \times r_2)\).

Note that \((−r_1) \times (−r_2) + (−(r_1 \times r_2)) = (−r_1) \times (−r_2) + r_1 \times (−r_2) = ((−r_1) + r_1) \times (−r_2) = 0 \times (−r_2) = 0\). Then, we have \((−r_1) \times (−r_2) = r_1 \times r_2\).

This completes the proof of the proposition. \(\Box\)

Definition 6.12 Let \((R, +, \times, 0)\) be a ring and \(S \subseteq R\). Then, \(S\) is said to be an ideal if \((S, +, 0)\) is a subgroup of \((R, +, 0)\) and \(r \times s, s \times r \in S\), \(\forall r \in R\) and \(\forall s \in S\).

Definition 6.13 Let \((R, +_R, \times_R, 0_R)\) and \((S, +_S, \times_S, 0_S)\) be two rings and \(T: R \to S\). \(T\) is said to be a ring-homomorphism if it is a homomorphism and \(T(r_1) \times_T T(r_2) = T(r_1 \times_R r_2), \forall r_1, r_2 \in R\). \(T\) is said to be a ring-isomorphism if it is a bijective ring-homomorphism, in this case, the two rings are said to be isomorphic.
6.3 Field

Definition 6.14 Let \((F, +, \times, 0)\) be a commutative ring with identity element \(1\). Then, the quintuple \((F, +, \times, 0, 1)\) is said to be a field if \((F \setminus \{0\}, \times, 1)\) form an abelian group.

Proposition 6.15 Let \((F, +, \times, 0, 1)\) be a field, \(A \subseteq F\), and \(0, 1 \in A\). Then, \((A, +, \times, 0, 1)\) is a field (which will be called a subfield) if, and only if, \(\forall a_1, a_2 \in A, a_1 + (-a_2) \in A\), and \(a_1 \times a_2^{-1} \in A\) when \(a_2 \neq 0\), where \(a_2^{-1}\) denotes the multiplicative inverse of \(a_2\).

Proof “Necessity” This is straightforward.

“Sufficiency” By Proposition 6.2, \((A, +, 0)\) is a group. Since \((F, +, 0)\) is abelian, then \((A, +, 0)\) is also an abelian group. \(\forall a_1 \in A\) with \(a_1 \neq 0\). By the assumption of the proposition, we have \(1 \in A\), by the property of field, we have \(1 \neq 0\). Then, \(a_1^{-1} = 1 \times a_1^{-1} \in A\). \(\forall a_1, a_2 \in A\). If \(a_2 = 0\), then, by Proposition 6.11, \(a_1 \times a_2 = 0 \in A\). On the other hand, if \(a_2 \neq 0\), we have \(a_2^{-1} \in A\) and \(a_1 \times a_2 = a_1 \times (a_2^{-1})^{-1} \in A\). Hence, by Proposition 6.10, \((A, +, \times, 0)\) is a ring. Since \((F, +, \times, 0)\) is a commutative ring with identity element \(1\) and \(1 \in A\), then \((A, +, \times, 0)\) is also a commutative ring with identity element \(1\). Note that \(1 \in A \setminus \{0\}\), then, \(A \setminus \{0\} \neq \emptyset\) and \(A \setminus \{0\} \subseteq F \setminus \{0\}\). \(\forall a_1, a_2 \in A \setminus \{0\}, a_1 \times a_2^{-1} \in A\). We claim that \(a_1 \times a_2^{-1} \neq 0\) since, otherwise, \(a_1 = a_1 \times 1 = a_1 \times (a_2^{-1} \times a_2) = (a_1 \times a_2^{-1}) \times a_2 = 0\) by Proposition 6.11, which is a contradiction. Hence, \(a_1 \times a_2^{-1} \in A \setminus \{0\}\). Since \((F \setminus \{0\}, \times, 1)\) is an abelian group, then, by Proposition 6.2, \((A \setminus \{0\}, \times, 1)\) is also a group and is further abelian. Therefore, \((A, +, \times, 0, 1)\) is a field.

This completes the proof of the proposition. \(\square\)

6.4 Vector Spaces

Associated with every vector space is a set of scalars. This set of scalars can be any algebraic field \(F := (F, +, \cdot, 0, 1)\). Examples of fields are the rational numbers \(\mathbb{Q}\), the real numbers \(\mathbb{R}\), and the complex numbers \(\mathbb{C}\). Here, we will abuse the notation to say \(x \in F\) when \(x \in F\).

Definition 6.16 A vector space \(X\) over a field \(F := (F, +, \cdot, 0, 1)\) is a set \(X\) of elements called vectors together with two operations \(\oplus\) and \(\otimes\).

\(\oplus : X \times X \to X\) is called vector addition. It associates any two vectors \(x, y \in X\) with a vector \(x \oplus y \in X\), the sum of \(x\) and \(y\). \(\otimes : F \times X \to X\) is called scalar multiplication. It associates a scalar \(\alpha \in F\) and a vector \(x \in X\) with a vector \(\alpha \otimes x \in X\), the scalar multiple of \(x\) by \(\alpha\). Furthermore, the following properties hold for \(\forall x, y, z \in X\) and \(\forall \alpha, \beta \in F\)

\[(i)\] \(x \oplus y = y \oplus x\); \hspace{1em} \text{(commutative law)}

\[(ii)\] \((x \oplus y) \oplus z = x \oplus (y \oplus z);\hspace{1em} \text{(associative law)}\]
We thus denote the quadruple \((X, \oplus, \ominus, \vartheta)\) by \(X\). The vector space is denoted by \((X, \mathcal{F})\).

For convenience, \((-1) \otimes x := \ominus x\) and called the negative of the vector \(x\). Note that
\[
\ominus(x \oplus y) = \ominus x \oplus \ominus y
\]
We will also denote \(x \ominus y := x \ominus (\ominus y)\). We will abuse the notation to say \(x \in X\) when \(x \in X\). Note that \((X, \oplus, \ominus, \vartheta)\) forms an abelian group.

**Proposition 6.17** Let \(X := (X, \otimes, \ominus, \vartheta)\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\). \(\forall x, y, z \in X, \forall \alpha, \beta \in F\), we have
1. \(x \ominus y = x \oplus z \Rightarrow y = z\); (cancellation law)
2. \(\alpha \otimes x = \alpha \otimes y\) and \(\alpha \neq 0 \Rightarrow x = y\); (cancellation law)
3. \(\alpha \otimes x = \beta \otimes x\) and \(x \neq \vartheta \Rightarrow \alpha = \beta\); (cancellation law)
4. \((\alpha - \beta) \otimes x = \alpha \otimes x \ominus \beta \otimes x\); (distributive law)
5. \(\alpha \otimes (x \ominus y) = \alpha \otimes x \ominus \alpha \otimes y\); (distributive law)
6. \(\alpha \otimes \vartheta = \vartheta\).

We will call \(\vartheta\) the origin.

**Example 6.18** \(X = \{\vartheta\}\) with \(\vartheta \oplus \vartheta = \vartheta\) and \(\alpha \otimes \vartheta = \vartheta, \forall \alpha \in F\). Then, \((X, \oplus, \ominus, \vartheta)\) is a vector space over \(\mathcal{F}\).

**Example 6.19** Let \(\mathcal{F} := (F, +, \cdot, 0, 1)\) be a field. Then, \((F, +, \cdot, 0)\) is a vector space over \(\mathcal{F}\). We will abuse the notation and say that \(\mathcal{F}\) is a vector space over \(F\).

**Example 6.20** Let \(\mathcal{F} := (F, +, \cdot, 0, 1)\) be a field, \(\mathcal{Y} := (Y, \ominus_Y, \ominus_Y, \vartheta_Y)\) be a vector space over \(\mathcal{F}\), and \(A\) be a set. \(X = \{f : A \rightarrow \mathcal{Y}\}\), that is, \(X\) is the set of all \(\mathcal{Y}\)-valued functions on \(A\). Define vector addition and scalar multiplication by, \(\forall x, y \in X, \forall \alpha \in \mathcal{F}, z_a := x \oplus y \in X\) is given by \(z_a(u) = x(u) \ominus_Y y(u), \forall u \in A, z_a := \alpha \otimes x \in X\) is given by \(z_a(u) = \alpha \otimes_Y x(u), \forall u \in A\). Let \(\vartheta \in X\) be given by \(\vartheta(u) = \vartheta_Y, \forall u \in A\).

Now, we will show that \((X, \oplus, \ominus, \vartheta)\) is a vector space over \(\mathcal{F}\). \(\forall x, y, z \in X, \forall \alpha, \beta \in \mathcal{F}, \forall u \in A\), we have
CHAPTER 6. VECTOR SPACES

Then, it is straightforward to check that
\[ F \]
Therefore,
\[ X \]
that
\[ F \]
Define vector addition and scalar multiplication by
\[ x \]
\[ F \]
Example 6.22
\[ X \]
\[ F \]
\[ F \]
\[ F \]

\[ \begin{align*}
(i) \quad (x \oplus y)(u) &= x(u) \oplus_Y y(u) = y(u) \oplus_Y x(u) = (y \oplus x)(u) \Rightarrow x \oplus y = y \oplus x; \\
(ii) \quad ((x \oplus y) \oplus z)(u) &= (x \oplus y)(u) \oplus_Y z(u) = (x(u) \oplus_Y y(u)) \oplus_Y z(u) = x(u) \oplus_Y (y(u) \oplus_Y z(u)) = (x \oplus (y \oplus z))(u) \Rightarrow (x \oplus y) \oplus z = x \oplus (y \oplus z); \\
(iii) \quad (x \oplus \vartheta)(u) &= x(u) \oplus_Y \vartheta_Y = x(u) \Rightarrow x \oplus \vartheta = x; \\
(iv) \quad (\alpha \otimes (x \oplus y))(u) &= \alpha \otimes_Y (x \oplus y)(u) = \alpha \otimes_Y (x(u) \oplus_Y y(u)) = \alpha \otimes_Y x(u) \oplus_Y \alpha \otimes_Y y(u) = (\alpha \otimes x)(u) \oplus_Y (\alpha \otimes y)(u) \Rightarrow \alpha \otimes (x \oplus y) = \alpha \otimes x \oplus \alpha \otimes y; \\
(v) \quad ((\alpha + \beta) \otimes x)(u) &= (\alpha + \beta) \otimes_Y x(u) = \alpha \otimes_Y x(u) \oplus_Y \beta \otimes_Y x(u) = (\alpha \otimes x)(u) \oplus_Y (\beta \otimes x)(u) \Rightarrow (\alpha + \beta) \otimes x = \alpha \otimes x \oplus \beta \otimes x; \\
(vi) \quad ((\alpha \otimes x)(u) &= (\alpha \otimes y)(u) = \alpha \otimes_Y (\beta \otimes x)(u) \Rightarrow (\alpha \otimes x) = \alpha \otimes (\beta \otimes x); \\
(vii) \quad (0 \otimes x)(u) &= 0 \otimes_Y x(u) = \vartheta_Y = \vartheta(u) \Rightarrow 0 \otimes x = \vartheta; \quad (1 \otimes x)(u) = 1 \otimes_Y x(u) = x(u) \Rightarrow 1 \otimes x = x.
\end{align*} \]

Therefore, \( X \) is a vector space over \( F \). This vector space will be denoted by \((M(A,\mathbb{Y}),\mathbb{F})\).

Example 6.21
Let \( F := (F,+,\cdot,0,1) \) be a field. \( X = F^n \) with \( n \in \mathbb{N} \). Define vector addition and scalar multiplication by \( x \oplus y := (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n) \in X, \alpha \otimes x := (\alpha \xi_1, \ldots, \alpha \xi_n) \in X, \forall \alpha \in F. \) Let \( \vartheta := (0,0,\ldots,0) \in X. \) Then, it is straightforward to check that \( F^n := (X,\oplus,\otimes,\vartheta) \) is a vector space over \( F \).

Example 6.22
Let \( F := (F,+,\cdot,0,1) \) be a field. \( X = F^{m \times n} := \{m \times n \text{-dimensional F valued matrices}\} \) with \( m, n \in \mathbb{N} \). Define vector addition and scalar multiplication by \( x \oplus y := (\xi_{ij} + \eta_{ij})_{m \times n} \in X, \alpha \otimes x := (\alpha \xi_{ij})_{m \times n} \in X, \forall \alpha \in F, \forall x = (\xi_{ij}), y = (\eta_{ij})_{m \times n} \in X, \forall \alpha \in F, \forall x = (\xi_{ij}), y = (\eta_{ij}). \) Let \( \vartheta := (0)_{m \times n} \in X. \) Then, it is straightforward to check that \( F^{m \times n} := (X,\oplus,\otimes,\vartheta) \) is a vector space over \( F \).

Example 6.23
Let \( F := (F,+,\cdot,0,1) \) be a field. \( X = \{(\xi_k)_{k=1}^\infty \mid \xi_k \in F, \forall k \in \mathbb{N}\}. \) Define vector addition and scalar multiplication by \( x \oplus y := (\xi_k + \eta_k)_{k=1}^\infty \in X, \alpha \otimes x := (\alpha \xi_k)_{k=1}^\infty \in X, \forall x = (\xi_k)_{k=1}^\infty \in X, \forall \alpha \in F. \) Let \( \vartheta := (0,0,\ldots) \in X. \) Then, it is straightforward to check that \( (X,\oplus,\otimes,\vartheta) \) is a vector space over \( F \).

6.5 Product Spaces

Proposition 6.24 Let \( \mathcal{X} := (X,\oplus_X,\otimes_X,\vartheta_X) \) and \( \mathcal{Y} := (Y,\oplus_Y,\otimes_Y,\vartheta_Y) \) be vector spaces over the field \( F := (F,+,\cdot,0,1). \) The Cartesian product of \( \mathcal{X} \) and \( \mathcal{Y}, \) denoted by \( X \times \mathcal{Y}, \) is the quadruple \( (X \times \mathcal{Y},\oplus,\otimes,(\vartheta_X,\vartheta_Y)). \)
where the vector addition \( \oplus : (X \times Y) \times (X \times Y) \to X \times Y \) and the scalar multiplication \( \otimes : \mathcal{F} \times (X \times Y) \to X \times Y \) are given by, \( \forall (x_1, y_1), (x_2, y_2) \in X \times Y, \forall \alpha \in \mathcal{F}, (x_1, y_1) \oplus (x_2, y_2) := (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) \) and \( \alpha \otimes (x_1, y_1) := (\alpha \otimes_X x_1, \alpha \otimes_Y y_1) \). Then, \( (X \times \mathcal{Y}, \mathcal{F}) \) is a vector space.

**Proof** Let \( \vartheta := (\vartheta_X, \vartheta_Y) \in X \times Y. \forall (x_1, y_1), (x_2, y_2), (x_3, y_3) \in X \times Y, \forall \alpha, \beta \in \mathcal{F}.

(i) \((x_1, y_1) \oplus (x_2, y_2) = (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) = (x_2, y_2) \oplus (x_1, y_1)\);

(ii) \((x_1, y_1) \oplus (x_2, y_2) \oplus (x_3, y_3) = (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) \oplus (x_3, y_3) = ((x_1 \oplus_X x_2) \oplus_X x_3, (y_1 \oplus_Y y_2) \oplus_Y y_3) = (x_1 \oplus_X (x_2 \oplus_X x_3), y_1 \oplus_Y (y_2 \oplus_Y y_3)) = (x_1, y_1) \oplus ((x_2, y_2) \oplus (x_3, y_3))\);

(iii) \((x_1, y_1) \vartheta = (x_1 \vartheta_X, x_1 \vartheta_Y) = (x_1, y_1)\);

(iv) \(\alpha \otimes ((x_1, y_1) \oplus (x_2, y_2)) = \alpha \otimes (x_1 \oplus_X x_2, y_1 \oplus_Y y_2) = (\alpha \otimes_X x_1 \oplus_X x_2, \alpha \otimes_Y y_1 \oplus_Y y_2) = (\alpha \otimes_X x_1, \alpha \otimes_Y y_1) \oplus (\alpha \otimes_X x_2, \alpha \otimes_Y y_2) = \alpha \otimes (x_1, y_1) \oplus \alpha \otimes (x_2, y_2)\);

(v) \((\alpha + \beta) \otimes (x_1, y_1) = ((\alpha + \beta) \otimes_X x_1, (\alpha + \beta) \otimes_Y y_1) = (\alpha \otimes_X x_1 \oplus_X x_2, \alpha \otimes_Y y_1 \oplus_Y y_2) = (\alpha \otimes_X x_1, \alpha \otimes_Y y_1) \oplus (\beta \otimes_X x_2, \beta \otimes_Y y_1) = \alpha \otimes (x_1, y_1) \oplus \beta \otimes (x_1, y_1)\);

(vi) \((\alpha \beta) \otimes (x_1, y_1) = ((\alpha \beta) \otimes_X x_1, (\alpha \beta) \otimes_Y y_1) = (\alpha \otimes_X (\beta \otimes_X x_1), \alpha \otimes_Y (\beta \otimes_Y y_1)) = \alpha \otimes (\beta \otimes_X x_1, \beta \otimes_Y y_1) = \alpha \otimes (\beta \otimes (x_1, y_1))\);

(vii) \(0 \otimes (x_1, y_1) = (0 \otimes_X x_1, 0 \otimes_Y y_1) = (\vartheta_X, \vartheta_Y) = \vartheta, \ 1 \otimes (x_1, y_1) = (1 \otimes_X x_1, 1 \otimes_Y y_1) = (x_1, y_1)\).

Hence, \(X \times \mathcal{Y}\) is a vector space over \(\mathcal{F}\). \(\square\)

With the above definition, it is easy to generalize to \(X_1 \times X_2 \times \cdots \times X_n\), where \(n \in \mathbb{N}\). We will also write \(X^n = \underbrace{X \times \cdots \times X}_n\), where \(n \in \mathbb{N}\). When \(n = 0\), \(\prod_{i=1}^{n} X_i\) is given by the vector space defined in Example 6.18.

### 6.6 Subspaces

**Proposition 6.25** Let \(X := (X, \oplus, \otimes, \vartheta)\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\), and \(M \subseteq X\) with \(M \neq \emptyset\). Then, \(M := (M, \oplus, \otimes, \vartheta)\) is a vector space over \(\mathcal{F}\) (which will be called a subspace of \((X, \mathcal{F})\)) if, and only if, \(\forall x, y \in M, \forall \alpha, \beta \in \mathcal{F}, we have \alpha \otimes x + \beta \otimes y \in M\). We will also abuse the notation to say \(M\) is a subspace of \((X, \mathcal{F})\). \(M\) is said to be a proper subspace of \((X, \mathcal{F})\) if \(M \subset X\).
The range space \( \beta A \) is said to be affine if \( y \) is an operator and \( \forall x, y \in M, \forall \alpha, \beta \in F, x \oplus y = 1 \otimes x \oplus 1 \otimes y \in M \). Hence, \( M \) is closed under vector addition. \( \alpha \otimes x = (\alpha + 0) \otimes x = \alpha \otimes x \oplus 0 \otimes x \in M \). Hence, \( M \) is closed under scalar multiplication. Then, it is straightforward to show that \( \mathcal{M} \) is a vector space over \( F \). “Only if” This is straightforward. This completes the proof of the proposition.

Example 6.26 We present the following list of examples of subspaces.

1. Let \((\mathcal{X}, F)\) be a vector space. Then, the singleton set \( M = \{\varnothing\} \) is a subspace.

2. Consider the vector space \((\mathbb{R}^3, \mathbb{R})\). Any straight line or plane that passes through the origin is a subspace.

3. Consider the vector space \((\mathbb{R}^n, \mathbb{R})\). Let \( a \in \mathbb{R}^n \). The set \( M := \{x \in \mathbb{R}^n \mid \langle a, x \rangle = 0\} \) is a subspace.

4. Let \( X := \{(\xi_k)_{k=1}^\infty \mid \xi_k \in \mathbb{R}, k \in \mathbb{N}\}, \oplus \) and \( \otimes \) be the usual addition and scalar multiplication, and \( \vartheta = (0, 0, \ldots) \). By Example 6.23, \( \mathcal{X} := (X, \oplus, \otimes, \vartheta) \) is a vector space over \( \mathbb{R} \). Let \( M := \{\xi_k\}_{k=1}^\infty \in \mathcal{X} \mid \lim_{k \to \infty} \xi_k \in \mathbb{R}\} \) is a subspace.

5. Let \( X := \{f : (0, 1] \to \mathbb{R}^n\}, \oplus \) and \( \otimes \) be the usual addition and scalar multiplication, and \( \vartheta : (0, 1] \to \mathbb{R}^n \) be given by \( \vartheta(t) = \vartheta_{\mathbb{R}^n}, \forall t \in (0, 1] \). By Example 6.20, \( \mathcal{X} := (X, \oplus, \otimes, \vartheta) \) is a vector space over \( \mathbb{R} \). Let \( M := \{f \in \mathcal{X} \mid f \) is continuous} \) is a subspace.

To simplify notation in the theory, we will later simply discuss a vector space \((\mathcal{X}, F)\) without further reference to components of \( \mathcal{X} \), where the operations are understood to be \( \oplus_X \) and \( \otimes_X \) and the null vector is understood to be \( \vartheta_X \). When it is clear from the context, we will neglect the subscript \( \mathcal{X} \). Also, we will write \( x_1 + x_2 \) for \( x_1 \oplus x_2 \) and \( \alpha x_1 \) for \( \alpha \otimes x_1 \), \( \forall x_1, x_2 \in \mathcal{X}, \forall \alpha \in F \).

Definition 6.27 Let \( \mathcal{X} \) be a vector space over the field \( F \). \( f : \mathcal{X} \to F \) is said to be a functional.

Definition 6.28 Let \( \mathcal{X} \) and \( \mathcal{Y} \) be vector spaces over the field \( F \). \( A : \mathcal{X} \to \mathcal{Y} \) is said to be linear if \( \forall x_1, x_2 \in \mathcal{X}, \forall \alpha, \beta \in F, A(\alpha x_1 + \beta x_2) = \alpha A(x_1) + \beta A(x_2) \). Then, \( A \) is called a (vector space) homomorphism or a linear operator. Furthermore, if it is bijective, then \( A \) is said to be a (vector space) isomorphism. The null space of \( A \) is \( \mathcal{N}(A) := \{x \in \mathcal{X} \mid A(x) = 0\} \). The range space of \( A \) is \( \mathcal{R}(A) := \text{range}(A) \). \( B : \mathcal{X} \to \mathcal{Y} \) is said to be an affine operator if \( B(x) = A(x) + y_0, \forall x \in \mathcal{X}, \) where \( A : \mathcal{X} \to \mathcal{Y} \) is a linear operator and \( y_0 \in \mathcal{Y} \).
Example 6.29 A row vector \( v \in \mathbb{R}^{1 \times n} \) is a linear functional on \( \mathbb{R}^n \). A matrix \( A \in \mathbb{R}^{m \times n} \) is a linear function of \( \mathbb{R}^n \) to \( \mathbb{R}^m \).

For linear operators, we will adopt the following convention. Let \( A : \mathcal{X} \to \mathcal{Y} \), \( B : \mathcal{Y} \to \mathcal{Z} \), and \( x \in \mathcal{X} \), where \( \mathcal{X}, \mathcal{Y}, \mathcal{Z} \) are vector spaces and \( A \) and \( B \) are linear operators. We will write \( Ax \) for \( A(x) \) and \( BA \) for \( B \circ A \). Clearly, \( \mathcal{N}(A) \) is a subspace of \( \mathcal{X} \) and \( \mathcal{R}(A) \) is a subspace of \( \mathcal{Y} \). When \( A \) and \( B \) are bijective, we will denote \( A^{-1} \) and \( B^{-1} \) as \( \text{inv} \) operators, then \( BA \) is also bijective and \( (BA)^{-1} = A^{-1}B^{-1} \).

Definition 6.30 Let \( (\mathcal{X}, \mathcal{F}) \) be a vector space, \( \alpha \in \mathcal{F} \), and \( S, T \subseteq \mathcal{X} \). The sets \( \alpha S \) and \( S + T \) are defined by

\[
\alpha S := \{ \alpha s \in \mathcal{X} \mid s \in S \}; \quad S + T := \{ s + t \in \mathcal{X} \mid s \in S, t \in T \}
\]

This concept is illustrated in Figure 6.1.

![Figure 6.1: The sum of two sets.](image)

We should note that \( S + T = T + S \), \( \emptyset + S = \emptyset \), \( \{ \varnothing \} + S = S \), and \( \alpha \emptyset = \emptyset \). Thus, \( S - T := S + (-T) \).

Proposition 6.31 Let \( M \) and \( N \) be subspaces of a vector space \( (\mathcal{X}, \mathcal{F}) \) and \( \alpha \in \mathcal{F} \). Then, \( M \cap N \), \( M + N \), and \( \alpha M \) are subspaces of \( (\mathcal{X}, \mathcal{F}) \).

Proof Since \( M \) and \( N \) are subspaces, then \( \varnothing \in M \) and \( \varnothing \in N \). Hence, \( \varnothing \in M \cap N \neq \emptyset \), \( \varnothing \in M + N \neq \emptyset \), and \( \varnothing = \varnothing \in \alpha M \neq \emptyset \).

\( \forall x, y \in M \cap N, \forall \alpha, \beta \in \mathcal{F} \), \( \alpha x + \beta y \in M \) and \( \alpha x + \beta y \in N \). Then, \( \alpha x + \beta y \in M \cap N \). Hence, \( M \cap N \) is a subspace.

\( \forall x, y \in M + N, \forall \alpha, \beta \in \mathcal{F} \), we have \( x = x_1 + x_2 \) and \( y = y_1 + y_2 \), where \( x_1, y_1 \in M \) and \( x_2, y_2 \in N \). Then, \( \alpha x_1 + \beta y_1 \in M \) and \( \alpha x_2 + \beta y_2 \in N \). This implies that \( \alpha x + \beta y = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \in M + N \). Hence, \( M + N \) is a subspace.
\( \forall x, y \in \alpha M, \forall \alpha, \beta \in \mathcal{F}, \) we have \( x = \alpha \bar{x} \) and \( y = \alpha \bar{y}, \) where \( \bar{x}, \bar{y} \in M. \)

Then, \( \alpha x + \beta y = \alpha \bar{x} + \beta \bar{y} = \bar{\alpha (\bar{x} + \bar{y})} \in \alpha M. \) Hence, \( \alpha M \) is a subspace.

This completes the proof of the proposition. \( \square \)

We should note that \( M \cup N \) is in general not a subspace.

**Definition 6.32** A linear combination of vectors \( x_1, \ldots, x_n, \) where \( n \in \mathbb{Z}_+ \), in a vector space \((X, \mathcal{F})\) is a sum of the form \( \sum_{i=1}^{n} \alpha_i x_i := \alpha_1 x_1 + \cdots + \alpha_n x_n, \) where \( \alpha_1, \ldots, \alpha_n \in \mathcal{F}. \)

Note that + is defined for two vectors. To sum \( n \) vectors, one must add two at a time. By the definition of vector space, the simplified notion is not ambiguous. When \( n = 0, \) we take the sum to be \( \emptyset_X. \) A linear combination can only involve finitely many vectors.

**Definition 6.33** Let \((X, \mathcal{F})\) be a vector space and \( S \subseteq X. \)

\[
\text{span}(S) := \{ \sum_{i=1}^{n} \alpha_i x_i \mid x_i \in S, \alpha_i \in \mathcal{F}, i = 1, \ldots, n, n \in \mathbb{Z}_+ \}
\]

is called the subspace generated by \( S. \)

**Proposition 6.34** Let \((X, \mathcal{F})\) be a vector space and \( S \subseteq X. \) Then, \( \text{span}(S) \) is the smallest subspace containing \( S. \)

**Proof** Clearly, \( \emptyset \in \text{span}(S) \neq \emptyset. \) \( \forall x, y \in \text{span}(S), \forall \alpha, \beta \in \mathcal{F}. \) It is easy to show that \( \alpha x + \beta y \in \text{span}(S). \) Hence, \( \text{span}(S) \) is a subspace.

Clearly \( \text{span}(S) \supseteq S. \)

\( \forall M \subseteq X \) such that \( S \subseteq M \) and \( M \) is a subspace. Clearly, \( \emptyset \in M \) by the proof of Proposition 6.25. \( \forall y \) equals to a linear combination of vectors in \( S. \) Then, \( y = \sum_{i=1}^{n} \alpha_i x_i \) with \( n \in \mathbb{Z}_+, \) \( x_1, \ldots, x_n \in S, \) and \( \alpha_1, \ldots, \alpha_n \in \mathcal{F}. \) This implies that \( x_1, \ldots, x_n \in M \) and \( y \in M. \) Hence, \( \text{span}(S) \subseteq M. \)

Hence, \( \text{span}(S) \) is the smallest subspace containing \( S. \) This completes the proof of the proposition. \( \square \)

**Definition 6.35** Let \((X, \mathcal{F})\) be a vector space, \( M \subseteq X \) be a subspace, and \( x_0 \in X. \) Then, \( V := \{x_0\} + M \) is called a linear variety.

The translation of a subspace is a linear variety. We will abuse the notation to write \( x_0 + M \) for \( \{x_0\} + M. \) \( \forall \bar{x} \in V, \) \( V - \bar{x} := V - \{\bar{x}\} \) is a subspace.

**Definition 6.36** Let \((X, \mathcal{F})\) be a vector space, \( S \subseteq X, \) and \( S \neq \emptyset. \) The linear variety generated by \( S, \) denoted by \( \nu(S), \) is defined as the intersection of all linear varieties in \( X \) that contain \( S. \)

**Proposition 6.37** Let \((X, \mathcal{F})\) be a vector space and \( \emptyset \neq S \subseteq X. \) Then, \( \nu(S) \) is a linear variety given by \( \nu(S) = x_0 + \text{span}(S - x_0), \) where \( x_0 \) is any vector in \( S. \)
Proposition 6.39 \[ \text{Definition 6.38} \]

Proof \[ \forall \text{ that } \ x \in \mathbb{R} \text{ or } \mathbb{C}. \]

Definition 6.38 Let \((\mathcal{X}, K)\) be a vector space and \(C \subseteq \mathcal{X}\). \(C\) is said to be convex if, \(\forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subseteq \mathbb{R}, \text{ we have } \alpha x_1 + (1-\alpha)x_2 \in C. \)

Subspaces and linear varieties are convex, so is \(\emptyset. \)

Proposition 6.39 Let \((\mathcal{X}, K)\) be a vector space and \(K, G \subseteq \mathcal{X}\) be convex sets. Then,

1. \(\lambda K\) is convex, \(\forall \lambda \in K; \)

2. \(K + G\) is convex.

Proof \[ \forall \lambda \in K. \forall x_1, x_2 \in \lambda K, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \exists k_1, k_2 \in K \text{ such that } x_1 = \lambda k_1 \text{ and } x_2 = \lambda k_2. \text{ Then, } \alpha x_1 + (1-\alpha)x_2 = \alpha \lambda k_1 + (1-\alpha)\lambda k_2 = \lambda (\alpha k_1 + (1-\alpha)k_2). \text{ Since } K \text{ is convex, } \alpha k_1 + (1-\alpha)k_2 \in K. \text{ This implies that } \alpha x_1 + (1-\alpha)x_2 \in \lambda K. \text{ Hence, } \lambda K \text{ is convex.} \]

\[ \forall x_1, x_2 \in K + G, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \exists k_1, k_2 \in K \text{ and } \exists g_1, g_2 \in G \text{ such that } x_i = k_i + g_i, i = 1, 2. \text{ Note that } \]

\[ \alpha x_1 + (1-\alpha)x_2 = \alpha(k_1 + g_1) + (1-\alpha)(k_2 + g_2) = \alpha k_1 + \alpha g_1 + (1-\alpha)k_2 + (1-\alpha)g_2 = (\alpha k_1 + (1-\alpha)k_2) + (\alpha g_1 + (1-\alpha)g_2) \]

Since \(K\) and \(G\) are convex, then \((\alpha k_1 + (1-\alpha)k_2) \in K\) and \((\alpha g_1 + (1-\alpha)g_2) \in G. \text{ This implies that } \alpha x_1 + (1-\alpha)x_2 \in K + G. \text{ Hence, } K + G \text{ is convex.} \]

This completes the proof of the proposition. \(\square\)

Proposition 6.40 Let \((\mathcal{X}, K)\) be a vector space and \(\{C_\lambda\}_{\lambda \in \Lambda}\) be a collection of convex subsets of \(\mathcal{X}\). Then, \(C := \bigcap_{\lambda \in \Lambda} C_\lambda\) is convex.

Proof \[ \forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \forall \lambda \in \Lambda, x_1, x_2 \in C_\lambda. \text{ Since } C_\lambda \text{ is convex, } \alpha x_1 + (1-\alpha)x_2 \in C_\lambda. \text{ This implies that } \alpha x_1 + (1-\alpha)x_2 \in C. \text{ Hence, } C \text{ is convex.} \text{ This completes the proof of the proposition.} \(\square\)

Definition 6.41 Let \((\mathcal{X}, K)\) be a vector space and \(S \subseteq \mathcal{X}\). The convex hull generated by \(S\), denoted by \(\text{co}(S)\), is the smallest convex set containing \(S\).
CHAPTER 6. VECTOR SPACES

Definition 6.42 Let \((X, K)\) be a vector space. A convex combination of vectors \(x_1, \ldots, x_n \in X\), where \(n \in \mathbb{N}\), is a linear combination \(\sum_{i=1}^{n} \alpha_i x_i\) with \(\alpha_i \in [0, 1] \subset \mathbb{R}\), \(\forall i = 1, \ldots, n\), and \(\sum_{i=1}^{n} \alpha_i = 1\).

Proposition 6.43 Let \((X, K)\) be a vector space and \(S \subseteq X\). Then,
\[
\text{co}(S) = \{ \text{convex combinations of vectors in } S \}
\]

Proof We need the follow result.

Claim 6.43.1 Let \(G \subseteq X\) be a convex subset. Then any convex combination of vectors in \(G\) belongs to \(G\).

Proof of claim: We need to show: \(\forall n \in \mathbb{N}, \forall x_1, \ldots, x_n \in G\), \(\forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}\), such that \(\alpha_i \in [0, 1] \subset \mathbb{R}\), \(i = 1, \ldots, n\), \(\sum_{i=1}^{n} \alpha_i = 1\) implies \(\sum_{i=1}^{n} \alpha_i x_i \in G\). We will prove this by mathematical induction on \(n\).

1° Consider \(n = 1\). Then, \(\alpha_1 = 1\) and \(\sum_{i=1}^{1} \alpha_i x_i = x_1 \in G\). The result holds.

2° Assume that the result holds for \(n = k \in \mathbb{N}\).

3° Consider the case \(n = k+1\). Without loss of generality, assume \(\alpha_1 > 0\).

By the induction hypothesis, we have \(\sum_{i=1}^{k} \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_k} x_i \in G\). Then,
\[
\sum_{i=1}^{k+1} \alpha_i x_i = (\alpha_1 + \cdots + \alpha_k) \sum_{i=1}^{k} \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_k} x_i + \alpha_{k+1} x_{k+1}
\]
6.7. CONVEX SETS

\[ = \alpha_{k+1}x_{k+1} + (1 - \alpha_{k+1}) \sum_{i=1}^{k} \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_k} x_i \]

By the convexity of \( G \), we have \( \sum_{i=1}^{k+1} \alpha_i x_i \in G \). Hence, the result holds for \( n = k + 1 \).

This completes the induction process and the proof of the claim. \( \square \)

\( \forall x_1, x_2 \in K := \{ \text{convex combinations of vectors in } S \}, \forall \alpha \in [0, 1] \subset \mathbb{R} \). By the definition of \( K \), \( \forall i = 1, 2, \exists n_i \in \mathbb{N}, \exists y_{i,1}, \ldots, y_{i,n_i} \in S, \exists \alpha_{i,1}, \ldots, \alpha_{i,n_i} \in [0, 1] \subset \mathbb{R} \), such that \( \sum_{j=1}^{n_i} \alpha_{i,j} = 1 \) and \( x_i = \sum_{j=1}^{n_i} \alpha_{i,j} y_{i,j} \). Then,

\[ \alpha x_1 + (1 - \alpha) x_2 = \alpha \sum_{j=1}^{n_1} \alpha_{1,j} y_{1,j} + (1 - \alpha) \sum_{j=1}^{n_2} \alpha_{2,j} y_{2,j} \]

Note that \( \alpha \alpha_{1,j} \geq 0, j = 1, \ldots, n_1 \), and \( (1 - \alpha) \alpha_{2,j} \geq 0, j = 1, \ldots, n_2 \), and \( \sum_{j=1}^{n_1} \alpha_{1,j} + \sum_{j=1}^{n_2} (1 - \alpha) \alpha_{2,j} = \alpha \sum_{j=1}^{n_1} \alpha_{1,j} + (1 - \alpha) \sum_{j=1}^{n_2} \alpha_{2,j} = \alpha + (1 - \alpha) = 1 \).

Hence, \( \alpha x_1 + (1 - \alpha) x_2 \in K \). This shows that \( K \) is convex. Clearly, \( S \subseteq K \).

On the other hand, fix any convex set \( G \) in the vector space, satisfying \( S \subseteq G \). \( \forall \beta \in K \), by Claim 6.43.1, \( \beta \in G \) since \( S \subseteq G \). Then, \( K \subseteq G \).

The above implies that \( K \) is the smallest convex set containing \( S \). Hence, \( K = \text{co}(S) \). This completes the proof of the proposition. \( \square \)

**Definition 6.44** Let \( (\mathcal{X}, \mathbb{K}) \) be a vector space and \( C \subseteq \mathcal{X} \). \( C \) is said to be a cone with vertex at origin if \( \vartheta \in C \) and, \( \forall x \in C, \forall \alpha \in [0, \infty) \subset \mathbb{R} \), we have \( \alpha x \in C \). \( C \) is said to be a cone with vertex \( p \in \mathcal{X} \) if \( C = p + D \), where \( D \) is a cone with vertex at origin. \( C \) is said to be a conic segment if \( \vartheta \in C \) and, \( \forall x \in C, \forall \alpha \in [0, 1] \subset \mathbb{R} \), we have \( \alpha x \in C \).

If vertex is not explicitly mentioned, it is assumed to be at origin.

Convex cones: arises in connection with positive vectors. In \( \mathbb{R}^n \) with \( n \in \mathbb{N} \), the positive cone may be defined as

\[ P = \{ x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n \mid \xi_i \geq 0, \ i = 1, \ldots, n \} \]
6.8 Linear Independence and Dimensions

Definition 6.45 Let \((\mathcal{X}, \mathcal{F})\) be a vector space, \(x \in \mathcal{X}\), and \(S \subseteq \mathcal{X}\). The vector \(x\) is said to be linearly dependent upon \(S\) if \(x \in \text{span}(S)\). Otherwise, \(x\) is said to be linearly independent of \(S\). \(S\) is said to be a linearly independent set if, \(\forall y \in S\), \(y\) is linearly independent of \(S \setminus \{y\}\).

Note that \(\emptyset\) is a linearly independent set; \(\{x\}\) is a linearly independent set if, and only if, \(x \neq \emptyset\); and \(\{x_1, x_2\}\) is a linearly independent set if, and only if, \(x_1\) and \(x_2\) do not lie on a common line through the origin.

Theorem 6.46 Let \(\mathcal{X}\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\) and \(S \subseteq \mathcal{X}\). Then, \(S\) is a linearly independent set if, and only if, \(\forall n \in \mathbb{N}, \forall \alpha_1, \ldots, \alpha_n \in \mathcal{F}, \forall x_1, \ldots, x_n \in \mathcal{X}\) which are distinct, we have \(\sum_{i=1}^{n} \alpha_i x_i = \emptyset\) implies that \(\alpha_1 = 0, i = 1, \ldots, n\).

Proof “Sufficiency” We will prove it using an argument of contradiction. Suppose \(S\) is not a linearly independent set. Then, \(\exists y \in S\) such that \(y\) is linearly dependent upon \(S \setminus \{y\}\). So, \(y \in \text{span}(S \setminus \{y\})\). \(\exists n \in \mathbb{N}, \exists x_2, \ldots, x_n \in S \setminus \{y\}\), and \(\exists \alpha_2, \ldots, \alpha_n \in \mathcal{F}\) such that \(y = \sum_{i=2}^{n} \alpha_i x_i\) (when \(n = 1\), then \(y = \emptyset\)). Without loss of generality, we may assume that \(x_2, \ldots, x_n\) are distinct. Let \(x_1 = y\) and \(\alpha_1 = -1 \neq 0\). Then, we have \(\sum_{i=1}^{n} \alpha_i x_i = \emptyset\) with \(\alpha_1 \neq 0\) and \(x_1, \ldots, x_n\) are distinct. This is a contradiction. Hence, the sufficiency result holds.

“Necessity” We again prove this by an argument of contradiction. Suppose the result does not hold. \(\exists n \in \mathbb{N}, \exists \alpha_1, \ldots, \alpha_n \in \mathcal{F}\), and \(\exists x_1, \ldots, x_n \in S\) which are distinct such that \(\sum_{i=1}^{n} \alpha_i x_i = \emptyset\) and \(\exists i_0 \in \{1, \ldots, n\}\) such that \(\alpha_{i_0} \neq 0\). Without loss of generality, we may assume \(i_0 = 1\). Then, we have \(\alpha_1 x_1 = -\sum_{i=2}^{n} \alpha_i x_i\). Hence, \(x_1 = \sum_{i=2}^{n} (-\alpha_{i-1}^{-1} \alpha_i) x_i\). Note that \(x_1 \neq x_i\) implies that \(x_i \in S \setminus \{x_1\}, i = 2, \ldots, n\). Hence, \(x_1\) is linearly dependent upon \(S \setminus \{x_1\}\). This is a contradiction. Then, the necessity result holds.

This completes the proof of this theorem. \(\Box\)

Corollary 6.47 Let \(\mathcal{X}\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\), \(\{x_1, \ldots, x_n\} \subseteq \mathcal{X}\) be a linearly independent set, \(x_i\)'s are distinct, \(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathcal{F}\), and \(n \in \mathbb{Z}_+\). Assume that \(\sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} \beta_i x_i\). Then, \(\alpha_i = \beta_i, i = 1, \ldots, n\).

Proof The assumption implies \(\sum_{i=1}^{n} (\alpha_i - \beta_i) x_i = \emptyset\). When \(n = 0\), clearly the result holds. When \(n \in \mathbb{N}\), by Theorem 6.46, we have \(\alpha_i - \beta_i = 0, i = 1, \ldots, n\). This completes the proof of the corollary. \(\Box\)

Definition 6.48 Let \(\mathcal{X}\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\), \(n \in \mathbb{Z}_+, x_1, \ldots, x_n \in \mathcal{X}\). The vectors \(x_1, \ldots, x_n\) are linearly independent if, \(\forall \alpha_1, \ldots, \alpha_n \in \mathcal{F}\), \(\sum_{i=1}^{n} \alpha_i x_i = \emptyset\) implies \(\alpha_1 = \cdots = \alpha_n = 0\). Otherwise, these vectors are said to be linearly dependent.
Lemma 6.49 Let $\mathcal{X}$ be a vector space over the field $\mathcal{F} := (F, +, \cdot, 0, 1)$. Then,

1. $x_1, \ldots, x_n \in \mathcal{X}$ are linearly independent, where $n \in \mathbb{Z}_+$, if, and only if, they are distinct and the set $\{x_1, \ldots, x_n\}$ is a linearly independent set.

2. $S \subseteq \mathcal{X}$ is a linearly independent set if, and only if, $\forall n \in \mathbb{N}$, $\forall x_1, \ldots, x_n \in S$ which are distinct implies that $x_1, \ldots, x_n$ are linearly independent.

Proof 1. “Sufficiency” When $n = 0$, $x_1, \ldots, x_n$ are linearly independent and $\emptyset$ is a linearly independent set. Hence, the result holds. When $n \in \mathbb{N}$, this is straightforward from Theorem 6.46. “Necessity” We will prove this using an argument of contradiction. Suppose the result does not hold. We will distinguish two exhaustive cases: Case 1: $x_1, \ldots, x_n$ are not distinct; Case 2: the set $\{x_1, \ldots, x_n\}$ is not a linearly independent set. Case 1. Without loss of generality, assume $x_1 = x_2$. Set $\alpha_1 = 1$, $\alpha_2 = -1$ and the rest of $\alpha_i$’s to 0. Then, we have $\sum_{i=1}^{n} \alpha_i x_i = 0$ and hence $x_1, \ldots, x_n$ is linearly dependent. This is a contradiction. Case 2. By Theorem 6.46, $\exists m \in \mathbb{N}, \exists \alpha_1, \ldots, \alpha_m \in \mathcal{F}$ which are not all 0’s, $\exists y_1, \ldots, y_m \in \{x_1, \ldots, x_n\}$ which are distinct, such that $\sum_{i=1}^{m} \alpha_i y_i = 0$. Clearly, $m \leq n$, otherwise $y_i$’s are not distinct. Without loss of generality, assume $y_1 = x_1, \ldots, y_m = x_m$. Set $\alpha_{m+1} = \cdots = \alpha_n = 0$. Then, we have $\sum_{i=1}^{n} \alpha_i x_i = 0$ with $\alpha_1, \ldots, \alpha_n$ not all 0’s. This is a contradiction. Thus, we have arrived at a contradiction in every case. Hence, the necessity result holds.

2. This is straightforward from Theorem 6.46. This completes the proof of the lemma.

Definition 6.50 Let $(\mathcal{X}, \mathcal{F})$ be a vector space and $S \subseteq \mathcal{X}$ be a linearly independent set with $n \in \mathbb{Z}_+$ elements. $S$ is said to be a basis of the vector space if $\text{span}(S) = \mathcal{X}$. In this case, the vector space is said to be finite dimensional with dimension $n$. All other vector spaces are said to be infinite dimensional.

Theorem 6.51 Let $\mathcal{X}$ be a finite dimensional vector space over the field $\mathcal{F} := (F, +, \cdot, 0, 1)$. Then, the dimension $n \in \{0\} \cup \mathbb{N}$ is unique. Furthermore, any linearly independent set of $n$ vectors form a basis of the vector space.

Proof Let $n \in \mathbb{Z}_+$ be the minimum of dimensions for $\mathcal{X}$. Then, there exists a set $S_1 \subseteq \mathcal{X}$ with $n$ elements such that $S_1$ is a basis for the vector space. We will show that $\forall y_1, \ldots, y_m \in \mathcal{X}$ with $m > n$, then $y_1, \ldots, y_m$ are linearly dependent. This implies that any subset with more than $n$ elements cannot be a linearly independent set by Lemma 6.49. Henceforth, the dimension of the vector space is unique.
∀y₁, ..., yₘ ∈ X with m > n. There are two exhaustive and mutually exclusive cases: Case 1: n = 0; Case 2: n > 0. Then, S₁ = ∅ and X contains a single vector ϑ. Hence, y₁ = ⋯ = yₘ = ϑ. Clearly, y₁, ..., yₘ are linearly dependent. This case is proven. Case 2: n > 0. Take S₁ = {x₁, ..., xₙ}. Since S₁ is a basis for the vector space, then, ∀i ∈ {1, ..., m}, yᵢ ∈ span(S₁), that is, ∃αᵢ𝑗 ∈ F, j = 1, ..., n, such that yᵢ = ∑ᵢ₌₁⁻ⁿ αᵢ𝑗xⱼ. Consider the m × n-dimensional matrix A := (αᵢⱼ)ₘ×ₙ. Since m > n, then rank(A) < m, which implies that the row vectors of A are linearly dependent. ∃β₁, ..., βₙ ∈ F, which are not all 0’s, such that ∑ᵢ₌₁ᵐ βᵢαᵢⱼ = 0, j = 1, ..., n. This implies ∑ᵢ₌₁ᵐ βᵢyᵢ = ∑ᵢ₌₁ᵐ ∑ⱼ₌₁ⁿ βᵢαᵢⱼxⱼ = ∑ⱼ₌₁ⁿ(∑ᵢ₌₁ᵐ βᵢαᵢⱼ)xⱼ = ϑ. Hence, y₁, ..., yₘ are linearly dependent. This case is also proven.

Hence, the dimension of the vector space is unique.

∀S₂ ⊆ X with n vectors that is a linearly independent set, we will show that S₂ is a basis for the vector space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: n = 0; Case 2: n > 0. Case 1: n = 0. Then, S₂ = ∅ and X is a singleton set consisting of the null vector, which equals to span(S₂). Hence, S₂ is a basis for the vector space. Case 2: n > 0. Take S₂ = {z₁, ..., zₙ}. ∀x ∈ X, x, z₁, ..., zₙ (for a total of n + 1 vectors), by the preceding proof, are linearly dependent. Then, ∃α, β₁, ..., βₙ ∈ F, which are not all 0’s, such that αx + ∑ᵢ₌₁ⁿ βᵢzᵢ = ϑ. Suppose α = 0. Then, ∑ᵢ₌₁ⁿ βᵢzᵢ = ϑ. Since S₂ is a linearly independent set, by Theorem 6.46, then β₁ = ⋯ = βₙ = 0. This contradicts with the fact that α, β₁, ..., βₙ are not all 0’s. Hence, α ≠ 0. Then, x = ∑ᵢ₌₁ⁿ(−α⁻¹βᵢ)zᵢ and x ∈ span(S₂). Therefore, we have X = span(S₂). Hence, S₂ is a basis of the vector space.

This completes the proof of the theorem. □

Finite-dimensional spaces are simpler to analyze. Many results of finite-dimensional spaces may be generalized to infinite-dimensional spaces. We endeavor to stress the similarity between the finite- and infinite-dimensional spaces.
Chapter 7

Banach Spaces

7.1 Normed Linear Spaces

Vector spaces admit algebraic properties, but they lack topological properties.

Denote \( K \) to be either \( \mathbb{R} \) or \( \mathbb{C} \).

**Definition 7.1** Let \((X, K)\) be a vector space. A **norm** on the vector space is a real-valued function \( \| \cdot \| : X \to [0, \infty) \subset \mathbb{R} \) that satisfies the following properties: \( \forall x, y \in X, \forall \alpha \in K, \)

\[
\begin{align*}
(i) & \quad 0 \leq \| x \| < +\infty \text{ and } \| x \| = 0 \iff x = \vartheta; \\
(ii) & \quad \| x + y \| \leq \| x \| + \| y \|; \quad \text{(triangle inequality)} \\
(iii) & \quad \| \alpha x \| = |\alpha| \| x \| .
\end{align*}
\]

A **real (complex) normed linear space** is a vector space over the field \( \mathbb{R} \) (or \( \mathbb{C} \)) together with a norm defined on it. A normed linear space consisting of the triple \((X, K, \| \cdot \|)\).

To simplify notation in the theory, we may later simply discuss a normed linear space \( X := (X, K, \| \cdot \|) \) without further reference to components of \( X \), where the operations are understood to be \( \oplus_X \) and \( \otimes_X \), the null vector is understood to be \( \vartheta_X \), and the norm is understood to be \( \| \cdot \|_X \). When it is clear from the context, we will neglect the subscript \( X \).

**Example 7.2** Let \( n \in \mathbb{Z}_+ \). The Euclidean space \((\mathbb{R}^n, \mathbb{R}, |\cdot|)\) is a normed linear space, where the norm is defined by \(|(\xi_1, \ldots, \xi_n)| = \sqrt{\xi_1^2 + \cdots + \xi_n^2}, \forall(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n\). Note that we specifically denote the Euclidean norm as \(|\cdot|\), rather than \(\| \cdot \|\), to distinguish it from other norms on \(\mathbb{R}^n\).

**Example 7.3** Let \( n \in \mathbb{Z}_+ \). The space \((\mathbb{C}^n, \mathbb{C}, |\cdot|)\) is a normed linear space, where the norm is defined by \(|(\xi_1, \ldots, \xi_n)| = \sqrt{\sum_{i=1}^{n} |\xi_i|^2}, \forall(\xi_1, \ldots, \xi_n) \in \mathbb{C}^n\).
\(C^n\). Note that we specifically denote the norm as \(| \cdot |\), rather than \(\| \cdot \|\), to distinguish it from other norms on \(C^n\).

**Example 7.4** Let \(X := \{ f : [a, b] \to \mathbb{R} \}\), where \(a, b \in \mathbb{R}\) with \(a \leq b\), \(\oplus\) and \(\otimes\) be the usual vector addition and scalar multiplication, and \(\vartheta : [a, b] \to \mathbb{R}\) be given by \(\vartheta(t) = 0\), \(\forall t \in [a, b]\). By Example 6.20, \(X := (\mathbb{R}, \oplus, \otimes, \vartheta)\) is a vector space over \(\mathbb{R}\). Let \(M := \{ f \in X \mid f\) is continuous\} is a subspace by Proposition 6.25. Let \(\mathcal{M} := (M, \oplus, \otimes, \vartheta)\). Then, \((\mathcal{M}, \mathbb{R})\) is a vector space. Introduce a norm on this space by \(\| x \| = \max_{a \leq t \leq b} \{ | x(t) |, \forall x \in \mathcal{M}\). Now, we verify the properties of the norm. \(\forall x, y \in \mathcal{M}, \forall \alpha \in \mathbb{R}\).

(i) Since \([a, b]\) is a nonempty compact set and \(x\) is continuous, then, \(\| x \| = | x(t) | \in [0, \infty) \subset \mathbb{R}\), for some \(t \in [a, b]\). Clearly, \(\| x \| = 0 \Leftrightarrow x = \vartheta\).

(ii) \(\| x + y \| = \max_{a \leq t \leq b} \{ | x(t) + y(t) | \} \leq \max_{a \leq t \leq b} \{ | x(t) | + | y(t) | \} = \max_{a \leq t \leq b} \{ | x(t) | \} + \max_{a \leq t \leq b} \{ | y(t) | \} = \| x \| + \| y \|\).

(iii) \(\| \alpha x \| = \max_{a \leq t \leq b} \{ | \alpha x(t) | \} = | \alpha | \max_{a \leq t \leq b} \{ | x(t) | \} = | \alpha | \| x \|\), where have made use of Proposition 3.81 and the fact that \(\| x \| \in \mathbb{R}\) in the third equality.

Hence, \(C([a, b]) := (\mathcal{M}, \mathbb{R}, \| \cdot \|)\) is a normed linear space.

**Example 7.5** Let \(X := \{ f : [a, b] \to \mathbb{R} \}\), where \(a, b \in \mathbb{R}\) with \(a < b\), \(\oplus\) and \(\otimes\) be the usual vector addition and scalar multiplication, and \(\vartheta : [a, b] \to \mathbb{R}\) be given by \(\vartheta(t) = 0\), \(\forall t \in [a, b]\). By Example 6.20, \(X := (\mathbb{R}, \oplus, \otimes, \vartheta)\) is a vector space over \(\mathbb{R}\). Let \(M := \{ f \in X \mid f\) is continuously differentiable\} is a subspace by Proposition 6.25. Let \(\mathcal{M} := (M, \oplus, \otimes, \vartheta)\). Then, \((\mathcal{M}, \mathbb{R})\) is a vector space. Introduce a norm on this space by \(\| x \| = \max_{a \leq t \leq b} \{ | x(t) | + \max_{a \leq t \leq b} \{ | x^{(1)}(t) | \} \}, \forall x \in \mathcal{M}\).

Now, we verify the properties of the norm. \(\forall x, y \in \mathcal{M}, \forall \alpha \in \mathbb{R}\).

(i) Since \([a, b]\) is a nonempty compact set and \(x\) is continuously differentiable, then, \(\| x \| = | x(t_1) | + | x^{(1)}(t_2) | \in [0, \infty) \subset \mathbb{R}\), for some \(t_1, t_2 \in [a, b]\). Clearly, \(\| x \| = 0 \Leftrightarrow x = \vartheta\).

(ii) \(\| x + y \| = \max_{a \leq t \leq b} \{ | x(t) + y(t) | + \max_{a \leq t \leq b} \{ | x^{(1)}(t) + y^{(1)}(t) | \} \leq \max_{a \leq t \leq b} \{ | x(t) | + \max_{a \leq t \leq b} \{ | y(t) | + \max_{a \leq t \leq b} \{ | x^{(1)}(t) | + \max_{a \leq t \leq b} \{ | y^{(1)}(t) | \} = \| x \| + \| y \|\).

(iii) \(\| \alpha x \| = \max_{a \leq t \leq b} \{ | \alpha x(t) | \} + \max_{a \leq t \leq b} \{ | \alpha x^{(1)}(t) | \} = | \alpha | \| x \|\).

Hence, \(C_1([a, b]) := (\mathcal{M}, \mathbb{R}, \| \cdot \|)\) is a normed linear space.

**Example 7.6** Let \(X := \{ (\xi_k)_{k=1}^\infty \mid \xi_k \in \mathbb{R}, \forall k \in \mathbb{N} \}\), \(\oplus\) and \(\otimes\) be the usual vector addition and scalar multiplication, and \(\vartheta := (0, 0, \ldots) \in X\). By Example 6.23, \(X := (\mathbb{R}, \oplus, \otimes, \vartheta)\) is a vector space over \(\mathbb{R}\). For \(p \in [1, \infty) \subset \mathbb{R}\).
7.1. NORMED LINEAR SPACES

Let \( M_p := \{ (\xi_k)_{k=1}^{\infty} \in X \mid \sum_{k=1}^{\infty} |\xi_k|^p < +\infty \} \). Define the norm \( \| x \|_p = \left( \sum_{k=1}^{\infty} |\xi_k|^p \right)^{1/p}, \forall x = (\xi_k)_{k=1}^{\infty} \in M_p \). Let \( M_\infty := \{ (\xi_k)_{k=1}^{\infty} \in X \mid \sup_{k \geq 1} |\xi_k|, \forall x = (\xi_k)_{k=1}^{\infty} \in M_\infty \}. \) Define the norm \( \| x \|_\infty = \sup_{k \geq 1} |\xi_k|, \forall x = (\xi_k)_{k=1}^{\infty} \in M_\infty \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: Equality holds if, and only if, \( a \beta y \leq (a \alpha x + b \beta y)_{p} \leq (|a| x + |\beta| y)_{p} < +\infty \). Then, we have \( \| a \alpha x + b \beta y \|_p = |a| \| x \|_p + |\beta| \| y \|_p < +\infty \). This implies that \( a \alpha x + b \beta y \in M_p \). Hence, \( M_p \) is a subspace in \( (X, \mathbb{R}) \) and \( l_p := (M_p, \mathbb{R}, \| \cdot \|_p) \) is a normed linear space.

Lemma 7.7 \forall a, b \in [0, \infty) \subset \mathbb{R}, \forall \lambda \in (0, 1) \subset \mathbb{R}, we have \[
\alpha^{\lambda}b^{1-\lambda} \leq \lambda a + (1-\lambda)b
\]
where equality holds if, and only if, \( a = b \).

Proof Define \( f : [0, \infty) \to \mathbb{R} \) by \( f(t) = t^\lambda - \lambda t + \lambda - 1, \forall t \in [0, \infty) \subset \mathbb{R} \). Then, \( f \) is continuous, and is differentiable on \( (0, \infty) \). \( f(1) = \lambda t^\lambda - \lambda, \forall t \in (0, \infty) \). Then, \( f^{(1)}(t) > 0, \forall t \in (0, 1) \) and \( f^{(1)}(t) < 0, \forall t \in (1, \infty) \). Then, \( f(t) \leq f(1) = 0, \forall t \in [0, \infty) \), where equality holds if, and only if, \( t = 1 \).

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( b = 0 \); Case 2: \( b > 0 \). Case 1: \( b = 0 \). We have \( a \beta y \leq (a \alpha x + b \beta y)_{p} \leq (|a| x + |\beta| y)_{p} < +\infty \). Then, we immediately obtain the desired inequality, where equality holds if, and only if, \( a/b = 1 \Leftrightarrow a = b \). This case is also proved.

This completes the proof of the lemma.

\[\square\]

Theorem 7.8 (Hölder’s Inequality) Let \( p \in [1, +\infty) \subset \mathbb{R} \) and \( q \in (1, +\infty) \subset \mathbb{R}_+ \) with \( 1/p + 1/q = 1 \). Then, \( \forall x = (\xi_k)_{k=1}^{\infty} \in l_p, \forall y = (\eta_k)_{k=1}^{\infty} \in l_q, \) we have
\[
\sum_{k=1}^{\infty} |\xi_k\eta_k| \leq \| x \|_p \| y \|_q
\]
When \( q < \infty \), equality holds if, and only if, \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( a |\xi_k|^p = \beta |\eta_k|^q \), \( k = 1, 2, \ldots \). When \( q = \infty \), equality holds if, and only if, \( |\eta_k| = \| y \|_\infty \) for any \( k \in \mathbb{N} \) with \( |\xi_k| > 0 \).

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( q = \infty \); Case 2: \( 1 < q < +\infty \). Case 1: \( q = \infty \). Then, \( p = 1 \). We have
\[
\sum_{k=1}^{\infty} |\xi_k\eta_k| = \sum_{k=1}^{\infty} |\xi_k| \| \eta_k \| \leq \sum_{k=1}^{\infty} |\xi_k| \| y \|_\infty = \| x \|_1 \| y \|_\infty
\]
where equality holds if, and only if, \( |\eta_k| = \|y\|_\infty \) for any \( k \in \mathbb{N} \) with \( |\xi_k| > 0 \). This case is proved.

Case 2: \( 1 < q < +\infty \). Then, \( 1 < p < +\infty \). We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: \( \|x\|_p \|y\|_q = 0 \); Case 2b: \( \|x\|_p \|y\|_q > 0 \). Case 2a: \( \|x\|_p \|y\|_q = 0 \). Without loss of generality, assume \( \|y\|_q = 0 \). Then, \( y = 0 \) and

\[
\sum_{k=1}^{\infty} |\xi_k \eta_k| = 0 = \|x\|_p \|y\|_q
\]

Equality holds \( \Rightarrow \alpha = 0 \) and \( \beta = 1 \) and \( \alpha \|\xi_k\|^p = \beta |\eta_k|^q \), \( k = 1, 2, \ldots \). This subcase is proved. Case 2b: \( \|x\|_p \|y\|_q > 0 \). Then, \( \|x\|_p > 0 \) and \( \|y\|_q > 0 \). \( \forall k \in \mathbb{N} \), by Lemma 7.7 with \( a = \left( \frac{|\xi_k|}{\|x\|_p} \right)^p \), \( b = \left( \frac{|\eta_k|}{\|y\|_q} \right)^q \), and \( \lambda = 1/p \), we have

\[
\frac{|\xi_k \eta_k|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} \frac{|\xi_k|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|\eta_k|^q}{\|y\|_q^q}
\]

with equality if, and only if, \( \frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q} \). Summing the above inequality for all \( k \in \mathbb{N} \), we have

\[
\sum_{k=1}^{\infty} \frac{|\xi_k \eta_k|}{\|x\|_p \|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1
\]

with equality if, and only if, \( \frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q} \), \( k = 1, 2, \ldots \). Equality \( \Rightarrow \alpha = 1/\|x\|_p^p \) and \( \beta = 1/\|y\|_q^q \) and \( \alpha \|\xi_k\|^p = \beta |\eta_k|^q \), \( k = 1, 2, \ldots \). On the other hand, if \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( \alpha \|\xi_k\|^p = \beta |\eta_k|^q \), \( k = 1, 2, \ldots \), then, without loss of generality, assume \( \beta \neq 0 \). Let \( \alpha_1 = \alpha/\beta \). Then, \( \alpha_1 \|\xi_k\|^p = |\eta_k|^q \), \( k = 1, 2, \ldots \). Hence, \( \alpha_1 \|x\|_p^p = \|y\|_q^q \), which further implies that \( \alpha_1 = \|y\|_q^q / \|x\|_p^p \). Hence, \( \frac{|\xi_k|^p}{\|x\|_p^p} = \frac{|\eta_k|^q}{\|y\|_q^q} \), \( k = 1, 2, \ldots \). This implies equality. Therefore, equality if, and only if, \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( \alpha \|\xi_k\|^p = \beta |\eta_k|^q \), \( k = 1, 2, \ldots \). This subcase is proved.

This completes the proof of the theorem. \( \square \)

When \( p = 2 = q \), the Hölder’s inequality becomes the well-known Cauchy-Schwarz Inequality:

\[
\sum_{k=1}^{\infty} |\xi_k \eta_k| \leq \left( \sum_{k=1}^{\infty} |\xi_k|^2 \right)^{1/2} \left( \sum_{k=1}^{\infty} |\eta_k|^2 \right)^{1/2}
\]
7.1. NORMED LINEAR SPACES

Theorem 7.9 (Minkowski’s Inequality) Let \( p \in [1, \infty] \subset \mathbb{R} \). \( \forall x, y \in l_p \), then, \( x + y \in l_p \) and

\[
\|x + y\|_p \leq \|x\|_p + \|y\|_p
\]

When \( 1 < p < \infty \), equality holds if, and only if, \( \exists \alpha, \beta \in [0, \infty) \subset \mathbb{R} \), which are not both zeros, such that \( \alpha x = \beta y \).

Proof \( \forall x = (\xi_k)_{k=1}^\infty, y = (\eta_k)_{k=1}^\infty \subset l_p \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( p = 1 \); Case 2: \( p = \infty \); Case 3: \( 1 < p < \infty \).

Case 1: \( p = 1 \). \( \|x + y\|_1 = \sum_{k=1}^\infty |\xi_k + \eta_k| \leq \sum_{k=1}^\infty (|\xi_k| + |\eta_k|) = \|x\|_1 + \|y\|_1 \). Equality holds if, and only if, \( \xi_k, \eta_k \geq 0 \), \( k = 1, 2, \ldots \). Hence, \( x + y \in l_1 \).

Case 2: \( p = \infty \). \( \|x + y\|_\infty = \sup_{k \in \mathbb{N}} |\xi_k + \eta_k| \leq \sup_{k \in \mathbb{N}} (|\xi_k| + |\eta_k|) \leq \sup_{k \in \mathbb{N}} |\xi_k| + \sup_{k \in \mathbb{N}} |\eta_k| = \|x\|_\infty + \|y\|_\infty < +\infty \). Hence, \( x + y \in l_\infty \).

Case 3: \( 1 < p < \infty \). We will further distinguish two exhaustive and mutually exclusive cases: Case 3a: \( \|x\|_p, \|y\|_p = 0 \); Case 3b: \( \|x\|_p, \|y\|_p > 0 \). Without loss of generality, assume \( \|y\|_p = 0 \). Then, \( y = \emptyset \) and \( \eta_k = 0 \), \( k = 1, 2, \ldots \). Hence, \( x + y = x \in l_p \) and \( \|x + y\|_p = \|x\|_p = \|x\|_p + \|y\|_p \). Let \( \alpha = 0 \) and \( \beta = 1 \), we have \( \alpha x = \beta y \).

Hence, equality holds if, and only if, \( \exists \alpha, \beta \in [0, \infty) \subset \mathbb{R} \), which are not both zeros, such that \( \alpha x = \beta y \). This subcase is proved.

Case 3b: \( \|x\|_p, \|y\|_p > 0 \). Then, \( \|x\|_p > 0 \) and \( \|y\|_p > 0 \). Let \( \lambda = \frac{\|x\|_p}{\|x\|_p + \|y\|_p} \in (0, 1) \subset \mathbb{R} \). \( \forall k \in \mathbb{N} \).

\[
\frac{|\xi_k + \eta_k|^p}{(||x||_p + ||y||_p)^p} \leq \left( \frac{|\xi_k| + |\eta_k|}{||x||_p + ||y||_p} \right)^p = \left( \frac{|\xi_k|}{||x||_p} + (1 - \lambda) \frac{|\eta_k|}{||y||_p} \right)^p
\]

Since \( 1 < p < \infty \), then the function \( t^p \) is strictly convex on \( [0, \infty) \subset \mathbb{R} \).

Then, we have

\[
\frac{|\xi_k + \eta_k|^p}{(||x||_p + ||y||_p)^p} \leq \frac{\lambda |\xi_k|^p}{||x||_p^p} + (1 - \lambda) \frac{\eta_k|^p}{||y||_p^p}
\]

where equality holds \( \iff \xi_k, \eta_k \geq 0 \) and \( \frac{|\xi_k|}{||x||_p} = \frac{|\eta_k|}{||y||_p} \iff \frac{\xi_k}{||x||_p} = \frac{\eta_k}{||y||_p} \).

Summing the above inequalities for all \( k \in \mathbb{N} \), we have

\[
\frac{\|x + y\|_p^p}{(||x||_p + ||y||_p)^p} = \sum_{k=1}^\infty \frac{|\xi_k + \eta_k|^p}{(||x||_p + ||y||_p)^p} \leq \lambda \sum_{k=1}^\infty \frac{|\xi_k|^p}{||x||_p^p} + (1 - \lambda) \sum_{k=1}^\infty \frac{|\eta_k|^p}{||y||_p^p} = 1
\]
Therefore,
\[\|x + y\|_p \leq \|x\|_p + \|y\|_p < +\infty\]
where equality holds \(\Longleftrightarrow \frac{\xi_k}{\|x\|_p} = \frac{\eta_k}{\|y\|_p}, \ k = 1, 2, \ldots \ \Longleftrightarrow (1/\|x\|_p)x = (1/\|y\|_p)y\). Equality implies that \(\alpha = 1/\|x\|_p\) and \(\beta = 1/\|y\|_p\) and \(\alpha x = \beta y\). On the other hand, if \(\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}\), which are not both zeros, such that \(\alpha x = \beta y\), then, without loss of generality, assume \(\beta \neq 0\).

Then, \(y = \alpha_1 x\) with \(\alpha_1 = \alpha/\beta\). Easy to show that \(\|y\|_p = \alpha_1 \|x\|_p\). Then, \(\alpha_1 = \|y\|_p/\|x\|_p\). This implies that \((1/\|x\|_p)x = (1/\|y\|_p)y\) which further implies equality. Hence, equality holds if, and only if, \(\exists \alpha, \beta \in [0, \infty),\) which are not both zeros, such that \(\alpha x = \beta y\). Clearly, \(x + y \in l_p\). This subcase is proved.

This completes the proof of the theorem. \(\square\)

**Example 7.10** Let \(X\) be a real (complex) normed linear space. Let \(Y := \{(\xi_k)_{k=1}^{\infty} \mid \xi_k \in X, \forall k \in \mathbb{N}\}\). By Example 6.20, \((Y, \oplus, \otimes, \vartheta)\) is a vector space over \(\mathbb{K}\), where \(\oplus, \otimes,\) and \(\vartheta\) are defined as in the example.

For \(p \in [1, \infty) \subset \mathbb{R}\), let \(M_p := \{(\xi_k)_{k=1}^{\infty} \mid \sum_{k=1}^{\infty} \|\xi_k\|^p_X < +\infty\}\). Define the norm \(\|x\|_p = (\sum_{k=1}^{\infty} \|\xi_k\|^p_X)^{1/p}, \forall x = (\xi_k)_{k=1}^{\infty} \in M_p\). Let \(M_\infty := \{(\xi_k)_{k=1}^{\infty} \mid \sup_{k \geq 1} \|\xi_k\|_X < +\infty\}\). Define the norm \(\|x\|_\infty = \sup_{k \geq 1} \|\xi_k\|_X, \forall x = (\xi_k)_{k=1}^{\infty} \in M_\infty\).

\[\|x + y\|_p = \begin{cases} \left(\sum_{k=1}^{\infty} \|\xi_k + \eta_k\|^p_X\right)^{1/p} \leq \left(\sum_{k=1}^{\infty} (\|\xi_k\|_X + \|\eta_k\|_X)^p\right)^{1/p}, & p \in [1, \infty) \\
\sup_{k \geq 1} \|\xi_k + \eta_k\|_X \leq \sup_{k \geq 1} (\|\xi_k\|_X + \|\eta_k\|_X) & p = \infty \end{cases}\]

\[\leq \begin{cases} \left(\sum_{k=1}^{\infty} \|\xi_k\|^p_X\right)^{1/p} + \left(\sum_{k=1}^{\infty} \|\eta_k\|^p_X\right)^{1/p}, & p \in [1, \infty) \\
\sup_{k \geq 1} \|\xi_k\|_X + \sup_{k \geq 1} \|\eta_k\|_X & p = \infty \end{cases}\]

where we have made use of the Minkowski’s Inequality. In the preceding inequality, when \(1 < p < \infty\), equality holds if, and only if, \(\|\xi_k + \eta_k\|_X = \|\xi_k\|_X + \|\eta_k\|_X, \forall k \in \mathbb{N}\), and \(\exists \alpha, \beta \in [0, \infty) \subset \mathbb{R}\), which are not both zeros, such that \(\alpha \|\xi_k\|_X = \beta \|\eta_k\|_X, \forall k \in \mathbb{N}\).

\(\forall p \in [1, \infty] \subset \mathbb{R}\). Note that \(\vartheta = (\vartheta_X, \vartheta_X, \ldots) \in M_p \neq \emptyset\). \(\forall x = (\xi_k)_{k=1}^{\infty} \in M_p, \forall \alpha \in \mathbb{K}\), we have

\[\|\alpha x\|_p = \begin{cases} \left(\sum_{k=1}^{\infty} \|\alpha \xi_k\|^p_X\right)^{1/p} = \left(\sum_{k=1}^{\infty} |\alpha|^p \|\xi_k\|^p_X\right)^{1/p}, & p \in [1, \infty) \\
\sup_{k \geq 1} \|\alpha \xi_k\|_X = \sup_{k \geq 1} |\alpha| \|\xi_k\|_X & p = \infty \end{cases}\]
Example 7.11 Let \( X, y \in M_p \), \( \forall \alpha, \beta \in \mathbb{K} \), we have \( \| \alpha x + \beta y \|_p \leq \| \alpha x \|_p + \| \beta y \|_p = |\alpha| \| x \|_p + |\beta| \| y \|_p < +\infty \). Then, \( \alpha x + \beta y \in M_p \). Hence, \( \mathcal{M}_p := (M_p, \oplus, \odot, \vartheta) \) is a subspace in \( (Y, \oplus, \odot, \vartheta) \).

It is easy to check that, \( \forall x \in \mathcal{M}_p \), \( \| x \|_p \in [0, \infty) \subset \mathbb{R} \), and \( \| x \|_p = 0 \) \( \iff \) \( x = (\vartheta \chi, \vartheta \chi, \ldots) = \vartheta \). Therefore, \( l_p(X) := (\mathcal{M}_p, \mathcal{K}, \| \cdot \|_p) \) is a real (complex) normed linear space, \( \forall p \in [1, \infty) \subset \mathbb{R}_e \).

\[ \| \cdot \|_p = \begin{cases} |\alpha| \left( \sum_{k=1}^{\infty} \| \xi_k \|_X^p \right)^{1/p} & p \in [1, \infty) \\ |\alpha| \sup_{k \geq 1} \| \xi_k \|_X & p = \infty \end{cases} \]

where we have made use of Proposition 3.81. Then, \( \forall x, y \in M_p \), \( \forall \alpha, \beta \in \mathbb{K} \),

\[ \alpha x + \beta y = \alpha x + \beta y \] is a subspace by Proposition 6.25. Let

\[ \mathcal{N} \subseteq X \]

be a subspace. Then, \( \mathcal{N} \) is also a normed linear space, \( \forall s \in S \).

**Definition 7.12** Let \( X \) be a set, \( Y \) be a normed linear space, \( f : X \to Y \). \( f \) is said to be bounded if \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| f(x) \| \leq M \), \( \forall x \in X \).

\( S \subseteq Y \) is said to be bounded if \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| s \| \leq M \), \( \forall s \in S \).

**Proposition 7.13** Let \( X := (\mathcal{X}, \mathcal{K}, \| \cdot \|) \) be a normed linear space and \( M \subseteq X \) be a subspace. Then, \( (\mathcal{M}, \mathcal{K}, \| \cdot \|) \) is also a normed linear space.

**Proof** This is straightforward and is therefore omitted.

### 7.2 The Natural Metric

A normed linear space is actually a metric space.
Proposition 7.14 A normed linear space \( \mathcal{X} := (\mathcal{X}, \mathbb{K}, \| \cdot \|) \) admits the natural metric \( \rho : \mathcal{X} \times \mathcal{X} \to [0, \infty) \subset \mathbb{R} \) given by \( \rho(x, y) = \|x - y\| \), \( \forall x, y \in \mathcal{X} \).

Proof \( \forall x, y, z \in \mathcal{X} \), we have (i) \( \rho(x, y) \in [0, \infty) \); (ii) \( \rho(x, y) = \|x - y\| = 0 \iff x - y = 0 \iff x = y \); (iii) \( \rho(x, y) = \|x - y\| = \|(-1)(x - y)\| = \|y - x\| = \rho(y, x) \); and (iv) Note that

\[
\rho(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\|
\]

Thus, \( \rho(x, z) = \rho(x, y) + \rho(y, z) \)

Hence, \( \rho \) is a metric. This completes the proof of the proposition. \( \Box \)

Now, we can talk about topological properties and metric properties of a normed linear space. When we refer to these properties of a normed linear space, we are referring to the above metric specifically.

Proposition 7.15 Let \( \mathcal{X} \) be a normed linear space and \( C \subseteq \mathcal{X} \) be convex. Then, \( \overline{C} \) and \( C^\circ \) are convex.

Proof \( \forall x_1, x_2 \in \overline{C}, \forall \alpha \in [0, 1] \subset \mathbb{R} \), we need to show that \( \alpha x_1 + (1 - \alpha)x_2 \in \overline{C} \). By Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R} \), \( \mathcal{B}(x_i, r) \cap C \neq \emptyset \), \( i = 1, 2 \). Let \( p_i \in \mathcal{B}(x_i, r) \cap C, i = 1, 2 \). Then, by convexity of \( C \), we have \( \alpha p_1 + (1 - \alpha)p_2 \in C \). Note that

\[
\| \alpha (x_1 + (1 - \alpha)x_2) - (\alpha p_1 + (1 - \alpha)p_2) \| = \| \alpha (x_1 - p_1) + (1 - \alpha)(x_2 - p_2) \| \leq \alpha \| x_1 - p_1 \| + (1 - \alpha) \| x_2 - p_2 \| < r
\]

Thus, \( \alpha x_1 + (1 - \alpha)x_2 \in \mathcal{B}(x_1 + (1 - \alpha)x_2, r) \cap C \neq \emptyset \). Hence, \( \alpha x_1 + (1 - \alpha)x_2 \in \overline{C} \) by the arbitrariness of \( r \) and Proposition 3.3. Then, \( \overline{C} \) is convex.

\( \forall x_1, x_2 \in C^\circ, \forall \alpha \in [0, 1] \subset \mathbb{R} \), we need to show that \( \alpha x_1 + (1 - \alpha)x_2 \in C^\circ \). \( \exists r_i \in (0, \infty) \subset \mathbb{R} \) such that \( \mathcal{B}(x_i, r_i) \subseteq C \), \( i = 1, 2 \). Let \( r := \min\{r_1, r_2\} > 0 \). \( \forall r \in \mathcal{B}(\alpha x_1 + (1 - \alpha)x_2, r) \), let \( w := p - \alpha x_1 - (1 - \alpha)x_2 \). Then, \( \| w \| < r \) and \( x_i + w \in \mathcal{B}(x_i, r_i) \subseteq C, i = 1, 2 \). By the convexity of \( C \), we have \( C \ni \alpha(x_1 + w) + (1 - \alpha)(x_2 + w) = p \). Hence, we have \( \mathcal{B}(\alpha x_1 + (1 - \alpha)x_2, r) \subseteq C \) and \( \alpha x_1 + (1 - \alpha)x_2 \in C^\circ \). Hence, \( C^\circ \) is convex.

This completes the proof of the proposition. \( \Box \)

Proposition 7.16 Let \( \mathcal{X} \) be a normed linear space, \( x_0 \in \mathcal{X}, S \subseteq \mathcal{X}, \) and \( P = x_0 + S \). Then, \( \overline{P} = x_0 + \overline{S} \) and \( P^\circ = x_0 + S^\circ \).

Proof \( \forall x \in \overline{P} \), by Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R} \), \( \exists p_0 \in P \cap \mathcal{B}(x, r) \). Then, \( p_0 - x_0 \in S \cap \mathcal{B}(x - x_0, r) \neq \emptyset \). Hence, by Proposition 3.3, \( x - x_0 \in \overline{S} \) and \( x \in x_0 + \overline{S} \). Hence, \( \overline{P} \subseteq x_0 + \overline{S} \).

On the other hand, \( \forall x \in x_0 + \overline{S} \), we have \( x - x_0 \in \overline{S} \) and, by Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R} \), \( \exists s_0 \in S \cap \mathcal{B}(x - x_0, r) \neq \emptyset \). Then, \( x_0 + s_0 \in P \cap \mathcal{B}(x, r) \neq \emptyset \). By Proposition 3.3, \( x \in \overline{P} \). Hence, \( x_0 + \overline{S} \subseteq \overline{P} \).
Therefore, we have $\overline{P} = x_0 + \overline{S}$.

\[ \forall x \in P^o, \exists r \in (0, \infty) \subset \mathbb{R} \text{ such that } B(x, r) \subseteq P. \text{ Then, } B(x, r) = B(x, r) - x_0 \subseteq P - x_0 = S. \text{ Hence, } x - x_0 \in S^o, x \in x_0 + S^o \text{ and } P^o \subseteq x_0 + S^o. \]

On the other hand, $\forall x \in x_0 + S^o$, we have $x - x_0 \in S^o$. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $B(x - x_0, r) \subseteq S$. Hence, $B(x, r) = B(x - x_0, r) + x_0 \subseteq x_0 + S = P$.

Hence, we have $P^o = x_0 + S^o$. This completes the proof of the proposition. $\square$

**Proposition 7.17** Let $X$ be a normed linear space over $\mathbb{K}$. Then, the following statements hold.

(i) If $M \subseteq X$ is a subspace, then $\overline{M}$ is a subspace.

(ii) If $V \subseteq X$ is a linear variety, then $\overline{V}$ is a linear variety.

(iii) If $C \subseteq X$ is a cone with vertex $p \in X$, then $\overline{C}$ is a cone with vertex $p$.

(iv) Let $x_0 \in X$ and $r \in (0, \infty) \subset \mathbb{R}$. Then, $\overline{B}(x_0, r) = \overline{B}(x_0, r)$.

**Proof**

(i) Clearly, we have $\emptyset \in M \subseteq \overline{M} \neq \emptyset$. $\forall x, y \in \overline{M}$, $\forall \alpha \in \mathbb{K}$, we will show that $x + y, \alpha x \in \overline{M}$. Then, $\overline{M}$ is a subspace. $\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in B(x, r/2) \cap M \neq \emptyset$ and $\exists y_0 \in B(y, r/2) \cap M \neq \emptyset$.

Since $M$ is a subspace, then $x_0 + y_0 \in M$ and $x_0 + y_0 \in B(x + y, r)$. Hence, we have $M \cap B(x + y, r) \neq \emptyset$, which implies that $x + y \in \overline{M}$.

$\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in B(x, r/(1 + |\alpha|)) \cap M \neq \emptyset$. Since $M$ is a subspace, then $\alpha x_0 \in M$ and $\alpha x_0 \in B(x, r)$.

Hence, we have $M \cap B(\alpha x, r) \neq \emptyset$, which implies that $\alpha x \in \overline{M}$. Hence, $\overline{M}$ is a subspace.

(ii) Note that $V = x_0 + M$, where $x_0 \in X$ and $M \subseteq X$ is a subspace.

By Proposition 7.16, $\overline{V} = x_0 + \overline{M}$. By (i), $\overline{M}$ is a subspace. Then, $\overline{V}$ is a linear variety.

(iii) First, we will prove the special case $p = \emptyset$. Clearly, $\emptyset \in C \subseteq \overline{C}$. $\forall x \in \overline{C}$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, we will show that $\alpha x \in \overline{C}$. $\forall r \in (0, \infty) \subset \mathbb{R}$, by Proposition 3.3, $\exists x_0 \in C \cap B(x, r/(1 + |\alpha|)) \neq \emptyset$. Then, we have $\alpha x_0 \in C \cap B(\alpha x, r) \neq \emptyset$. By Proposition 3.3, $\alpha x \in \overline{C}$. Hence, $\overline{C}$ is a cone with vertex at the origin.

For general $p \in X$, we have $C = p + C_0$, where $C_0$ is a cone with vertex at the origin. By Proposition 7.16, $\overline{C} = p + \overline{C_0}$. By the special case we have shown, $\overline{C_0}$ is a cone with vertex at the origin. Hence, $\overline{C}$ is a cone with vertex $p$.

(iv) First, we will prove the special case $x_0 = \emptyset$. By Proposition 4.3, $\overline{B}(\emptyset, r)$ is a closed set and $\overline{B}(\emptyset, r) \supseteq B(\emptyset, r)$. Then, we have $\overline{B}(\emptyset, r) \subseteq \overline{B}(\emptyset, r)$. On the other hand, $\forall x \in B(\emptyset, r)$, define $(x_k)_{k=1}^{\infty} \subseteq X$ by $x_k = \frac{1}{kn}r$, $\forall k \in \mathbb{N}$. Then, $\|x_k\| = \frac{1}{kn}r \leq \frac{1}{(k+1)n}r < r$, $\forall k \in \mathbb{N}$.

Therefore, $(x_k)_{k=1}^{\infty} \subseteq B(\emptyset, r)$. It is obvious that $\lim_{k \in X} x_k = x$, then, by
Proposition 4.13, \( x \in \overline{B(\emptyset, r)} \). Then, we have \( \overline{B(\emptyset, r)} \subseteq \overline{B(\emptyset, r)} \). Hence, the result holds for \( x_0 = \emptyset \).

For arbitrary \( x_0 \in X \), by Proposition 7.16, \( \overline{B(x_0, r)} = x_0 + \overline{B(\emptyset, r)} = x_0 + \overline{B(\emptyset, r)} = \overline{B(x_0, r)} \). Hence, the result holds.

This completes the proof of the proposition. \( \square \)

**Proposition 7.18** Let \( X \) be a normed linear space, \( P \subseteq X \), and \( P \neq \emptyset \). The closed linear variety generated by \( P \), denoted by \( V(\overline{P}) \), is the intersection of all closed linear varieties containing \( P \). Then, \( V(\overline{P}) = \overline{\bigcap \{ V() \}} \).

**Proof** By Proposition 7.17, \( \overline{V(\overline{P})} \) is a closed linear variety containing \( P \). Then, we have \( V(\overline{P}) \subseteq \overline{V(\overline{P})} \). On the other hand, \( V(\overline{P}) \) is a closed linear variety containing \( P \), and then \( \overline{V(\overline{P})} \subseteq V(\overline{P}) \). Hence, \( \overline{V(\overline{P})} \subseteq V(\overline{P}) \). Therefore, the result holds. This completes the proof of the proposition. \( \square \)

The justification of the definition of \( V(\overline{P}) \) is that intersection of linear varieties is a linear variety when the intersection is nonempty.

**Definition 7.19** Let \( X \) be a normed linear space and \( P \subseteq X \) be nonempty. \( x \in P \) is said to be a relative interior point of \( P \) if it is an interior point of \( P \) relative to the subset topology of \( V(\overline{P}) \). The set of all relative interior points of \( P \) is called the relative interior of \( P \), and denoted by \( ^{o}P \). \( P \) is said to be relatively open if \( P \) is open in the subset topology of \( V(\overline{P}) \).

**Proposition 7.20** Let \( X \) be a normed linear space and \( (x_\alpha)_{\alpha \in A} \subseteq X \) be a net. Then, \( \lim_{\alpha \in A} x_\alpha = x \in X \) if, and only if, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists x_0 \in A, \forall \alpha \in A \) with \( \alpha_0 < \alpha \), we have \( \| x - x_\alpha \| < \epsilon \).

**Proposition 7.21** Let \( X \) be a normed linear space. Then, \( \| \cdot \| \) is a uniformly continuous function on \( X \).

**Proof** This follows directly from Propositions 4.30 and 7.14. \( \square \)

### 7.3 Product Spaces

**Proposition 7.22** Let \( X := (X, K, \| \cdot \|_X) \) and \( Y := (Y, K, \| \cdot \|_Y) \) be normed linear spaces. By Proposition 6.24, \( X \times Y \) is a vector space over \( K \). Define a function \( \| \cdot \| : X \times Y \rightarrow [0, \infty) \subseteq \mathbb{R} \) by \( \| (x, y) \| := (\| x \|_X^2 + \| y \|_Y^2)^{1/2}, \forall (x, y) \in X \times Y \). Then, \( (X \times Y, K, \| \cdot \|) \) is a normed linear space. This normed linear space will be called the Cartesian product of \( X \) and \( Y \) and be denoted by \( X \times Y \).

**Proof** \( \forall (x_1, y_1), (x_2, y_2) \in X \times Y, \forall \alpha \in K, \) we have \( \| (x_1, y_1) \| = (\| x_1 \|_X^2 + \| y_1 \|_Y^2)^{1/2} \in [0, \infty) \subseteq \mathbb{R} \) and \( \| (x_1, y_1) \| = 0 \Leftrightarrow \| x_1 \|_X = 0 \) and \( \| y_1 \|_Y = 0 \Leftrightarrow x_1 = \emptyset_X \) and \( y_1 = \emptyset_Y \Leftrightarrow (x_1, y_1) = \emptyset; \| (x_1, y_1) + (x_2, y_2) \| = \| (x_1 + x_2, y_1 + y_2) \| = (\| x_1 + x_2 \|_X^2 + \| y_1 + y_2 \|_Y^2)^{1/2} \leq (\| x_1 \|_X^2 + \| x_2 \|_X^2 + \| y_1 \|_Y^2 + \| y_2 \|_Y^2)^{1/2} \leq \sqrt{2} \)
7.3. PRODUCT SPACES

\[(\|y_1\|_y + \|y_2\|_y)^2 \leq (\|x_1\|_X + \|y_1\|_y)^2 + (\|x_2\|_X + \|y_2\|_y)^2 = \|(x_1, y_1)\| + \|(x_2, y_2)\|,\]

where the first inequality follows from the fact that \(X\) and \(Y\) are normed linear spaces and the second inequality follows from

the Minkowski’s Inequality: \(\|\alpha(x_1, y_1)\| = \|\alpha x_1, \alpha y_1\| = (\|\alpha x_1\|_X^2 + \|\alpha y_1\|_Y^2)^{1/2} = (\|\alpha\|^2 \|x_1\|_X^2 + \|\alpha\|^2 \|y_1\|_Y^2)^{1/2} = |\alpha| \|x_1\|_X^2 + \|y_1\|_Y^2)^{1/2} = |\alpha| \|(x_1, y_1)\|\). Hence, \(\|\cdot\|\) is a norm on \(X\times Y\). This completes the proof of the proposition.

Clearly, the natural metric for the Cartesian product \(X \times Y\) is the Cartesian metric defined in Definition 4.28.

The above proposition may also be generalized to the case of \(X_1 \times X_2 \times \cdots \times X_n\), where \(n \in \mathbb{N}\) and \(X_i\)'s are normed linear spaces over the same field \(K\). When \(n = 0\), it should be noted that \(\prod_{\alpha \in \varnothing} X_\alpha = (\{0 = \varnothing\}, K, \|\cdot\|)\), where \(\|\varnothing\| = 0\).

**Proposition 7.23** Let \(X := (X, K, \|\cdot\|)\) be a normed linear space. Then, the vector addition \(\oplus_X : X \times X \to X\) is uniformly continuous; and the scalar multiplication \(\otimes_X : K \times X \to X\) is continuous.

**Proof** \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall(x_1, x_2), (y_1, y_2) \in X \times X\) with

\[\|(x_1, x_2) - (y_1, y_2)\|_{\times X} < \epsilon/\sqrt{2}\]

we have

\[\|(x_1 + x_2) - (y_1 + y_2)\| \leq \|x_1 - y_1\| + \|x_2 - y_2\|\]

\[\leq \sqrt{2}(\|x_1 - y_1\|^2 + \|x_2 - y_2\|^2)^{1/2} = \sqrt{2}\|(x_1, x_2) - (y_1, y_2)\|_{\times X} < \epsilon\]

Hence, the vector addition \(\oplus_X\) is uniformly continuous.

\(\forall \alpha_0 \in K, \forall x_0 \in X, \forall \epsilon \in (0, \infty) \subset \mathbb{R},\) let \(\delta = \epsilon/(1 + \epsilon + |\alpha_0| + \|x_0\|) \in (0, 1) \subset \mathbb{R} \). \(\forall (\alpha, x) \in B_{K \times X}((\alpha_0, x_0), \delta)\), we have

\[\|\alpha x - \alpha_0 x_0\| \leq \|\alpha x - \alpha_0 x_0\| + \|\alpha x_0 - \alpha_0 x_0\|\]

\[= |\alpha| \|x - x_0\| + \|x_0\| |\alpha - \alpha_0|\]

\[\leq \left(|\alpha|^2 + \|x_0\|^2\right)^{1/2} \left(\|x - x_0\|^2 + |\alpha - \alpha_0|^2\right)^{1/2}\]

\[\leq (|\alpha| + \|x_0\|)\delta \leq (1 + \|\alpha_0\| + \|x_0\|)\delta < \epsilon\]

where we have made use of Cauchy-Schwarz Inequality in the second inequality. Hence, \(\otimes_X\) is continuous at \((\alpha_0, x_0)\). By the arbitrariness of \((\alpha_0, x_0)\), we have that \(\otimes_X\) is continuous.

This completes the proof of the proposition. \(\square\)

**Definition 7.24** Let \(X := (X, K, \|\cdot\|_X)\) and \(Y := (Y, K, \|\cdot\|_Y)\) be two normed linear spaces over the same field \(K\) and \(A : X \to Y\) be a vector space isomorphism. \(A\) is said to be an isometrical isomorphism if \(\|Ax\|_Y = \|x\|_X, \forall x \in X\). The \(X\) and \(Y\) are said to be isometrically isomorphic.
Let $A : X \to Y$ be an isometrical isomorphism between $X$ and $Y$. Then, $A$ is an isometry between $X$ and $Y$. Both $A$ and $A_{\text{inv}}$ are uniformly continuous. $X$ and $Y$ are equal to each other up to a relabeling of their vectors.

**Proposition 7.25**  
Let $X_\alpha, \alpha \in \Lambda$, be normed linear spaces over $\mathbb{K}$, where $\Lambda$ is a finite index set. Let $\Lambda = \bigcup_{\beta \in \Gamma} \Lambda_\beta$, where $\Lambda_\beta$'s are pairwise disjoint and finite and $\Gamma$ is also finite. $\forall \beta \in \Gamma$, let $X^{(\beta)} := \prod_{\alpha \in \Lambda_\beta} X_\alpha$ be the Cartesian product space. Let $X^{(\Gamma)} := \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha$ be the Cartesian product space of product spaces, and $X := \prod_{\alpha \in \Lambda} X_\alpha$ be the Cartesian product space. Then, $X$ and $X^{(\Gamma)}$ are isometrically isomorphic.

**Proof**  
Define $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \to \prod_{\alpha \in \Lambda} X_\alpha$ by, $\forall x \in X^{(\Gamma)}$, $\forall \alpha \in \Lambda, \exists ! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}, \pi_\alpha(E(x)) = \pi_\alpha^{(\beta_\alpha)}(\pi_\beta^{(\Gamma)}(x))$. By Proposition 4.32, $E$ is a isometry. It is clear that $E$ is a linear operator since the projection functions are linear for vector spaces. Hence, $E$ is a vector space isomorphism. Then, $E$ is an isometrical isomorphism. This completes the proof of the proposition. $\square$

### 7.4 Banach Spaces

**Definition 7.26**  
If a normed linear space is complete with respect to the natural metric, then it is called a Banach space.

Clearly, a Cauchy sequence in a normed linear space is bounded.

**Proposition 7.27**  
A normed linear space $X$ is complete if, and only if, every absolutely summable series is summable, that is $\forall (x_n)_{n=1}^{\infty} \subseteq X$, $\sum_{n=1}^{\infty} \|x_n\| \in \mathbb{R} \Rightarrow \sum_{n=1}^{\infty} x_n \in X$.

**Proof**  
"Only if" Let $X$ be complete and $(x_n)_{n=1}^{\infty} \subseteq X$ be absolutely summable. Then, $\sum_{n=1}^{\infty} \|x_n\| \in \mathbb{R}$ and $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}$ such that, $\forall n, m \in \mathbb{N}$ with $n \geq m$, we have $\sum_{i=m}^{n} \|x_i\| < \epsilon$. Then, $\|\sum_{i=m}^{n} x_i\| < \epsilon$. Let $s_n := \sum_{i=1}^{n} x_i, \forall n \in \mathbb{N}$, be the partial sum. Then, $(s_n)_{n=1}^{\infty} \subseteq X$ is a Cauchy sequence. By the completeness of $X$, $\sum_{n=1}^{\infty} x_n = \lim_{n \to \infty} s_n \in X$.

"If" Let $(x_n)_{n=1}^{\infty} \subseteq X$ be a Cauchy sequence. Let $n_0 = 0. \forall k \in \mathbb{N}$, $\exists n_k \in \mathbb{N}$ with $n_k > n_{k-1}$ such that $\|x_n - x_m\| < 2^{-k}, \forall n, m \geq n_k$. Then, $(x_{n_k})_{k=1}^{\infty}$ is a subsequence of $(x_n)_{n=1}^{\infty}$. Let $y_1 := x_{n_1}$ and $y_k := x_{n_k} - x_{n_{k-1}}, \forall k \geq 2$. Then, the sequence $(y_k)_{k=1}^{\infty} \subseteq X$ and its $k$th partial sum is $x_{n_k}$. Note that $\|y_k\| < 2^{-k+1}, \forall k \geq 2$. Then, $\sum_{k=1}^{\infty} \|y_k\| < \|y_1\| + 1$, which implies that $\sum_{k=1}^{\infty} y_k = x_0 \in X$. Hence, we have $\lim_{k \to \infty} x_{n_k} = x_0 \in X$.

We will now show that $\lim_{n \to \infty} x_n = x_0. \forall k \in \mathbb{N}, \exists N \in \mathbb{N}$ with $N \geq k$ such that $\|x_n - x_0\| < 2^{-k}, \forall i \geq N, \forall n \geq N$, we have $\|x_n - x_0\| \leq \|x_n - x_{n_N}\| + \|x_{n_N} - x_0\| < 2^{-N} + 2^{-k} \leq 2^{-k+1}$. Hence, $\lim_{n \to \infty} x_n = x_0$. Therefore, $X$ is complete.
This completes the proof of the proposition.

We frequently take great care to formulate problems arising in applications as equivalent problems in Banach spaces rather than other incomplete spaces. The principal advantage of Banach spaces in optimization problems is that when seeking an optimal vector, we often construct a sequence (net) of vectors, each member of which is superior to preceding ones. The desired optimal vector is then the limit of the sequence (net). In order for this scheme to be effective, there must be available a test for convergence which can be applied when the limit is unknown. The Cauchy criterion for convergence meets this requirement provided the space is complete.

**Example 7.28** Consider Example 7.11 with \(a = 0\) and \(b = 1\). We will show that the space \((M, \mathbb{R}, \| \cdot \|)\) is incomplete. Take \(\{x_n\}_{n=1}^{\infty} \subseteq M\) to be,

\[
x_n(t) = \begin{cases} 
0 & 0 \leq t \leq \frac{1}{2} - \frac{1}{n+1} \\
(n+1)t - \frac{n+1}{2} + 1 & \frac{1}{2} - \frac{1}{n+1} < t \leq \frac{1}{2} \\
1 & t > \frac{1}{2}
\end{cases}
\]

![Figure 7.1: Sequence for Example 7.28.](image)

Clearly, \(x_n\) is continuous and \(\|x_n\| < 1, \forall n \in \mathbb{N}\). \(\forall n, m \in \mathbb{N}\), we have

\[
\|x_n - x_m\| = \int_0^1 |x_n(t) - x_m(t)| \, dt = \left| \int_0^1 (x_n(t) - x_m(t)) \, dt \right| \\
= \left| \frac{1}{2(n+1)} - \frac{1}{2(m+1)} \right|
\]

Then, \((x_n)_{n=1}^{\infty}\) is Cauchy. Yet, it is obvious that there is no continuous function \(x \in M\) such that \(\lim_{n\in\mathbb{N}} x_n = x\), i.e., \(\lim_{n\in\mathbb{N}} \int_0^1 |x_n(t) - x(t)| \, dt = 0\). Hence, the space is incomplete.

**Example 7.29** \(\mathbb{R}^n\) with norm defined as in Example 7.2 or Example 7.3 is a Banach space.

**Example 7.30** Consider the real normed linear space \(C([a, b])\) defined in Example 7.4, where \(a, b \in \mathbb{R}\) and \(a \leq b\). We will show that it is complete.
Take a Cauchy sequence \((x_n)_{n=1}^{\infty} \subseteq C([a,b])\). \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}\) such that \(\forall n, m \geq N, 0 \leq |x_n(t) - x_m(t)| \leq \|x_n - x_m\| < \epsilon, \forall t \in [a,b]\). This shows that, \(\forall t \in [a,b], (x_n(t))_{n=1}^{\infty} \subseteq \mathbb{R}\) is a Cauchy sequence, which converges to \(x(t) \in \mathbb{R}\) since \(\mathbb{R}\) is complete. This defines a function \(x : [a,b] \rightarrow \mathbb{R}\). It is easy to show that \((x_n)_{n=1}^{\infty}\), viewed as a sequence of functions of \([a,b]\) to \(\mathbb{R}\), converges uniformly to \(x\). By Proposition 4.26, \(x\) is continuous. Hence, \(x \in C([a,b])\). It is easy to see that \(\lim_{n \in \mathbb{N}} \|x_n - x\| = 0\). Hence, \(C([a,b])\) is a Banach space.

Example 7.31  Let \(K\) be a countably compact topological space, \(X\) be a normed linear space over the field \(K\), \(Y := \{f : K \rightarrow X\}\). Define the usual vector addition \(\oplus\) and scalar multiplication \(\otimes\) and null vector \(\vartheta\) on \(Y\) as in Example 6.20. Then, \(Y := (Y, \oplus, \otimes, \vartheta)\) is a vector space over \(K\). Let \(M := \{f \in Y \mid f\) is continuous\}. Then, by Propositions 3.32, 6.25, and 7.23, \(M := (M, \oplus, \vartheta, \varnothing)\) is a subspace of \((Y, K)\). Define a function \(\| \cdot \| : M \rightarrow (0, \infty) \subset \mathbb{R}\) by \(\|f\| = \max \{\sup_{k \in K} \|f(k)\|_X, 0\}, \forall f \in M\). This function is well defined by Propositions 7.21, 3.12, and 5.29. We will show that \(\| \cdot \|\) defines a norm on \(M\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(K = \emptyset\); Case 2: \(K \neq \emptyset\). Case 1: \(K = \emptyset\). Then, \(M\) is a singleton set \(\{\emptyset\}\) and \(\|\emptyset\| = 0\). Clearly, \((M, K, \| \cdot \|)\) is a normed linear space. Case 2: \(K \neq \emptyset\). \(\forall f, g \in M, \forall \alpha \in K, \|f\| = \max_{k \in K} \|f(k)\|_X\) by Propositions 7.21, 3.12, and 5.29. \(\|f\| = 0 \iff \|f(k)\|_X = 0, \forall k \in K \iff f(k) = \vartheta_X, \forall k \in K \iff f = \vartheta\). \(\|f + g\| = \max_{k \in K} \|f(k) + g(k)\|_X \leq \max_{k \in K} \|f(k)\|_X + \max_{k \in K} \|g(k)\|_X = \|f\| + \|g\|. \|\alpha f\| = \max_{k \in K} \|\alpha f(k)\|_X = \max_{k \in K} |\alpha| \|f(k)\|_X = |\alpha| \|f\|\). Hence, \((M, K, \| \cdot \|)\) is a normed linear space. In both cases, we have shown that \(C(K, X) := (M, K, \| \cdot \|)\) is a normed linear space.

Example 7.32  Let \(K\) be a countably compact topological space, \(X\) be a Banach space over the field \(K\) (with norm \(\| \cdot \|_X\)). Consider the normed linear space \(C(K, X)\) (with norm \(\| \cdot \|\)) defined in Example 7.31. We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(K = \emptyset\); Case 2: \(K \neq \emptyset\). Case 1: \(K = \emptyset\). Then, \(C(K, X)\) is a singleton set. Hence, any Cauchy sequence must converge. Thus, \(C(K, X)\) is a Banach space. Case 2: \(K \neq \emptyset\). Take a Cauchy sequence \((x_n)_{n=1}^{\infty} \subseteq C(K, X)\). \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}\) such that \(\forall n, m \geq N, 0 \leq \|x_n(k) - x_m(k)\|_X \leq \|x_n - x_m\| < \epsilon, \forall k \in K\). This shows that, \(\forall k \in K, (x_n(k))_{n=1}^{\infty} \subseteq X\) is a Cauchy sequence, which converges to \(x(k) \in X\) since \(X\) is complete. This defines a function \(x : K \rightarrow X\). It is easy to show that \((x_n)_{n=1}^{\infty}\), viewed as a sequence of functions of \(K\) to \(X\), converges uniformly to \(x\). By Proposition 4.26, \(x\) is continuous. Hence, \(x \in C(K, X)\). It is easy to see that \(\lim_{n \in \mathbb{N}} \|x_n - x\| = 0\). Hence, \(C(K, X)\) is a Banach space. In both cases, we have shown that \(C(K, X)\) is a Banach space when \(X\) is a Banach space.

Example 7.33  Let \(X\) be a Banach space over the field \(K\). Consider the normed linear space \(L_p(X)\) (with norm \(\| \cdot \|_{L_p}\)) defined in Example 7.10,
7.4. BANACH SPACES

where \( p \in [1, \infty] \subset \mathbb{R}_e \). We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( p \in (1, \infty) \); Case 2: \( p = \infty \).

Case 1: \( p \in [1, \infty) \). Take a Cauchy sequence \((x_n)_{n=1}^{\infty} \subseteq l_p(X)\), where \( x_n = (\xi_n,k)_{k=1}^{\infty} \subseteq X, \forall n \in \mathbb{N} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}_e, \exists N \in \mathbb{N} \) such that \( \forall n, m \geq N, \) we have \( \|x_n - x_m\|_p < \epsilon \). \( \forall k \in \mathbb{N}, \) \( \|\xi_{n,k} - \xi_{m,k}\|_X \leq \left( \sum_{i=1}^{\infty} \|\xi_{n,i} - \xi_{m,i}\|_X^p \right)^{1/p} = \|x_n - x_m\|_p < \epsilon \). Hence, \( \forall k \in \mathbb{N}, \) \( (\xi_{n,k})_{n=1}^{\infty} \subseteq X \) is a Cauchy sequence, which converges to some \( \xi_k \in X \) since \( X \) is a Banach space. Let \( x := (\xi_k)_{k=1}^{\infty} \subseteq X \). We will show that \( x \in l_p(X) \) and \( \lim_{n \in \mathbb{N}} x_n = x \). Since \((x_n)_{n=1}^{\infty} \) is a Cauchy sequence, then it is bounded, that is \( \exists M \in [0, \infty) \subset \mathbb{R}_e \) such that \( \|x_n\|_p \leq M, \forall n \in \mathbb{N} \). Then, we have \( \|x_n\|_p^p = \sum_{i=1}^{\infty} \|\xi_{n,i}\|_X^p \leq M^p, \forall n \in \mathbb{N} \). Then, \( \forall n, k \in \mathbb{N} \), we have \( \sum_{i=1}^{k} \|\xi_{n,i}\|_X^p \leq M^p \). By Propositions 7.21 and 3.66, we have \( \sum_{i=1}^{k} \|\xi_{i,k}\|_X^p = \lim_{n \in \mathbb{N}} \sum_{i=1}^{k} \|\xi_{n,i}\|_X^p \leq M^p \). Hence, we have \( \sum_{i=1}^{k} \|\xi_{i,k}\|_X^p = \lim_{n \in \mathbb{N}} \sum_{i=1}^{k} \|\xi_{n,i} - \xi_{n,i}\|_X^p \leq M^p \). Taking limit as \( n \to \infty \), by Propositions 7.21, 3.66, and 7.23, we have \( \sum_{i=1}^{k} \|\xi_{n,i} - \xi_{i,k}\|_X^p = \lim_{n \in \mathbb{N}} \sum_{i=1}^{k} \|\xi_{n,i} - \xi_{i,k}\|_X^p \leq \epsilon^p \). Hence, we have \( \|x_n - x\|_p^p = \left( \sum_{i=1}^{\infty} \|\xi_{n,i} - \xi_i\|_X^p \right)^{1/p} \leq \epsilon \). This shows that \( \lim_{n \in \mathbb{N}} x_n = x \). Hence, \( l_p(X) \) is complete and therefore a Banach space.

Case 2: \( p = \infty \). Take a Cauchy sequence \((x_n)_{n=1}^{\infty} \subseteq l_\infty(X)\), where \( x_n = (\xi_{n,k})_{k=1}^{\infty} \subseteq X, \forall n \in \mathbb{N} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}_e, \exists N \in \mathbb{N} \) such that \( \forall n, m \geq N, \) we have \( \|x_n - x_m\|_\infty < \epsilon \). \( \forall k \in \mathbb{N}, \) \( \|\xi_{n,k} - \xi_{m,k}\|_X = \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_{m,i}\|_X < \epsilon \). Hence, \( \forall k \in \mathbb{N}, (\xi_{n,k})_{n=1}^{\infty} \subseteq X \) is a Cauchy sequence, which converges to some \( \xi_k \in \mathbb{R}_e \) since \( X \) is a Banach space. Let \( x := (\xi_k)_{k=1}^{\infty} \subseteq X \). We will show that \( x \in l_\infty(X) \) and \( \lim_{n \in \mathbb{N}} x_n = x \). Since \((x_n)_{n=1}^{\infty} \) is a Cauchy sequence, then it is bounded, that is \( \exists M \in [0, \infty) \subset \mathbb{R}_e \) such that \( \forall n \in \mathbb{N}, \sup_{i \in \mathbb{N}} \|\xi_{n,i}\|_X \leq M \). Then, we have \( \|\xi_{n,k}\|_X = \lim_{n \in \mathbb{N}} \|\xi_{n,k}\|_X \leq M \). Hence, \( \|x\|_\infty = \sup_{i \in \mathbb{N}} \|\xi_i\|_X \leq M \). Therefore, \( x \in l_\infty(X) \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}_e, \forall n, m \in \mathbb{N} \) with \( n, m \geq N, \) we have \( \|x_n - x_m\|_\infty < \epsilon \). Then, \( \forall k \in \mathbb{N}, \|\xi_{n,k} - \xi_{m,k}\|_X < \epsilon \). Taking limit as \( m \to \infty \), by Propositions 7.21, 3.66, and 7.23, we have \( \|\xi_{n,k} - \xi_k\|_X = \lim_{n \in \mathbb{N}} \|\xi_{n,k} - \xi_{m,k}\|_X \leq \epsilon \). Hence, \( \|x_n - x\|_\infty = \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_i\|_X \leq \epsilon \). This shows that \( \lim_{n \in \mathbb{N}} x_n = x \). Hence, \( l_\infty(X) \) is complete and therefore a Banach space.

In summary, we have shown that \( l_p(X) \) is a Banach space, \( \forall p \in [1, \infty] \subset \mathbb{R}_e \), when \( X \) is a Banach space.

**Definition 7.34** Let \( X \) be a normed linear space and \( S \subseteq X \). \( S \) is said to be complete if \( S \) with the natural metric forms a complete metric space.
By Proposition 4.39, a subset in a Banach space is complete if, and only if, it is closed.

**Proposition 7.35** Let \( Y \) be a normed linear space over \( K \), \( S_1, S_2 \subseteq Y \) be separable subsets, and \( \alpha \in K \). Then, \( \text{span}(S_1), \text{span}(S_1), S_1 + S_2, \alpha S_1, S_1 \cap S_2, \) and \( S_1 \cup S_2 \) are separable subsets of \( Y \).

**Proof** Let \( K_Q := Q \) if \( K = \mathbb{R} \); and \( K_Q := \{a + ib \in \mathbb{C} \mid a, b \in Q\} \), if \( K = \mathbb{C} \). Clearly, \( K_Q \) is a countable dense set in \( K \). Let \( D \subseteq S_1 \) be a countable dense subset. Let \( \hat{D} := \{\sum_{i=1}^{n} \alpha_i y_i \mid n \in \mathbb{Z}^+, \alpha_i \in K_Q, y_i \in D, i = 1, \ldots, n\} \), which is a countable set. Clearly, \( \hat{D} \subseteq \text{span}(S_1) \) is a dense subset. Hence, \( \text{span}(S_1) \subseteq Y \) is separable. By Proposition 4.38, \( \text{span}(S_1) \) is separable. It is straightforward to show that \( S_1 + S_2, \alpha S_1, S_1 \cap S_2, \) and \( S_1 \cup S_2 \) are separable. \( \square \)

**Theorem 7.36** In a normed linear space \( X \), any finite-dimensional subspace \( M \subseteq X \) is complete.

**Proof** Let \( X \) be a normed linear space over the field \( K \). Let \( n \in \mathbb{Z}^+ \) be the dimension of \( M \), which is well defined by Theorem 6.51. We will prove the theorem by mathematical induction on \( n \).

1° \( n = 0 \). Then, \( M = \{\theta\} \). Clearly, any Cauchy sequence in \( M \) must converge to \( \theta \in M \). Hence, \( M \) is complete. \( \bar{n} = 1 \). Let \( \{e_1\} \subseteq M \) be a basis for \( M \). Clearly, \( e_1 \neq \theta \) and \( \|e_1\| > 0 \). Fix any Cauchy sequence \( (x_n)_{n=1}^{\infty} \subseteq M \). Then, \( x_n = \alpha_n e_1 \) for some \( \alpha_n \in K \), \( \forall n \in \mathbb{N} \). Since \( (x_n)_{n=1}^{\infty} \) is Cauchy, then \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \forall n, m \geq N, \) we have \( \|x_n - x_m\| < \epsilon \). Then, we have \( |\alpha_n - \alpha_m| = \|x_n - x_m\|/\|e_1\| < \epsilon \). Hence, \( (\alpha_n)_{n=1}^{\infty} \) is a Cauchy sequence in \( K \). Then, \( \lim_{n \in \mathbb{N}} \alpha_n = \alpha \in K \) since \( K \) is complete. It is easy to show that \( \lim_{n \in \mathbb{N}} x_n = \alpha e_1 \in M \). Hence, \( M \) is complete.

2° Assume \( M \) is complete when \( n = k - 1 \in \mathbb{N} \).

3° Consider the case \( n = k \in \{2, 3, \ldots\} \subset \mathbb{N} \). Let \( \{e_1, \ldots, e_k\} \subseteq M \) be a basis for \( M \). Define \( M_i := \text{span}(\{e_1, \ldots, e_k\} \setminus \{e_i\}) \) and \( \delta_i := \text{dist}(e_i, M_i) \), \( i = 1, \ldots, k \). \( \forall i = 1, \ldots, k \), \( \forall i = 1, \ldots, k \), we have \( \delta_i \in [0, \infty) \subset \mathbb{R} \). \( M_i \), which is a \( k - 1 \)-dimensional subspace of \( X \). By inductive assumption \( M_i \) is complete. By Proposition 4.39, \( M_i \) is closed. Clearly, \( e_i \notin M_i \). By Proposition 4.10, \( \delta_i > 0 \).

Let \( \delta := \min\{\delta_1, \ldots, \delta_k\} > 0 \). Fix any Cauchy sequence \( (x_n)_{n=1}^{\infty} \subseteq M \). \( \forall n \in \mathbb{N} \), \( x_n \) admits a unique representation \( x_n = \sum_{i=1}^{k} \lambda_{n,i} e_i \) where \( \lambda_{n,i} \in K, i = 1, \ldots, k \), by Definition 6.50 and Corollary 6.47. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N} \) such that \( \forall n, m \geq N, \) we have \( \|x_n - x_m\| < \delta \). \( \forall i \in \{1, \ldots, k\}, \) we have \( \epsilon > \|x_n - x_m\|/\delta = \left(\sum_{j=1}^{k} (\lambda_{n,j} - \lambda_{m,j}) e_j \right)/\delta \geq |\lambda_{n,i} - \lambda_{m,i}| \delta/\delta \geq |\lambda_{n,i} - \lambda_{m,i}|. \) Hence, \( (\lambda_{n,i})_{n=1}^{\infty} \subseteq K \) is a Cauchy sequence, \( \forall i \in \{1, \ldots, k\} \). Then, \( \lim_{n \in \mathbb{N}} \lambda_{n,i} = \lambda_i \in K \). Let \( x := \sum_{i=1}^{k} \lambda_i e_i \in M \).
Now, it is straightforward to show that $\lim_{n \in \mathbb{N}} \| x_n - x \| = 0$. Then, $\lim_{n \in \mathbb{N}} x_n = x \in M$. Hence, $M$ is complete.

This completes the induction process. This completes the proof of the theorem. \hfill \Box

**Definition 7.37** Let $\mathcal{X}$ be a vector space over the field $\mathbb{K}$ and $\| \cdot \|_1$ and $\| \cdot \|_2$ be two norms defined on $\mathcal{X}$. These two norms are said to be equivalent if $\exists K \in (0, \infty) \subset \mathbb{R}$ such that $\| x \|_1 / K \leq \| x \|_2 \leq K \| x \|_1$, $\forall x \in \mathcal{X}$.

Clearly, two norms are equivalent implies that the natural metrics are uniformly equivalent.

**Theorem 7.38** Let $\mathcal{X}$ be a finite-dimensional vector space over the field $\mathbb{K}$. Any two norms on $\mathcal{X}$ are equivalent.

**Proof** Let $n \in \mathbb{Z}_+$ be the dimension of $\mathcal{X}$, which is well defined by Theorem 6.51. Let $\| \cdot \|_1$ and $\| \cdot \|_2$ be two norms defined on $\mathcal{X}$. We will distinguish three exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n = 1$; Case 3: $n \geq 2$. Case 1: $n = 0$. Then, $\mathcal{X} = \{0\}$ and $\|0\|_1 = 0 = \|0\|_2$. Hence, the two norms are equivalent. Case 2: $n = 1$. Let $\{e_1\}$ be a basis for $\mathcal{X}$. Clearly, $e_1 \neq 0$. Let $\delta := \| e_1 \|_1 / \| e_1 \|_2 \in (0, \infty) \subset \mathbb{R}$ and $K = \max\{\delta, \delta^{-1}\} \in (0, \infty) \subset \mathbb{R}$. $\forall x \in \mathcal{X}$, $\exists \alpha \in \mathbb{K}$ such that $x = \alpha e_1$. Then, $\|x\|_1 = |\alpha| \| e_1 \|_1 = \| e_1 \|_2 \delta = \delta \| x \|_2$ and $\|x\|_1 / K \leq \| x \|_2 \leq K \| x \|_1$. Hence, the two norms are equivalent. Case 3: $n \geq 2$. Let $\{e_1, \ldots, e_n\} \subseteq \mathcal{X}$ be a basis for $\mathcal{X}$. $\forall x \in \mathcal{X}$, by Definition 6.50 and Corollary 6.47, $\exists \alpha_1, \ldots, \alpha_n \in \mathbb{K}$ such that $x = \sum_{i=1}^n \alpha_i e_i$. We will show that $\exists K_1 \in (0, \infty) \subset \mathbb{R}$ such that $\left(\sum_{i=1}^n |\alpha_i|\right)/K_1 \leq \|x\|_1 \leq K_1 \left(\sum_{i=1}^n |\alpha_i|\right)$. By a similar argument, $\exists K_2 \in (0, \infty) \subset \mathbb{R}$ such that $\left(\sum_{i=1}^n |\alpha_i|\right)/K_2 \leq \|x\|_2 \leq K_2 \left(\sum_{i=1}^n |\alpha_i|\right)$. Then, $\|x\|_1 / K \leq \| x \|_2 \leq K \| x \|_1$, where $K := K_1 K_2$. Hence, the two norms are equivalent. Clearly, $e_i \neq 0$ and $\| e_i \|_1 > 0$, $i = 1, \ldots, n$. Let $\delta := \max_{1 \leq i \leq n} \| e_i \|_1 \in (0, \infty) \subset \mathbb{R}$. Then, $\|x\|_1 \leq \sum_{i=1}^n |\alpha_i| \| e_i \|_1 \leq \delta \left(\sum_{i=1}^n |\alpha_i|\right)$. Define $M_i = \text{span}\{e_1, \ldots, e_n \setminus \{e_i\}\}$, $i = 1, \ldots, n$, and $\delta_i = \text{dist}(e_i, M_i)$ (with respect to $\| \cdot \|_1$). Since $\{e_1, \ldots, e_n\}$ is a basis for $\mathcal{X}$, then $e_i \notin M_i$, $\forall i \in \{1, \ldots, k\}$, by Theorem 7.36, $M_i$ is complete (with respect to $\| \cdot \|_1$). By Proposition 4.39, $M_i$ is closed (with respect to $\| \cdot \|_1$). By Proposition 4.10, we have $\delta_i \in (0, \infty) \subset \mathbb{R}$. Let $\delta := \min\{\delta_1, \ldots, \delta_n\} \in (0, \infty) \subset \mathbb{R}$. Then, $\|x\|_1 = \|\sum_{i=1}^n \alpha_i e_i\|_1 \geq |\alpha_i| \delta_i \geq |\alpha_i| \delta$, $\forall i \in \{1, \ldots, n\}$. Hence, we have $\|x\|_1 \geq (\delta / n) \left(\sum_{i=1}^n |\alpha_i|\right)$. Let $K_1 = \max\{\delta / n, \delta\} > 0$. Then, we have $\left(\sum_{i=1}^n |\alpha_i|\right)/K_1 \leq \|x\|_1 \leq K_1 \left(\sum_{i=1}^n |\alpha_i|\right)$. This completes the proof of the theorem. \hfill \Box

### 7.5 Compactness

**Definition 7.39** Let $X$ be a normed linear space and $S \subseteq X$. $S$ is said to be compact if $S$ together with the natural metric forms a compact metric space.
Proposition 5.29 says that a continuous function achieves its minimum and maximum on nonempty countably compact spaces. This has immediate generalization to infinite-dimensional spaces. Yet, the compactness restriction is so severe in infinite-dimensional spaces that it is applicable in minority of problems.

Lemma 7.40 Let \((\mathcal{X}, \mathbb{C})\) be a vector space. Then \(\mathcal{X}\) is also a vector space over the field \(\mathbb{R}\). Furthermore, if \(\mathcal{X} := (\mathcal{X}, \mathbb{C}, \| \cdot \|)\) is a normed linear space, then \(\mathcal{X}_\mathbb{R} := (\mathcal{X}, \mathbb{R}, \| \cdot \|)\) is also a normed linear space. \(\mathcal{X}\) and \(\mathcal{X}_\mathbb{R}\) are isometric and admit the same metric space properties. In particular, \(\mathcal{X}\) is a Banach space if, and only if, \(\mathcal{X}_\mathbb{R}\) is a Banach space.

Proof Let \((\mathcal{X}, \mathbb{C})\) be a vector space. Note that \(\mathbb{R} \subset \mathbb{C}\). Then, \(\forall x, y, z \in \mathcal{X}\), \(\forall \alpha, \beta \in \mathbb{R}\), we have (i) \(x + y = y + x\); (ii) \((x + y) + z = x + (y + z)\); (iii) \(\vartheta x = x\); (iv) \(\alpha (x + y) = \alpha x + \alpha y\); (v) \((\alpha + \beta)x = \alpha x + \beta x\); (vi) \((\alpha \beta)x = \alpha (\beta x)\); (vii) \(0 \cdot x = \vartheta x\) and \(1 \cdot x = x\). Hence, \((\mathcal{X}, \mathbb{R})\) is a vector space.

Let \(\mathcal{X} := (\mathcal{X}, \mathbb{C}, \| \cdot \|)\) be a normed linear space. Then, \((\mathcal{X}, \mathbb{R})\) is a vector space. \(\forall x, y \in \mathcal{X}\), \(\forall \alpha \in \mathbb{R}\), we have (i) \(\| x \| \in [0, \infty) \subset \mathbb{R}\) and \(\| x \| = 0 \iff x = \vartheta X\); (ii) \(\| x + y \| \leq \| x \| + \| y \|\); (iii) \(\| \alpha x \| = |\alpha| \| x \|\). Hence, \(\mathcal{X}_\mathbb{R} := (\mathcal{X}, \mathbb{R}, \| \cdot \|)\) is a normed linear space.

Clearly, \(id_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}_\mathbb{R}\) is an isometry and the natural metrics induced by \(\mathcal{X}\) and \(\mathcal{X}_\mathbb{R}\) on \(\mathcal{X}\) are identical. Hence, \(\mathcal{X}\) and \(\mathcal{X}_\mathbb{R}\) admits the same metric space properties. Then, \(\mathcal{X}\) is a Banach space if, and only if, \(\mathcal{X}\) is complete if, and only if, \(\mathcal{X}_\mathbb{R}\) is complete if, and only if, \(\mathcal{X}_\mathbb{R}\) is a Banach space.

This completes the proof of the lemma.

Proposition 7.41 Let \(K \subseteq \mathbb{C}^n\) with \(n \in \mathbb{Z}_+\). Then, \(K\) is compact if, and only if, \(K\) is closed and bounded.

Proof “Necessity” By Proposition 5.38, \(K\) is complete and totally bounded. Then, \(K\) is bounded. By Proposition 4.39, \(K\) is closed.

“Sufficiency” Let \(\| \cdot \|\) be the norm on \(\mathbb{C}^n\). By Lemma 7.40, \(\mathcal{X} := (\mathbb{C}^n, \mathbb{C}, \| \cdot \|)\) admits the same metric space property as \(\mathbb{C}^n\). Then, \(K \subseteq \mathcal{X}\) is closed and bounded. Note that \(\mathcal{X}\) is isometrically isomorphic to \(\mathbb{R}^{2n}\). Hence, \(K \subseteq \mathcal{X}\) is compact. Then, \(K \subseteq \mathbb{C}^n\) is compact. This completes the proof of the proposition.

Proposition 7.42 Let \(\mathcal{X}\) be a finite-dimensional normed linear space over the field \(\mathbb{K}\). \(K \subseteq \mathcal{X}\) is compact if, and only if, \(K\) is closed and bounded.

Proof “Necessity” By Proposition 5.38, \(K\) is complete and totally bounded. Then, \(K\) is bounded. By Proposition 4.39, \(K\) is closed.

“Sufficiency” Let \(S \subseteq \mathcal{X}\) be a basis for \(\mathcal{X}\), that is, \(S\) is linearly independent and \(\text{span}(S) = \mathcal{X}\). Since \(\mathcal{X}\) is finite dimensional, then let \(n \in \mathbb{Z}_+\) be the dimension of \(\mathcal{X}\), which is well-defined by Theorem 6.51. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(n = 0\); Case 2:
7.6. QUOTIENT SPACES

Proposition 7.43 Let $M \subseteq X$ be a subspace of a vector space $X$ over a field $F := (\mathbb{F}, +, \cdot, 0, 1)$. $x_1, x_2 \in X$ are said to be equivalent modulo $M$ if $x_1 - x_2 \in M$. This equivalence relationship (easy to show) partitions the space $X$ into disjoint subsets, or classes, of equivalent elements: namely, the linear varieties that are distinct translations of the subspace $M$. These classes are often called the cosets of $M$. $\forall x \in X$, there is a unique coset of $M$, $[x] := M + [x]$, such that $x \in [x]$. The quotient of $X$ modulo $M$ is defined to be the set $N$ of all cosets of $M$. Define vector addition and scalar multiplication on $N$ by $\forall [x_1], [x_2] \in N$, $\forall a \in F$, $[x_1] \{ x_2 \} := [x_1 + x_2]$ and $a [x_1] := a [x_1]$. Let the null vector on $N$ be $[ \emptyset ] = [x_1]$. Then, $N$ together with the vector addition, scalar multiplication, and the null vector form a vector space over $F$. This vector space will be called the quotient space of $X$ modulo $M$ and denoted by $X/M$. Define a function $\phi : X \to X/M$ by $\phi(x) = [x]$, $\forall x \in X$. Then, $\phi$ is a linear function and will be called the natural homomorphism.

Proof Fix any $[x_1], [x_2], [x_3] \in X/M$ and any $\alpha, \beta \in F$. We will first show that vector addition and scalar multiplication are uniquely defined. $\forall x_1 \in [x_1], \forall x_2 \in [x_2], \forall x_3 \in [x_3], \forall x_1 \in [x_1] + [x_2] + [x_3] = [x_1 + x_2 + x_3] = [x_1] + ([x_2] + [x_3]) = [x_1] + ([x_2] + [x_3]);$ \item (ii) $([x_1] + [x_2]) + [x_3] = [x_1 + x_2 + x_3] = [x_1] + ([x_2] + [x_3]);$ \item (iii) $[x_1] + [\emptyset ] = [x_1 + \emptyset ] = [x_1];$ \item (iv) $a ([x_1] + [x_2]) = [ax_1 + ax_2] = [ax_1] + [ax_2] = a [x_1] + a [x_2];$ \item (v) $(a + \beta) [x_1] = [(a + \beta) x_1] = [ax_1 + \beta x_1] = [ax_1] + [\beta x_1] = a [x_1] + \beta [x_1];$ \item (vi) $(\alpha \beta) [x_1] = [(\alpha \beta) x_1] = [\alpha (\beta x_1)] = \alpha [\beta x_1] = \alpha (\beta [x_1]));$ \item (vii) $[x_1] = 0 [x_1] = [\emptyset ];$ \item (vii) $1 [x_1] = [x_1].$ Hence, $X/M$ is a vector space over $F$. 

$n \in \mathbb{N}$. Case 1: $n = 0$. Then, $X$ is a singleton set. Clearly $K$ is compact. Case 2: $n \in \mathbb{N}$. Let $S = \{ e_1, \ldots, e_n \}$. $\forall x \in X$, $x$ can be uniquely expressed as $\sum_{i=1}^{n} \alpha_i e_i$ for some $\alpha_1, \ldots, \alpha_n \in K$. This allows us to define a bijective mapping $\psi : X \to \mathbb{K}^n$ by $\psi(x) = (\alpha_1, \ldots, \alpha_n), \forall x = \sum_{i=1}^{n} \alpha_i e_i \in X$. Define a natural homomorphism $\psi : X \to X$ by $\psi(x) = (\alpha_1, \ldots, \alpha_n), \forall x = \sum_{i=1}^{n} \alpha_i e_i \in X$.
Clearly, $\phi$ is a linear function. This completes the proof of the proposition. \hfill $\blacksquare$

**Proposition 7.44** Let $X := (X, K, \| \cdot \|_X)$ be a normed linear space and $M \subseteq X$ be a closed subspace, and $X/M$ be the quotient space of $X$ modulo $M$. Define a norm $\| \cdot \|$ on $X/M$ by, $\forall [x] \in X/M, \| [x] \| := \inf_{m \in M} \| x - m \|_X = \text{dist}(x, M)$. Then, $X/M := (X/M, K, \| \cdot \|)$ is a normed linear space, which will be called the quotient space of $X$ modulo $M$.

**Proof** By Proposition 7.43, $X/M$ is a vector space over $K$. Here, we only need to show that $\| \cdot \|$ defines a norm on $X/M$. First, we show that $\| \cdot \| : X/M \to [0, \infty) \subset R$ is uniquely defined. $\forall [x] \in X/M, \forall y \in [x], y - x \in M$. Then, we have $\| [y] \| = \inf_{m \in M} \| y - m \|_X = \inf_{m \in M} \| x - y + y - m \|_X = \inf_{m \in M} \| x - m \|_X = \| [x] \| \leq \| x \|_X < +\infty$. Hence, $\| \cdot \|$ is uniquely defined.

Next, we show that $\| \cdot \|$ defines a norm on $X/M$. $\forall [x_1], [x_2] \in X/M, \forall \alpha \in K$, we have (i) $\| [x_1] \| \subseteq [0, \infty) \subset R$, since $\| [x_1] \| \leq \| x_1 \|_X < +\infty$. [i] $= \| [y] \| = \inf_{m \in M} \| x_1 - x_2 - m \|_X$. Note that $\| [x_1] \| = \inf_{m \in M} \| x_1 - m \|_X$, $\forall \epsilon \in (0, \infty) \subset R, \forall i = 1, 2, \exists m_i \in M$ such that $\| x_i - m_i \|_X < \| [x_i] \| + \epsilon$. Then, we have $\| x_1 + x_2 - m_1 - m_2 \|_X \leq \| x_1 - m_1 \|_X + \| x_2 - m_2 \|_X \leq \| [x_1] \| + \| [x_2] \| + 2\epsilon$. Hence, we have $\| [x_1] \| + \| [x_2] \| \leq \| [x_1] \| + \| [x_2] \| + 2\epsilon$. By the arbitrariness of $\epsilon$, we have $\| [x_1] \| + \| [x_2] \| \leq \| [x_1] \| + \| [x_2] \|$. (iii) $\| \alpha [x_1] \| = \| \alpha \|_X [x_1] = \| \alpha x_1 \| = \inf_{m \in M} \| \alpha x_1 - m \|_X$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$; Case 2: $\alpha \neq 0$. Case 1: $\alpha = 0$. $\| \alpha [x_1] \| = \inf_{m \in M} \| m \|_X = 0 = \| \alpha \| \| [x_1] \|$. Case 2: $\alpha \neq 0$. $\| \alpha [x_1] \| = \inf_{m \in M} \| \alpha x_1 - m \|_X = \inf_{m \in M} \| x_1 - m \|_X = \| \alpha \| \| [x_1] \|$. Hence, in both cases, we have $\| \alpha [x_1] \| = \| \alpha \| \| [x_1] \|$. This shows that $\| \cdot \|$ is a norm on $X/M$. Hence, $X/M$ is a normed linear space. This completes the proof of the proposition. \hfill $\blacksquare$

**Proposition 7.45** Let $X$ be a Banach space and $M \subseteq X$ be a closed subspace. Then, the quotient space $X/M$ is a Banach space.

**Proof** By Proposition 7.44, $X/M$ is a normed linear space. Here, we need only to show that $X/M$ is complete. Let $\| \cdot \|_X$ be the norm on $X$ and $\| \cdot \|$ be the norm on $X/M$. We will prove this using Proposition 7.27. Fix any absolutely summable series $(\{x_n\})_{n=1}^\infty \subseteq X/M$. Then, $\sum_{n=1}^\infty \| x_n \|_X \in R, \forall n \in N, \exists y_n \in \{x_n\}$ such that $\| y_n \|_X < \| x_n \| + 2^{-n}$. Then, $\sum_{i=1}^\infty \| y_i \|_X < \sum_{i=1}^\infty (\| x_i \| + 2^{-i}) = \sum_{i=1}^\infty \| x_i \| + 1$. Then, $(y_n)_{n=1}^\infty \subseteq X$ is absolutely summable. By Proposition 7.27, $\sum_{n=1}^\infty y_n = y \in X$, since $X$ is complete. Now, it is easy to show that $\sum_{n=1}^\infty x_n = \sum_{n=1}^\infty y_n = [y] \in X/M$. Hence, $X/M$ is complete by Proposition 7.27. This completes the proof of the proposition. \hfill $\blacksquare$
7.7. **THE STONE-WEIERSTRASS THEOREM**

**Definition 7.46** Let \((\mathcal{X}, \mathbb{K})\) be a vector space and \(\|\cdot\| : \mathcal{X} \to [0, \infty) \subseteq \mathbb{R}\). Assume that \(\|\cdot\|\) satisfies (ii) and (iii) of Definition 7.1 and \(\|\varnothing\| = 0\), but not necessarily (i). Then, \(\|\cdot\|\) is called a pseudo-norm.

**Proposition 7.47** Let \((\mathcal{X}, \mathbb{K})\) be a vector space and \(\|\cdot\| : \mathcal{X} \to [0, \infty) \subseteq \mathbb{R}\) be a pseudo-norm on \(\mathcal{X}\). Then, the set \(M := \{x \in \mathcal{X} \mid \|x\| = 0\}\) is a subspace of \((\mathcal{X}, \mathbb{K})\). On the quotient space \(\mathcal{X}/M\), define a norm \(\|\cdot\|_{1} : \mathcal{X}/M \to [0, \infty) \subseteq \mathbb{R}\) by \(\|x\|_{1} = \|x\|, \forall x \in \mathcal{X}/M\). Then, the space \((\mathcal{X}/M, \mathbb{K}, \|\cdot\|_{1})\) is a normed linear space, which will be called the quotient space of \((\mathcal{X}, \mathbb{K})\) modulo \(\|\cdot\|\).

**Proof** We need the following claim.

**Claim 7.47.1** \(\forall x \in \mathcal{X}, \forall m \in M\), we have \(\|x\| = \|x + m\|\).

**Proof of claim:** Note that \(\|x\| \leq \|x + m\| + \|−m\| = \|x + m\| \leq \|x\| + \|m\| = \|x\|\). This completes the proof of the claim. 

Clearly, \(\varnothing \in M \neq \emptyset\). \(\forall m_{1}, m_{2} \in M, \forall \alpha \in \mathbb{K}\), we have \(\alpha m_{1} \in M\), by the properties of the pseudo-norm, and \(m_{1} + m_{2} \in M\) by Claim 7.47.1. Hence, \(M\) is a subspace of \((\mathcal{X}, \mathbb{K})\). By Proposition 7.43, \(\mathcal{X}/M\) is a vector space over \(\mathbb{K}\).

By Claim 7.47.1, \(\|\cdot\|_{1} : \mathcal{X}/M \to [0, \infty) \subseteq \mathbb{R}\) is uniquely defined.

Next, we show that \(\|\cdot\|_{1}\) is a norm on \(\mathcal{X}/M\). \(\forall [x_{1}], [x_{2}] \in \mathcal{X}/M, \forall \alpha \in \mathbb{K}\), \(\|x_{1}\|_{1} = \|x_{1}\| \in [0, \infty) \subseteq \mathbb{R}\). \([x_{1}] = [\vartheta_{\mathcal{X}}]\) implies that \(\|x_{1}\|_{1} = \|\vartheta_{\mathcal{X}}\|_{1} = 0\). \(\|x_{1}\|_{1} = 0\) implies that \(\|x_{1}\| = 0\) and \(x_{1} \in M\), which further implies that \(x_{1} \in [\vartheta_{\mathcal{X}}]\) and \([x_{1}] = [\vartheta_{\mathcal{X}}]\). \(\|x_{1} + x_{2}\|_{1} = \|x_{1} + x_{2}\| \leq \|x_{1}\| + \|x_{2}\| = \|x_{1}\|_{1} + \|x_{2}\|_{1}\). \(\alpha [x_{1}]_{1} = \|\alpha x_{1}\|_{1} = \|\alpha x_{1}\| = \|\alpha\| \|x_{1}\|_{1}\). \(\|x_{1}\|_{1}\). \(\|\cdot\|_{1}\) is a norm on \(\mathcal{X}/M\). Therefore, \((\mathcal{X}/M, \mathbb{K}, \|\cdot\|_{1})\) is a normed linear space. This completes the proof of the proposition.

---

**7.7 The Stone-Weierstrass Theorem**

**Definition 7.48** Let \(X\) be a set, \(f : X \to \mathbb{R}\), and \(g : X \to \mathbb{R}\). \(f \vee g : X \to \mathbb{R}\) and \(f \wedge g : X \to \mathbb{R}\) are defined by \((f \vee g)(x) = \max\{f(x), g(x)\}\) and \((f \wedge g)(x) = \min\{f(x), g(x)\}\), \(\forall x \in \mathcal{X}\).

**Proposition 7.49** Let \(\mathcal{X}\) be a topological space and \(f : \mathcal{X} \to \mathbb{R}\) and \(g : \mathcal{X} \to \mathbb{R}\) be continuous at \(x_{0} \in \mathcal{X}\). Then, \(f \vee g\) and \(f \wedge g\) are continuous at \(x_{0}\). If furthermore \(f\) and \(g\) are continuous, then \(f \vee g\) and \(f \wedge g\) are continuous.

**Proof** This is straightforward. 

**Example 7.50** Let \(\mathcal{X}\) be a topological space and \(\mathcal{Y}\) be a normed linear space over the field \(\mathbb{K}\). Let \((\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})\) be the vector space defined in
Example 6.20. Let $V := \{ f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid f \text{ is continuous} \}$. Then, by Propositions 3.12, 3.32, 6.25, and 7.23, $V$ is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})$. This subspace will be denoted by $\mathcal{C}_v(\mathcal{X}, \mathcal{Y})$.

**Definition 7.51** Let $\mathcal{X}$ be a topological space, $\mathcal{C}_v(\mathcal{X}, \mathbb{R})$ be the vector space defined in Example 7.50. $L \subseteq \mathcal{C}_v(\mathcal{X}, \mathbb{R})$ is said to be a lattice if $L \neq \emptyset$ and $\forall f, g \in L, f \lor g \in L$ and $f \land g \in L$. A subspace $M \subseteq \mathcal{C}_v(\mathcal{X}, \mathbb{R})$ is said to be an algebra if $\forall f, g \in M, fg \in M$, where $fg$ is the product of functions $f$ and $g$.

**Proposition 7.52** Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space, $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be a lattice. Assume $h : \mathcal{X} \to \mathbb{R}$ defined by $h(x) = \inf_{f \in L} f(x)$, $\forall x \in \mathcal{X}$ is continuous. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists g \in L$ such that $0 \leq g(x) - h(x) < \epsilon$, $\forall x \in \mathcal{X}$.

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since $L$ is a lattice and therefore nonempty. Then, the result holds. Case 2: $\mathcal{X} \neq \emptyset$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{X}$, since $h(x) \in \mathbb{R}$, then $\exists f_x \in L$ such that $0 \leq f_x(x) - h(x) < \epsilon$. Since $f_x$ and $h$ are continuous, then $\exists O_x \in \mathcal{O}$ with $x \in O_x$ such that $\forall y \in O_x$, we have $|f_x(y) - f_x(x)| < \epsilon/3$ and $|h(y) - h(x)| < \epsilon/3$. Then, $|f_x(y) - h(y)| \leq |f_x(y) - f_x(x)| + |f_x(x) - h(x)| + |h(x) - h(y)| < \epsilon$, $\forall y \in O_x$. Then, $\mathcal{X} \subseteq \bigcup_{x \in \mathcal{X}} O_x$. Since $\mathcal{X}$ is compact, there then exists a finite set $X_N \subseteq \mathcal{X}$ such that $\mathcal{X} \subseteq \bigcup_{x \in X_N} O_x$. Since $\mathcal{X} \neq \emptyset$ then $X_N$ must be nonempty. Let $g := \bigwedge_{x \in X_N} f_x \in L$. $\forall x \in \mathcal{X}$, $\exists x_0 \in X_N$ such that $x \in O_{x_0}$ and $0 \leq g(x) - h(x) \leq f_{x_0}(x) - h(x) < \epsilon$. This completes the proof of the proposition.

**Proposition 7.53** Let $\mathcal{X} := (X, \mathcal{O})$ be a compact space, $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ be a lattice satisfying the following conditions:

(i) $L$ separates points, that is $\forall x, y \in \mathcal{X}$ with $x \neq y$, $\exists f \in L$ such that $f(x) \neq f(y)$;

(ii) $\forall f \in L, \forall c \in \mathbb{R}$, $c + f, cf \in L$.

Then, $\forall h \in \mathcal{C}(\mathcal{X}, \mathbb{R})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists $g \in L$ such that $0 \leq g(x) - h(x) < \epsilon$, $\forall x \in \mathcal{X}$.

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $\mathcal{C}(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = \mathcal{C}(\mathcal{X}, \mathbb{R})$ since $L$ is a lattice and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \neq \emptyset$. We need the following two results.
Claim 7.53.1 ∀a, b ∈ ℝ, ∀x₁, x₂ ∈ ℱ with x₁ ≠ x₂, ∃f ∈ L such that f(x₁) = a and f(x₂) = b.

Proof of claim: Let g ∈ L be such that g(x₁) ≠ g(x₂), which exists by (i). Let

\[ f = \frac{a - b}{g(x_1) - g(x_2)}g + \frac{bg(x_1) - ag(x_2)}{g(x_1) - g(x_2)} \in L \]

Then, f is the desired function. □

Claim 7.53.2 ∀a, b ∈ ℝ with a ≤ b, ∀ closed set F ⊆ ℱ, ∀x₀ ∈ ℱ with x₀ ∉ F, ∃f ∈ L such that f(x₀) = a, f(x) ≥ a, ∀x ∈ ℱ, and f(x) > b, ∀x ∈ F.

Proof of claim: ∀x ∈ F, we have x ≠ x₀. By Claim 7.53.1, ∃fₓ ∈ L such that fₓ(x₀) = a and fₓ(x) = b + 1. Let Oₓ := {x ∈ ℱ | fₓ(x) > b}. Then, Oₓ ∈ ℰ since fₓ is continuous. Clearly, x ∈ Oₓ. F ⊆ ∪ₓ∈F Oₓ. By Proposition 5.5, F is compact. Then, there exists a finite set Fₙ ⊆ F such that F ⊆ ∪ₓ∈Fₙ Oₓ. Take g ∈ L ∉ 0. Then, a = 0g + a ∈ L by (ii). Let f := a ∨ (∩ₓ∈Fₙ fₓ) ∈ L. Clearly, f(x₀) = a and f(x) ≥ a, ∀x ∈ ℱ. ∀x ∈ F, ∃x₀ ∈ Fₙ such that x ∈ Oₓ₀. Then, f(x) ≥ fₓ₀(x) > b. Hence, f is the desired function. This completes the proof of the claim. □

∀h ∈ ℂ(ℱ, ℝ), let Ļ := {f ∈ L | f(x) ≥ h(x), ∀x ∈ ℱ}. By the continuity of h, the compactness of ℱ, and Proposition 5.29, ∃b ∈ ℝ such that h(x) ≤ b, ∀x ∈ ℱ. Then, the constant function b ∈ Ļ ∉ 0 (b ∈ L by (iii)). ∀f₁, f₂ ∈ Ļ, it is easy to show that f₁ ∨ f₂ ∈ Ļ and f₁ ∧ f₂ ∈ Ļ. Hence, Ļ is a lattice. We will show that h(x) = infₓ∈Ľ f(x), ∀x ∈ ℱ. Then, the result follows from Proposition 7.52. ∀x₀ ∈ ℱ, ∀η ∈ (0, ∞) ⊂ ℝ, let Fₓ₀,η := {x ∈ ℱ | h(x) ≥ h(x₀) + η}. By the continuity of h and Proposition 3.10, Fₓ₀,η is closed. Clearly, x₀ ∉ Fₓ₀,η. Now, by Claim 7.53.2, ∃fₓ₀,η ∈ L such that fₓ₀,η(x₀) = h(x₀) + η, fₓ₀,η(x) ≥ h(x₀) + η, ∀x ∈ ℱ, and fₓ₀,η(x) > b, ∀x ∈ Fₓ₀,η. It is clear that fₓ₀,η(x) > h(x), ∀x ∈ ℱ. Hence, fₓ₀,η ∈ Ļ. Then, infₓ∈Ľ f(x₀) ≤ fₓ₀,η(x₀) = h(x₀) + η. By the definition of Ļ and the arbitrariness of η, we have h(x₀) ≤ infₓ∈Ľ f(x₀) ≤ h(x₀). Hence, we have h(x₀) = infₓ∈Ľ f(x₀).

This completes the proof of the proposition. □

Lemma 7.54 ∀ε ∈ (0, ∞) ⊂ ℝ, there exists a polynomial P(s) in one variable such that |P(s) − |s|| < ε, ∀s ∈ [−1, 1] ⊂ ℝ.

Proof This result is a special case of Bernstein Approximation Theorem (Bartle, 1976, pg. 171). For an generalization of the Bernstein Approximation Theorem to multi-variable case, see Page 531. □

Lemma 7.55 Let ℱ be a nonempty compact space and ℱ ⊆ ℂ(ℱ, ℝ) be an algebra. Then, ℱ is an algebra.
Then, \( \| \cdot \| \) is an algebra. Furthermore, let \( h = f - g \in \mathcal{A} \). By Lemma 7.55, \( \mathcal{A} \) is an algebra and satisfies

(i) \( \mathcal{A} \) separates points, that is, \( \forall \bar{x}, \bar{y} \in \mathcal{X} \) with \( \bar{x} \neq \bar{y} \), then \( \exists f \in \mathcal{A} \) such that \( \bar{x} f = \bar{y} f \);

(ii) \( \mathcal{A} \) contains all constant functions.

Then, \( \mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R}) \), that is \( \mathcal{A} \) is dense in \( \mathcal{C}(\mathcal{X}, \mathbb{R}) \).

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{X} = \emptyset \); Case 2: \( \mathcal{X} \neq \emptyset \). Case 1: \( \mathcal{X} = \emptyset \). \( \mathcal{C}(\mathcal{X}, \mathbb{R}) \) is a singleton set and \( \mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R}) \) since \( \mathcal{A} \) is a subspace and therefore nonempty.

Then, the result holds.

Case 2: \( \mathcal{X} \neq \emptyset \). By Lemma 7.55, \( \mathcal{A} \) is an algebra. We need the following claim.

**Claim 7.56.1** \( \forall f \in \mathcal{A} \), then \( \| f \| \in \mathcal{A} \).

**Proof of claim:** Fix \( f \in \mathcal{A} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists g \in \mathcal{A} \) such that \( \| f - g \| < \epsilon / 2 \). By Proposition 5.29 and Definitions 5.1 and 5.21, \( \exists N \in (0, \infty) \subset \mathbb{R} \) such that \( \| g \| < N \). Then, \( g / N \in \mathcal{A} \) since \( \mathcal{A} \) is an algebra. By Lemma 7.54, \( \exists \) a polynomial \( P(s) \) such that \( | P(s) - s | < \frac{\epsilon}{2N} \), \( \forall s \in [-1, 1] \subset \mathbb{R} \). Hence, we have \( P \circ (g / N) \in \mathcal{A} \) since \( \mathcal{A} \) is an algebra. Furthermore, \( | P(g(x) / N) - g(x) / N | < \frac{\epsilon}{2N} \), \( \forall x \in \mathcal{X} \). Let \( h = NP \circ (g / N) \in \mathcal{A} \). Then, we have \( \| h - g \| < \epsilon / 2 \). Note that \( \| f - h \| \leq \| f - g \| + \| g - h \| \leq \| f - g \| + \| g - h \| < \epsilon \). Hence, \( \| f \| \in \mathcal{A} \), by the arbitrariness of \( \epsilon \). This completes the proof of the claim. \( \square \)

\( \forall f, g \in \mathcal{A} \), we have

\[
\begin{align*}
f \vee g &= \frac{1}{2}(f + g) + \frac{1}{2} | f - g | \in \mathcal{A} \\
f \wedge g &= \frac{1}{2}(f + g) - \frac{1}{2} | f - g | \in \mathcal{A}
\end{align*}
\]
Hence, $\mathcal{A}$ is a lattice. Clearly, $\mathcal{A}$ separates points since $\mathcal{A}$ separates points.

By Proposition 7.53, $\forall h \in C(X, \mathbb{R})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists $g \in \mathcal{A}$ such that $0 \leq g(x) - h(x) < \epsilon$, $\forall x \in X$. Then, $\|g - h\| < \epsilon$, and $h \in \overline{\mathcal{A}} = \mathcal{A}$, where the last equality follows from Proposition 3.3. Hence, we have $\overline{\mathcal{A}} = C(X, \mathbb{R})$. This completes the proof of the theorem. 

**Corollary 7.57** Let $X \subseteq \mathbb{R}^n$ be a closed and bounded set with the subset topology $\mathcal{O}$, where $n \in \mathbb{Z}^+$, $X := (X, \mathcal{O})$, and $f : X \to \mathbb{R}$ be a continuous function. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P : X \to \mathbb{R}$ in $n$ variables such that $|f(x) - P(x)| < \epsilon$, $\forall x \in X$.

**Proof** By Proposition 5.40, $X$ is a compact space. Let $C(X, \mathbb{R})$ be the Banach space defined in Example 7.32. Then, $f \in C(X, \mathbb{R})$. Let $\mathcal{A}$ be all polynomial in $n$ variables on $X$. Clearly, $\mathcal{A} \subseteq C(X, \mathbb{R})$ and is a linear subspace of $C(X, \mathbb{R})$. Then, it is easy to show that $\mathcal{A}$ is an algebra. $\forall x_1, x_2 \in X$ with $x_1 \neq x_2$, there exists a coordinate $i_0 \in \{1, \ldots, n\}$ such that $\pi_{i_0}(x_1) \neq \pi_{i_0}(x_2)$. Then, the polynomial $f \in \mathcal{A}$ given by $f(x) = \pi_{i_0}(x)$, $\forall x \in X$, separates $x_1$ and $x_2$. Hence, $\mathcal{A}$ separates points. Clearly, $\mathcal{A}$ contains all the constant functions. By the Stone-Weierstrass Theorem, $\overline{\mathcal{A}} = C(X, \mathbb{R})$. This completes the proof of the corollary. 

**Corollary 7.58** Let $S_1 = [0, 2\pi] \subset \mathbb{R}$ and $f \in C(S_1, \mathbb{R})$. Assume that $f(0) = f(2\pi)$. Let $\mathcal{M} = \{g \in C(S_1, \mathbb{R}) \mid g(x) = \cos(nx), \forall x \in S_1$ or $g(x) = \sin(nx), \forall x \in S_1, \text{ where } n \in \mathbb{Z}^+\}$. Let $\mathcal{A} = \text{ span } (\mathcal{M}) \subseteq C(S_1, \mathbb{R})$. Then, $f \in \overline{\mathcal{A}}$.

**Proof** Let $S_2 \subset \mathbb{R}^2$ be the unit circle: $S_2 := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 = 1\}$. Define a mapping $\Psi : S_1 \to S_2$ by $\Psi(x) = (\cos(x), \sin(x))$, $\forall x \in S_1$. Clearly, $\Psi$ is surjective and continuous. Note that $S_1$ is compact and $S_2$ is Hausdorff. It is obvious that we may define a function $\Phi : S_2 \to \mathbb{R}$ such that $\Phi \circ \Psi = f$, which is continuous. By Proposition 5.18, we have $\Phi$ is continuous. Note that $S_2$ is closed and bounded in $\mathbb{R}^2$, then $S_2$ is compact by Proposition 5.40. Hence, $\Phi \in C(S_2, \mathbb{R})$. By Corollary 7.57, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, there exists a polynomial $P : S_2 \to \mathbb{R}$ such that $|\Phi(y_1, y_2) - P(y_1, y_2)| < \epsilon$, $\forall y_1, y_2 \in S_2$. Then, $|f(x) - P \circ \Psi(x)| = |\Phi \circ \Psi(x) - P \circ \Psi(x)| < \epsilon$, $\forall x \in S_1$. Note that $P \circ \Psi \in \mathcal{A}$, since, $\forall \gamma, \theta \in \mathbb{R}$,

\[
\begin{align*}
(sin(\theta))^2 &= \frac{1}{2} (1 - cos(2\theta)); \\
(cos(\theta))^2 &= \frac{1}{2} (1 + cos(2\theta)); \\
\sin(\gamma) \cos(\theta) &= \frac{1}{2} (\sin(\gamma + \theta) + \sin(\gamma - \theta)); \\
\sin(\gamma) \sin(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) - \cos(\gamma + \theta)); \\
\cos(\gamma) \cos(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) + \cos(\gamma + \theta))
\end{align*}
\]

Hence, $f \in \overline{\mathcal{A}}$, by the arbitrariness of $\epsilon$. This completes the proof of the corollary. 

□
For ease of presentation below, we will define two functions: \( \text{sqr} : \mathbb{R} \rightarrow \mathbb{R} \) by \( \text{sqr}(x) = x^2, \forall x \in \mathbb{R} \), and \( \sqrt{\cdot} : [0, \infty) \rightarrow [0, \infty) \) by \( \sqrt{x} = \sqrt{x}, \forall x \in [0, \infty) \).

**Corollary 7.59** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a compact space and \( \mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R}) \) be an algebra that satisfies

(i) \( \mathcal{A} \) separates points;

(ii) \( \exists f_0 \in \mathcal{A} \text{ such that } f_0(x) \neq 0, \forall x \in \mathcal{X}. \)

Then, the constant function \( 1 \in \overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R}). \)

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{X} = \emptyset \); Case 2: \( \mathcal{X} \neq \emptyset \). Case 1: \( \mathcal{X} = \emptyset \). \( \mathcal{C}(\mathcal{X}, \mathbb{R}) \) is a singleton set and \( \mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R}) \) since \( \mathcal{A} \) is a subspace and therefore nonempty. Then, the result holds.

Case 2: \( \mathcal{X} \neq \emptyset \). By Lemma 7.55, \( \overline{\mathcal{A}} \) is an algebra.

Since \( \mathcal{A} \) is an algebra, then, \( \text{sqr} \circ f_0 \in \mathcal{A} \), which satisfies that \( \text{sqr} \circ f_0(x) > 0, \forall x \in \mathcal{X} \). Furthermore, \( g_0 := \text{sqr} \circ f_0/\|\text{sqr} \circ f_0\| \in \mathcal{A} \), which satisfies \( g_0 : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R} \).

By Corollary 7.57, \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), there exists a polynomial \( Q_\varepsilon \) in one variable such that \( |\sqrt{s} - Q_\varepsilon(s)| < \varepsilon, \forall s \in [0, 1] \subset \mathbb{R} \), and \( Q_\varepsilon(0) = 0 \).

\( \forall f : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R} \) with \( f \in \overline{\mathcal{A}}, Q_\varepsilon \circ f \in \overline{\mathcal{A}} \) since \( \overline{\mathcal{A}} \) is an algebra and \( Q_\varepsilon(0) = 0 \). Then, \( \text{sqr} \circ f \in \overline{\mathcal{A}} = \overline{\mathcal{A}} \). Recursively, we may conclude that \( \text{sqr}^n \circ g_0 \in \overline{\mathcal{A}}, \forall n \in \mathbb{N} \).

By Proposition 5.29, \( \exists x_m \in \mathcal{X} \) such that \( g_0(x) \geq g_0(x_m) := \gamma > 0, \forall x \in \mathcal{X} \). \( \forall \varepsilon \in (0, 1) \subset \mathbb{R}, \exists n_0 \in \mathbb{N} \) such that \( |1 - \gamma^{2^{-n_0}}| < \varepsilon \). Then, \( \|1 - \text{sqr}^{n_0} \circ g_0\| < \varepsilon \). Hence, the constant function \( 1 \in \overline{\mathcal{A}} = \overline{\mathcal{A}} \).

Clearly, \( \overline{\mathcal{A}} \) separates points since \( \mathcal{A} \) does. Hence, by Theorem 7.56, \( \mathcal{C}(\mathcal{X}, \mathbb{R}) = \overline{\mathcal{A}} = \overline{\mathcal{A}} \). This completes the proof of the corollary. \( \square \)

**Proposition 7.60** Let \( \mathcal{X} \) be a compact space and \( \mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R}) \) be an algebra that separates points on \( \mathcal{X} \). Then, either \( \overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R}) \) or \( \exists x \in \mathcal{X} \) such that \( \overline{\mathcal{A}} = \{ f \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid f(x) = 0 \} \).

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{X} = \emptyset \); Case 2: \( \mathcal{X} \neq \emptyset \). Case 1: \( \mathcal{X} = \emptyset \). \( \mathcal{C}(\mathcal{X}, \mathbb{R}) \) is a singleton set and \( \mathcal{A} = \mathcal{C}(\mathcal{X}, \mathbb{R}) \) since \( \mathcal{A} \) is a subspace and therefore nonempty. Then, the result holds.

Case 2: \( \mathcal{X} \neq \emptyset \). By Lemma 7.55, \( \overline{\mathcal{A}} \) is an algebra. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: \( 1 \in \overline{\mathcal{A}} \); Case 2b: \( 1 \notin \overline{\mathcal{A}} \). Case 2a: \( 1 \in \overline{\mathcal{A}} \). Clearly, \( \overline{\mathcal{A}} \) is an algebra that separates points. By Stone-Weierstrass Theorem, \( \mathcal{C}(\mathcal{X}, \mathbb{R}) = \overline{\mathcal{A}} = \overline{\mathcal{A}} \). Then, the result holds.

Case 2b: \( 1 \notin \overline{\mathcal{A}} \). Then, \( \overline{\mathcal{A}} \subset \mathcal{C}(\mathcal{X}, \mathbb{R}) \). By Corollary 7.59, \( \forall f \in \overline{\mathcal{A}} \), there exists \( x \in \mathcal{X} \) such that \( f(x) = 0 \). Then, we have the following claims.
7.7. THE STONE-WEIERSTRASS THEOREM

Claim 7.60.1 \( \forall f \in \mathcal{A}, |f| \in \mathcal{A} \). Hence, \( \mathcal{A} \) is a lattice.

Proof of claim: \( \forall f \in \mathcal{A} \), by the compactness of \( \mathcal{X} \) and Proposition 5.29, \( \exists M \in (0, \infty) \subset \mathbb{R} \) such that \( \| f \| \leq M \). Let \( g := f/M \in \mathcal{A} \). Then, \( g : \mathcal{X} \to [-1, 1] \subset \mathbb{R}, \forall x \in (0, \infty) \subset \mathbb{R} \), by Lemma 7.54, there exists a polynomial \( P : [-1, 1] \to \mathbb{R} \) with \( P(0) = 0 \) such that \( |P(s) - s| < \epsilon/M, \forall s \in [-1, 1] \subset \mathbb{R} \). Since \( \mathcal{A} \) is an algebra, then \( M \cdot P \circ g \in \mathcal{A} \) and \( \| f - M \cdot P \circ g \| = M \| g - P \circ g \| < \epsilon \). Hence, \( |f| \in \mathcal{A} = \mathcal{A} \).

\( \forall f, g \in \mathcal{A} \), we have

\[
\begin{align*}
\forall g \in \mathcal{A}:
 f \land g &= \frac{1}{2}(f + g) - \frac{1}{2}|f - g| \in \mathcal{A} \\
 f \lor g &= \frac{1}{2}(f + g) + \frac{1}{2}|f - g| \in \mathcal{A}
\end{align*}
\]

Hence, \( \mathcal{A} \) is a lattice. This completes the proof of the claim. \( \square \)

Claim 7.60.2 \( \exists x_0 \in \mathcal{X} \) such that \( f(x_0) = 0, \forall f \in \mathcal{A} \).

Proof of claim: We will prove this using an argument of contradiction. Suppose the claim is false. \( \forall x \in \mathcal{X}, \exists f_x \in \mathcal{A} \) such that \( f_x(x) \neq 0 \). By Claim 7.60.1, \( |f_x| \in \mathcal{A} \) and \( |f_x(x)| > 0 \). Let \( O_x := \{ x \in \mathcal{X} \mid |f_x(x)| > 0 \} \). Since \( f_x \in \mathcal{A} \subseteq C(\mathcal{X}, \mathbb{R}) \), then \( x \in O_x \in \mathcal{O}_X \). Hence, we have \( \mathcal{X} \subseteq \bigcup_{x \in \mathcal{X}} O_x \). By the compactness of \( \mathcal{X} \), there exists a finite set \( X_N \subseteq \mathcal{X} \) such that \( \mathcal{X} \subseteq \bigcup_{x \in X_N} O_x \). Clearly, \( X_N \neq \emptyset \) since \( \mathcal{X} \neq \emptyset \). Let \( f := \sum_{x \in X_N} |f_x| \in \mathcal{A} \). Hence, \( f(x) > 0, \forall x \in \mathcal{X} \). By Corollary 7.59, we have \( 1 \in \mathcal{A} = \mathcal{A} \), which is a contradiction. Hence, the claim is true. This completes the proof of the claim. \( \square \)

Claim 7.60.3 \( \forall \) closed set \( F \subseteq \mathcal{X} \) with \( x_0 \notin F \), we have

(i) \( \text{let } A_F := \{ h \in C(F, \mathbb{R}) \mid \exists f \in \mathcal{A} \text{ such that } h = f|_F \}, \text{ then } \overline{A_F} = C(F, \mathbb{R}) \);

(ii) \( \forall \epsilon \in (0, 1) \subset \mathbb{R}, \exists g \in \mathcal{A} \text{ such that } g : \mathcal{X} \to [0, 1] \subset \mathbb{R} \text{ and } g(x) \geq 1 - \epsilon, \forall x \in F \).

Proof of claim: \( \text{Clearly, } A_F \text{ is an algebra since } \mathcal{A} \text{ is an algebra.} \)

Furthermore, \( A_F \) separate points on \( F \) since \( \mathcal{A} \) separates points on \( \mathcal{X} \).

By Proposition 5.5, \( F \) with the subset topology is compact.

\( \forall x \in F, \) we have \( x \neq x_0 \). \( \exists f_x \in \mathcal{A} \) such that \( f_x(x) \neq f_x(x_0) = 0 \). By Claim 7.60.1, \( |f_x| \in \mathcal{A} \) and \( |f_x(x)| > 0 \). Let \( O_x := \{ x \in \mathcal{X} \mid |f_x(x)| > 0 \} \). Then, we have \( x \in O_x \in \mathcal{O}_X \). Hence, \( F \subseteq \bigcup_{x \in F} O_x \). By the compactness of \( F \), there exists a finite set \( F_N \subseteq F \) such that \( F \subseteq \bigcup_{x \in F_N} O_x \). Let \( f := \sum_{x \in F_N} |f_x| \). Then, \( f \in \mathcal{A} \) and \( f(x) > 0, \forall x \in F \).

Then, \( h := f|_F \in A_F \) and \( h(x) > 0, \forall x \in F \). By Corollary 7.59, we have \( \overline{A_F} = C(F, \mathbb{R}) \). Hence, (i) is true.
∃M ∈ (0, ∞) ⊂ R such that ∥f∥ ≤ M. Then, go := f/M ∈ A and go : X → [0, 1] ⊂ R.

By Corollary 7.57, ∀ε ∈ (0, ∞) ⊂ R, there exists a polynomial Qε in one variable such that |√s − Qε(s)| < ε, ∀s ∈ [0, 1] ⊂ R, and Qε(0) = 0.

∀h : X → [0, 1] ⊂ R with h ∈ A, Qε ◦ h ∈ A since A is an algebra and Qε(0) = 0. Then, sqrt ◦ h ∈ A. Recursively, we may conclude that sqrtn ◦ go ∈ A, ∀n ∈ N.

By the compactness of F and Proposition 5.29, ∃γ ∈ (0, 1] ⊂ R such that g0(x) ≥ γ, ∀x ∈ F. ∀ε ∈ (0, 1] ⊂ R, ∃γ0 ∈ N such that γεγ0 ≥ 1 − ε. Then, g := sqrtn ◦ go ∈ A, g : X → [0, 1] ⊂ R and g(x) ≥ 1 − ε, ∀x ∈ F. Hence, (ii) is true.

This completes the proof of the claim. □

Claim 7.60.4 ∀g ∈ C(X, R) with g(x0) = 0 and g(x) ≥ 0, ∀x ∈ X, we have g ∈ A.

Proof of claim:
∀ε ∈ (0, ∞) ⊂ R, let O := {x ∈ X | g(x) < ε/2}. Then, x0 ∈ O ∈ OX since g is continuous. Let M := ∥g∥ ∈ [0, ∞) ⊂ R.

By Claim 7.60.3, ∃h1 ∈ A such that h1 : X → [0, 1] ⊂ R and h1(x) ≥ 2/3, ∀x ∈ O. Note that h1(x0) = 0 by Claim 7.60.2. Let h2 := Mh1/2 ∈ A. Then, we have 0 ≤ h2(x) ≤ 3M/2, ∀x ∈ X, and h2(x) ≥ M ≥ g(x), ∀x ∈ O.

Let U := {x ∈ O | h2(x) < ε/2}. Then, x0 ∈ U ∈ OX since O ∈ OX, h2 is continuous, and h2(x0) = 0. Furthermore, g|U ∈ C(U, R). By Claim 7.60.3, ∃h3 ∈ A such that |h3|U (x) − g|U (x)| < ε, ∀x ∈ U. Define h4 := (h3 ∨ 0) ∧ h2. Then, h4 ∈ A by Claim 7.60.1. ∀x ∈ X, X, we will show that |g(x) − h4(x)| < ε by distinguishing three exhaustive and mutually exclusive cases: Case A: x ∈ U; Case B: x ∈ O \ U; Case C: x ∈ O.

Case A: x ∈ U. Then, we have 0 ≤ h2(x) < ε/2, 0 ≤ g(x) < ε/2. Then, 0 ≤ h4(x) = (h3(x) ∨ 0) ∧ h2(x) ≤ h2(x) < ε/2. Therefore, |g(x) − h4(x)| < ε.

Case B: x ∈ O \ U. Then, we have |h3(x) − g(x)| < ε, h2(x) ≥ 0, and 0 ≤ g(x) < ε/2. Hence, 0 ≤ h3(x) ∨ 0 < g(x) + ε. Then, 0 ≤ h4(x) < g(x) + ε and −ε/2 < −g(x) ≤ h4(x) − g(x) < ε. Therefore, we have |g(x) − h4(x)| < ε.

Case C: x ∈ O. Then, we have |h3(x) − g(x)| < ε and h2(x) ≥ M ≥ g(x) ≥ 0. Hence, g(x) − ε < h3(x) < g(x) + ε, g(x) − ε < h3(x) ∨ 0 < g(x) + ε, and g(x) − ε < h4(x) < g(x) + ε. Therefore, |g(x) − h4(x)| < ε.

Hence, in all three cases, we have |g(x) − h4(x)| < ε. Hence ∥g − h4∥ < ε.

By the arbitrariness of ε, we have g ∈ A = A. □

∀f ∈ C(X, R) with f(x0) = 0, let f+ := f ∨ 0 and f− := (−f) ∨ 0. Clearly, f+, f− ∈ C(X, R), f+(x0) = 0, f−(x0) = 0, and f = f+ − f−. By Claim 7.60.4, f+ ∈ A since A is an algebra. This coupled with Claim 7.60.2, we have that A = {f ∈ C(X, R) | f(x0) = 0}. Hence, the result holds in this case.
This completes the proof of the proposition. □

7.8 Linear Operators

Definition 7.61 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$. A linear operator $A : \mathcal{X} \to \mathcal{Y}$ is said to be bounded if $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_y \leq M \|x\|_x$, $\forall x \in \mathcal{X}$.

Proposition 7.62 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$ and $A : \mathcal{X} \to \mathcal{Y}$ be a linear operator. Then,

(i) if $A$ is bounded then $A$ is uniformly continuous;

(ii) if $A$ is continuous at some $x_0 \in \mathcal{X}$ then $A$ is bounded.

Proof

(i). $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_y \leq M \|x\|_x$, $\forall x \in \mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon/(1 + M) \in (0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in \mathcal{X}$ with $\|x_1 - x_2\|_X < \delta$, we have $\|Ax_1 - Ax_2\|_y = \|A(x_1 - x_2)\|_y \leq M \|x_1 - x_2\|_X < \epsilon$. Hence, $A$ is uniformly continuous.

(ii). $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in B_X(0, \delta)$, we have $\|Ax\|_y < \epsilon$. $\forall \bar{x} \in B_X(0, \delta)$, we have $\|Ax\|_y = \|A(x + \bar{x}) - A\|_y = \max \{\delta \epsilon, \epsilon \} \leq \epsilon$. Hence, $A$ is continuous at $\bar{x}$. Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $\|Ax\|_y < 1$, $\forall x \in B_X(0, \delta_0)$. Then, $\forall x \in \mathcal{X}$, we have two possibilities: (a) $x = \bar{x}$, then $\|Ax\|_y = 0 \leq \frac{\epsilon}{\delta_0} \|x\|_X$; (b) $x \neq \bar{x}$, then $\frac{\delta_0}{2\|x\|_X} x \in B_X(0, \delta_0)$ and we have $\|Ax\|_y = \frac{\delta_0}{2\|x\|_X} A \left( \frac{\delta_0}{2\|x\|_X} x \right) \leq \frac{\delta_0}{2\|x\|_X} \|x\|_X$. Hence, $A$ is bounded.

This completes the proof of the proposition. □

Proposition 7.63 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$. Let $(\mathcal{M}(X, Y), \mathbb{K})$ be the vector space defined in Example 6.20 with the null vector $\emptyset$. Let $P := \{ A \in \mathcal{M}(X, Y) \mid A$ is linear $\}$. Then, $P$ is a subspace of $\mathcal{M}(X, Y)$. Define a functional $\| \cdot \|$ on $P$ by $\| A \| := \inf \{ M \in [0, \infty) \subset \mathbb{R} \mid \|Ax\|_Y \leq M \|x\|_X, \forall x \in \mathcal{X} \}$, $\forall A \in P$. Then, $\forall A \in P$, $\| A \|$ may be equivalently expressed as

$$
\| A \| = \max \left\{ \sup_{x \in \mathcal{X}} \frac{\|Ax\|_Y}{\|x\|_X}, 0 \right\} 
= \max \left\{ \sup_{x \in \mathcal{X}} \frac{\|Ax\|_Y}{\|x\|_X}, 0 \right\} 
\leq \sup_{x \in \mathcal{X}} \frac{\|Ax\|_Y}{\|x\|_X}, 0
$$

Let $N := \{ A \in P \mid \| A \| < +\infty \}$. Then, $(N, \mathbb{K}, \| \cdot \|) =: \mathcal{B}(X, Y)$ is a normed linear space.
Proof

It is easy to check that $P$ is closed under vector addition and scalar multiplication in $\mathcal{M}(\mathcal{X}, \mathcal{Y})$. Clearly, $\emptyset \in P \neq \emptyset$. Hence, $P$ is a subspace of $\mathcal{M}(\mathcal{X}, \mathcal{Y})$.

Next, we prove the four equivalent definitions of $\| A \|$. $\forall A \in P$, define the set $B_A := \{ M \in [0, \infty) \in \mathbb{R} \mid \| A x \|_y \leq M \| x \|_X \}, \forall x \in \mathcal{X} \}. $ Let $\delta_{A1} := \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| A x \|_y$. Since $\| \vartheta x \|_X = 0$ and $\| A \vartheta x \|_y = \| \vartheta y \|_y = 0$, then $\delta_{A1} \geq 0$. $\forall M \in B_A$, we have $\| A x \|_y \leq M \| x \|_X \leq M, \forall x \in \mathcal{X}$ with $\| x \|_X \leq 1$. Hence, we have $\delta_{A1} \leq M$ and $\delta_{A1} \leq \| A \|.$

Let $\delta_{A2} := \max \left\{ \sup_{x \in \mathcal{X}, \| x \|_X = 1} \| A x \|_y, 0 \right\}$. $\forall x \in \mathcal{X}$ with $x \neq \vartheta_x$, we have $\delta_{A1} \geq \| A \left( \frac{x}{\| x \|_X} \right) \|_y = \frac{\| A x \|_y}{\| x \|_X}$. Hence, we have $\delta_{A2} \leq \delta_{A1}$. Let $\delta_{A3} := \max \left\{ \sup_{x \in \mathcal{X}, \| x \|_X = 1} \| A x \|_y, 0 \right\}$. Clearly, we have $0 \leq \delta_{A3} \leq \delta_{A2}$.

On the other hand, $\forall x \in \mathcal{X}$, we have either $x = \vartheta_x$, then $\| A x \|_y = 0 \leq \delta_{A3} \| x \|_X$; or $x \neq \vartheta_x$, then $\| x \|_X > 0$ and $\delta_{A3} \geq \| A \left( \frac{x}{\| x \|_X} \right) \|_y = \frac{\| A x \|_y}{\| x \|_X}$, and hence, $\| A x \|_y \leq \delta_{A3} \| x \|_X$. Thus, $\forall x \in \mathcal{X}$, we have $\| A x \|_y \leq \delta_{A3} \| x \|_X$. If $\delta_{A3} = +\infty$, then $\| A \| \leq \delta_{A3}$. If $\delta_{A3} < +\infty$, then $\delta_{A3} \in B_A$ and $\| A \| \leq \delta_{A3}$. Therefore, $\| A \| \leq \delta_{A3}$.

Hence, $\| A \| = \delta_{A1} = \delta_{A2} = \delta_{A3}$. It is easy to check that $N$ is closed under vector addition and scalar multiplication in $P$. Clearly, $\emptyset \in N \neq \emptyset$. Hence, $N$ is a subspace of $P$.

Finally, we will show that $\| \cdot \|_\varnothing$ defines a norm on $N$. $\forall A, B \in N, \forall \alpha \in \mathbb{K}$.
(i) $\| A \| = 0 \iff A = \emptyset$.
(ii) $\| A + B \| = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| A x \|_y + \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| B x \|_y = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| A x + B x \|_y = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| (A + B) x \|_y$.
(iii) $\| \alpha A \| = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| \alpha A x \|_y = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| A x \|_y = \| A \|$. We have made use of Proposition 3.81 in the fourth equality.

Hence, $B(\mathcal{X}, \mathcal{Y})$ is a normed linear space. \hfill \Box

**Proposition 7.64** Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over the field $\mathbb{K}$, $A \in B(\mathcal{X}, \mathcal{Y})$, and $B \in B(\mathcal{Y}, \mathcal{Z})$. Then, $\forall x \in \mathcal{X}$, we have $\| A x \| \leq \| A \| \| x \|$, where the three norms are over three different normed linear spaces. Furthermore, $\| B A \| \leq \| B \| \| A \|$, where the three norms are over three different normed linear spaces.

**Proof** This is straightforward, and is therefore omitted. \hfill \Box

**Proposition 7.65** Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over the field $\mathbb{K}$ and $B(\mathcal{X}, \mathcal{Y})$ be the normed linear space of bounded linear operators of $\mathcal{X}$
to $y$. Define an operator $\psi : B(\mathcal{X}, Y) \times \mathcal{X} \to Y$ by $\psi(A, x) = Ax$, $\forall A \in B(\mathcal{X}, Y)$, $\forall x \in \mathcal{X}$. Then, $\psi$ is continuous.

**Proof** \quad $\forall (A_0, x_0) \in B(\mathcal{X}, Y) \times \mathcal{X}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon \frac{\epsilon}{\epsilon + \|A_0\|}$, then $\|\psi(A, x) - \psi(A_0, x_0)\| = \|Ax - A_0x_0\| = \|Ax - Ax_0 + Ax_0 - A_0x_0\| \leq \|A(x - x_0)\| + \|A - A_0\|\|x_0\| \leq \|A\|\|x - x_0\| + \|A - A_0\||x_0\| \leq (\|A\| + \|A - A_0\|)\delta + \|x_0\| \delta \leq (\delta + \|A_0\|)\delta + \|x_0\| \delta \leq \epsilon + \epsilon = 2\epsilon$. Hence, $\psi$ is continuous. This completes the proof of the proposition. \hfill $\square$

**Proposition 7.66** Let $\mathcal{X}$ and $Y$ be normed linear spaces over the field $K$. Let $B(\mathcal{X}, Y)$ be the normed linear space defined in Proposition 7.63. If $Y$ is a Banach space, then $B(\mathcal{X}, Y)$ is also a Banach space.

**Proof** \quad All we need to show is that $B(\mathcal{X}, Y)$ is complete. Take a Cauchy sequence $(A_n)_{n=1}^{\infty} \subseteq B(\mathcal{X}, Y)$. $\forall x \in \mathcal{X}$, we have $\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \epsilon \|x\|$. By Proposition 7.64, $(A_n)_{n=1}^{\infty}$ is a Cauchy sequence. Since $Y$ is a Banach space, then $\exists! y \in Y$ such that $\lim_{n \to \infty} A_n x = y$. Hence, we may define a function $f : \mathcal{X} \to Y$ by $f(x) = y$, $\forall x \in \mathcal{X}$.

Let $x_1, x_2 \in \mathcal{X}$, $\forall \alpha, \beta \in K$, by Propositions 7.23, 4.15, and 3.67, we have $f(\alpha x_1 + \beta x_2) = \lim_{n \to \infty} A_n (\alpha x_1 + \beta x_2) = \lim_{n \to \infty} (\alpha A_n x_1 + \beta A_n x_2) = \alpha \lim_{n \to \infty} A_n x_1 + \beta \lim_{n \to \infty} A_n x_2 = \alpha f(x_1) + \beta f(x_2)$. Hence, $f$ is linear.

Since $(A_n)_{n=1}^{\infty}$ is Cauchy, then $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$, $\forall n, m \geq N$, $\|A_n - A_m\| \leq \epsilon$. $\forall x \in \mathcal{X}$, $\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \epsilon \|x\|$. Then, by Propositions 3.66, 7.21, and 7.23, $\|f(x) - A_m x\| = \lim_{n \to \infty} \|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \epsilon \|x\|$. Therefore, $\|f(x)\| \leq \|f(x) - A_m x\| + \|A_n x\| \leq \epsilon \|x\| + \|A_n\| \|x\| = (\epsilon + \|A_n\|) \|x\|$. This implies that $f$ is bounded. Hence, $f \in B(\mathcal{X}, Y)$. By the arbitrariness of $\epsilon$, we have $\lim_{n \to \infty} \|A_n - f\| = 0$ and hence $\lim_{n \to \infty} A_n = f \in B(\mathcal{X}, Y)$. Hence, $B(\mathcal{X}, Y)$ is complete.

This completes the proof of the proposition. \hfill $\square$

**Proposition 7.67** Let $\mathcal{X}$ and $Y$ be normed linear spaces over the field $K$, $\mathcal{X}$ be finite dimensional with dimension $n \in \mathbb{Z}_+$, and $A : \mathcal{X} \to Y$ be a linear operator. Then, $A \in B(\mathcal{X}, Y)$.

**Proof** \quad Let $X_N \subseteq \mathcal{X}$ be a basis for $\mathcal{X}$, which then contains exactly $n$ vectors. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $n = 0$; Case 2: $n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X} = \{0\}$. Then, by Proposition 7.63, we have $\|A\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|Ax\| = 0$. Hence, $A \in B(\mathcal{X}, Y)$.

Case 2: $n \in \mathbb{N}$. Let $X_N = \{x_1, \ldots, x_n\}$. $\forall x \in \mathcal{X}$, by Corollary 6.47 and Definition 6.50, $\exists! \alpha_1, \ldots, \alpha_n \in K$ such that $x = \sum_{i=1}^{n} \alpha_i x_i$. Then, we may define an alternative norm $\|\cdot\|_1$ on $\mathcal{X}$ by $\|x\|_1 = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}$. Then,
It is easy to check that \( \| \cdot \|_1 \) defines a norm on \( \mathcal{X} \). By Theorem 7.38, \( \exists M \in (0, \infty) \subset \mathbb{R} \) such that \( \| x \| / M \leq \| x \|_1 \leq M \| x \|, \forall x \in \mathcal{X} \). Define \( r_i := \| Ax_i \|, i = 1, \ldots, n \). Then, by Proposition 7.63,

\[
\| A \| = \sup_{x \in \mathcal{X}, \| x \| \leq 1} \| Ax \| = \sup_{\alpha_1, \ldots, \alpha_n \in \mathbb{R}, \| \sum_{i=1}^n \alpha_i x_i \| \leq 1} \left\| A \left( \sum_{i=1}^n \alpha_i x_i \right) \right\| \\
\leq \sup_{\alpha_1, \ldots, \alpha_n \in \mathbb{R}} \left\| \sum_{i=1}^n \alpha_i \right\| = \left( \sum_{i=1}^n r_i^2 \right)^{1/2} \left( \sum_{i=1}^n \| x_i \|^2 \right)^{1/2} \\
= \sup_{x \in \mathcal{X}, \| x \| \leq 1} \left( \sum_{i=1}^n r_i^2 \right)^{1/2} \| x \| \leq \sup_{x \in \mathcal{X}, \| x \| \leq 1} \left( \sum_{i=1}^n r_i^2 \right)^{1/2} M \| x \| \\
\leq M \left( \sum_{i=1}^n r_i^2 \right)^{1/2}
\]

where we have applied the Cauchy-Schwarz Inequality in the second inequality. Hence, \( A \in \mathbb{B}(\mathcal{X}, \mathcal{Y}) \).

This completes the proof of the proposition. \( \square \)

**Proposition 7.68** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over the field \( \mathbb{K} \) and \( A : \mathcal{X} \to \mathcal{Y} \) be a linear operator. If \( A \) is bounded, then \( \mathcal{N}(A) \subseteq \mathcal{X} \) is closed. On the other hand, if \( \mathcal{N}(A) \) is closed and \( \mathcal{R}(A) \subseteq \mathcal{Y} \) is finite dimensional, then \( A \) is bounded.

**Proof** Let \( A \) be bounded. By Proposition 7.62, \( A \) is continuous. Note that the set \( \{ \partial y \} \subseteq \mathcal{Y} \) is closed by Proposition 3.34. By Proposition 3.10, \( \mathcal{N}(A) = A_{\text{inv}}(\partial y) \) is closed.

Let \( \mathcal{N}(A) \) be closed and \( \mathcal{R}(A) \) be finite dimensional. Let \( n \in \mathbb{Z}_+ \) be the dimension of \( \mathcal{R}(A) \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( n = 0 \); Case 2: \( n \in \mathbb{N} \). Case 1: \( n = 0 \). Then, \( \mathcal{R}(A) = \{ \partial y \} \). Hence, \( A = \partial_{\mathcal{B}(\mathcal{X}, \mathcal{Y})} \), which is clearly bounded.

Case 2: \( n \in \mathbb{N} \). Then, \( \exists X_N = \{ x_1, \ldots, x_n \} \subseteq \mathcal{X} \) with exactly \( n \) elements such that \( A(X_N) \subseteq \mathcal{Y} \) is a basis of \( \mathcal{R}(A) \). Let \( \mathcal{X}/\mathcal{N}(A) \) be the quotient normed linear space as defined in Proposition 7.44. \( \forall x \in \mathcal{X} \), \( Ax \in \mathcal{R}(A) = \text{span}(A(X_N)) \). Then, by Definition 6.50 and Corollary 6.47, \( \exists! \alpha_1, \ldots, \alpha_n \in \mathbb{K} \) such that \( Ax = \sum_{i=1}^n \alpha_i x_i \). Then, we have \( A(x - \sum_{i=1}^n \alpha_i x_i) = \partial y \) and \( x - \sum_{i=1}^n \alpha_i x_i \in \mathcal{N}(A) \). Hence, \( \{ x \} = \{ \sum_{i=1}^n \alpha_i x_i \} = \sum_{i=1}^n \alpha_i [x]_i \). This implies that \( \mathcal{X}/\mathcal{N}(A) \subseteq \text{span}(\{ [x_1], \ldots, [x_n] \}) \). Hence, \( \mathcal{X}/\mathcal{N}(A) \) is finite dimensional.

Define a mapping \( \tilde{A} : \mathcal{X}/\mathcal{N}(A) \to \mathcal{Y} \) by \( \tilde{A} [x] = Ax \), \( \forall [x] \in \mathcal{X}/\mathcal{N}(A) \). Note that \( \forall \bar{x} \in [x] \), we have \( x - \bar{x} \in \mathcal{N}(A) \) and \( Ax = \bar{Ax} \). Hence, \( A \) is uniquely defined. \( \forall [x_1], [x_2] \in \mathcal{X}/\mathcal{N}(A), \forall \alpha, \beta \in \mathbb{K} \), we have \( \tilde{A}(\alpha [x_1] + \beta [x_2]) = \tilde{A}(\alpha [x_1] + \beta [x_2]) = \alpha \tilde{A} [x_1] + \beta \tilde{A} [x_2] = \alpha Ax + \beta Ax = (\alpha + \beta) Ax \).
\[ \beta(x_2) = \bar{A}(\alpha x_1 + \beta x_2) = A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \alpha \bar{A}(x_1) + \beta \bar{A}(x_2). \] Hence, \( \bar{A} \) is a linear operator. By Proposition 7.67, \( \bar{A} \) is bounded.

This completes the proof of the proposition. \( \square \)

**Proposition 7.69** Let \( X \) be a normed linear space and \( M \subseteq X \) be a closed subspace. Let \( X/M \) be the quotient normed linear space and \( \phi : X \to X/M \) be the natural homomorphism. Then, \( \| \phi \| \leq 1 \).

**Proof** By Proposition 7.43, \( \phi \) is linear. \( \forall x \in X, \| \phi(x) \| = \inf_{m \in M} \| x - m \| \leq \| x \| \). Then, \( \phi \in B(X, X/M) \) and \( \| \phi \| \leq 1 \). This completes the proof of the proposition. \( \square \)

**Proposition 7.70** Let \( X \) and \( Y \) be normed linear spaces over the field \( \mathbb{K} \), \( A \in B(X, Y) \), and \( \phi : X \to X/N(A) \) be the natural homomorphism. Then, \( \exists B \in B(X/N(A), Y) \) such that \( A = B \circ \phi \), \( B \) is injective, and \( \| A \| = \| B \| \).

**Proof** By Proposition 7.68, \( N(A) \) is closed. Then, by Proposition 7.44, \( X/N(A) \) is a normed linear space. Define \( B : X/N(A) \to Y \) by \( B([x]) = Ax, \forall [x] \in X/N(A) \). Note that \( \forall \bar{x} \in [x] \), we have \( x - \bar{x} \in N(A) \) and \( Ax = A\bar{x} \). Hence, \( B \) is uniquely defined. \( \forall [x_1], [x_2] \in X/N(A), \forall \alpha, \beta \in \mathbb{K} \), we have \( B(\alpha [x_1] + \beta [x_2]) = B(\overline{\alpha x_1 + \beta x_2}) = A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \alpha B(x_1) + \beta B(x_2) \). Hence, \( B \) is linear. \( \forall [x_1], [x_2] \in X/N(A), B(x_1) = B(x_2) \) implies that \( Ax_1 = Ax_2 \), \( x_1 - x_2 \in N(A) \), and hence, \( x_1 = x_2 \). This shows that \( B \) is injective.

\( \forall x \in X, B \circ \phi(x) = B([x]) = Ax \). Hence, \( A = B \circ \phi \).

\( \forall [x] \in X/N(A), \forall m \in N(A), \| B[x] \| = \| Ax \| = \| A(x - m) \| \leq \| A \| \| x - m \|. \) This implies that \( \| B[x] \| \leq \| A \| \| x \|. \) Hence, \( B \in B(X/N(A), Y) \) and \( \| B \| \leq \| A \|. \) By Propositions 7.64 and 7.69, \( \| A \| = \| B \circ \phi \| \leq \| B \| \| \phi \| \leq \| B \|. \) Then, \( \| A \| = \| B \|. \) This completes the proof of the proposition. \( \square \)

### 7.9 Dual Spaces

#### 7.9.1 Basic concepts

**Definition 7.71** Let \( X \) be a normed linear space over the field \( \mathbb{K} \). Then, the space \( B(X, \mathbb{K}) \), which consists of all bounded linear functional on \( X \), is called the dual of \( X \) and denoted by \( X^* \). We will denote the vectors in \( X^* \) by \( x \), and denote \( x_*(x) \) by \( \langle x_*, x \rangle \).

**Proposition 7.72** Let \( X \) be a normed linear space over the field \( \mathbb{K} \) and \( X^* \) be its dual. Then, the following statements hold.

(i) If a linear functional \( f : X \to \mathbb{K} \) is continuous at some \( x_0 \in X \), then \( f \in X^* \).
(ii) \( \forall x_* \in X^* \), \( x_* \) is uniformly continuous.

(iii) For any linear functional \( f : X \rightarrow \mathbb{K} \), we have \( \| f \| = \inf \{ M \in [0, \infty) \subset \mathbb{R} \mid | f(x) | \leq M \| x \|_X, \forall x \in X \} = \sup_{x \in X} \{ | f(x) | \} \) = \max \{ \sup_{x \in X, x \neq \varnothing} | f(x) | / \| x \|_X \} = \max \{ \sup_{x \in X, \| x \|_X = 1} | f(x) |, 0 \}.

(iv) \( | \langle x_*, x \rangle | \leq \| x_* \| \| x \|, \forall x \in X, \forall x_* \in X^* \).

(v) \( \langle \cdot, \cdot \rangle \) is a continuous function on \( X^* \times X \).

(vi) \( X^* \) is a Banach space.

(vii) A linear functional \( f : X \rightarrow \mathbb{K} \) is bounded if, and only if, \( N(f) \) is closed.

Proof (i) and (ii) are direct consequences of Proposition 7.62. (iii) is a direct consequence of Proposition 7.63. (iv) follows from Proposition 7.64. (v) follows from Proposition 7.65. (vi) follows from Proposition 7.66 since \( \mathbb{K} \) is a Banach space. Finally, (vii) follows from Proposition 7.68. This completes the proof of the proposition. \( \square \)

7.9.2 Duals of some common Banach spaces

Example 7.73 Let us consider the dual of \( X := \{ \varnothing \}_{\mathbb{N}}, \mathbb{K}, \| \cdot \| \) \}. Clearly, there is a single linear functional on \( X \) given by \( f : X \rightarrow \mathbb{K} \) and \( f(\varnothing) = 0 \). Clearly, this linear functional is bounded with norm 0. Hence, \( X^* = \{ f \mid \mathbb{K}, \| \cdot \|_1 \} \), which is isometrically isomorphic to \( X \). \( \diamond \)

Example 7.74 Let us consider the dual of \( X := \mathbb{K}^n, n \in \mathbb{N} \), \( \forall x := (\xi_1, \ldots, \xi_n) \in \mathbb{K}^n \), any functional of the form \( f(x) = \sum_{i=1}^{n} \eta_i \xi_i \) with \( \eta_1, \ldots, \eta_n \in \mathbb{K} \) is clearly linear. By Cauchy-Schwarz Inequality, we have \( | f(x) | = | \sum_{i=1}^{n} \eta_i \xi_i | \leq \sum_{i=1}^{n} | \eta_i | \| \xi_i \| \leq \sqrt{\sum_{i=1}^{n} | \eta_i |^2} \sqrt{\sum_{i=1}^{n} | \xi_i |^2} = \sqrt{\sum_{i=1}^{n} | \eta_i |^2} \|. \) Hence, \( f \) is bounded and \( \| f \| \leq \sqrt{\sum_{i=1}^{n} | \eta_i |^2} \). Since the equality is achieved at \( x = (\overline{\eta_1}, \ldots, \overline{\eta_n}) \) in the above inequality, where \( \overline{\eta_i} \) denotes the complex conjugate of \( \eta_i \), then \( \| f \| = \sqrt{\sum_{i=1}^{n} | \eta_i |^2} \). Clearly, for different \( n \)-tuple \( (\eta_1, \ldots, \eta_n) \), the linear functional \( f \) is distinct. Now, let \( f \) be a bounded linear functional on \( X \). Let \( e_i \in X \) be the \( i \)th unit vector (all components of \( e_i \) are zero except a 1 at the \( i \)th component), \( i = 1, \ldots, n \). Let \( \eta_i = f(e_i) \in \mathbb{K}, i = 1, \ldots, n \). Then, \( \forall x := (\xi_1, \ldots, \xi_n) \in X \), we have \( x = \sum_{i=1}^{n} \xi_i e_i \) and \( f(x) = \sum_{i=1}^{n} \eta_i \xi_i \). Define a function \( \Psi : X^* \rightarrow \mathbb{K}^n \) by \( \Psi(f) = (\eta_1, \ldots, \eta_n), \forall f \in X^* \). Clearly, \( \Psi \) is linear, bijective and norm preserving. Hence, the dual of \( \mathbb{K}^n \) is (isometrically isomorphic to) \( \mathbb{K}^n \). \( \diamond \)
Lemma 7.75 Let $X$ be a normed linear space over $K$, $X^*$ be its dual, and $x_* \in X^*$. Then, $\forall \epsilon \in (0,1) \subset R$, $\exists x \in X$ with $\|x\| \leq 1$ such that $\langle \langle x_*, x \rangle \rangle \in R$ and $\langle \langle x_*, x \rangle \rangle \geq (1 - \epsilon) \|x_*\|$. 

Proof $\forall \epsilon \in (0,1) \subset R$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|x_*\| = 0$; Case 2: $\|x_*\| > 0$. Case 1: $\|x_*\| = 0$. Take $x = \vartheta$ and the result holds. Case 2: $\|x_*\| > 0$. Note that $\|x_*\| = \sup_{x \in X, \|x\| \leq 1} |\langle \langle x_*, x \rangle \rangle|$. Then, $\exists \bar{x} \in X$ with $\|\bar{x}\| \leq 1$ such that $|\langle \langle x_*, \bar{x} \rangle \rangle| \geq (1 - \epsilon) \|x_*\| > 0$. Take $x := \frac{\langle \langle x_*, x \rangle \rangle}{\langle \langle x_*, \bar{x} \rangle \rangle}$, $\bar{x} \in X$. Then, $\|x\| = \|\bar{x}\| \leq 1$ and $\langle \langle x_*, x \rangle \rangle = \frac{\langle \langle x_*, \bar{x} \rangle \rangle}{\langle \langle x_*, \bar{x} \rangle \rangle} = |\langle \langle x_*, \bar{x} \rangle \rangle| \in R$ and $\langle \langle x_*, x \rangle \rangle \geq (1 - \epsilon) \|x_*\|$. This completes the proof of the lemma. $\square$

Example 7.76 Let $Y_i$ be a normed linear space over the field $K$, $Y_i^*$ be its dual, $i = 1, \ldots, n$, $n \in N$. We will study the dual of $X := \prod_{i=1}^{n} Y_i$. We will show that $X^* = \prod_{i=1}^{n} Y_i^*$ isometrically isomorphically.

Let $f \in X^*$. $\forall i = 1, \ldots, n$, define $A_i : \prod_{i=1}^{n} Y_i \to \bar{X}$ by $x = (y_1, \ldots, y_n) = A_i y_i, \forall y_i \in Y_i$, and $y_j = \vartheta y_j, \forall j \in 1, \ldots, n$ with $j \neq i$, and $y_i = y$. Clearly, $A_i$ is well-defined and linear and bounded with $\|A_i\| \leq 1$. Define $f_i : \prod_{i=1}^{n} Y_i \to K$ by $f_i = f \circ A_i$. Then, $f_i$ is a bounded linear functional on $\prod_{i=1}^{n} Y_i$, with $\|f_i\| \leq \|f\| \cdot |A_i| \leq \|f\|$. Hence, $f_i \in Y_i^*$. Denote $f_i := y_i$. $\forall x = (y_1, \ldots, y_n) \in X$, we have $y_i \in Y_i, i = 1, \ldots, n$. By the linearity of $f_i$, we have

$$f(x) = f \left( \sum_{i=1}^{n} A_i y_i \right) = \sum_{i=1}^{n} f(A_i y_i) = \sum_{i=1}^{n} \langle\langle y_i, y \rangle\rangle$$  \hspace{1cm} (7.1)

Let $x_* := (y_{*1}, \ldots, y_{*n}) \in \prod_{i=1}^{n} Y_i^*$. $\forall \epsilon \in (0,1) \subset R$, $\forall i = 1, \ldots, n$, by Lemma 7.75, $\exists \hat{y}_i \in Y_i$ with $\|\hat{y}_i\| \leq 1$ such that $\langle\langle y_i, \hat{y}_i \rangle\rangle \geq (1 - \epsilon) \|y_{*i}\|$. Let $\hat{y}_i := y_{i} \cdot \hat{y}_i$. Then, $\langle\langle y_{*i}, \hat{y}_i \rangle\rangle \geq (1 - \epsilon) \|y_{*i}\|$. Then, we have $|f(\sum_{i=1}^{n} A_i \hat{y}_i)| = \left| \sum_{i=1}^{n} \langle\langle y_{*i}, \hat{y}_i \rangle\rangle \right| \geq \sum_{i=1}^{n} \langle\langle y_{*i}, \hat{y}_i \rangle\rangle \geq (1 - \epsilon) \sum_{i=1}^{n} \|y_{*i}\|^2 = (1 - \epsilon) \|x_*\|^2$. Since $f \in X^*$, then, by Proposition 7.72, we have $(1 - \epsilon) \|x_*\|^2 \leq \|f\| \sum_{i=1}^{n} A_i \hat{y}_i = \|f\| \left( \sum_{i=1}^{n} \|y_{*i}\|^2 \right)^{1/2} \leq \|f\| \left( \sum_{i=1}^{n} \|y_{*i}\|^2 \right)^{1/2} = \|f\| \|x_*\|$. By the arbitrariness of $\epsilon$, we have $\|x_*\| \leq \|f\|$. 

Based on the preceding analysis, we may define a function $\psi : X^* \to \prod_{i=1}^{n} Y_i^*$ by $\psi(f) = (f \circ A_1, \ldots, f \circ A_n)$, $\forall f \in X^*$. Then, $\|\psi(f)\| \leq \|f\|$. Clearly, $\psi$ is linear. $\forall f_1, f_2 \in X^*$ with $\psi(f_1) = \psi(f_2)$. Then, we have $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in N$. $\forall x := (y_1, \ldots, y_n) \in X$, by (7.1), we have $f_1(x) = \sum_{i=1}^{n} f_1(A_i y_i) = \sum_{i=1}^{n} f_2(A_i y_i) = f_2(x)$. Then, we have $f_1 = f_2$. Hence, $\psi$ is injective. $\forall x := (y_1, \ldots, y_n) \in \prod_{i=1}^{n} Y_i^*$, define $\phi(x) : X \to K$ by $\phi(x)(x) = \sum_{i=1}^{n} \langle\langle y_{*i}, y_i \rangle\rangle$, $\forall x := (y_1, \ldots, y_n) \in X$. Note that, $\forall x := (y_1, \ldots, y_n) \in X$, $\sum_{i=1}^{n} \langle\langle y_{*i}, y_i \rangle\rangle \leq \sum_{i=1}^{n} \|y_{*i}\| \|y_i\| \leq \|x\| \|x\|$, where we have made use of Proposition 7.72 in the first inequality and
Cauchy-Schwarz Inequality in the second inequality. Clearly, \( \phi(x_*) \) is linear and bounded with

\[
\| \phi(x_*) \| = \sup_{x \in X, \|x\| \leq 1} |\phi(x_*)(x)| \leq \sup_{x=(y_1, \ldots, y_n) \in X, \|x\| \leq 1} \sum_{i=1}^n |\langle y_{i*}, y_i \rangle| \leq \sup_{x \in X, \|x\| \leq 1} \| x_* \| \| x \| \leq \| x_* \| \|
\]

(7.2)

Hence, \( \phi(x_*) \in X^* \) and \( \psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \ldots, \phi(x_*) \circ A_n) = (y_1, \ldots, y_n) = x_* \). Then, \( \psi \circ \phi = \id_{\prod_{i=1}^n y_i^*} \). Hence, \( \psi \) is surjective. Therefore, \( \psi \) is bijective and admits inverse \( \psi_{\text{inv}} \). By Proposition 2.4, \( \phi = \psi_{\text{inv}} \).

\( \forall f \in X^* \), by (7.2), we have \( \| f \| = \| \phi(\psi(f)) \| \leq \| \psi(f) \| \leq \| f \| \). Then, \( \| \psi(f) \| = \| f \| \) and \( \psi \) is an isometry. Hence, \( \psi \) is an isometrical isomorphism.

\( \diamond \)

Example 7.77 Let us consider the dual of \( X := l_p(y), p \in [1, \infty) \subset \mathbb{R} \), \( y \) is a normed linear space over \( K \). Let \( y^* \) be the dual of \( y \) and \( q \in (1, \infty) \subset \mathbb{R} \) with \( 1/p + 1/q = 1 \). We will show that \( X^* \) is isometrically isomorphic to \( l_q(y^*) \).

Let \( f \in X^* \) \( \forall i \in \mathbb{N} \), define \( A_i : y \rightarrow X \) by \( x = (y_1, y_2, \ldots, y_i) \) with \( y_i \in y \), and \( y_j = \delta y \forall j \in \mathbb{N} \) with \( j \neq i \), and \( y_i = y \). Clearly, \( A_i \) is well-defined, linear, and bounded with \( \| A_i \| \leq 1 \). Define \( f_i : y \rightarrow K \) by \( f_i = f \circ A_i \). Then, \( f_i \) is a bounded linear functional on \( y \) with \( \| f_i \| \leq \| f \| \| A_i \| \leq \| f \| \).

Hence, \( f_i \in y^* \). Denote \( f_i = y_{i*} \). \( \forall x = (y_1, y_2, \ldots) \in X \), we have \( y_i \in y \), \( \forall i \in \mathbb{N} \), and \( \sum_{i=1}^{\infty} \| y_i \|^p < +\infty \). Then, \( \lim_{n \in \mathbb{N}} \sum_{i=n+1}^{\infty} \| y_i \|^p = 0 \). This implies that \( \lim_{n \in \mathbb{N}} \| x - \sum_{i=1}^{n} A_i y_i \| = \lim_{n \in \mathbb{N}} \left( \sum_{i=n+1}^{\infty} \| y_i \|^p \right)^{1/p} = 0 \).

Hence, \( \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} A_i y_i = x \). By the continuity of \( f \) and Proposition 3.66, we have

\[
f(x) = \lim_{n \in \mathbb{N}} f \left( \sum_{i=1}^{n} A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f(A_i y_i) = \lim_{n \in \mathbb{R}} \sum_{i=1}^{n} \langle y_{i*}, y_i \rangle \quad (7.3)
\]

Claim 7.77.1 \( x_* := (y_{1*}, y_{2*}, \ldots) \in l_q(y^*) \) and \( \| x_* \| \leq \| f \| \).

Proof of claim: We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( 1 < p < +\infty \); Case 2: \( p = 1 \).

Case 1: \( 1 < p < +\infty \). Then, \( 1 < q < +\infty \). \( \forall \epsilon \in (0, 1) \subset \mathbb{R} \), \( \forall i \in \mathbb{N} \), by Lemma 7.75, \( \exists \bar{y}_i \in y \) with \( \| \bar{y}_i \| \leq 1 \) such that \( \langle y_{i*}, \bar{y}_i \rangle \in \mathbb{R} \) and \( \langle y_{i*}, \bar{y}_i \rangle \geq (1 - \epsilon) \| y_{i*} \| \). Let \( \bar{y}_i = \| y_{i*} \|^q / \| \bar{y}_i \| \| y_{i*} \|^q / p^{1/p} \). Then, \( \langle \bar{y}_i \cdot \bar{y}_i \rangle \geq (1 - \epsilon) \| y_{i*} \|^q / p^{1/p} + 1 \). Then, \( \forall n \in \mathbb{N} \), \( f(\sum_{i=1}^{n} A_i \bar{y}_i) \geq \| \sum_{i=1}^{n} \langle y_{i*}, \bar{y}_i \rangle \| \geq (1 - \epsilon) \left( \sum_{i=1}^{n} \| y_{i*} \|^q / p^{1/p} \right) \leq \| f \| \left( \sum_{i=1}^{n} \| y_{i*} \|^q / p^{1/p} \right) \leq \| f \| \left( \sum_{i=1}^{n} \| y_{i*} \|^q \right)^{1/p} \leq \| f \| / (1 - \epsilon) \). By the arbitrariness of \( n \), we have \( \left( \sum_{i=1}^{\infty} \| y_{i*} \|^q \right)^{1/q} \leq \| f \| / (1 - \epsilon) \). By the
arbitrariness of $\epsilon$, we have $(\sum_{i=1}^{\infty} \| y_{si} \|_{A}^q)^{1/q} \leq \| f \|$. Hence, the result holds in this case.

Case 2: $p = 1$. Then, $q = +\infty$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, by Lemma 7.75, $\exists \hat{y}_i \in \mathcal{Y}$ with $\| \hat{y}_i \| \leq 1$ such that $\langle \langle y_{si}, \hat{y}_i \rangle \rangle \in \mathbb{R}$ and $\langle \langle y_{si}, \hat{y}_i \rangle \rangle \geq (1 - \epsilon) \| y_{si} \|$. Then, $(1 - \epsilon) \| y_{si} \| \leq \| \langle \langle y_{si}, \hat{y}_i \rangle \rangle \| \leq \| f \| \| A_i \| \| \hat{y}_i \| \leq \| f \|$. By the arbitrariness of $n$, we have $\sup_{i \geq 1} \| y_{si} \| \leq \| f \| / (1 - \epsilon)$. By the arbitrariness of $\epsilon$, we have $\sup_{i \geq 1} \| y_{si} \| \leq \| f \|$. Hence, the result holds in this case.

This completes the proof of the claim.

The preceding analysis shows that we may define a function $\psi: \mathcal{X}^* \rightarrow l_q(\mathcal{Y}^*)$ by $\psi(f) = (f \circ A_1, f \circ A_2, \ldots)$, $\forall f \in \mathcal{X}^*$. Clearly, $\psi$ is linear. $\forall f_1, f_2 \in \mathcal{X}^*$ with $\psi(f_1) = \psi(f_2)$. Then, we have $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in \mathbb{N}$. $\forall x := (y_1, y_2, \ldots) \in \mathcal{X}$, by (7.3), we have $f_1(x) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_1(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_2(A_i y_i) = f_2(x)$. Then, we have $f_1 = f_2$. Hence, $\psi$ is injective. $\forall x := (y_1, y_2, \ldots) \in \mathcal{X}$, define $\phi(x_*) = \psi(x)$. Then, we have $\| \phi(x_*) \| = \sum_{i=1}^{\infty} \| \langle \langle y_{si}, y_i \rangle \rangle \| \leq \sum_{i=1}^{\infty} \| y_i \| \| y_i \| \leq \| x_* \| \| x \|$, where we have made use of Proposition 7.72 in the first inequality and Hölder's Inequality in the second inequality. Hence, $\phi(x_*)(x)$ is well-defined. Hence, $\phi(x_*)$ is well-defined. Clearly, $\phi(x_*)$ is linear and bounded with

$$\| \phi(x_*) \| = \sup_{x \in \mathcal{X}, \| x \| \leq 1} \| \phi(x_*)(x) \| \leq \sup_{x = (y_1, y_2, \ldots) \in \mathcal{X}, \| x \| \leq 1} \sum_{i=1}^{\infty} \| \langle \langle y_{si}, y_i \rangle \rangle \| \leq \sum_{i=1}^{\infty} \| y_i \| \| y_i \| \leq \| x_* \| \| x \|$$  \hspace{1cm} (7.4)

Hence, $\phi(x_*) \in \mathcal{X}^*$ and $\psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \phi(x_*) \circ A_2, \ldots) = (y_1, y_2, \ldots) = x_*$. Then, $\psi \circ \phi = \text{id}_{l_q(\mathcal{Y}^*)}$. Hence, $\psi$ is surjective. Therefore, $\psi$ is bijective and admits inverse $\psi_{\text{inv}}$. By Proposition 2.4, $\phi = \psi_{\text{inv}}$. $\forall f \in \mathcal{X}^*$, by Claim 7.77.1 and (7.4), we have $\| f \| = \| \phi(\psi(f)) \| \leq \| \psi(f) \| \leq \| f \|$. Then, $\| \psi(f) \| = \| f \|$ and $\psi$ is an isometry. Hence, $\psi$ is an isometrical isomorphism.

**Example 7.78** Let $\mathcal{Y}$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{Y}^*$ be its dual, and $M = \{ x = (y_1, y_2, \ldots) \in l_1(\mathcal{Y}) \mid \lim_{n \in \mathbb{N}} y_n = \vartheta_y \}$. Clearly, $M$ is a subspace of $l_1(\mathcal{Y})$. By Proposition 7.13, $M$ is a normed linear space over $\mathbb{K}$, which will be denoted by $c_0(\mathcal{Y})$. Next, we will study the dual of $\mathcal{X} := c_0(\mathcal{Y})$. We will show that $\mathcal{X}^* = l_1(\mathcal{Y}^*)$ isometrically isomorphically.

Let $f \in \mathcal{X}^*$. $\forall i \in \mathbb{N}$, define $A_i : \mathcal{Y} \rightarrow \mathcal{X}$ by $x = (y_1, y_2, \ldots) = A_i y$, $\forall y \in \mathcal{Y}$, and $y_j = \vartheta_y$, $\forall j \in \mathbb{N}$ with $j \neq i$, and $y_i = y$. Clearly, $A_i$ is well-defined and linear and bounded with $\| A_i \| \leq 1$. Define $f_i : \mathcal{Y} \rightarrow \mathcal{K}$ by $f_i = f \circ A_i$. Then, $f_i$ is a bounded linear functional on $\mathcal{Y}$ with $\| f_i \| \leq \| f \| \| A_i \| \leq \| f \|$. Hence, $f_i \in \mathcal{Y}^*$. Denote $f_i := y_{si}$. $\forall x = (y_1, y_2, \ldots) \in \mathcal{X}$, we have $y_i \in \mathcal{Y}$, $\forall i \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} y_i = \vartheta_y$, and $\lim_{n \in \mathbb{N}} \| y_i \| = 0$ (by Proposition 7.21). Then, $\lim_{n \in \mathbb{N}} \sup_{1 \leq k \leq n+1} \| y_k \| = 0$. This implies that $\lim_{n \in \mathbb{N}} \| x - \sum_{i=1}^{n} A_i y_i \| = \lim_{n \in \mathbb{N}} \sup_{k \geq n+1} \| y_k \| = 0$. Hence, $f_i$ is a bounded linear functional on $\mathcal{Y}$ with $\| f_i \| \leq \| f \| \| A_i \| \leq \| f \|$.
\[
\lim_{n \in \mathbb{N}} \max_{k \geq n+1} \|y_k\| = 0. \quad \text{Hence, } \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} A_i y_i = x. \quad \text{By the continuity of } f \text{ and Proposition 3.66, we have}
\]

\[
f(x) = \lim_{n \in \mathbb{N}} f \left( \sum_{i=1}^{n} A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f \left( A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} \langle y_{i+1}, y_i \rangle \quad (7.5)
\]

Let \( x_* := (y_{i_1}, y_{i_2}, \ldots) \), \( \forall \epsilon \in (0, 1) \subset \mathbb{R} \), \( \forall i \in \mathbb{N} \), by Lemma 7.75, \( \exists \hat{y}_i \in \mathcal{Y} \) with \( \| \hat{y}_i \| \leq 1 \) such that \( \langle \langle y_{i+1}, \hat{y}_i \rangle \rangle \in \mathbb{R} \) and \( \langle \langle y_{i+1}, \hat{y}_i \rangle \rangle \geq (1 - \epsilon) \| y_{i+1} \| \).

Then, \( \forall n \in \mathbb{N}, (1 - \epsilon) \sum_{i=1}^{n} \| y_{i+1} \| \leq \sum_{i=1}^{n} \| \hat{y}_i \| = \sum_{i=1}^{n} \| f \left( A_i y_i \right) \| = \sum_{i=1}^{n} \| \langle y_{i+1}, \hat{y}_i \rangle \| \leq \| f \| \). By the arbitrariness of \( n \), we have \( \sum_{i=1}^{\infty} \| y_{i+1} \| \leq \| f \| / (1 - \epsilon) \). By the arbitrariness of \( \epsilon \), we have

\[
\sum_{i=1}^{\infty} \| y_{i+1} \| \leq \| f \| \quad (7.6)
\]

Hence, \( x_* \in l_1(\mathcal{Y}^*) \).

Based on the preceding analysis, we may define a function \( \psi : \mathbb{X}^* \rightarrow l_1(\mathcal{Y}^*) \) by \( \psi(f) = (f \circ A_1, f \circ A_2, \ldots) \), \( \forall f \in \mathbb{X}^* \). Clearly, \( \psi \) is linear. \( \forall f_1, f_2 \in \mathbb{X}^* \) with \( \psi(f_1) = \psi(f_2) \). Then, we have \( f_1 \circ A_1 = f_2 \circ A_1 \), \( \forall i \in \mathbb{N} \).

\( \forall x := (y_{i_1}, y_{i_2}, \ldots) \in \mathbb{X} \), by (7.5), we have \( f_1(x) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_1(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_2(A_i y_i) = f_2(x) \). Then, we have \( f_1 = f_2 \). Hence, \( \psi \) is injective. \( \forall x_* := (y_{i_1}, y_{i_2}, \ldots) \in l_1(\mathcal{Y}^*) \), define \( \phi(x_*) : \mathbb{X} \rightarrow \mathbb{K} \) by \( \phi(x_*)(x) := \sum_{i=1}^{\infty} \langle \langle y_{i+1}, y_i \rangle \rangle \), \( \forall x_* := (y_{i_1}, y_{i_2}, \ldots) \in \mathbb{X} \). Note that, \( \forall x := (y_{i_1}, y_{i_2}, \ldots) \in \mathbb{X} \),

\[
\sum_{i=1}^{\infty} \langle \langle y_{i+1}, y_i \rangle \rangle \leq \sum_{i=1}^{\infty} \| y_{i+1} \| \| y_i \| \leq \| x_* \| \| x \| \quad \text{where we have made use of Proposition 7.72 in the first inequality and Hölder’s inequality in the second inequality. Hence, } \phi(x_*)(x) \text{ is well-defined. Hence, } \phi(x_*) \text{ is well-defined. Clearly, } \phi(x_*) \text{ is linear and bounded with}
\]

\[
\| \phi(x_*) \| = \sup_{x \in \mathbb{X}, \| x \| \leq 1} | \phi(x_*)(x) | \leq \sup_{x \in \mathbb{X}, \| x \| \leq 1} \| x \| \| x_* \| \quad (7.7)
\]

Hence, \( \phi(x_*) \in \mathbb{X}^* \) and \( \psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \phi(x_*) \circ A_2, \ldots) = (y_{i_1}, y_{i_2}, \ldots) = x_* \). Then, \( \psi \circ \phi = \text{id}_{l_1(\mathcal{Y}^*)} \). Hence, \( \psi \) is surjective. Therefore, \( \psi \) is bijective and admits inverse \( \psi_{\text{inv}} \). By Proposition 2.4, \( \phi = \psi_{\text{inv}} \).

(7.6) and (7.7), we have \( \| f \| = \| \phi(f) \| \leq \| f \| \). Then, \( \| \psi(f) \| = \| f \| \) and \( \psi \) is an isometry. Hence, \( \psi \) is an isometrical isomorphism.

**7.9.3 Extension form of Hahn-Banach Theorem**

**Definition 7.79** Let \( \mathcal{X} \) be a vector space over \( \mathbb{K} \). A sublinear functional is \( p : \mathcal{X} \rightarrow \mathbb{R} \) satisfying, \( \forall x_1, x_2 \in \mathcal{X}, \forall \alpha \in \mathbb{R} \) with \( \alpha \geq 0 \), (i) \( p(x_1 + x_2) \leq p(x_1) + p(x_2) \); and (ii) \( p(\alpha x_1) = \alpha p(x_1) \).
7.9. DUAL SPACES

Note that any \( \| \cdot \| \) on \( \mathcal{X} \) is a sublinear functional. We introduce the above definition to illustrate the full generality of the Hahn-Banach Theorem.

**Theorem 7.80 (Extension Form of Hahn-Banach Theorem)** Let \( \mathcal{X} \) be a vector space over the field \( \mathbb{R} \), \( p : \mathcal{X} \to \mathbb{R} \) be a sublinear functional, \( M \subseteq \mathcal{X} \) be a subspace, and \( f : M \to \mathbb{R} \) be a linear functional on \( M \) satisfying \( f(m) \leq p(m) \), \( \forall m \in M \). Then, \( \exists \) a linear functional \( F : \mathcal{X} \to \mathbb{R} \) such that \( F|M = f \) and \( F(x) \leq p(x) \), \( \forall x \in \mathcal{X} \). Furthermore, if \( \mathcal{X} \) is normed with \( \| \cdot \| \) and \( p \) is continuous at \( \vartheta_x \), then \( F \) is continuous.

**Proof** We will prove the theorem using Zorn’s Lemma. Define a collection of extensions of \( f, \mathcal{E} \), by

\[
\mathcal{E} := \{ (g, N) \mid N \subseteq \mathcal{X} \text{ is a subspace}, M \subseteq N, \ g : N \to \mathbb{R} \text{ is a linear functional, such that } g|_M = f \text{ and } g(n) \leq p(n), \forall n \in N \}
\]

Clearly, \( (f, M) \in \mathcal{E} \neq \emptyset \). Define a relation \( \preceq \) on \( \mathcal{E} \) by \( \forall (g_1, N_1), (g_2, N_2) \in \mathcal{E} \), we say \( (g_1, N_1) \preceq (g_2, N_2) \) if \( N_1 \subseteq N_2 \) and \( g_2|_{N_1} = g_1 \). Clearly, \( \preceq \) is reflexive, antisymmetric, and transitive. Hence, \( \preceq \) is an antisymmetric partial ordering on \( \mathcal{E} \).

For any nonempty subcollection \( \mathcal{C} \subseteq \mathcal{E} \) such that \( \preceq \) is a total ordering on \( \mathcal{C} \), let \( N_c = \bigcup_{(g, N) \in \mathcal{C}} N. \) Then, \( \vartheta_x \in M \subseteq N_c \subseteq \mathcal{X} \). \( \forall x_1, x_2 \in N_c, \forall \alpha, \beta \in \mathbb{R}, \exists (g_1, N_1), (g_2, N_2) \in \mathcal{C} \) such that \( x_1 \in N_1 \) and \( x_2 \in N_2 \). Since \( \mathcal{C} \) is totally ordered, by Proposition 2.12, then, without loss of generality, we may assume that \( (g_1, N_1) \preceq (g_2, N_2) \). Then, \( N_1 \subseteq N_2 \) and \( x_1, x_2 \in N_2 \). This implies that \( \alpha x_1 + \beta x_2 \in N_2 \subseteq N_c \), since \( N_2 \) is a subspace. The above shows that \( N_c \) is a subspace of \( \mathcal{X} \).

Define a functional \( g_c : N_c \to \mathbb{R} \) by \( \forall x \in N_c, \exists (g, N) \in \mathcal{C} \) such that \( x \in N, \) we assign \( g_c(x) := g(x). \) Such a functional is uniquely defined because of the following reasoning. \( \forall x \in N_c, \forall (g_1, N_1), (g_2, N_2) \in \mathcal{C} \) such that \( x \in N_1 \cap N_2 \). By the total ordering of \( \mathcal{C} \) and Proposition 2.12, we may assume that, without loss of generality, \( (g_1, N_1) \preceq (g_2, N_2) \). Then, \( x \in N_1 \cap N_2 \) and \( g_c(\alpha x_1 + \beta x_2) = g_2(\alpha x_1 + \beta x_2) = \alpha g_2(x_1) + \beta g_2(x_2) = \alpha g_1(x_1) + \beta g_1(x_2). \) Hence, \( g_c \) is a linear functional. \( g_c(x_1) = g_2(x_1) \leq p(x_1). \) \( \forall m \in M, g_c(m) = g_2(m) = f(m). \) Therefore, \( (g_c, N_c) \in \mathcal{E}. \)

\( \forall (g, N) \in \mathcal{C}, \) we have \( N \subseteq N_c \) and \( \forall n \in N, g_c(n) = g(n). \) Then, \( (g, N) \preceq (g_c, N_c) \). Hence, \( (g_c, N_c) \) is an upper bound of \( \mathcal{C}. \)

By Zorn’s Lemma, \( \exists (g_M, N_M) \in \mathcal{E}, \) which is maximal with respect to \( \preceq \). Now, we are going to show that \( N_M = \mathcal{X}. \) Suppose \( N_M \subset \mathcal{X}. \) Then, \( \exists x_0 \in \mathcal{X} \setminus N_M. \) Let \( N_c = \{ x \in \mathcal{X} \mid \exists a \in \mathbb{R}, \exists n \in N_M : x = ax_0 + n \}. \) Clearly, \( N_M \subset N_c, x_0 \in N_c, \) and \( N_c \) is a subspace of \( \mathcal{X}. \) \( \forall x \in N_c, \exists ! \alpha \in \mathbb{R} \)
and \( \exists \| n \in N_M \) such that \( x = \alpha x_0 + n \) (otherwise, we may deduce that \( x_0 \in N_M \)). Define \( g_e : N_e \to \mathbb{R} \) by \( g_e(x) = g_e(\alpha x_0 + n) = g_M(n) + \alpha g_e(x_0) \), \( \forall x = \alpha x_0 + n \in N_e, \alpha \in \mathbb{R} \), and \( n \in N_M \), where \( g_e(x_0) \in \mathbb{R} \) is a constant to be determined. Clearly, \( g_e \) is well-defined and is a linear functional on \( N_e \). \( \forall n \in N_M \), we have \( g_e(n) = g_M(n) \), and hence, \( g_e|_{N_M} = g_M \). Then, \( g_e|_{M} = f \).

We will show that \((g_e, N_e) \in E\) by finding an admissible constant \( g_e(x_0) \) such that \( g_e(x) \leq p(x), \forall x \in N_e \). \( \forall n_1, n_2 \in N_M \), we have \( g_M(n_1) + g_M(n_2) = g_M(n_1 + n_2) \leq p(n_1 + n_2) \leq p(n_1 + x_0) + p(n_2 - x_0) \). Then, \( g_M(n_2) - p(n_2 - x_0) \leq p(n_1 + x_0) - g_M(n_1) \). By the arbitrariness of \( n_1 \) and \( n_2 \), we have \( \sup_{n_2 \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq \inf_{n_1 \in N_M} (p(n_1 + x_0) - g_M(n_1)) \).

Since \( \vartheta_X \in N_M \), then \( \exists c \in \mathbb{R} \) such that

\[
\begin{align*}
-\infty &< \sup_{n_2 \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq c \\
&\leq \inf_{n_1 \in N_M} (p(n_1 + x_0) - g_M(n_1)) < +\infty
\end{align*}
\]

Let \( g_e(x_0) := c \).

\( \forall x \in N_e, x = \alpha x_0 + n \), where \( \alpha \in \mathbb{R} \) and \( n \in N_M \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( \alpha > 0 \); Case 2: \( \alpha = 0 \); Case 3: \( \alpha < 0 \). Case 1: \( \alpha > 0 \). \( g_e(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(p(n/\alpha + x_0) - g_M(n/\alpha) + g_M(n/\alpha)) = \alpha p(x/\alpha) = p(x) \).

Case 2: \( \alpha = 0 \). Then, \( x \in N_M \) and \( g_e(x) = g_M(n) \leq p(n) = p(x) \).

Case 3: \( \alpha < 0 \). \( g_e(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(g_M(-n/\alpha) - p(-n/\alpha - x_0) + g_M(n/\alpha)) = -\alpha p(-x/\alpha) = p(x) \).

Hence, we have \( g_e(x) \leq p(x) \) in all three cases. Therefore, \((g_e, N_e) \in E\). We have shown that \((g_M, N_M) \in (g_e, N_e) \) and \((g_M, N_M) \neq (g_e, N_e) \). This contradicts the fact that \((g_M, N_M) \) is maximal in \( E \) by Proposition 2.12.

Hence, \( N_M = X \) and \( F = g_M \).

If, in addition, \( X \) is normed with \( \| \cdot \| \) and \( p \) is continuous at \( \vartheta_X \), then, \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( |p(x)| < \varepsilon, \forall x \in B(\vartheta_X, \delta) \).

\( \forall x \in B(\vartheta_X, \delta) \), we will distinguish two exhaustive and mutually exclusive cases: Case 1: \( F(x) \geq 0 \); Case 2: \( F(x) < 0 \). Case 1: \( F(x) \geq 0 \). Then, \( 0 \leq F(x) \leq p(x) < \varepsilon \) and \( |F(x)| < \varepsilon \). Case 2: \( F(x) < 0 \). Then, \( |F(x)| = -F(x) = F(-x) \leq p(-x) < \varepsilon \). Hence, in both cases, we have \( |F(x)| < \varepsilon \).

This completes the proof of the theorem. \( \square \)

To obtain the counterpart result for complex vector spaces, we need the following result.

**Lemma 7.81** Let \((X, \mathbb{C})\) be a vector space, \( f : X \to \mathbb{C} \) be a functional on \((X, \mathbb{C})\), and \( g : X \to \mathbb{R} \) be given by \( g(x) = \text{Re}(f(x)), \forall x \in X \). Then, \( f \) is a linear functional on \((X, \mathbb{C})\) if, and only if, \( g \) is a linear functional on \((X, \mathbb{R})\) and \( f(x) = g(x) - ig(ix), \forall x \in X \).
Theorem 7.82 (Extension Form of Hahn-Banach Theorem)

Let $F$ be a vector space. Furthermore, if $X \subseteq X$ is a subspace (of $X$), $\alpha$, $\beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = \alpha g(x_1) + \beta g(x_2)$. Hence, $g$ is a linear functional on $(X, \mathbb{R})$. By Lemma 7.60, $(X, \mathbb{R})$ is a vector space. “Necessity” $\forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = \alpha g(x_1) + \beta g(x_2)$. Hence, $g$ is a linear functional on $(X, \mathbb{R})$. Therefore, $g$ is continuous at $x_1, x_2 \in X$.

Proof

By Lemma 7.40, $(X, \mathbb{R})$ is a vector space. “Necessity” $\forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = \alpha g(x_1) + \beta g(x_2)$. Hence, $g$ is a linear functional on $(X, \mathbb{R})$. By Lemma 7.60, $(X, \mathbb{R})$ is a vector space. “Necessity” $\forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = \alpha g(x_1) + \beta g(x_2)$. Hence, $g$ is a linear functional on $(X, \mathbb{R})$. Therefore, $g$ is continuous at $x_1, x_2 \in X$.

Theorem 7.82 (Extension Form of Hahn-Banach Theorem) Let $X$ be a vector space over the field $\mathbb{C}$, $p : X \to \mathbb{R}$ be a sublinear functional, $M \subseteq X$ be a subspace (of $(X, \mathbb{C})$), and $f : M \to \mathbb{C}$ be a linear functional satisfying $\text{Re}(f(m)) \leq p(m)$, $\forall m \in M$. Then, there exists a linear functional $F : X \to \mathbb{C}$, such that $F|_M = f$ and $\text{Re}(F(x)) \leq p(x)$, $\forall x \in X$. Furthermore, if $X$ is normed with $\| \cdot \|$ and $p$ is continuous at $\vartheta_X$, then $F$ is continuous.

Proof

By Lemma 7.40, $(X, \mathbb{R})$ is a vector space and $(M, \mathbb{R})$ is also a vector space. It is easy to show that $(M, \mathbb{R})$ is a subspace of $(X, \mathbb{R})$. Define $g : M \to \mathbb{R}$ by $g(m) = \text{Re}(f(m)), \forall m \in M$. By Lemma 7.81, $g$ is a linear functional on $(M, \mathbb{R})$ and $f(m) = g(m) - ig(im), \forall m \in M$. Note that $p$ is a sublinear functional on $(X, \mathbb{C})$, then it is a sublinear functional on $(X, \mathbb{R})$. Furthermore, $g(m) = \text{Re}(f(m)) \leq p(m), \forall m \in M$. Then, by Hahn-Banach Theorem, Theorem 7.80, $\exists$ a linear functional $G$ on $(X, \mathbb{R})$ such that $G|_M = g$ and $G(x) \leq p(x), \forall x \in X$.

Define a functional $F$ on $(X, \mathbb{C})$ by $F(x) = G(x) - iG(ix), \forall x \in X$. By Lemma 7.81, $F$ is a linear functional on $(X, \mathbb{C})$. $\forall m \in M, im \in M$, since $M$ is a subspace of $(X, \mathbb{C})$. Then, $F(m) = G(m) - iG(im) = g(m) - ig(im) = f(m)$. Hence, we have $F|_M = f$. $\forall x \in X$, $\text{Re}(F(x)) = G(x) \leq p(x)$. Hence, $F$ is the functional we seek.

Furthermore, if $(X, \mathbb{C})$ is normed with $\| \cdot \|$ and $p$ is continuous at $\vartheta_X$, then, by Lemma 7.40, $X_{\mathbb{R}} := (X, \mathbb{R}, \| \cdot \|)$ is a normed linear space. By Theorem 7.80, $G$ is continuous. Then, $G : X_{\mathbb{R}} \to \mathbb{C}$ is continuous. By Propositions 3.12, 3.32, and 7.23, $F : X_{\mathbb{R}} \to \mathbb{C}$ is continuous. By Lemma 7.40, $F$ is continuous on $X := (X, \mathbb{C}, \| \cdot \|)$.

This completes the proof of the theorem. □
Theorem 7.83 (Simple Version of Hahn-Banach Theorem) Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( M \subseteq \mathcal{X} \) be a subspace, \( f : M \to \mathbb{K} \) be a linear functional which is bounded (on \( M \)), that is \( \|f\|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| \in [0, \infty) \subseteq \mathbb{R} \). Then, there exists a linear functional \( F \) on \( \mathcal{X} \) such that \( F|_M = f \) and \( \|F\| = \|f\|_M \).

**Proof** Define a functional \( p : \mathcal{X} \to \mathbb{R} \) by \( p(x) = \|f\|_M \|x\|, \forall x \in \mathcal{X} \). It is easy to check that \( p \) is a sublinear functional. We are going to distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathbb{K} = \mathbb{R} \); Case 2: \( \mathbb{K} = \mathbb{C} \).

Case 1: \( \mathbb{K} = \mathbb{R} \). Clearly, by Proposition 7.72, we have \( f(m) \leq |f(m)| \leq p(m), \forall m \in M \), and \( p \) is continuous. By Hahn-Banach Theorem, Theorem 7.80, \( \exists \) a linear functional \( F : \mathcal{X} \to \mathbb{R} \) such that \( F|_M = f \) and \( F(x) \leq p(x), \forall x \in \mathcal{X} \). Furthermore, \( F \) is continuous. Hence, \( F \in \mathcal{X}^* \), by Proposition 7.72. Note that \( \|F\| = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\| \). On the other hand, we have, \( \forall x \in \mathcal{X} \),

\[
F(x) = F(-x) \leq p(-x) = \|f\|_M \|x\| \text{ and } F(x) \leq p(x) = \|f\|_M \|x\|.
\]

Hence, \( |F(x)| \leq \|f\|_M \|x\| \). Then, \( \|F\| \leq \|f\|_M \). Hence, \( \|F\| = \|f\|_M \).

Case 2: \( \mathbb{K} = \mathbb{C} \). Note that, \( \forall m \in M \), \( \Re (f(m)) \leq |\Re (f(m))| \leq |f(m)| \leq p(m) \), by Proposition 7.72. By Hahn-Banach Theorem, Theorem 7.82, there exists a linear functional \( F : \mathcal{X} \to \mathbb{C} \) such that \( F|_M = f \) and \( \Re (F(x)) \leq p(x), \forall x \in \mathcal{X} \). Furthermore, since \( p \) is continuous, then \( F \) is continuous. By Proposition 7.72, \( F \in \mathcal{X}^* \). Note that \( \|F\| = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\| \). On the other hand, \( \forall x \in \mathcal{X} \), we have either \( F(x) = 0 \), then \( |F(x)| \leq \|f\|_M \|x\| \); or \( F(x) \neq 0 \), then \( |F(x)| = \frac{|f(x)|}{F(x)} \cdot F(x) = F\left(\frac{|f(x)|}{F(x)}ight) \).

Hence, \( \|F\| \leq \|f\|_M \). Therefore, \( \|F\| = \|f\|_M \).

This completes the proof of the theorem. \( \square \)

**Corollary 7.84** Let \( \mathcal{X} \) be a vector space over the field \( \mathbb{K} \), \( M \subseteq \mathcal{X} \) be a subspace, \( p : \mathcal{X} \to \mathbb{R} \) be a sublinear functional, \( f : M \to \mathbb{K} \) be a linear functional on \( M \), and \( (\mathcal{G}, \circ, E) \) be an abelian semigroup of linear operators of \( \mathcal{X} \) to \( \mathcal{X} \). Assume that the following conditions hold.

(i) \( Am \in M, \forall A \in \mathcal{G} \) and \( \forall m \in M \).

(ii) \( p(Ax) \leq p(x), \forall x \in \mathcal{X} \) and \( \forall A \in \mathcal{G} \).

(iii) \( f(Am) = f(m), \forall A \in \mathcal{G} \) and \( \forall m \in M \).

(iv) \( \Re (f(m)) \leq p(m), \forall m \in M \).

Then, there exists a linear functional \( F : \mathcal{X} \to \mathbb{K} \) such that \( F|_M = f \), \( F(Ax) = F(x), \forall A \in \mathcal{G} \) and \( \forall x \in \mathcal{X} \), and \( \Re (F(x)) \leq p(x), \forall x \in \mathcal{X} \).
Furthermore, if, in addition, \( X \) is normed with \( \| \cdot \| \) and \( p \) is continuous at \( \vartheta_X \), then \( F \) is continuous.

**Proof** Define a functional \( q : X \to \mathbb{R} \) by

\[
q(x) = \inf_{A_1, \ldots, A_n \in G} \frac{1}{n} p \left( \sum_{i=1}^{n} A_i x \right); \quad \forall x \in X
\]

Since \( p(\vartheta_X) = 0 \), then \( q(\vartheta_X) = 0 \). Since \( E \in G \), then, \( \forall x \in X \), \( q(x) \leq p(Ex) \leq p(x) < +\infty \). \( \forall x_1, x_2 \in X \), \( \forall n, k \in \mathbb{N} \), \( \forall A_1, \ldots, A_n \in G \), \( \forall B_1, \ldots, B_k \in G \), we have

\[
q(x_1 + x_2) \leq \frac{1}{nk} p \left( \sum_{i=1}^{n} \sum_{j=1}^{k} A_i \circ B_j(x_1 + x_2) \right)
\]

\[
= \frac{1}{nk} p \left( \sum_{i=1}^{n} \sum_{j=1}^{k} A_i \circ B_j(x_1) + \sum_{i=1}^{n} \sum_{j=1}^{k} A_i \circ B_j(x_2) \right)
\]

\[
= \frac{1}{nk} p \left( \sum_{i=1}^{n} A_i \left( \sum_{j=1}^{k} B_j(x_1) \right) + \sum_{j=1}^{k} B_j \left( \sum_{i=1}^{n} A_i(x_2) \right) \right)
\]

\[
\leq \frac{1}{nk} p \left( \sum_{i=1}^{n} A_i \left( \sum_{j=1}^{k} B_j(x_1) \right) \right) + \frac{1}{nk} p \left( \sum_{j=1}^{k} B_j \left( \sum_{i=1}^{n} A_i(x_2) \right) \right)
\]

\[
\leq \frac{1}{nk} \sum_{i=1}^{n} p \left( A_i \left( \sum_{j=1}^{k} B_j(x_1) \right) \right) + \frac{1}{nk} \sum_{j=1}^{k} p \left( B_j \left( \sum_{i=1}^{n} A_i(x_2) \right) \right)
\]

\[
= \frac{1}{k} p \left( \sum_{j=1}^{k} B_j(x_1) \right) + \frac{1}{n} p \left( \sum_{i=1}^{n} A_i(x_2) \right)
\]

where the first two equalities follow from the fact that \( G \) is an abelian semigroup of linear operators on \( X \), the second and third inequalities follow from the fact that \( p \) is sublinear, the fourth inequality follows from (ii) in the assumption. Then, by the definition of \( q \), we have

\[
q(x_1 + x_2) \leq q(x_1) + q(x_2)
\]

and the right-hand-side makes sense since \( q(x_1) < +\infty \) and \( q(x_2) < +\infty \). Note that \( 0 = q(\vartheta_X) \leq q(x_1) + q(-x_1) \). Then, \( q(x_1) > -\infty \). Hence, \( q : X \to \mathbb{R} \).
On the other hand, we have a linear functional on \( X \) such that

\[
q \left( \sum_{i=1}^{n} A_i(x) \right) = \frac{1}{n} \sum_{i=1}^{n} A_i(x)
\]

for \( \alpha, p \in \mathbb{R} \). Hence, by the arbitrari ness of \( q \) where the first inequality follows from the definition of \( p \), the second equality follows from (ii) in the assumption, and the third inequality follows from Proposition 3.81 and the fact that \( q(x) \in \mathbb{R} \). Hence, we have shown that \( q \) is a sublinear functional.

∀ \( m \in M, \forall n \in \mathbb{N}, \forall A_1, \ldots, A_n \in \mathcal{G} \), we have

\[
\text{Re} \left( f(m) \right) = \frac{1}{n} \sum_{i=1}^{n} \text{Re} \left( f(A_i, m) \right) = \frac{1}{n} \text{Re} \left( f \left( \sum_{i=1}^{n} A_i, m \right) \right)
\]

where the first equality follows from (i) and (iii) in the assumptions, the second equality follows from the linearity of \( f \), and the first inequality follows from (iv) in the assumptions. Hence, we have \( \text{Re} \left( f(m) \right) \leq q(m) \), ∀ \( m \in M \).

By the extension forms of Hahn-Banach Theorem (Theorem 7.80 for \( \mathbb{K} = \mathbb{R} \) and Theorem 7.82 for \( \mathbb{K} = \mathbb{C} \)), there exists a linear functional \( F : \mathcal{X} \rightarrow \mathbb{K} \) such that \( F|_M = f \) and \( \text{Re} \left( F(x) \right) \leq q(x) \), ∀ \( x \in \mathcal{X} \).

∀ \( x \in \mathcal{X}, \forall A \in \mathcal{G}, \forall n \in \mathbb{N} \), we have

\[
q(x - Ax) \leq \frac{1}{n} \left( E(x - Ax) + A(x - Ax) + \cdots + A^{n-1}(x - Ax) \right)
\]

where the first inequality follows from the definition of \( q \), the second inequality follows from the fact that \( p \) is sublinear, and the third inequality follows from (ii) in the assumption. Hence, by the arbitrariness of \( n \), we have \( q(x - Ax) \leq 0 \). We will show that \( F(Ax) = F(x) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \mathbb{K} = \mathbb{R} \); Case 2: \( \mathbb{K} = \mathbb{C} \).

Case 1: \( \mathbb{K} = \mathbb{R} \). We have \( F(x) - F(Ax) = F(x - Ax) \leq q(x - Ax) \leq 0 \). On the other hand, we have \( F(Ax) - F(x) = F(-x) - F(A(-x)) = F((-x) - A(-x)) \leq q((-x) - A(-x)) \leq 0 \). Hence, \( F(x) = F(Ax) \).

Case 2: \( \mathbb{K} = \mathbb{C} \). Define \( H : \mathcal{X} \rightarrow \mathbb{R} \) by \( H(x) = \text{Re} \left( F(x) \right) \), ∀ \( x \in \mathcal{X} \). By Lemma 7.40, \( (\mathcal{X}, \mathbb{R}) \) is a vector space. By Lemma 7.81, \( H \) is a linear functional on \( (\mathcal{X}, \mathbb{R}) \) and \( F(x) = H(x) - iH(ix) \), ∀ \( x \in \mathcal{X} \). Note that \( H(x) - H(Ax) = H(x - Ax) = \text{Re} \left( F(x - Ax) \right) \leq q(x - Ax) \leq 0 \). On the other hand, we have \( H(Ax) - H(x) = H(-x) - H(A(-x)) = \).
DUAL SPACES

7.9. DUAL SPACES

\[ H((-x) - A(-x)) = \Re (F((-x) - A(-x))) \leq q((-x) - A(-x)) \leq 0. \]

Hence, \( H(x) = H(Ax) \). Then, \( F(Ax) = H(Ax) - iH(Aix) = H(Ax) - iH(A(Aix)) = H(x) - iH(ix) = F(x) \).

Furthermore, if, in addition, \( \mathcal{X} \) is normed with \( \| \cdot \| \) and \( p \) is continuous at \( \varphi_x \), then \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall x \in \mathcal{B}(\varphi_x, \delta) \), we have \( |p(x) - p(\varphi_x)| = |p(x)| < \varepsilon \). Note that \( q(x) \leq p(x) < \varepsilon \). Then, \( q(-x) \leq p(-x) < \varepsilon \). Since \( q \) is a sublinear functional, we have \( 0 = q(\varphi_x) \leq q(x) + q(-x) \). Hence, we have \( q(x) \geq -q(-x) > -\varepsilon \). Therefore, \( |q(x) - q(\varphi_x)| = |q(x)| < \varepsilon \). Hence, \( q \) is continuous at \( \varphi_x \). Therefore, by the extension forms of Hahn-Banach Theorem, \( F \) is continuous.

This completes the proof of the corollary.

\[\square\]

Proposition 7.85 Let \( \mathcal{X} \) be a normed linear space over the field \( \mathbb{K} \) and \( x_0 \in X \) with \( x_0 \neq \varphi_x \). Then, \( \exists x_* \in \mathcal{X}^* \) with \( \| x_* \| = 1 \) such that \( \langle x_*, x_0 \rangle = \| x_* \| \| x_0 \| \).

Proof Consider the subspace \( M := \text{span}(\{x_0\}) \). Define \( f : M \to \mathbb{K} \) by \( f(\alpha x_0) = \alpha \| x_0 \|, \forall \alpha \in \mathbb{K} \). Clearly, \( f \) is uniquely defined by Proposition 6.17. Clearly \( f \) is a linear functional on \( M \) and \( \| f \|_M := \sup_{m \in M, \| m \| \leq 1} |f(m)| = \sup_{\alpha \in \mathbb{K}, \| \alpha x_0 \| \leq 1} |\alpha \| x_0 \| | = 1 \). By the simple version of Hahn-Banach Theorem, there exists \( x_* \in \mathcal{X}^* \) such that \( x_* \mathcal{M} = f \) and \( \| x_* \| = \| f \|_M = 1 \). Then, \( \langle x_*, x_0 \rangle = f(x_0) = \| x_0 \| = \| x_* \| \| x_0 \| \). This completes the proof of the proposition.

\[\square\]

Example 7.86 Let \( Y \) be a normed linear space over the field \( \mathbb{K} \), \( \mathcal{X} := l_\infty(Y) \) be the normed linear space as defined in Example 7.10, and \( y_0 \in Y \) with \( y_0 \neq \varphi_y \). We will show that \( \mathcal{X}^* \neq l_1(Y^*) \) by Hahn-Banach Theorem.

Let \( M := \{ x := (\xi_1, \xi_2, \ldots) \in \mathcal{X} \mid \lim_{n \to \infty} \xi_n \in Y \} \). Clearly, \( M \) is a subspace of \( \mathcal{X} \). By Proposition 7.85, there exists \( y_* \in \mathcal{Y}^* \) such that \( \| y_* \| = 1 \) and \( \langle y_*, y_0 \rangle = \| y_0 \| > 0 \). Define \( f : M \to \mathbb{K} \) by \( f(m) = \langle y_*, \lim_{n \to \infty} \xi_n \rangle, \forall m := (\xi_1, \xi_2, \ldots) \in M \). It is easy to show that \( f \) is a linear functional on \( M \). \( \forall m = (\xi_1, \xi_2, \ldots) \in M \) with \( \| m \| \leq 1 \), we have

\[ |f(m)| = |\langle y_*, \lim_{n \to \infty} \xi_n \rangle| \leq \| \lim_{n \to \infty} \xi_n \| = \lim_{n \to \infty} \| \xi_n \| \leq \| y_* \| \leq 1 \]

where we have applied Propositions 7.72, 7.21, and 3.66. Hence, we have \( \| f \|_M := \sup_{m \in M, \| m \| \leq 1} |f(m)| \leq 1 \). Consider \( m_0 = (y_0, y_0, \ldots) \in M \), we have \( |f(m_0)| = |\langle y_*, y_0 \rangle| = \| y_0 \| = \| m_0 \| > 0 \). Hence, by Proposition 7.72, \( \| f \|_M = 1 \). By the simple version of Hahn-Banach Theorem, there exists \( x_* \in \mathcal{X}^* \) such that \( x_* \mathcal{M} = f \) and \( \| x_* \| = \| f \|_M = 1 \). Clearly, there does not exist \( (\eta_1, \eta_2, \ldots) \in l_1(Y^*) \) such that \( x_* \) is given by \( \langle x_*, x \rangle = \sum_{n=1}^{\infty} \langle \eta_n, \xi_n \rangle, \forall x = (\xi_1, \xi_2, \ldots) \in \mathcal{X} \). Hence, \( \mathcal{X}^* \neq l_1(Y^*) \). \(\diamond\)
7.9.4 Second dual space

**Definition 7.87** Let $\mathcal{X}$ be a normed linear space over the field $K$ and $\mathcal{X}^*$ be its dual. The dual space of $\mathcal{X}^*$ is called the second dual of $\mathcal{X}$ and is denoted by $\mathcal{X}^{**}$.

**Remark 7.88** Let $\mathcal{X}$ be a normed linear space over the field $K$, $\mathcal{X}^*$ be its dual, and $\mathcal{X}^{**}$ be its second dual. Then, $\mathcal{X}$ is isometrically isomorphic to a dense subset of a Banach space, which can be taken as a subspace of $\mathcal{X}^{**}$. This can be proved as follows.

By Proposition 7.72, $\mathcal{X}^{**}$ is a Banach space over $K$. $\forall x \in \mathcal{X}$, define a functional $f : \mathcal{X}^* \to K$ by $f(x_*) = \langle x_*, x \rangle$, $\forall x_* \in \mathcal{X}^*$. Clearly, $f$ is a linear functional on $\mathcal{X}^*$. Note that $|f(x_*)| = |\langle x_*, x \rangle| \leq \|x\| \|x_*\|$, $\forall x_* \in \mathcal{X}^*$, by Proposition 7.72. Then, $\|f\| \leq \|x\|$. When $x = \vartheta_{\mathcal{X}}$, then $\|f\| = 0 = \|x\|$. When $x \neq \vartheta_{\mathcal{X}}$, then, by Proposition 7.85, $\exists x_0 \in \mathcal{X}^*$ with $\|x_0\| = 1$ such that $\langle x_0, x \rangle = \|x\|$. Then, $\|f\| = \sup_{x_* \in \mathcal{X}^*, \|x_*\| \leq 1} |f(x_*)| \geq |f(x_0)| = \|x\|$. Hence, we have $\|f\| = \|x\|$. Hence, $f \in \mathcal{X}^{**}$. Thus, we may define a natural mapping $\phi : \mathcal{X} \to \mathcal{X}^{**}$ by $\phi(x) = f$, $\forall x \in \mathcal{X}$.

$\forall x_1, x_2 \in \mathcal{X}$, $\forall \alpha, \beta \in K$, $\forall x_* \in \mathcal{X}^*$, we have $\phi(\alpha x_1 + \beta x_2)(x_*) = \langle \alpha x_1 + \beta x_2, x_* \rangle = \alpha \langle x_1, x_* \rangle + \beta \langle x_2, x_* \rangle = \alpha \phi(x_1)(x_*) + \beta \phi(x_2)(x_*)$. Hence, $\phi$ is a linear function. $\forall x \in \mathcal{X}$, we have $\|\phi(x)\| = \|x\|$. Hence, $\phi$ is an isometry and $\phi$ is injective. Therefore, $\phi$ is an isometrical isomorphism between $\mathcal{X}$ and $\phi(\mathcal{X})$. Clearly, $\phi(\mathcal{X}) \subseteq \mathcal{X}^{**}$ is a subspace. Then, by Proposition 7.17, $\phi(\mathcal{X}) \subseteq \mathcal{X}^{**}$ is a closed subspace. By Proposition 4.39, $\phi(\mathcal{X})$ is complete. Hence, $\phi(\mathcal{X})$ is a Banach space. By Proposition 3.5, $\phi(\mathcal{X})$ is dense in $\phi(\mathcal{X})$. This completes the proof of the remark.

**Definition 7.89** Let $\mathcal{X}$ be a normed linear space over the field $K$, $\mathcal{X}^{**}$ be its second dual, and $\phi : \mathcal{X} \to \mathcal{X}^{**}$ be the natural mapping as defined in Remark 7.88. $\mathcal{X}$ is said to be reflexive if $\phi(\mathcal{X}) = \mathcal{X}^{**}$, that is $\mathcal{X}$ and $\mathcal{X}^{**}$ are isometrically isomorphic. Then, we may label $\mathcal{X}^{**}$ such that $\mathcal{X}^{**} = \mathcal{X}$ and $\phi = \text{id}_\mathcal{X}$.

A reflexive normed linear space is clearly a Banach space.

**Proposition 7.90** Let $\mathcal{X}$ be a Banach space over the field $K$ and $\mathcal{X}^*$ be its dual. Then, $\mathcal{X}$ is reflexive if, and only if, $\mathcal{X}^*$ is reflexive.

**Proof** Let $\mathcal{X}^{**}$ be the second dual of $\mathcal{X}$ and $\mathcal{X}^{***}$ be the dual of $\mathcal{X}^{**}$.

"Necessity" Let $\mathcal{X}$ be reflexive. Then, $\mathcal{X}^{**} = \mathcal{X}$ isometrically isomorphically. Then, $\mathcal{X}^* = (\mathcal{X}^{**})^* = \mathcal{X}^{***}$ isometrically isomorphically. Hence, $\mathcal{X}^*$ is reflexive.

"Sufficiency" Let $\mathcal{X}^*$ be reflexive. Then, $\mathcal{X}^{***} = \phi_s(\mathcal{X}^*)$, where $\phi_s : \mathcal{X}^* \to \mathcal{X}^{***}$ is the natural mapping on $\mathcal{X}^*$. We will show that $\mathcal{X}$ is reflexive by an argument of contradiction. Suppose $\mathcal{X}$ is not reflexive. Let $\phi : \mathcal{X} \to \mathcal{X}^{**}$ be the natural mapping. $\phi$ is an isometrical isomorphism between
7.9. DUAL SPACES

\(X\) and \(ϕ(X)\). Since \(X\) is complete, then \(ϕ(X)\) is also complete. Then, by Proposition 4.39, \(ϕ(X)\) is a closed subspace of \(X^{**}\). Since \(X\) is not reflexive, then there exists \(y_{**} \in X^{**} \setminus ϕ(X)\). Then, \(y_{**} \neq ϕ(x)\).

Let \(δ := \inf_{m_{**} \in ϕ(X)} \| y_{**} - m_{**} \|\). By Proposition 4.10, \(δ \in (0, \infty) \subset \mathbb{R}\).

Consider the subspace \(N := \{ m_{**} + αy_{**} \in X^{**} \mid α \in \mathbb{K}, m_{**} \in ϕ(X) \}\).

Since \(y_{**} \notin ϕ(X)\), then \(\forall n_{**} \in N, \exists α \in \mathbb{K}\) and \(∃! m_{**} \in ϕ(X)\) such that \(n_{**} = αy_{**} + m_{**}\). Define a linear functional \(f : N \rightarrow \mathbb{K}\) by \(f(n_{**}) = α, ∀n_{**} = αy_{**} + m_{**} \in N\), where \(m_{**} \in ϕ(X)\) and \(α \in \mathbb{K}\). \(∀n_{**} \in N\), let \(n_{**} = αy_{**} + m_{**}\), where \(m_{**} \in ϕ(X)\) and \(α \in \mathbb{K}\). We have \(f(n_{**}) = |α| = |αδ| ≤ 1/δ \| αy_{**} + m_{**} \| = 1/δ \| n_{**} \|. \) Hence, \(f(M) ≤ 1/δ \). By the simple version of Hahn-Banach Theorem, there exists \(F \in X^{**}\) such that \(F|_M = f\) and \(\| F \| = \| f \|_M\).

Since \(X^*\) is reflexive, then \(ϕ(X^*) = X^{**}\).

Then, \(∃ y_{*} \in X^*\) such that \(F_1 = ϕ(y_{*})\). Then, \(∀ x_{**} \in X^{**}\), we have \(F(x_{**}) = ϕ(y_{*})(x_{**}) = \langle x_{**}, y_{*} \rangle\). \(∀ y_{**} \in ϕ(X)\), we have \(\langle m_{**}, y_{**} \rangle = F(m_{**}) = f(m_{**}) = 0\). Hence, \(∀ x \in X, \langle y_{*}, x \rangle = ϕ(x)(y_{*}) = \langle ϕ(x), y_{*} \rangle = 0\).

Hence, \(y_{*} = \vartheta \). Then, \(F = ϕ(\vartheta) = \vartheta_{**}\). This contradicts with the fact that \(F(y_{**}) = f(y_{**}) = 1 \neq 0\). Therefore, \(X\) must be reflexive.

This completes the proof of the proposition. □

Example 7.91 Clearly, the normed linear space \((\mathcal{V}_X, \mathbb{K}, \| \cdot \|)\) as defined in Example 7.73 is reflexive. Clearly \(\mathbb{K}^n, n \in \mathbb{N}\), are reflexive. Let \(\mathcal{V}_y\) be a reflexive Banach space. Then, \(l_p(\mathcal{V}_y), p \in (1, \infty) \subset \mathbb{R}\), are reflexive. \(l_1(\mathcal{V}_y)\) is not reflexive by Example 7.86. Then, by Proposition 7.90, \(l_\infty(\mathcal{V}_y)\) is not reflexive either. □

Proposition 7.92 Let \(X\) be a finite-dimensional normed linear space over the field \(\mathbb{K}\). Then, \(X\) is reflexive.

Proof Let \(n \in \mathbb{Z}_+\) be the dimension of \(X\) and \(ϕ : X \rightarrow X^{**}\) be the natural mapping. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(n = 0\); Case 2: \(n \in \mathbb{N}\). Case 1: \(n = 0\). Then, \(X = (\{ 0 \}, \mathbb{K}, \| \cdot \|)\) and \(X^* = (\{ 0 \}, \mathbb{K}, \| \cdot \|)\).

Clearly, \(ϕ(0) = 0_{**}\). Hence, \(ϕ(X) = X^{**}\) and \(X\) is reflexive.

Case 2: \(n \in \mathbb{N}\). Let \(\{ e_1, ..., e_n \} \subseteq X\) be a basis of \(X\). \(∀ x \in X\), by Corollary 6.47, \(x\) can be uniquely expressed as \(\sum_{j=1}^{n} α_j e_j\) for some \(α_1, ..., α_n \in \mathbb{K}\). \(∀ i = 1, ..., n\), define \(f_i : X \rightarrow \mathbb{K}\) by \(f_i(x) = α_i\), \(∀ x = \sum_{j=1}^{n} α_j e_j \in X\). Clearly, \(f_i\) is well-defined and a linear functional. By Proposition 7.97, \(f_i\) is continuous and \(f_i \in X^*\). Denote \(f_i\) by \(e_i \in X^*\). \(∀ x_{**} \in X^*, ∀ i = 1, ..., n\), let \(β_i = \langle x_{**}, e_i \rangle \in \mathbb{K}\). \(∀ x = \sum_{j=1}^{n} α_j e_j \in X\), we have

\[\langle \langle x_{**}, x \rangle \rangle = \sum_{j=1}^{n} α_j \langle \langle x_{**}, e_j \rangle \rangle = \sum_{j=1}^{n} β_j α_j = \sum_{k=1}^{n} β_j \langle \langle e_{**}, x \rangle \rangle = \langle \langle \sum_{k=1}^{n} β_j e_{**}, x \rangle \rangle\]
Hence, \( x_\star = \sum_{j=1}^{n} \beta_j e_{\star j} \). Therefore, \( X^\star = \text{span}\left(\{e_{\star 1}, \ldots, e_{\star n}\}\right) \). \( \forall x_{\star \star} \in X^{\star \star}, \forall i = 1, \ldots, n \), let \( \gamma_i = \langle \langle x_{\star \star}, e_{\star i} \rangle \rangle \in \mathbb{K} \). \( \forall x_\star = \sum_{j=1}^{n} \beta_j e_{\star j} \in X^\star \), we have

\[
\langle \langle x_{\star \star}, x_\star \rangle \rangle = \sum_{j=1}^{n} \beta_j \langle \langle x_{\star \star}, e_{\star j} \rangle \rangle = \sum_{j=1}^{n} \beta_j \gamma_j = \sum_{j=1}^{n} \beta_j \langle \langle e_{\star j}, \sum_{k=1}^{n} \gamma_k e_k \rangle \rangle
\]

\[
= \langle \langle x_\star, \sum_{k=1}^{n} \gamma_k e_k \rangle \rangle = \langle \langle \phi \left( \sum_{k=1}^{n} \gamma_k e_k \right), x_\star \rangle \rangle
\]

Hence, \( x_{\star \star} = \phi \left( \sum_{k=1}^{n} \gamma_k e_k \right) \) and \( X^{\star \star} \subseteq \phi(X) \). This shows that \( X \) is reflexive.

This completes the proof of the proposition. \( \square \)

### 7.9.5 Alignment and orthogonal complements

**Definition 7.93** Let \( X \) be a normed linear space, \( x \in X \), \( X^\star \) be the dual, and \( x_\star \in X^\star \). We say that \( x_\star \) is aligned with \( x \) if \( \langle \langle x_\star, x \rangle \rangle = \| x_\star \| \| x \| \).

Clearly, \( \vartheta \) aligns with any vector in the dual and \( \vartheta_\star \) is aligned with any vector in \( X \). By Proposition 7.85, \( \forall x \in X \) with \( x \neq \vartheta \), there exists a \( x_\star \in X^\star \) with \( x_\star \neq \vartheta_\star \) that is aligned with it.

**Definition 7.94** Let \( X \) be a normed linear space, \( x \in X \), \( X^\star \) be the dual, and \( x_\star \in X^\star \). We say that \( x \) and \( x_\star \) are orthogonal if \( \langle \langle x_\star, x \rangle \rangle = 0 \).

**Definition 7.95** Let \( X \) be a normed linear space, \( S \subseteq X \), and \( X^\star \) be the dual. The orthogonal complement of \( S \), denoted by \( S^\perp \), consists of all \( x_\star \in X^\star \) that is orthogonal to every vector in \( S \), that is \( S^\perp := \{ x_\star \in X^\star \mid \langle \langle x_\star, x \rangle \rangle = 0, \forall x \in S \} \).

**Definition 7.96** Let \( X \) be a normed linear space, \( X^\star \) be the dual, and \( S \subseteq X^\star \). The pre-orthogonal complement of \( S \), denoted by \( ^\perp S \), consists of all \( x \in X \) that is orthogonal to every vector in \( S \), that is \( ^\perp S := \{ x \in X \mid \langle \langle x_\star, x \rangle \rangle = 0, \forall x_\star \in S \} \).

**Proposition 7.97** Let \( X \) be a normed linear space over the field \( \mathbb{K} \), \( M \subseteq X \) be a subspace, and \( y \in \mathbb{K} \). Then, the following statements hold.

1. \( \delta := \inf_{m \in M} \| y - m \| = \max_{x_\star \in M^\perp, \| x_\star \| \leq 1} \text{Re} \left( \langle \langle x_\star, y \rangle \rangle \right) \), where the maximum is achieved at some \( x_{\star 0} \in M^\perp \) with \( \| x_{\star 0} \| \leq 1 \). If the infimum is achieved at \( m_0 \in M \) then \( y - m_0 \) is aligned with \( x_{\star 0} \).

2. If \( \exists m_0 \in M \) and \( \exists x_{\star 0} \in M^\perp \) with \( \| x_{\star 0} \| = 1 \) such that \( y - m_0 \) is aligned with \( x_{\star 0} \), then the infimum is achieved at \( m_0 \) and the maximum is achieved at \( x_{\star 0} \), that is \( \delta = \| y - m_0 \| = \langle \langle x_{\star 0}, y \rangle \rangle = \text{Re} \left( \langle \langle x_{\star 0}, y \rangle \rangle \right) \).
Proof  
(i) \( \forall x_\ast \in M^\perp \) with \( \| x_\ast \| \leq 1, \forall m \in M \), we have

\[
\Re ( \langle x_\ast, y \rangle ) = \Re ( \langle x_\ast, y - m \rangle ) \leq \| \Re ( \langle x_\ast, y - m \rangle ) \|
\leq \| \langle x_\ast, y - m \rangle \| \leq \| x_\ast \| \| y - m \| \leq \| y \|
\]

where we have applied Proposition 7.72 in the third inequality. Then, we have

\[
\sup_{x_\ast \in M^\perp, \| x_\ast \| \leq 1} \Re ( \langle x_\ast, y \rangle ) \leq \inf_{m \in M} \| y - m \| = \delta
\]

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \delta = 0 \); Case 2: \( \delta > 0 \). Take \( x_{00} := \partial_\ast \).

Clearly, \( x_{00} \in M^\perp \) and \( \| x_{00} \| = 0 \leq 1 \). Then, \( \delta = \Re ( \langle x_{00}, y \rangle ) \leq \sup_{x_\ast \in M^\perp, \| x_\ast \| \leq 1} \Re ( \langle x_\ast, y \rangle ) \). Therefore, we have \( \delta = \inf_{m \in M} \| y - m \| = \max_{x_\ast \in M^\perp, \| x_\ast \| \leq 1} \Re ( \langle x_\ast, y \rangle ) \) and the maximum is achieved at some \( x_{00} \in M^\perp \) with \( \| x_{00} \| \leq 1 \). If the infimum is achieved at \( m_0 \in M \), then, \( 0 = \delta = \| y - m_0 \| \) and \( y = m_0 \). Then, clearly \( y - m_0 \) is aligned with \( x_{00} \). This case is proved.

Case 2: \( \delta > 0 \). Then, \( y \in X \setminus M \). Consider the subspace \( \text{span} ( M \cup \{ y \} ) =: \bar{M} \). \( \forall x \in M \), there exists a unique \( \alpha \in K \) and a unique \( m \in M \) such that \( x = \alpha y + m \). Then, we may define a functional \( f : \bar{M} \to K \) by \( f(x) = f(\alpha y + m) = \alpha \delta \). Clearly, \( f \) is a linear functional on \( M \).

\[
\| f \|_{\bar{M}} := \sup_{\| x \| \leq 1} | f(x) | \leq \sup_{\alpha \in K, m \in M, \| \alpha y + m \| \leq 1} | \alpha | \| \delta \| \leq \sup_{\alpha \in K, m \in M, \| \alpha y + m \| \leq 1} \| \alpha y + m \| \leq 1.
\]

By the simple version of Hahn-Banach Theorem, \( \exists x_{00} \in X^\ast \) such that \( x_{00}|_{\bar{M}} = f \) and \( \| x_{00} \| = \| f \|_{\bar{M}} \leq 1 \).

\( \forall m \in M \), we have \( \langle x_{00}, m \rangle = f(m) = 0 \). Hence, \( x_{00} \in M^\perp \).

Note that \( \delta = f(y) = \langle x_{00}, y \rangle \). Hence, \( \delta = \inf_{m \in M} \| y - m \| = \max_{x_\ast \in M^\perp, \| x_\ast \| \leq 1} \Re ( \langle x_\ast, y \rangle ) \) and the maximum is achieved at some \( x_{00} \in M^\perp \) with \( \| x_{00} \| \leq 1 \). If the infimum is achieved at \( m_0 \in M \), then \( \delta = \| y - m_0 \| = \langle x_{00}, y \rangle = \langle x_{00}, y - m_0 \rangle \leq \| x_{00} \| \| y - m_0 \| \leq \| y - m_0 \| \), where the first inequality follows from Proposition 7.72. Hence, \( \langle x_{00}, y - m_0 \rangle = \| x_{00} \| \| y - m_0 \| \) and \( x_{00} \) is aligned with \( y - m_0 \). This case is proved.

(ii) Note that \( \delta \leq \| y - m_0 \| = \| x_{00} \| \| y - m_0 \| = \langle x_{00}, y - m_0 \rangle = \langle x_{00}, y \rangle = \Re ( \langle x_{00}, y \rangle ) \leq \delta \). Then, the result follows.

This completes the proof of the proposition. \( \square \)

Proposition 7.98 Let \( X \) be a normed linear space over \( K \), \( A, C \subseteq X \), and \( B, D \subseteq X^\ast \). Then, the following statements holds.

(i) If \( A \subseteq C \), then \( C^\perp \subseteq A^\perp \). Similarly, if \( B \subseteq D \), then \( D^\perp \subseteq B^\perp \).

(ii) \( A^\perp \subseteq X^\ast \) is a closed subspace.

(iii) \( B^\perp \subseteq X \) is a closed subspace.

(iv) \( \overline{\{ A^\perp \}} = \text{span} ( A ) \).
Proof. (i) This follows directly from Definition 7.95.

(ii) \( \vartheta_\ast \in A^\perp \neq \emptyset \). \( \forall x_{\ast 1}, x_{\ast 2} \in A^\perp, \forall \alpha, \beta \in K, \forall x \in A \), we have 
\[ \langle \langle \alpha x_{\ast 1} + \beta x_{\ast 2}, x \rangle \rangle = \alpha \langle \langle x_{\ast 1}, x \rangle \rangle + \beta \langle \langle x_{\ast 2}, x \rangle \rangle = 0. \]
Hence, \( \alpha x_{\ast 1} + \beta x_{\ast 2} \in A^\perp \) and \( A^\perp \) is a subspace. By Proposition 4.10, we have 
\[ \langle \langle x_{\ast}, x \rangle \rangle \leq \delta \] 
where we have applied Proposition 7.72 in the third inequality. Then, we
have \( \langle \langle x_{\ast}, x \rangle \rangle \leq \delta \). By Proposition 3.3, \( \langle \langle x_{\ast}, x \rangle \rangle \leq \delta \) is a subspace. On the other hand, \( A \subseteq A^\perp \) and \( A^\perp \) is closed. Hence, \( \langle \langle x_{\ast}, x \rangle \rangle \leq \delta \). Then, we have
\[ \langle \langle x_{\ast}, x \rangle \rangle = \lim_{n \to \infty} \langle \langle x_{\ast}, x \rangle \rangle = 0. \]
Then, \( x_{\ast} \in A^\perp \) and \( A^\perp \subseteq A^\perp \).

By Proposition 3.3, \( A^\perp = A^\perp \) and \( A^\perp \) is closed. Hence, (ii) follows.

(iii) This can be shown by a similar argument as (ii).

(iv) Clearly, \( A \subseteq \overline{\langle \langle A \rangle \rangle} \). Then, by (iii), we have \( M := \overline{\langle \langle A \rangle \rangle} \subseteq \langle \langle A \rangle \rangle \). On the other hand, \( A \subseteq M \), then, by (i), we have \( A^\perp \supseteq M^\perp \) and 
\( \langle \langle A \rangle \rangle \subseteq \langle \langle M \rangle \rangle \). \( \forall x_0 \in \langle \langle M \rangle \rangle \), by Proposition 7.97, we have 
\[ \inf_{m \in M} \| x_0 - m \| = \sup_{x_0 \in M^\perp, \| x_0 \| \leq 1} \Re \left( \langle \langle x_0, x_0 \rangle \rangle \right) = 0. \]

Since \( M \) is closed, by Proposition 4.10, \( x_0 \in M \). Hence, we have 
\( \langle \langle A \rangle \rangle \subseteq \langle \langle M \rangle \rangle \subseteq M \). Hence, \( M = \langle \langle A \rangle \rangle \). Hence, (iv) follows.

This completes the proof of the proposition. \( \square \)

Proposition 7.99 Let \( X \) be a normed linear space over the field \( K \), \( M \subseteq X \) be a subspace, and \( y_\ast \in X^\ast \). Then, the following statements holds.

(i) \( \delta := \min_{x_{\ast} \in M^\perp} \| y_\ast - x_{\ast} \| = \sup_{m \in M, \| m \| \leq 1} \Re \left( \langle \langle y_\ast, m \rangle \rangle \right) =: \delta \), where the minimum is achieved at some \( x_{\ast 0} \in M^\perp \). If the supremum is achieved at \( m_0 \in M \) with \( \| m_0 \| \leq 1 \), then \( m_0 \) is aligned with \( y_\ast - x_{\ast 0} \).

(ii) If \( \exists m_0 \in M \) with \( \| m_0 \| = 1 \) and \( \exists x_{\ast 0} \in M^\perp \) such that \( y_\ast - x_{\ast 0} \) is aligned with \( m_0 \), then the minimum is achieved at \( x_{\ast 0} \) and the supremum is achieved at \( m_0 \), that is \( \delta = \| y_\ast - x_{\ast 0} \| = \langle \langle y_\ast, m_0 \rangle \rangle = \Re \left( \langle \langle y_\ast, m_0 \rangle \rangle \right) \).

Proof. (i) \( \forall x_{\ast} \in M^\perp, \forall m \in M \) with \( \| m \| \leq 1 \), we have 
\[ \Re \left( \langle \langle y_\ast, m \rangle \rangle \right) = \Re \left( \langle \langle y_\ast - x_{\ast}, m \rangle \rangle \right) \leq | \Re \left( \langle \langle y_\ast - x_{\ast}, m \rangle \rangle \right) | \leq | \langle \langle y_\ast - x_{\ast}, m \rangle \rangle | \leq \| y_\ast - x_{\ast} \| \| m \| \leq \| y_\ast - x_{\ast} \| \]
where we have applied Proposition 7.72 in the third inequality. Then, we have
\[ \delta = \inf_{x_{\ast} \in M^\perp} \| y_\ast - x_{\ast} \| = \sup_{m \in M, \| m \| \leq 1} \Re \left( \langle \langle y_\ast, m \rangle \rangle \right) \]

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \delta = 0 \); Case 2: \( \delta > 0 \). Case 1: \( \delta = 0 \). By Proposition 7.98, \( M^\perp \) is a closed subspace. By Proposition 4.10, we have \( y_\ast \in M^\perp \) the minimum is achieved at a unique vector \( x_{\ast 0} := y_\ast \). Take \( m_0 = \vartheta \) with \( \| m_0 \| = 0 \leq 1 \). Then, \( \delta = \Re \left( \langle \langle y_\ast, m_0 \rangle \rangle \right) \leq \sup_{m \in M, \| m \| \leq 1} \Re \left( \langle \langle y_\ast, m \rangle \rangle \right) = \delta \leq \delta \). Hence,
7.9. DUAL SPACES

\[ \delta = \min_{x \in M^\perp} \| y_s - x \| = \sup_{m \in M^\perp, \|m\| \leq 1} \Re \left( \langle \langle y_s, m \rangle \rangle \right) = \tilde{\delta} \]
and the minimum is achieved at \( x_{s_0} = y_s \in M^\perp \). If the supremum is achieved at \( m_0 \in M \) with \( \| m_0 \| \leq 1 \). Then, clearly \( m_0 \) is aligned with \( y_s - x_{s_0} \). This case is proved.

Case 2: \( \delta > 0 \). Then, \( y_s \not\in M^\perp \). Consider the subspace \( M \). Then, we may define a functional \( f : M \to K \) by \( f(m) = \langle \langle y_s, m \rangle \rangle \). Clearly, \( f \) is a linear functional on \( M \). \( \forall m \in M \) with \( \| m \| \leq 1 \), we have either \( f(m) = 0 \), then \( \| f(m) \| = 0 \leq \sup_{m \in M^\perp, \|m\| \leq 1} \Re \left( \langle \langle y_s, m \rangle \rangle \right) = \delta \); or \( f(m) \neq 0 \), then \( \| f(m) \| = \frac{f(m)}{f(m)} f(m) = \Re \left( \frac{f(m)}{f(m)} f(m) \right) = \Re \left( \frac{\langle \langle y_s, m \rangle \rangle}{f(m)} f(m) \right) \leq \delta \). Hence, \( \| f \|_M := \sup_{m \in M^\perp, \|m\| \leq 1} \| f(m) \| \leq \delta \). By the simple version of Hahn-Banach Theorem, \( \exists y_{s_0} \in X^\ast \) such that \( y_{s_0}|_M = f \) and \( \| y_{s_0} \| = \| f \|_M \leq \delta \). Let \( x_{s_0} := y_s - y_{s_0} \). \( \forall m \in M \), we have \( \langle \langle x_{s_0}, m \rangle \rangle = \langle \langle y_s, m \rangle \rangle - \langle \langle y_{s_0}, m \rangle \rangle = 0 \). Hence, \( x_{s_0} \in M^\perp \). Note that \( \delta \geq \delta \geq \| f \|_M = \| y_s \| = \| y_s - x_{s_0} \| \geq \inf_{x \in M^\perp} \| y_s - x \| = \delta \).

Hence, \( \delta = \min_{x \in M^\perp} \| y_s - x \| = \sup_{m \in M^\perp, \|m\| \leq 1} \Re \left( \langle \langle y_s, m \rangle \rangle \right) \) and the minimum is achieved at some \( x_{s_0} \in M^\perp \). If the supremum is achieved at \( m_0 \in M \) with \( \| m_0 \| \leq 1 \), then \( \delta = \| y_s - x_{s_0} \| \geq \| y_s - x_{s_0} \| \| m_0 \| \| \leq \| \langle \langle y_s, m_0 \rangle \rangle \| = \| \langle \langle y_s, m_0 \rangle \rangle \| \geq \Re \left( \langle \langle y_s, m_0 \rangle \rangle \right) = \delta \), where the second inequality follows from Proposition 7.72. Hence, \( \langle \langle y_s - x_{s_0}, m_0 \rangle \rangle = \langle \langle y_s - x_{s_0}, m_0 \rangle \rangle = \langle \langle y_s, m_0 \rangle \rangle = \Re \left( \langle \langle y_s, m_0 \rangle \rangle \right) \). Hence, the result follows.

This completes the proof of the proposition.

\[ \square \]

**Proposition 7.100** Let \( X \) be a normed linear space over the field \( K \) and \( S \subseteq X \) be a subspace. By Proposition 7.13, \( S \) is a normed linear space over \( K \). Then, the following statement holds.

(i) \( S^\ast \) is isometrically isomorphic to \( X^\ast / S^\perp \).

(ii) If \( X \) is reflexive and \( S \) is closed, then \( S \) is reflexive.

**Proof**

(i) Define a mapping \( A : X^\ast \to S^\ast \) by \( \langle \langle A(x_s), s \rangle \rangle = \langle \langle x_s, s \rangle \rangle \), \( \forall x_s \in X^\ast \) and \( \forall s \in S \). \( \forall x_s \in X^\ast \), we will show that \( A(x_s) \in S^\ast \). Clearly, \( A(x_s) \) is a linear functional on \( S \). \( \| A(x_s) \| := \sup_{s \in S, \|s\| \leq 1} \| \langle \langle x_s, s \rangle \rangle \| = \sup_{s \in S, \|s\| \leq 1} \| \langle \langle x_s, s \rangle \rangle \| \leq \sup_{s \in S, \|s\| \leq 1} \| x_s \| \| s \| \leq \| x_s \| < +\infty \). Hence, \( A(x_s) \in S^\ast \) and \( S \) is well-defined.

\( \forall x_{s_1}, x_{s_2} \in X^\ast \), \( \forall \alpha, \beta \in K \), \( \forall s \in S \), we have

\[ \langle \langle A(\alpha x_{s_1} + \beta x_{s_2}), s \rangle \rangle = \langle \langle \alpha x_{s_1} + \beta x_{s_2}, s \rangle \rangle = \alpha \langle \langle x_{s_1}, s \rangle \rangle + \beta \langle \langle x_{s_2}, s \rangle \rangle \]

\[ = \alpha \langle \langle A(x_{s_1}), s \rangle \rangle + \beta \langle \langle A(x_{s_2}), s \rangle \rangle = \langle \langle \alpha A(x_{s_1}) + \beta A(x_{s_2}), s \rangle \rangle \]

Hence, \( A \) is a linear function. Since \( \| A x_s \| \leq \| x_s \| \), \( \forall x_s \in X^\ast \), then \( A \in B(\mathcal{X}, S^\ast) \) with \( \| A \| \leq 1 \).
By Proposition 7.68, \( N(A) \subseteq X^\ast \) is a closed subspace. \( \forall x_\ast \in N(A), \) we have \( Ax_\ast = \vartheta_S \). \( \forall s \in S, 0 = \langle \langle Ax_\ast, s \rangle \rangle = \langle \langle x_\ast, s \rangle \rangle. \) Hence, \( x_\ast \in S^\perp. \) Therefore, \( N(A) \subseteq S^\perp. \) On the other hand, \( \forall x_\ast \in S^\perp, \forall s \in S, \) we have \( 0 = \langle \langle x_\ast, s \rangle \rangle = \langle \langle Ax_\ast, s \rangle \rangle. \) Then, \( Ax_\ast = \vartheta_S \ast. \) Hence, \( x_\ast \in N(A). \) Thus, \( S^\perp \subseteq N(A). \) In conclusion, \( S^\perp = N(A). \)

By Proposition 7.45, \( X^\ast/S^\perp \) is a Banach space. Let \( \phi : X^\ast \rightarrow X^\ast/S^\perp \) be the natural homomorphism. By Proposition 7.70, there exists \( A_D \in B (X^\ast/S^\perp, S^\ast) \) such that \( A = A_D \circ \phi, A_D \) is injective, and \( \| A_D \| = \| A \| \leq 1. \)

\( \forall s_\ast \in S^\ast, \) by the simple version of Hahn-Banach Theorem, there exists \( x_\ast \in X^\ast \) such that \( x_\ast|_S = s_\ast \) and \( \| x_\ast \| = \| s_\ast \|. \) \( \forall s \in S, \) we have \( \langle \langle s_\ast, s \rangle \rangle = \langle \langle Ax_\ast, s \rangle \rangle = \langle \langle s_\ast, s \rangle \rangle = \langle \langle A_D(\phi(x_\ast)), s \rangle \rangle = \langle \langle A_D([x_\ast]), s \rangle \rangle. \) Hence, \( s_\ast = A_D([x_\ast]). \) Then, \( A_D \) is surjective. Thus, \( A_D \) is bijective.

\( \forall [x_\ast] \in X^\ast/S^\perp, \| [x_\ast] \| = \inf_{y_\ast \in S^\perp} \| x_\ast - y_\ast \|, \) by Proposition 7.44. By Proposition 7.99, we have
\[
\| [x_\ast] \| = \min_{y_\ast \in S^\perp} \| x_\ast - y_\ast \| = \sup_{s \in S} \| s \| \leq 1
\]
\[
= \sup_{s \in S} \| A_D([x_\ast], s) \| \leq \sup_{s \in S} \| A_D(x_\ast, s) \| \leq \sup_{s \in S} \| A_D([x_\ast], s) \| \leq \| A_D([x_\ast], s) \| \leq \| A([x_\ast, y_\ast]) \| = \inf_{y_\ast \in N(A)} \| A([x_\ast, y_\ast]) \| \leq \inf_{y_\ast \in S^\perp} \| A([x_\ast, y_\ast]) \| \]
\]
where we have applied Proposition 7.72 in the third inequality and Proposition 7.64 in the fifth inequality. Therefore, we have \( \| [x_\ast] \| = \| A_D([x_\ast]) \| \) and \( A_D \) is an isometry. Thus, \( A_D \) is an isometrical isomorphism between \( X^\ast/S^\perp \) and \( S^\ast. \) Hence, \( (i) \) is established.

(ii) Let \( X \) be reflexive and \( S \) be a closed subspace. Let \( \psi : S \rightarrow S^{\ast\ast} \) be the natural mapping. All we need to show is that \( \psi(S) = S^{\ast\ast}. \) Fix a \( s_\ast \in S^{\ast\ast}. \) Define a functional \( \tau : X^\ast \rightarrow \mathbb{K} \) by \( \tau(x_\ast) = \langle \langle s_\ast, Ax_\ast \rangle \rangle, \) \( \forall x_\ast \in X^\ast. \) It is easy to show that \( \tau \) is a linear functional on \( X^\ast. \) \( \forall x_\ast \in X^\ast \) with \( \| x_\ast \| \leq 1, \) we have \( |\tau(x_\ast)| = |\langle \langle s_\ast, Ax_\ast \rangle \rangle| \leq \| s_\ast \| \| Ax_\ast \| \leq \| s_\ast \| \| A \| \| x_\ast \| \leq \| s_\ast \| \| A \| \| x_\ast \| \leq \| s_\ast \| < +\infty, \) where we have applied Propositions 7.72 and 7.64. Hence, \( \tau \in X^{\ast\ast}. \)

Since \( X \) is reflexive, then, by Remark 7.88 and Definition 7.89, \( \exists x_0 \in X \) such that \( \tau(x_0) = \langle \langle x_0, x_0 \rangle \rangle, \) \( \forall x_\ast \in X^\ast. \) \( \forall y_\ast \in S^\perp, \) we have \( \langle \langle y_\ast, x_0 \rangle \rangle = \tau(y_\ast) = \langle \langle s_\ast, Ay_\ast \rangle \rangle = \langle \langle s_\ast, \vartheta_S \ast \rangle \rangle = 0, \) where the third equality follows from the fact that \( S^\perp = N(A). \) Hence, \( x_0 \in \perp(S^\perp) = S, \) by Proposition 7.98. Hence, \( \forall s_\ast \in S^\ast, \exists x_\ast \in X^\ast \) such that \( Ax_\ast = A_D([x_\ast]) = s_\ast. \)
since $A_D$ is an isometrical isomorphism. Then, we have \( \langle \langle s_*, s_* \rangle \rangle = \langle \langle s_*, Ax_* \rangle \rangle = \tau(x_*) = \langle \langle x_*, x_0 \rangle \rangle = \langle \langle Ax_*, x_0 \rangle \rangle = \langle \langle s_*, x_0 \rangle \rangle \). This implies that $s_* = \psi(x_0)$ and $\psi(S) = S^{**}$. Hence, (ii) is established.

This completes the proof of the proposition. \(\square\)

### 7.10 The Open Mapping Theorem

**Definition 7.101** Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces and $A : \mathcal{X} \rightarrow \mathcal{Y}$. $A$ is called an open mapping if $\forall O_X \in \mathcal{O}_X, \ we \ have \ A(O_X) \in \mathcal{O}_Y$, that is the image of each open set is open.

**Proposition 7.102** Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $S, T \subseteq \mathcal{X}$, and $\alpha \in \mathbb{K}$. Then, the following statements hold.

(i) $\overline{\alpha S} = \alpha \overline{S}$.

(ii) If $\alpha \neq 0$, then $\overline{\alpha S} = \alpha \overline{S}$.

(iii) If $\alpha \neq 0$, then $(\alpha S)^o = \alpha S^o$.

(iv) $\overline{S + T} \subseteq \overline{S} + \overline{T}$.

(v) $S^o + T^o \subseteq (S + T)^o$.

**Proof**

(i) We will distinguish three exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$ and $S = \emptyset$; Case 2: $\alpha = 0$ and $S \neq \emptyset$; Case 3: $\alpha \neq 0$. 

Case 1: $\alpha = 0$ and $S = \emptyset$. Then, $\overline{S} = \emptyset$ and $\overline{\alpha S} = \emptyset = \alpha \overline{S}$. Case 2: $\alpha = 0$ and $S \neq \emptyset$. Then, $\overline{S} \neq \emptyset$ and $\overline{\alpha S} = \{0\} = \alpha \overline{S}$.

Case 3: $\alpha \neq 0$. $\forall x \in \alpha S$, by Proposition 4.13, $\exists (x_n)_{n=1}^\infty \subseteq \alpha S$ such that $\lim_{n \to \infty} x_n = x$. Note that $(\alpha x_n)_{n=1}^\infty \subseteq S$ and $\lim_{n \to \infty} \alpha x_n = \alpha^{-1}x$ by Propositions 3.66 and 7.23. By Proposition 4.13, $\alpha^{-1}x \in \overline{S}$. Then, $x \in \alpha \overline{S}$. This shows that $\overline{\alpha S} \subseteq \alpha \overline{S}$.

On the other hand, $\forall x \in \overline{\alpha S}$, then $\alpha^{-1}x \in \overline{S}$. By Proposition 4.13, $\exists (x_n)_{n=1}^\infty \subseteq S$ such that $\lim_{n \to \infty} x_n = \alpha^{-1}x$. Note that $(\alpha x_n)_{n=1}^\infty \subseteq \alpha S$ and $\lim_{n \to \infty} \alpha x_n = x$ by Propositions 3.66 and 7.23. By Proposition 4.13, $x \in \overline{\alpha S}$. Then, $\overline{\alpha S} \subseteq \overline{\alpha S}$. Hence, $\overline{\alpha S} = \alpha \overline{S}$.

(ii) $\forall x \in \overline{\alpha S}$ if, and only if, $\alpha^{-1}x \in S$ if, and only if, $x \in \overline{\alpha S}$. Hence, $\overline{\alpha S} = \alpha \overline{S}$.

(iii) $\forall x \in (\alpha S)^o$, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $B(x, \delta) \subseteq \alpha S$. Then, $B(\alpha^{-1}x, \delta/|\alpha|) = \alpha^{-1}B(x, \delta) \subseteq S$. Hence, $\alpha^{-1}x \in S^o$. Then, $x \in \alpha S^o$. This shows that $(\alpha S)^o \subseteq \alpha S^o$.

On the other hand, $\forall x \in \alpha S^o$, we have $\alpha^{-1}x \in S^o$. $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $B(\alpha^{-1}x, \delta) \subseteq S$. Then, $B(x, |\alpha|\delta) = \alpha B(\alpha^{-1}x, \delta) \subseteq \alpha S$. Hence, $x \in (\alpha S)^o$. This shows that $\alpha S^o \subseteq (\alpha S)^o$. Hence, we have $(\alpha S)^o = \alpha S^o$. 


(iv) \( \forall \bar{x} \in \mathbb{S} + T, \exists \bar{s} \in \mathbb{S} \) and \( \exists \bar{t} \in T \) such that \( \bar{x} = \bar{s} + \bar{t} \). By Proposition 4.13, \( \exists (s_n)_{n=1}^{\infty} \subseteq S \) and \( \exists (t_n)_{n=1}^{\infty} \subseteq T \) such that \( \lim_{n \to \infty} s_n = \bar{s} \) and \( \lim_{n \to \infty} t_n = \bar{t} \). Then, \( (s_n + t_n)_{n=1}^{\infty} \subseteq S + T \) and \( \lim_{n \to \infty} (s_n + t_n) = \bar{s} + \bar{t} = \bar{x} \), by Propositions 7.23, 3.66, and 3.67. By Proposition 4.13, \( \bar{x} \in S + T \).

Hence, \( S + T \subseteq \mathbb{S} \).

(v) \( \forall x \in S^o + T^o, \exists s_0 \in S^o \) and \( \exists t_0 \in T^o \) such that \( x = s_0 + t_0 \). Then, \( \exists r_x, r_t \in (0, \infty) \subseteq \mathbb{R} \) such that \( \mathcal{B}(s_0, r_x) \subseteq S \) and \( \mathcal{B}(t_0, r_t) \subseteq T \). Thus, we have \( \mathcal{B}(x, r_x + r_t) = \mathcal{B}(s_0, r_x) + \mathcal{B}(t_0, r_t) \subseteq S + T \) and \( x \in (S + T)^o \). Hence, \( S^o + T^o \subseteq (S + T)^o \).

This completes the proof of the proposition. \( \square \)

Theorem 7.103 (Open Mapping Theorem) Let \( X \) and \( Y \) be Banach spaces over the field \( \mathbb{K} \) and \( A \in \mathcal{B}(X, Y) \) be surjective. Then, \( A \) is an open mapping. Furthermore, if \( A \) is injective, then \( A_{inv} \in \mathcal{B}(Y, X) \).

Proof We need the following claim.

Claim 7.103.1 The image of the unit ball in \( X \) under \( A \) contains an open ball centered at the origin in \( Y \), that is \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \mathcal{B}_Y(\vartheta_y, \delta) \subseteq A(\mathbb{B}_X(\vartheta_X, 1)) \).

Proof of claim: Let \( S_n := \mathcal{B}_X(\vartheta_X, 2^{-n}) \), \( \forall n \in \mathbb{Z}^+ \). Since \( A \) is linear and surjective and \( \bigcup_{k=1}^{\infty} kS_1 = \mathbb{X} \), then, by Proposition 2.5, we have \( Y = \bigcup_{k=1}^{\infty} A(kS_1) = \bigcup_{k=1}^{\infty} kA(S_1) \). Since \( Y \) is a complete metric space, then, by Baire Category Theorem, \( Y \) is second category everywhere. Then, \( Y \) is not of first category, that is \( Y \) is not countable union of nowhere dense sets. Then, by Proposition 7.102, \( A(S_1) \subseteq Y \) is not nowhere dense. Then, \( A(S_1) \subseteq Y \) is not nowhere dense. By Proposition 3.40, \( \exists \bar{y} \in Y \) and \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \mathcal{B}_Y(\bar{y}, \delta) \subseteq A(S_1) \). Then, we have \( \mathcal{B}_Y(\vartheta_y, \delta) = \mathcal{B}_Y(\bar{y}, \delta) \subseteq A(S_1) \subseteq A(S_1) - A(S_1) \). Note that, by Proposition 7.102 and the linearity of \( A \), \( -A(S_1) = -A(S_1) = A(-S_1) = A(S_1) \). Then, again by Proposition 7.102 and the linearity of \( A \), \( \mathcal{B}_Y(\vartheta_y, 2^{-n}\delta) = 2^{-n}\mathcal{B}_Y(\vartheta_y, \delta) \subseteq 2^{-n}A(S_0) = A(2^{-n}S_0) = A(S_n) \subseteq Y \), \( \forall n \in \mathbb{Z}^+ \).

Now, we proceed to show that \( \mathcal{B}_Y(\vartheta_y, \delta/2) \subseteq A(S_0) \). Fix an arbitrary vector \( y \in \mathcal{B}_Y(\vartheta_y, \delta/2) \). Then, \( y \in A(S_1) \) and \( \exists x_1 \in S_1 \), such that \( \| y - Ax_1 \| < 2^{-2}\delta \). This implies that \( y - Ax_1 \in \mathcal{B}_Y(\vartheta_y, 2^{-2}\delta) \subseteq A(S_2) \). Recursively, \( \forall n \in \mathbb{N} \) with \( n \geq 2 \), \( y - \sum_{k=1}^{n-1} Ax_k \in A(S_n) \). Then, \( \exists x_n \in S_n \) such that \( \| y - \sum_{k=1}^{n} Ax_k \| < 2^{-n-1}\delta \). This implies that \( y - \sum_{k=1}^{n} Ax_k \in \mathcal{B}_Y(\vartheta_y, 2^{-n-1}\delta) \subseteq A(S_{n+1}) \). Since \( x_n \in S_n \) and \( \| x_n \| < 2^{-n}, \forall n \in \mathbb{N} \), then \( \sum_{k=1}^{\infty} \| x_n \| < 1 \). By Proposition 7.27, we have \( \sum_{k=1}^{\infty} x_n = x \in Y \). It is easy to show that \( x \in S_0 \). Then, \( y = \lim_{n \to \infty} \sum_{k=1}^{n} Ax_k = \lim_{n \to \infty} A(\sum_{k=1}^{n} x_k) = Ax \), where we have applied Proposition 3.66. Thus, \( y \in A(S_0) \) and \( \mathcal{B}_Y(\vartheta_y, \delta/2) \subseteq A(S_0) \).
This completes the proof of the claim. □

Fix any open set \( O \subseteq X \) and any \( y \in A(O) \). Let \( x \in O \) be such that \( Ax = y \). Then, there exists \( r \in (0, \infty) \subset \mathbb{R} \) such that \( B_X (x, r) \subseteq O \). By Claim 7.103.1, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( B_Y (\vartheta y, \delta) \subseteq A(B_X (\vartheta x, r)) \).

By the linearity of \( A \), we have \( B_Y (y, \delta) = y + B_Y (\vartheta y, \delta) \subseteq Ax + A(B_X (\vartheta x, r)) = A(B_X (x, r)) \subseteq A(O) \). By the arbitrariness of \( y, A(O) \subseteq Y \) is open.

If, in addition, \( A \) is injective, then \( A \) is bijective and \( A_{\text{inv}} \) exists. Since \( A \) is open mapping, then \( A_{\text{inv}} \) is continuous. It is obvious that \( A_{\text{inv}} : Y \to X \) is linear since \( A \) is linear. Therefore, \( A_{\text{inv}} \in B(Y, X) \).

This completes the proof of the theorem. □

**Proposition 7.104** Let \( X \) be a vector space over \( K \) and \( \| \cdot \|_1 \) and \( \| \cdot \|_2 \) be two norms on \( X \) such that \( X_1 := (X, K, \| \cdot \|_1) \) and \( X_2 := (X, K, \| \cdot \|_2) \) are Banach spaces. If \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| x \|_2 \leq M \| x \|_1, \forall x \in X \). Then, the two norms are equivalent.

**Proof** Consider the mapping \( A := \text{id}_X : X_1 \to X_2 \). Clearly \( A \) is a linear bijective function. By the assumption of the proposition, \( A \in B(X_1, X_2) \). Then, \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| x \|_1 = \| Ax \|_1 \leq M \| x \|_2, \forall x \in X \). Now, take \( K = \max \{ M, \frac{1}{M} \} + 1 \in (0, \infty) \subset \mathbb{R} \). Then, we have \( \| Ax \|_1 / K \leq \| x \|_2 \leq K \| x \|_1, \forall x \in X \). Hence, the two norms are equivalent. This completes the proof of the proposition. □

**Theorem 7.105 (Closed Graph Theorem)** Let \( X := (X, K, \| \cdot \|) \) and \( Y \) be Banach spaces over the field \( K \) and \( A : X \to Y \) be a linear operator. Assume that \( A \) satisfies that \( \forall \{ x_n \}_{n=1}^{\infty} \subseteq X \) with \( \lim_{n \in \mathbb{N}} x_n = x_0 \in X \) and \( \lim_{n \in \mathbb{N}} Ax_n = y_0 \in Y \), we have \( y_0 = Ax_0 \). Then, \( A \in B(X, Y) \).

**Proof** Define a functional \( \| \cdot \|_1 : X \to \mathbb{R} \) by \( \| x \|_1 := \| x \| + \| Ax \|, \forall x \in X \). It is easy to show that \( \| \cdot \|_1 \) defines a norm on \( X \). Let \( X_1 := (X, K, \| \cdot \|_1) \) be the normed linear space. We will show that \( X_1 \) is complete. Fix any Cauchy sequence \( \{ x_n \}_{n=1}^{\infty} \subseteq X_1 \). It is easy to see that \( \{ x_n \}_{n=1}^{\infty} \subseteq X \) is a Cauchy sequence; and that \( \{ Ax_n \}_{n=1}^{\infty} \subseteq Y \) is a Cauchy sequence. By the completeness of \( X \) and \( Y \), there exist \( x_0 \in X \) and \( y_0 \in Y \) such that \( \lim_{n \in \mathbb{N}} \| x_n - x_0 \| = 0 \) and \( \lim_{n \in \mathbb{N}} \| Ax_n - y_0 \| = 0 \). By the assumption of the theorem, we have \( y_0 = Ax_0 \). Hence, we have \( \lim_{n \in \mathbb{N}} \| x_n - x_0 \| = 0 \).

Then, \( \lim_{n \in \mathbb{N}} x_n = x_0 \) in \( X_1 \). Thus, \( X_1 \) is a Banach space. Clearly, \( \forall x \in X \), we have \( \| x \| \leq \| x \|_1 \). By Proposition 7.104, \( \| \cdot \| \) and \( \| \cdot \|_1 \) are equivalent. Then, \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| x \|_1 \leq M \| x \|, \forall x \in X \). Then, we have \( \| Ax \| \leq M \| x \|, \forall x \in X \). Hence, \( A \in B(X, Y) \). This completes the proof of the theorem. □

Recall that the graph of a function \( f : X \to Y \) is the set \( \{ (x, y) \in X \times Y : x \in X, y = f(x) \} \). Then, we have the following alternative statement of the Closed Graph Theorem.
Theorem 7.106 (Closed Graph Theorem) Let \( X \) and \( Y \) be Banach spaces over the field \( K \) and \( A : X \to Y \) be a linear operator. Then, 
\( A \in \text{B}(X, Y) \) if, and only if, the graph of \( A \) is a closed set in \( X \times Y \).

Proof “Sufficiency” Let the graph of \( A \) be \( \text{graph}(A) \). Fix any sequence \( (x_n)_{n=1}^{\infty} \subseteq X \) with \( \lim_{n \to \infty} x_n = x_0 \in X \) and \( \lim_{n \to \infty} Ax_n = y_0 \in Y \). Note that \( ((x_n, Ax_n))_{n=1}^{\infty} \subseteq \text{graph}(A) \subseteq X \times Y \) and \( \lim_{n \to \infty}(x_n, Ax_n) = (x_0, y_0) \), by Proposition 3.67. Then, by Proposition 4.13, we have \( (x_0, y_0) \in \text{graph}(A) \), where the equality follows from Proposition 3.3. Hence, we have \( y_0 = Ax_0 \). By the Closed Graph Theorem, Theorem 7.105, we have \( A \in \text{B}(X, Y) \).

“Necessity” Let \( (x_0, y_0) \in \text{graph}(A) \), by Proposition 4.13, there exists \( ((x_n, Ax_n))_{n=1}^{\infty} \subseteq \text{graph}(A) \) such that \( \lim_{n \to \infty}(x_n, Ax_n) = (x_0, y_0) \). By Proposition 3.67, we have \( \lim_{n \to \infty} x_n = x_0 \) and \( \lim_{n \to \infty} Ax_n = y_0 \). By Proposition 3.66, we have \( y_0 = \lim_{n \to \infty} Ax_n = Ax_0 \). Hence, \( (x_0, y_0) \in \text{graph}(A) \). By Proposition 3.3, \( \text{graph}(A) \) is closed in \( X \times Y \).

This completes the proof of the theorem.

Proposition 7.107 Let \( X \) be a Banach space over the field \( K \), \( Y \) be a normed linear space over the same field, and \( F \subseteq \text{B}(X, Y) \). Assume that \( \forall x \in X, \exists M_x \in [0, \infty) \subseteq \mathbb{R} \) such that \( \|Tx\| \leq M_x, \forall T \in F \). Then, \( \exists M \in [0, \infty) \subseteq \mathbb{R} \), such that \( \|T\| \leq M, \forall T \in F \).

Proof By Baire Category Theorem, \( X \) is second category everywhere. \( \forall T \in F \), let \( f : X \to \mathbb{R} \) be given by \( f(x) = \|Tx\|, \forall x \in X \). By Propositions 7.21 and 3.12, \( f \) is a continuous real-valued function. By Uniform Boundedness Principle, there exist an open set \( O \subseteq X \) with \( O \neq \emptyset \) and \( M \in [0, \infty) \subseteq \mathbb{R} \) such that \( \|Tx\| \leq M, \forall T \in F, \forall x \in O \). Since \( O \) is nonempty and open, then \( \exists B_x(0, r) \subseteq O \) for some \( x_0 \in X \) and some \( r \in (0, \infty) \subseteq \mathbb{R} \). \( \forall x \in X \) with \( \|x\| < r, x + x_0 \in B_x(0, r) \subseteq O \) and \( \forall T \in F \), we have

\[
\|Tx\| = \|T(x + x_0) - T0\| \leq \|T(x + x_0)\| + \|T0\| \leq M + M_{x_0}
\]

Hence, \( \forall T \in F \), we have, \( \forall \epsilon \in (0, r) \subseteq \mathbb{R} \),

\[
\|T\| = \sup_{x \in X, \|x\| \leq 1} \|Tx\| = \sup_{x \in X, \|x\| \leq 1} (r - \epsilon)^{-1} \|T((r - \epsilon)x)\| \leq \frac{M + M_{x_0}}{r - \epsilon}
\]

By the arbitrariness of \( \epsilon \), we have \( \|T\| \leq (M + M_{x_0})/r = : M < +\infty \). This completes the proof of the proposition.

Proposition 7.108 Let \( X \) be a Banach space over the field \( K \), \( Y \) be a normed linear space over the same field, and \( (T_n)_{n=1}^{\infty} \subseteq \text{B}(X, Y) \). Assume that \( \forall x \in X, \lim_{n \to \infty} T_n x = T(x) \in Y \). Then, \( T \in \text{B}(X, Y) \).
Proof \( \forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{K}, \) we have
\[
T(\alpha x_1 + \beta x_2) = \lim_{n \to \infty} T_n(\alpha x_1 + \beta x_2) = \lim_{n \to \infty}(\alpha T_n x_1 + \beta T_n x_2) = \alpha T x_1 + \beta T x_2
\]
by Propositions 7.23 and 3.66. Hence, \( T \) is linear. \( \forall x \in X, \) by Propositions 7.21 and 3.66, \(
\lim_{n \to \infty} \| T_n x \| = \| T x \| \leq +\infty
\)
Then, \( \exists M \in (0, \infty) \subseteq \mathbb{R} \) such that \( \| T_n x \| \leq M, \forall n \in \mathbb{N} \). By Proposition 7.107, \( \exists M \in (0, \infty) \subseteq \mathbb{R} \) such that \( \| T_n x \| \leq M, \forall n \in \mathbb{N} \). \( \forall x \in X \) with \( \| x \| \leq 1 \), by Proposition 7.64, \( \| T x \| = \lim_{n \to \infty} \| T_n x \| \leq \limsup_{n \to \infty} \| T_n \| \| x \| \leq M < +\infty \).
Hence, \( \| T \| \leq M < +\infty \). Therefore, \( T \in B(X, Y) \). This completes the proof of the proposition.

\section{The Adjoints of Linear Operators}

\begin{proposition}
Let \( X \) and \( Y \) be normed linear spaces over the field \( \mathbb{K} \) and \( A \in B(X, Y) \). The adjoint operator of \( A \) is \( A^* : Y^* \to X^* \) defined by
\[
\langle \langle A^*(y_*) , x \rangle \rangle = \langle \langle y_* , Ax \rangle \rangle ; \ \forall x \in X, \ \forall y_* \in Y^*
\]
Then, \( A^* \in B(Y^*, X^*) \) with \( \| A^* \| = \| A \| . \)
\end{proposition}

\begin{proof}
First, we will show that \( A^* \) is well defined. \( \forall y_* \in Y^*, \forall x \in X, \langle \langle y_* , Ax \rangle \rangle \in \mathbb{K} \). Hence, \( f : X \to \mathbb{K} \) defined by \( f(x) = \langle \langle y_* , Ax \rangle \rangle \), \( \forall x \in X \), is a functional on \( X \). By the linearity of \( A \) and \( y_* \), \( f \) is a linear functional. \( \forall x \in X \), we have, by Proposition 7.64
\[
| f(x) | \leq \| y_* \| \| Ax \| \leq \| y_* \| \| A \| \| x \|
\]
Hence, \( f \) is a bounded linear functional with \( \| f \| \leq \| A \| \| y_* \| \). The above shows that \( A^*(y_*) = f \in X^* \). Hence, \( A^* \) is well-defined.

It is straightforward to show that \( A^* \) is a linear operator. By the fact that \( \| A^*(y_*) \| = \| f \| \leq \| A \| \| y_* \| , \forall y_* \in Y^* \), we have \( A^* \in B(Y^*, X^*) \) and \( \| A^* \| \leq \| A \| . \)

On the other hand, \( \forall x \in X \), we have either \( Ax = \vartheta \), then \( \| Ax \| = 0 \leq \| A^* \| \| x \| ; \) or \( Ax \neq \vartheta \), then, by Proposition 7.85, \( \exists y_* \in Y^* \) with \( \| y_* \| = 1 \) such that \( \| Ax \| = \langle \langle y_* , Ax \rangle \rangle \), which implies that \( \| Ax \| = \langle \langle A^* y_* , x \rangle \rangle \leq \| A^* y_* \| \| x \| \leq \| A^* \| \| y_* \| \| x \| = \| A \| \| x \| \). Hence, we must have \( \| Ax \| \leq \| A^* \| \| x \|. \) This implies that \( \| A \| \leq \| A^* \| . \)

Then, \( \| A \| = \| A^* \| . \) This completes the proof of the proposition.
\end{proof}

\begin{proposition}
Let \( X, Y, \) and \( Z \) be normed linear spaces over the field \( \mathbb{K} \). Then, the following statements hold.
\begin{enumerate}
\item \( \text{id}_X^* = \text{id}_X \).
\item If \( A_1, A_2 \in B(X, Y) \), then \( (A_1 + A_2)^* = A_1^* + A_2^* \).
\end{enumerate}
\end{proposition}
(iii) If \( A \in B (X, Y) \) and \( \alpha \in K \), then \( (\alpha A)' = \alpha A' \).

(iv) If \( A_1 \in B (X, Y) \) and \( A_2 \in B (Y, Z) \), then \( (A_2 A_1)' = A_1' A_2' \).

(v) If \( A \in B (X, Y) \) and \( A \) has a bounded inverse, then \( (A^{-1})' = (A')^{-1} \).

**Proof**

(i) \( \forall x_* \in X^\ast, \forall x \in X \), \( \langle \langle id^\ast_X (x_*), x \rangle \rangle = \langle \langle x_*, id_X (x) \rangle \rangle = \langle \langle x_*, x \rangle \rangle = \langle \langle id_X (x_*), x \rangle \rangle \). Hence, the result follows.

(ii) \( \forall y_* \in Y^\ast, \forall x \in X \), we have \( \langle \langle (A_1 + A_2)' y_*, x \rangle \rangle = \langle \langle (A_1 + A_2) y_*, x \rangle \rangle = \langle \langle (A_1 y_*) + (A_2 y_*) + (A_1' y_*), x \rangle \rangle = \langle \langle (A_2 y_*) + (A_1' y_*), x \rangle \rangle = \langle \langle (A_1' + A_2') y_*, x \rangle \rangle \). Hence, the result follows.

(iii) \( \forall y_* \in Y^\ast, \forall x \in X \), we have \( \langle \langle (\alpha A)' y_*, x \rangle \rangle = \langle \langle \alpha A y_*, x \rangle \rangle = \langle \langle \alpha (A' y_*), x \rangle \rangle = \langle \langle (\alpha A') y_*, x \rangle \rangle \). Hence, the result follows.

(iv) \( \forall z_* \in Z^\ast, \forall x \in X \), we have \( \langle \langle (A_2 A_1)' z_*, x \rangle \rangle = \langle \langle z_*, (A_2 A_1) x \rangle \rangle = \langle \langle z_*, A_2 (A_1 x) \rangle \rangle = \langle \langle A_2 z_*, A_1 x \rangle \rangle = \langle \langle A_1' (A_2' z_*), x \rangle \rangle = \langle \langle (A_1' A_2') z_* , x \rangle \rangle \). Hence, the result follows.

(v) By (i) and (iv), we have \( (A^{-1})' A' = (AA^{-1})' = id'_y = id_Y \ast \ast \), and \( A'(A^{-1})' = (A^{-1} A)' = id'_X = id_X \ast \ast \). By Proposition 2.4, we have \( (A^{-1})' = (A')^{-1} \).

This completes the proof of the proposition. \( \square \)

**Proposition 7.111** Let \( X \) and \( Y \) be normed linear spaces over the field \( K \), \( A \in B (X, Y) \), \( \phi_X : X \to X^\ast \) and \( \phi_Y : Y \to Y^\ast \) be the natural mappings on \( X \) and \( Y \), respectively, and \( A'' : X^\ast \to Y^\ast \) be the adjoint of the adjoint of \( A \). Then, we have \( A'' \circ \phi_X = \phi_Y \circ A \).

**Proof** \( \forall x \in X, \forall y_* \in Y^\ast, \) we have

\[
\langle \langle A'' (\phi_X (x)), y_* \rangle \rangle = \langle \langle \phi_X (x), A' y_* \rangle \rangle = \langle \langle A' y_*, x \rangle \rangle = \langle \langle y_*, A x \rangle \rangle
\]

Hence, the desired result follows. This completes the proof of the proposition. \( \square \)

By Proposition 7.111 and Remark 7.88, \( \forall A, B \in B (X, Y) \), we have \( A' = B' \) if, and only if, \( A = B \).

**Proposition 7.112** Let \( X \) and \( Y \) be normed linear spaces over the field \( K \) and \( A \in B (X, Y) \). Then,

\[
(R (A))^\perp = \mathcal{N} (A') ; \quad (R (A')) = \mathcal{N} (A)
\]

**Proof** Fix a vector \( y_* \in \mathcal{N} (A') \). \( \forall y \in R (A) \), \( \exists x \in X \) such that \( y = A x \). Then, \( \langle \langle y_*, y \rangle \rangle = \langle \langle y_*, A x \rangle \rangle = \langle \langle A' y_*, x \rangle \rangle = \langle \langle \partial_X (x), x \rangle \rangle = 0 \). Hence, \( y_* \in (R (A))^\perp \). This shows that \( \mathcal{N} (A') \subseteq (R (A))^\perp \).

On the other hand, fix a vector \( y_* \in (R (A))^\perp \). \( \forall x \in X \), we have \( \langle \langle A' y_*, x \rangle \rangle = \langle \langle y_*, A x \rangle \rangle = 0 \). Hence, we have \( A' y_* = \partial_X x \). This shows that \( (R (A))^\perp \subseteq \mathcal{N} (A') \).
Hence, we have \((\mathcal{R}(A))^\perp = \mathcal{N}(A')\).

Fix a vector \(x \in \mathcal{N}(A)\). \(\forall x_* \in \mathcal{R}(A')\), \(\exists y_* \in \mathcal{Y}^*\) such that \(x_* = A'y_*\). Then, \(\langle x_*, x \rangle = \langle y_*, Ax \rangle = \langle y_*, \vartheta y \rangle = 0\). Hence, \(x \in \perp(\mathcal{R}(A'))\).

This shows that \(\mathcal{N}(A) \subseteq \perp(\mathcal{R}(A'))\).

On the other hand, fix a vector \(x \in \perp(\mathcal{R}(A'))\). \(\forall y_* \in \mathcal{Y}^*\), we have \(0 = \langle A'y_*, x \rangle = \langle y_*, Ax \rangle\). Hence, by Proposition 7.85, we have \(Ax = \vartheta y\) and \(x \in \mathcal{N}(A)\). This shows that \(\perp(\mathcal{R}(A')) \subseteq \mathcal{N}(A)\).

Hence, we have \(\perp(\mathcal{R}(A')) = \mathcal{N}(A)\). This completes the proof of the proposition. \(\square\)

The dual version of the above proposition is deeper, which requires both Open Mapping Theorem and Hahn-Banach Theorem. Towards this end, we need the following result.

**Proposition 7.113** Let \(\mathcal{X}\) and \(\mathcal{Y}\) be Banach spaces over the field \(\mathbb{K}\) and \(A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})\). Assume that \(\mathcal{R}(A) \subseteq \mathcal{Y}\) is closed. Then, there exists \(K \in [0, \infty) \subset \mathbb{R}\) such that, \(\forall y \in \mathcal{R}(A)\), there exists \(x \in \mathcal{X}\) such that \(y = Ax\) and \(\|x\| \leq K\|y\|\).

**Proof** Since \(A\) is continuous, then \(\mathcal{N}(A) \subseteq \mathcal{X}\) is closed. By Proposition 7.45, \(\mathcal{X}/\mathcal{N}(A)\) is a Banach space. Let \(\phi : \mathcal{X} \rightarrow \mathcal{X}/\mathcal{N}(A)\) be the natural homomorphism, which is a bounded linear function by Proposition 7.69. By Proposition 7.70, there exists a bounded linear function \(A_D : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{R}(A)\) such that \(A = A_D \circ \phi\), \(\|A\| = \|A_D\|\) and \(A_D\) is injective. By Proposition 4.39, \(\mathcal{R}(A)\) is complete. Then, by Proposition 7.13, \(\mathcal{R}(A)\) is a Banach space. The mapping \(A_D\) is surjective to \(\mathcal{R}(A)\). Hence, \(A_D : \mathcal{X}/\mathcal{N}(A) \rightarrow \mathcal{R}(A)\) is a bijective bounded linear operator. By Open Mapping Theorem, \(A_D^{-1} \in \mathcal{B}(\mathcal{R}(A), \mathcal{X}/\mathcal{N}(A))\). \(\forall y \in \mathcal{R}(A)\), Let \([x] := A_D^{-1}y \in \mathcal{X}/\mathcal{N}(A)\). Then, by Proposition 7.64, \(\| [x] \| \leq \| A_D^{-1} \| \| y \|\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\| [x] \| = 0\); Case 2: \(\| [x] \| > 0\). Case 1: \(\| [x] \| = 0\). Then, \([x] = [\vartheta x]\). Take \(x = \vartheta x\), we have \(y = A_D [x] = Ax = \vartheta y\) and \(\|x\| = 0 = 2 \|A_D^{-1}\| \| y \|\). Case 2: \(\| [x] \| > 0\). Note that \(\| [x] \| = \inf_{x \in [x]} \| x \|\). Then, \(\exists x \in [x]\) such that \(\| x \| \leq 2 \| [x] \|\). Then, \(y = A_D [x] = Ax\) and \(\|x\| \leq 2 \| [x] \| \leq 2 \|A_D^{-1}\| \| y \|\). Hence, the desired result holds in both cases with \(K = 2 \|A_D^{-1}\|\). This completes the proof of the proposition. \(\square\)

**Proposition 7.114** Let \(\mathcal{X}\) and \(\mathcal{Y}\) be normed linear spaces over the field \(\mathbb{K}\) and \(A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})\). Assume that \(\mathcal{R}(A)\) is closed in \(\mathcal{Y}\)\). Then,

\[
\mathcal{R}(A) = \perp(\mathcal{N}(A'))
\]

If, in addition, \(\mathcal{X}\) and \(\mathcal{Y}\) are Banach spaces, then

\[
\mathcal{R}(A') = (\mathcal{N}(A))^\perp
\]
CHAPTER 7. BANACH SPACES

Proof. Fix \( y \in \mathcal{R}(A) \). Then \( \exists x \in X \) such that \( y = Ax \). \( \forall y_* \in \mathcal{N}(A') \), we have \( \langle y_*, y \rangle = \langle A'y_*, x \rangle = \langle (\vartheta^x)_*, x \rangle = 0 \). Hence, we have \( y \in \mathcal{N}(A') \) and \( \mathcal{R}(A) \subseteq \mathcal{N}(A') \).

On the other hand, fix \( y \in \mathcal{N}(A') \). By Proposition 7.97, \( \delta := \inf_{k \in \mathcal{R}(A)} \| y - k \| = \max_{y_* \in (\mathcal{R}(A))^\perp, \| y_* \| \leq 1} \text{Re}(\langle y_*, y \rangle) \), where the maximum is achieved at some \( y_0 \in (\mathcal{R}(A))^\perp \) with \( \| y_0 \| \leq 1 \). By Proposition 7.112, we have \( (\mathcal{R}(A))^\perp = \mathcal{N}(A') \) \( \supseteq y_0 \). Then, \( \delta = \text{Re}(\langle y_0, y \rangle) = 0 \). By Proposition 4.10, \( y \in \mathcal{R}(A) \) since \( \mathcal{R}(A) \subseteq y \) is closed. This shows that \( \mathcal{N}(A') \subseteq \mathcal{R}(A) \).

Hence, we have \( \mathcal{R}(A) = \mathcal{N}(A') \).

Let \( X \) and \( Y \) be Banach spaces. Fix an \( x_* \in \mathcal{R}(A') \), then \( \exists y_* \in Y^* \) such that \( x_* = A'y_* \). \( \forall x \in \mathcal{N}(A) \), we have \( \langle y_*, x \rangle = \langle A'y_*, x \rangle = \langle (\vartheta^x)_*, x \rangle = 0 \). Then, \( x_* \in \mathcal{N}(A) \)

On the other hand, fix an \( x_* \in (\mathcal{N}(A))^\perp \). Let \( K \subseteq [0, \infty) \subseteq \mathbb{R} \) be the constant described in Proposition 7.113. \( \forall y \in \mathcal{R}(A) \), \( \forall x \in X \) with \( Ax = y \), \( \langle x, x \rangle \) assumes a constant value that is dependent on \( y \) only. So, we may define a functional \( f : \mathcal{R}(A) \rightarrow \mathbb{K} \) by \( f(y) = \langle x, x \rangle \). \( \forall y \in \mathcal{R}(A) \) and \( \forall x \in X \) with \( Ax = y \). Clearly, \( f \) is a linear functional. By Proposition 7.113, \( \exists x_0 \in X \) with \( Ax_0 = y \), such that \( \| x_0 \| \leq K \| y \| \). Then, by Proposition 7.64, \( |f(y)| \leq \| y_* \| \| x \| \leq \| y_* \| K \| y \| \). Hence, \( \| f \|_{\mathcal{R}(A)} := \sup_{y \in \mathcal{R}(A), \| y \| \leq 1} |f(y)| \leq K \| x_* \| < +\infty \). By the simple version of Hahn-Banach Theorem 7.83, \( \exists y_* \in Y^* \) such that \( y_*|_{\mathcal{R}(A)} = f \) and \( \| y_* \| = \| f \|_{\mathcal{R}(A)} \). \( \forall x \in X \), we have

\[ \langle A'y_*, x \rangle = \langle y_*, Ax \rangle = f(Ax) = \langle x, x \rangle \]

This implies that \( A'y_* = x_* \in \mathcal{R}(A') \). Hence, we have \( (\mathcal{N}(A))^\perp \subseteq \mathcal{R}(A') \).

Therefore, we have \( \mathcal{R}(A') = (\mathcal{N}(A))^\perp \). This completes the proof of the proposition.

\[ \square \]

\subsection{7.12 Weak Topology}

\textbf{Definition 7.115} Let \( X \) be a normed linear space over the field \( \mathbb{K} \). The weak topology on \( X \), denoted by \( \mathcal{O}_X \subseteq X^* \), is the weak topology generated by \( X^* \), that is the weakest topology on \( X \) such that \( x_* : X \rightarrow \mathbb{K} \), \( \forall x_* \in X^* \), are continuous.

For a normed linear space \( X \), let \( \mathcal{O}_X \) be the natural topology induced by the norm on \( X \). Then, \( \mathcal{O}_X \subseteq \mathcal{O}_X \). We usually call the topology \( \mathcal{O}_X \) the strong topology. A set \( S \subseteq X \) is weakly open, that is \( S \in \mathcal{O}_X \), then \( S \) is strongly open, that is \( S \in \mathcal{O}_X \), and if \( S \) is weakly closed, then it
is strongly closed, but not conversely. We will denote the topological space 
\((X, \mathcal{O}_{\text{weak}}(X))\) by \(X_{\text{weak}}\).

**Proposition 7.116** Let \(X\) be a normed linear space over the field \(\mathbb{K}\). Then, the following statements hold.

(i) A basis for \(\mathcal{O}_{\text{weak}}(X)\) consists of all sets of the form

\[
\{ x \in X \mid \langle\langle x, x_i \rangle\rangle \in O_i, i = 1, \ldots, n \} \tag{7.8}
\]

where \(n \in \mathbb{N}\), \(O_i \in \mathcal{O}_{\mathbb{K}}\) and \(x_{i} \in X^*\), \(i = 1, \ldots, n\), and \(\mathcal{O}_{\mathbb{K}}\) is the natural topology on \(\mathbb{K}\). A basis at \(x_0 \in X\) for \(\mathcal{O}_{\text{weak}}(X)\) consists of all sets of the form

\[
\{ x \in X \mid |\langle\langle x, i, x - x_0 \rangle\rangle| < \epsilon, i = 1, \ldots, n \} \tag{7.9}
\]

where \(\epsilon \in (0, \infty) \subset \mathbb{R}\), \(n \in \mathbb{N}\), and \(x_{i} \in X^*\), \(i = 1, \ldots, n\).

(ii) \(X_{\text{weak}}\) is completely regular (T_{\text{T}_{\text{R}}}').

(iii) For a sequence \((x_k)_{k=1}^{\infty} \subseteq X_{\text{weak}}\), \(x_0 \in X\) is the limit point of the sequence in the weak topology if, and only if, \(\lim_{k \to \infty} \langle\langle x, x_k \rangle\rangle = \langle\langle x, x_0 \rangle\rangle\), \(\forall x \in X^*\). In this case, we will write \(\lim_{k \to \infty} x_k = x_0\) weakly and say that \((x_k)_{k=1}^{\infty}\) converges weakly to \(x_0\).

**Proof**

(i) By Definition 7.115, \(\mathcal{O}_{\text{weak}}(X)\) is the topology generated by sets of the form (7.8). We will show that these sets form a basis for the topology by Proposition 3.18. Take \(B = \{ x \in X \mid \langle\langle \vartheta, x \rangle\rangle \in \mathbb{K} \}\), which is of the form (7.8) and \(B = X\). \(\forall x \in X\), we have \(x \in B\). For any \(B_1, B_2 \subseteq X\), which are of the form (7.8), clearly, \(B_1 \cap B_2\) is again of the form (7.8). Hence, Proposition 3.18 applies and the sets of the form (7.8) form a basis for \(\mathcal{O}_{\text{weak}}(X)\). Let \(B \subseteq X\) be any set of the form (7.9). Clearly, \(x_0 \in B \subseteq \mathcal{O}_{\text{weak}}(X)\), \(\forall O \in \mathcal{O}_{\text{weak}}(X)\) with \(x_0 \in O\), by Definition 3.17, there exists a basis open set \(B_1 := \{ x \in X \mid \langle\langle x, x_i \rangle\rangle \in O_i, i = 1, \ldots, n \}\), for some \(n \in \mathbb{N}\) and some \(x_i \in X^*\) and \(O_i \in \mathcal{O}_{\mathbb{K}}\), \(i = 1, \ldots, n\), such that \(x_0 \in B_1 \subseteq O\). For each \(i = 1, \ldots, n\), \(c_i := \langle\langle x, x_i \rangle\rangle \in O_i \in \mathcal{O}_{\mathbb{K}}\). Then, \(\exists c_i \in (0, \infty) \subset \mathbb{R}\) such that \(\mathcal{B}_{\mathbb{K}}(c_i, \epsilon_i) \subseteq O_i\). Take \(\epsilon = \min_{1 \leq i \leq n} \epsilon_i \in (0, \infty) \subset \mathbb{R}\) and \(B_2 := \{ x \in X \mid |\langle\langle x, x_i \rangle \rangle| < \epsilon, i = 1, \ldots, n \}\). Clearly, \(B_2\) is of the form (7.9) and \(x_0 \in B_2 \subseteq B_1 \subseteq O\). Hence, sets of the form (7.9) form a basis at \(x_0\) for \(\mathcal{O}_{\text{weak}}(X)\). Thus, (i) holds.

(ii) \(\forall x_1, x_2 \in X\) with \(x_1 \neq x_2\), we have \(x_1 - x_2 \neq 0\). By Proposition 7.85, \(\exists x \in X^*\) with \(\|x\| = 1\) such that \(\langle\langle x, x_1 - x_2 \rangle\rangle = \|x_1 - x_2\| > 0\). Let \(O_1 := \{ a \in \mathbb{K} \mid \text{Re} (a) > \text{Re} (\langle\langle x, x_2 \rangle\rangle) + \|x_1 - x_2\|/2 \}\) and \(O_2 := \{ a \in \mathbb{K} \mid \text{Re} (a) < \text{Re} (\langle\langle x, x_2 \rangle\rangle) + \|x_1 - x_2\|/2 \}\). Then, \(O_1, O_2 \subseteq \mathcal{O}_{\mathbb{K}}\) and \(O_1 \cap O_2 = \emptyset\). Let \(B_1 := \{ x \in X \mid \langle\langle x, x \rangle\rangle \in O_1 \}\) and \(B_2 := \{ x \in X \mid \langle\langle x, x \rangle\rangle \in O_2 \}\). Then, \(B_1, B_2 \subseteq \mathcal{O}_{\text{weak}}(X)\), \(x_1 \in B_1, x_2 \in B_2\), and \(B_1 \cap B_2 = \emptyset\). This shows that \(X_{\text{weak}}\) is Hausdorff.
Next, we show that $\mathcal{X}_{\text{weak}}$ is completely regular. Fix any weakly closed set $F \subseteq \mathcal{X}_{\text{weak}}$ and $x_0 \in \overline{F} \in \mathcal{O}_{\text{weak}}(\mathcal{X})$. Then, there exists a basis open set $B = \{ x \in \mathcal{X} \mid \langle\langle x_{si}, x \rangle\rangle \in O_i, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$ and some $x_{si} \in \mathcal{X}^*$ and $O_i \in \mathcal{O}_K$, $i = 1, \ldots, n$, such that $x_0 \in B \subseteq \overline{F}$. Let $O := \prod_{i=1}^n O_i \subseteq \mathbb{K}^n$ which is open. Let $p_0 := (\langle\langle x_{s1}, x_0 \rangle\rangle, \ldots, \langle\langle x_{sn}, x_0 \rangle\rangle) \in \mathbb{K}^n$. Then, $p_0 \in O$. By Propositions 4.11 and 3.61, $\mathbb{K}^n$ is normal and therefore completely regular. Then, there exists a continuous function $f : \mathbb{K}^n \to [0, 1] \subset \mathbb{R}$ such that $f|_O = 0$ and $f(p_0) = 1$. Define $g : \mathcal{X}_{\text{weak}} \to [0, 1] \subset \mathbb{R}$ by $g(x) = f(\langle\langle x_{s1}, x \rangle\rangle, \ldots, \langle\langle x_{sn}, x \rangle\rangle), \forall x \in \mathcal{X}_{\text{weak}}$. By Propositions 3.12 and 3.32, $g$ is a continuous real-valued function on $\mathcal{X}_{\text{weak}}$, $g(x_0) = f(p_0) = 1$, and $g|_B = 0$. Hence, $\mathcal{X}_{\text{weak}}$ is completely regular. Thus, (ii) holds.

(iii) “Only if” $\forall x_s \in \mathcal{X}^*$, $x_s, x : \mathcal{X} \to \mathbb{K}$ is weakly continuous. By Proposition 3.66, we have $\lim_{n \in \mathbb{N}} \langle\langle x_s, x_k \rangle\rangle = \langle\langle x_s, x_0 \rangle\rangle$.

“$\Rightarrow$” Let $(x_k)_{k=1}^\infty$ satisfy that $\lim_{n \in \mathbb{N}} \langle\langle x_s, x_k \rangle\rangle = \langle\langle x_s, x_0 \rangle\rangle, \forall x_s \in \mathcal{X}^*$. For any basis open set $B = \{ x \in \mathcal{X} \mid \langle\langle x_{si}, x \rangle\rangle \in O_i, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$ and some $x_{si} \in \mathcal{X}^*$ and $O_i \in \mathcal{O}_K$, $i = 1, \ldots, n$, with $x_0 \in B$, we have that $\forall i = 1, \ldots, n, \exists N_i \in \mathbb{N}$ such that $\langle\langle x_{si}, x_k \rangle\rangle \in O_i, \forall k \geq N_i$. Take $N = \max_{1 \leq i \leq n} N_i \in \mathbb{N}$, $\forall k \geq N$, $x_k \in B$. This shows that $(x_k)_{k=1}^\infty$ converges weakly to $x_0$. Hence, (iii) holds.

This completes the proof of the proposition.

\begin{proposition}
Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $\mathcal{X}_{\text{weak}}$ be the topological space of $\mathcal{X}$ endowed with the weak topology. Then, vector addition $\oplus : \mathcal{X}_{\text{weak}} \times \mathcal{X}_{\text{weak}} \to \mathcal{X}_{\text{weak}}$ is continuous; and scalar multiplication $\otimes : \mathbb{K} \times \mathcal{X}_{\text{weak}} \to \mathcal{X}_{\text{weak}}$ is continuous.
\end{proposition}

\begin{proof}
Fix $(x_0, y_0) \in \mathcal{X}_{\text{weak}} \times \mathcal{X}_{\text{weak}}$. We will show that $\oplus$ is continuous at $(x_0, y_0)$. Fix a basis open set $B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$ with $x_0 + y_0 \in B$. Then, by Proposition 7.116, $B = \{ z \in \mathcal{X} \mid \langle\langle x_{si}, z - x_0 - y_0 \rangle\rangle < \epsilon, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_{si} \in \mathcal{X}^*, i = 1, \ldots, n$. Let $B_1 := \{ x \in \mathcal{X} \mid \langle\langle x_{si}, x - x_0 \rangle\rangle \mid < \epsilon/2, i = 1, \ldots, n \}$ and $B_2 := \{ y \in \mathcal{X} \mid \langle\langle x_{si}, y - y_0 \rangle\rangle \mid < \epsilon/2, i = 1, \ldots, n \}$. Clearly, $B_1 \times B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, $B_1 \times B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, and $(x_0, y_0) \in B_1 \times B_2$. $\forall (x, y) \in B_1 \times B_2$, $\forall i = 1, \ldots, n$, $\langle\langle x_{si}, x + y - x_0 - y_0 \rangle\rangle \mid \leq \langle\langle x_{si}, x - x_0 \rangle\rangle + \langle\langle x_{si}, y - y_0 \rangle\rangle < \epsilon$. Hence, $x + y \in B$. This shows that $\oplus$ is continuous at $(x_0, y_0)$. By the arbitrariness of $(x_0, y_0)$ and Proposition 3.9, $\oplus$ is continuous.

Fix $(\alpha_0, x_0) \in \mathbb{K} \times \mathcal{X}_{\text{weak}}$. We will show that $\otimes$ is continuous at $(\alpha_0, x_0)$. Fix a basis open set $B \in \mathcal{O}_{\text{weak}}(\mathcal{X})$ with $\alpha_0 x_0 \in B$. Then, by Proposition 7.116, $B = \{ z \in \mathcal{X} \mid \langle\langle x_{si}, z - \alpha_0 x_0 \rangle\rangle \mid < \epsilon, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_{si} \in \mathcal{X}^*, i = 1, \ldots, n$. Let $M := \max_{1 \leq i \leq n} \| x_{si} \| \in [0, \infty) \subset \mathbb{R}$. Let $B_1 := \{ \alpha \in \mathbb{K} \mid \langle\langle x_{si}, \alpha \rangle\rangle < \epsilon, i = 1, \ldots, n \}$, and $B_2 := \{ x \in \mathcal{X} \mid \langle\langle x_{si}, x - x_0 \rangle\rangle \leq \frac{\epsilon}{\| x_{si} \| (1 + \| x_{si} \| \| x_0 \| )}, i = 1, \ldots, n \}$. Clearly, $B_1 \in \mathcal{O}_K$ and $B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, $B_1 \times B_2 \in \mathcal{O}_{\text{weak}}(\mathcal{X})$, and $(\alpha_0, x_0) \in B_1 \times B_2$. $\forall (\alpha, x) \in B_1 \times B_2$, $\forall i = 1, \ldots, n$, $\langle\langle x_{si}, \alpha x - \alpha_0 x_0 \rangle\rangle \leq \frac{\epsilon}{\| x_{si} \| (1 + \| x_{si} \| \| x_0 \| )}$.
7.12. WEAK TOPOLOGY

| \langle x_i, \alpha x - \alpha x_0 \rangle | + | \langle x_i, \alpha x_0 - \alpha x_0 \rangle | = | \alpha | | \langle x_i, x - x_0 \rangle | + | \alpha - \alpha_0 | | \langle x_i, x - x_0 \rangle | + \alpha - \alpha_0 | \| x_i \| \| x_0 \| \leq (| \alpha_0 | + 1) | \langle x_i, x - x_0 \rangle | + | \alpha - \alpha_0 | \| x_0 \| < \epsilon. \] Hence, \( \alpha x \in B \). This shows that \( \circ \) is continuous at \((\alpha_0, x_0)\). By the arbitrariness of \((\alpha_0, x_0)\) and Proposition 3.9, \( \circ \) is continuous.

This completes the proof of the proposition.

\[ \textbf{Proposition 7.118} \] Let \( \mathcal{X} \) be a finite-dimensional normed linear space over the field \( \mathbb{K} \), \( \mathcal{O}_\mathcal{X} \) be the strong topology on \( \mathcal{X} \), and \( \mathcal{O}_{\text{weak}}(\mathcal{X}) \) be the weak topology on \( \mathcal{X} \). Then, \( \mathcal{O}_\mathcal{X} = \mathcal{O}_{\text{weak}}(\mathcal{X}) \).

**Proof** Clearly, \( \mathcal{O}_{\text{weak}}(\mathcal{X}) \subseteq \mathcal{O}_\mathcal{X} \). Fix any basis open set \( O = B_\mathcal{X}(x_0, r) \subseteq \mathcal{O}_\mathcal{X} \) with \( x_0 \in \mathcal{X} \) and \( r \in (0, \infty) \subset \mathbb{R} \). We will show that \( O \in \mathcal{O}_{\text{weak}}(\mathcal{X}) \).

Let \( n \in \mathbb{Z}_+ \) be the dimension of \( \mathcal{X} \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( n = 0 \); Case 2: \( n \in \mathbb{N} \).

Case 1: \( n = 0 \). Then, \( \mathcal{X} \) is a singleton set. \( O = \mathcal{X} \in \mathcal{O}_{\text{weak}}(\mathcal{X}) \). This case is proved.

Case 2: \( n \in \mathbb{N} \). Let \( \{ e_1, \ldots, e_n \} \subseteq \mathcal{X} \) be a basis of \( \mathcal{X} \) with \( \| e_i \| = 1 \), \( i = 1, \ldots, n \). Let \( f_i : \mathcal{X} \to \mathbb{K} \) be defined by \( f_i(x) = \alpha_i \), \( \forall x = \sum_{j=1}^{n} \alpha_j e_j \in \mathcal{X} \). Clearly, \( f_i \) is well-defined and is a linear functional. By Proposition 7.67, \( f_i \) is continuous. Denote \( f_i \) by \( e_i \in \mathcal{X}^* \). \( \forall x \in O \), let \( \delta = (r - \| x_1 - x_0 \|) / n \in (0, \infty) \subset \mathbb{R} \). Let \( B := \{ x \in \mathcal{X} \mid \| \langle e_i, x - x_1 \rangle \| < \delta, i = 1, \ldots, n \} \). By Proposition 7.116, \( B \in \mathcal{O}_{\text{weak}}(\mathcal{X}) \). \( \forall x \in B \), let \( x - x_1 = \sum_{j=1}^{n} \alpha_j e_j \). Then, \( \forall i = 1, \ldots, n \), \( | \alpha_i | = | \langle e_i, x - x_1 \rangle | < \delta \).

Thus, \( \| x - x_1 \| \leq \sum_{j=1}^{n} | \alpha_j | \| e_j \| \leq n \delta = r - \| x_1 - x_0 \| \). Then, we have \( \| x - x_0 \| \leq \| x - x_1 \| + \| x_1 - x_0 \| < r \) and \( x \in O \). Hence, \( x_1 \in B \subseteq O \).

Therefore, \( O \in \mathcal{O}_{\text{weak}}(\mathcal{X}) \).

In both cases, we have shown that \( O \in \mathcal{O}_{\text{weak}}(\mathcal{X}) \). Then, \( \mathcal{O}_\mathcal{X} \subseteq \mathcal{O}_{\text{weak}}(\mathcal{X}) \). Hence, \( \mathcal{O}_\mathcal{X} = \mathcal{O}_{\text{weak}}(\mathcal{X}) \). This completes the proof of the proposition.

Given a normed linear space \( \mathcal{X} \), its dual \( \mathcal{X}^* \) is a Banach space. On \( \mathcal{X}^* \), we can also talk about the notion of weak topology. The weak topology of \( \mathcal{X}^* \) is the weakest topology on \( \mathcal{X}^* \) such that all of the functional in \( \mathcal{X}^{**} \) are continuous. This topology turns out to be less useful than the weak topology for \( \mathcal{X}^* \) generated by \( \mathcal{X} \) (or more precisely by \( \phi(\mathcal{X}) \), where \( \phi : \mathcal{X} \to \mathcal{X}^{**} \) is the natural mapping). This leads us to the following definition.

\[ \textbf{Definition 7.119} \] Let \( \mathcal{X} \) be a normed linear space over the field \( \mathbb{K} \), \( \mathcal{X}^* \) be its dual, and \( \phi : \mathcal{X} \to \mathcal{X}^{**} \) be the natural mapping. The weak* topology on \( \mathcal{X}^* \), denoted by \( \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*) \subseteq \mathcal{X}^{**} \), is the weak topology generated by \( \phi(\mathcal{X}) \), that is the weakest topology on \( \mathcal{X}^* \) such that \( \phi(x) : \mathcal{X}^* \to \mathbb{K} \), \( \forall x \in \mathcal{X} \), are continuous.

For a normed linear space \( \mathcal{X} \), let \( \mathcal{O}_\mathcal{X} \) be the natural topology induced by the norm on \( \mathcal{X}^* \). Then, \( \mathcal{O}_{\text{weak}^*}(\mathcal{X}^*) \subseteq \mathcal{O}_{\text{weak}}(\mathcal{X}^*) \subseteq \mathcal{O}_{\mathcal{X}} \). Therefore, the
weak* topology is weaker than the weak topology on $X^*$, which is further weaker than the strong topology on $X^*$, $\mathcal{O}_{X^*}$. Clearly, if $X$ is reflexive, then the weak topology and the weak* topology coincide. We will denote the topological space $(X^*, \mathcal{O}_{\text{weak}*}(X^*))$ by $X_{\text{weak}*}$. We have the following two basic results for the weak* topology, which are counterpart results for Propositions 7.116 and 7.117.

**Proposition 7.120** Let $X$ be a normed linear space over the field $K$ and $\phi : X \to X^{**}$ be the natural mapping. Then, the following statements hold.

(i) A basis for $\mathcal{O}_{\text{weak}*}(X^*)$ consists of all sets of the form

$$\{ x_* \in X^* \mid \langle \langle x_*, x_i \rangle \rangle = \langle \langle \phi(x_i), x_* \rangle \rangle \in O_i, i = 1, \ldots, n \}$$

(7.10)

where $n \in \mathbb{N}$, $O_i \in \mathcal{O}_K$ and $x_i \in X$, $i = 1, \ldots, n$, and $\mathcal{O}_K$ is the natural topology on $K$. A basis at $x_{*0} \in X^*$ for $\mathcal{O}_{\text{weak}*}(X^*)$ consists of all sets of the form

$$\{ x_* \in X^* \mid \langle \langle x_* - x_{*0}, x_i \rangle \rangle = \langle \langle \phi(x_i), x_* - x_{*0} \rangle \rangle < \epsilon, i = 1, \ldots, n \}$$

(7.11)

where $\epsilon \in (0, \infty) \subset \mathbb{R}$, $n \in \mathbb{N}$, and $x_i \in X$, $i = 1, \ldots, n$.

(ii) $X_{\text{weak}*}$ is completely regular ($T_{3\frac{1}{2}}$).

(iii) For a sequence $(x_{*k})_{k=1}^{\infty} \subseteq X_{\text{weak}*}$, $x_{*0} \in X^*$ is the limit point of the sequence in the weak* topology if, and only if, $\lim_{k \in \mathbb{N}} \langle \langle x_{*k}, x \rangle \rangle = \langle \langle x_{*0}, x \rangle \rangle$, $\forall x \in X$. In this case, we will write $\lim_{k \in \mathbb{N}} x_{*k} = x_{*0}$ weak* and say that $(x_{*k})_{k=1}^{\infty}$ converges weak* to $x_{*0}$.

**Proof** (i) By Definition 7.119, $\mathcal{O}_{\text{weak}*}(X^*)$ is the topology generated by sets of the form (7.10). We will show that these sets form a basis for the topology by Proposition 3.18. Take $B = \{ x_* \in X^* \mid \langle \langle x_* , \emptyset \rangle \rangle \in K \}$, which is of the form (7.10) and $B = X^*$. $\forall x_* \in X^*$, we have $x_* \in B$. For any $B_1, B_2 \subseteq X^*$, which are of the form (7.10), clearly, $B_1 \cap B_2$ is again of the form (7.10). Hence, Proposition 3.18 applies and the sets of the form (7.10) form a basis for $\mathcal{O}_{\text{weak}*}(X^*)$.

Let $B \subseteq X^*$ be any set of the form (7.11). Clearly, $x_{*0} \in B \in \mathcal{O}_{\text{weak}*}(X^*)$. $\forall O \in \mathcal{O}_{\text{weak}*}(X^*)$ with $x_{*0} \in O$, by Definition 3.17, there exists a basis open set $B_1 := \{ x_* \in X^* \mid \langle \langle x_* , x_i \rangle \rangle \in O_i, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$ and some $x_i \in X$ and $O_i \in \mathcal{O}_K$, $i = 1, \ldots, n$, such that $x_{*0} \in B_1 \subseteq O$. For each $i = 1, \ldots, n$, $c_i := \langle \langle x_{*0}, x_i \rangle \rangle \in O_i \in \mathcal{O}_K$. Then, $\exists \epsilon_i \in (0, \infty) \subset \mathbb{R}$ such that $B_{K}(c_i, \epsilon_i) \subseteq O_i$. Take $\epsilon = \min_{1 \leq i \leq n} \epsilon_i \in (0, \infty) \subset \mathbb{R}$ and $B_2 := \{ x_* \in X^* \mid \langle \langle x_* - x_{*0}, x_i \rangle \rangle < \epsilon, i = 1, \ldots, n \}$. Clearly, $B_2$ is of the form (7.11) and $x_{*0} \in B_2 \subseteq B_1 \subseteq O$. Hence, sets of the form (7.11) form a basis at $x_{*0}$ for $\mathcal{O}_{\text{weak}*}(X^*)$. Thus, (i) holds.

(ii) $\forall x_{*1}, x_{*2} \in X^*$ with $x_{*1} \neq x_{*2}$, we have $x_{*1} - x_{*2} \neq \emptyset$. By Lemma 7.75, $\exists x \in X$ with $\| x \| \leq 1$ such that $\langle \langle x_{*1} - x_{*2}, x \rangle \rangle \in \mathbb{R}$ and
Theorem 7.12.2 Let $X$ be a normed linear space over the field $\mathbb{K}$ and $X^*$ be the topological space of $X^*$ endowed with the weak* topology. Then, vector addition $\oplus : X^* \times X^* \to X^*$ is continuous; and scalar multiplication $\odot : \mathbb{K} \times X^* \to X^*$ is continuous.

Proof. Fix $(x_0, y_0) \in X^* \times X^*$. We will show that $\oplus$ is continuous at $(x_0, y_0)$. Fix a basis open set $B \in O_{\text{weak}^*}(X^*)$ with $x_0 + y_0 \in B$. Then, by Proposition 7.12.1, $B = \{ x \in X^* \mid \| (z_s - x_0, y_0, x_i) \|_2 < \epsilon, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subseteq \mathbb{R}$, and for some $x_i \in X$, $i = 1, \ldots, n$. Let $B_1 := \{ x \in X^* \mid \| (x_i - x_0, x_i) \|_2 < \epsilon/2, i = 1, \ldots, n \}$ and $B_2 := \{ y \in X^* \mid \| (y_i - y_0, x_i) \|_2 < \epsilon/2, i = 1, \ldots, n \}$. Clearly, $B_1, B_2 \in O_{\text{weak}^*}(X^*)$, $B_1 \times B_2 \in O_{\text{weak}^* \times \text{weak}^*}$, and $(x_0, y_0) \in B_1 \times B_2$. Fix $(x, y) \in B_1 \times B_2$, $\forall i = 1, \ldots, n$. Then, $\| (x_i - x_0, y_i - y_0, x_i) \|_2 \leq \| (x_i - x_0, x_i) \|_2 + \| (y_i - y_0, x_i) \|_2 < \epsilon$. Hence, $x + y \in B$. This shows that $\oplus$ is continuous at $(x_0, y_0)$. By the arbitrariness of $(x_0, y_0)$ and Proposition 3.9, $\oplus$ is continuous.
Fix \((α_0, x_0) \in K \times X^*\). We will show that \(\otimes\) is continuous at \((α_0, x_0)\). Fix a basic open set \(B \in O_{weak^*}(X^*)\) with \(α_0 x_0 \in B\). Then, by Proposition 7.120, \(B = \{ z_s \in X^* \mid \langle \langle z_s - α_0 x_0, x_i \rangle \rangle < ε, i = 1, \ldots, n \}\), for some \(n \in N\), for some \(ε \in (0, \infty) \subset R\), and for some \(x_i \in X\), \(i = 1, \ldots, n\). Let \(M := \max_{1 \leq i \leq n} \| x_i \| \in [0, \infty) \subset R\). Let \(B_1 := \{ \alpha \in K \mid \alpha - α_0 < \frac{ε}{\| x_i \|} \}\) and \(B_2 := \{ x_s \in X^* \mid \| x_s - x_0, x_i \| < \frac{ε}{2(1 + \| x_0 \|)} \}\). Clearly, \(B_1 \in O_K\) and \(B_2 \in O_{weak^*}(X^*)\), \(B_1 \times B_2 \in O_K \times O_{weak^*}\), and \((α_0, x_0) \in B_1 \times B_2\). By Proposition 7.72, we have \(\| x_i \| \leq \| x_s \| \leq \| x_0 \| \leq r \) for some \(r \in Z^+\). Then, \(\langle \langle x_i, x \rangle \rangle \in I_x\). Let \(I_x \subseteq K\) be endowed with the subset topology \(O_x\). Denote the topological space \((I_x, O_x)\) by \(I_x\). By Proposition 5.40 or 7.41, \(I_x\) is compact. Let \(P := \prod_{x \in X} I_x\), whose topology is denoted by \(O_P\). By Tychonoff Theorem, \(P\) is compact. Define the equivalence map \(E : S \rightarrow P\) by \(π_x(E(x_s)) = \langle \langle x_s, x \rangle \rangle\), \(∀ x_s \in X\), \(∀ x \in X\). Then, \(E(α_0, x_0) = \langle \langle α_0 x_0, x \rangle \rangle = (\alpha_0, x_0)\)

This completes the proof of the proposition. □

The importance of the weak* topology stems from the following Theorem.

**Theorem 7.122 (Alephou Theorem)** Let \(X\) be a normed linear space over the field \(K\) and \(S = \overline{B} \times (θ, r)\), for some \(r \in Z^+\). Then, \(S\) is weak* compact.

**Proof** Let the weak* topology on \(S\) be \(O_{weak^*}(S)\). Denote the topological space \((S, O_{weak^*}(S))\) by \(S\). \(∀ x \in X, ∀ x_s \in S\), by Proposition 7.72, we have \(\| x_i \| \leq \| x_s \| \leq r \). Let \(I_x := \overline{B}(0, r \| x \|)\). Then, \(\langle \langle x_i, x \rangle \rangle \in I_x\). Let \(I_x \subseteq K\) be endowed with the subset topology \(O_x\). Denote the topological space \((I_x, O_x)\) by \(I_x\). By Proposition 5.40 or 7.41, \(I_x\) is compact. Let \(P := \prod_{x \in X} I_x\), whose topology is denoted by \(O_P\). By Tychonoff Theorem, \(P\) is compact. Define the equivalence map \(E : S \rightarrow P\) by \(π_x(E(x_s)) = \langle \langle x_s, x \rangle \rangle\), \(∀ x_s \in X\), \(∀ x \in X\).

**Claim 7.122.1** \(E : S \rightarrow E(S) \subseteq P\) is a homeomorphism.

**Proof of claim:** Clearly, \(E : S \rightarrow E(S)\) is surjective. \(∀ x_{s_1}, x_{s_2} \in S\) with \(x_{s_1} \neq x_{s_2}\), then \(∃ x \in X\) such that \(π_x(E(x_{s_1})) = \langle \langle x_{s_1}, x \rangle \rangle \neq \langle \langle x_{s_2}, x \rangle \rangle = π_x(E(x_{s_2}))\). Then, \(E(x_{s_1}) \neq E(x_{s_2})\). Hence, \(E\) is injective. Therefore, \(E : S \rightarrow E(S)\) is bijective and admits inverse \(E_{inv} : E(S) \rightarrow S\).

Fix any \(x_0 \in S\). Fix any basis open set \(O \in O_P\) with \(E(x_0) \in O\). By Proposition 3.25, \(O = \prod_{x \in X} O_x\), where \(O_x \in O_x\), \(∀ x \in X\), and \(O_x = I_x\) for all \(x\)’s except finitely many \(x\)’s, say \(x \in X_N\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(X_N = \emptyset\); Case 2: \(X_N \neq \emptyset\). Case 1: \(X_N = \emptyset\). Take \(x_1 \in θ\) and \(O_1 = K\). Let \(B := \{ x_s \in S \mid \langle \langle x_s, x_1 \rangle \rangle \in O_1 \} = S\). Clearly, \(x_0 \in B \in O_{weak^*}(S)\) and \(∀ x_s \in B, ∀ x \in X, π_x(E(x_s)) = \langle \langle x_s, x \rangle \rangle \in I_x = O_x\). Then, \(E(x_0) \in O\) and \(E(B) \subseteq O\). Case 2: \(X_N \neq \emptyset\). \(∀ x \in X_N\), by Proposition 3.4, let \(O_x \in O_K\) be such that \(O_x \cap I_x = O_x\). Let \(B := \{ x_s \in S \mid \langle \langle x_s, x \rangle \rangle \in O_x, ∀ x \in X_N \}\). Clearly, \(x_0 \in B \in O_{weak^*}(S)\). \(∀ x_s \in B, ∀ x \in X_N, π_x(E(x_s)) = \langle \langle x_s, x \rangle \rangle \in I_x = O_x\).
\(\bar{O}_x \cap I_x = O_x\). Hence, \(E(x_*) \in O\) and \(E(B) \subseteq O\). In both cases, we have shown that \(\overline{E}B \in \mathcal{O}_{weak}(S)\) with \(x_0 \in B\) such that \(E(B) \subseteq O\). Hence, \(E\) is continuous at \(x_0\). By the arbitrariness of \(x_0\) and Proposition 3.9, \(E\) is continuous.

Fix any \(p_0 \in E(S)\). Let \(x_0 = E_{inv}(p_0) \in S\). Fix any basis open set \(B \in \mathcal{O}_{weak}(S)\) with \(x_0 \in B\). By Proposition 7.120, \(B = \{x_* \in S \mid \langle\langle x_* , x_i \rangle\rangle \in O_i, i = 1, \ldots, n\}\), for some \(n \in \mathbb{N}\) and some \(x_i \in X\) and \(O_i \in \mathcal{O}_K, i = 1, \ldots, n\). Without loss of generality, we may assume that \(x_{1}, \ldots, x_{n}\) are distinct (otherwise, say \(x_{n} = x_{n-1}\), then set \(O_{n-1} = O_n \cap O_{n-1}\) and \(O_{i} = O_{i} \cap O_{1}, i = 1, \ldots, n - 2,\) and consider \(x_{1}, \ldots, x_{n-1}\) with \(O_{1}, \ldots, O_{n-1}\)). \(\forall x \in X\), if \(x = x_{i}\) for some \(i \in \{1, \ldots, n\}\), then let \(O_x = O_i \cap I_x \in O_x\); otherwise, let \(O_x = I_x \in O_x\). Let \(O := \prod_{x \in X} O_x\). Clearly, \(p_0 \in O\). \(\forall p \in O \cap E(S)\), let \(x_* = E_{inv}(p) \in S\). \(\forall i = 1, \ldots, n\), \(\langle\langle x_* , x_i \rangle\rangle = \pi_{x_i}(E(x_*)) \in O_z, O_i \cap I_x\). Then, \(x_* \in B\) and \(E_{inv}(O \cap E(S)) \subseteq B\). This shows that \(E_{inv}\) is continuous at \(p_0\). By the arbitrariness of \(p_0\) and Proposition 3.9, \(E_{inv}\) is continuous.

Hence, \(E : S \rightarrow E(S)\) is a homeomorphism. This completes the proof of the claim.

Next, we will show that \(E(S)\) is closed. \(\forall p_0 \in \overline{E(S)} \subseteq P\), define \(f_0 : X \rightarrow K\) by \(f_0(x) = \pi_x(p_0)\), \(\forall x \in X\). \(\forall x_1, x_2 \in X\), \(\forall \alpha, \beta \in K\), let \(z = \alpha x_1 + \beta x_2\). \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\forall x \in X\), if \(x \in \{x_1, x_2, z\}\), let \(O_x = B_K(\pi_x(p_0), \epsilon) \cap I_x \in O_x\), otherwise, let \(O_x = I_x \in O_x\). Let \(O := \prod_{x \in X} O_x \in O_P\). Clearly, \(p_0 \in O\). By Proposition 3.3, \(\exists p_1 \in E(S) \cap O\). Let \(x_* = E_{inv}(p_1) \in S\). \(\forall x \in \{x_1, x_2, z\}\), \(\langle\langle x_* , x \rangle\rangle = \pi_x(p_1) \in O_z = B_K(\pi_x(p_0), \epsilon) \cap I_z \subseteq B_K(\pi_0(x), \epsilon)\). Hence, we have \(\langle\langle x_* , x_1 \rangle\rangle - f_0(x_1) < \epsilon, \langle\langle x_* , x_2 \rangle\rangle - f_0(x_2) < \epsilon,\) and \(\langle\langle x_* , z \rangle\rangle - f_0(z) < \epsilon\). This implies that

\[
| f_0(z) - \alpha f_0(x_1) - \beta f_0(x_2) | = | f_0(z) - \langle\langle x_* , z \rangle\rangle - \alpha f_0(x_1) \\
+ \alpha \langle\langle x_* , x_1 \rangle\rangle - \beta f_0(x_2) + \beta \langle\langle x_* , x_2 \rangle\rangle | < \epsilon + | \alpha | \epsilon + | \beta | \epsilon
\]

By the arbitrariness of \(\epsilon\), we have \(f_0(z) = \alpha f_0(x_1) + \beta f_0(x_2)\). Hence, \(f_0\) is linear. \(\forall x \in X, f_0(x) = \pi_x(p_0) \in I_x\), then \(|f_0(x)| \leq r \parallel x \parallel\). Therefore, \(f_0 \in X^\prime\) and \(\parallel f_0 \parallel \leq r\). Hence, \(f_0 \in S\). \(\forall x \in X, \pi_x(\pi_{x}(f_0)) = \pi_0(x) = \pi_{x}(p_0)\). Hence, \(E(f_0) = p_0\) and \(p_0 \in E(S)\). This shows that \(E(S) \subseteq E(S)\). By Proposition 3.3, \(E(S)\) is closed.

By Proposition 5.5 and the compactness of \(P\), we have \(E(S)\) is compact. By Proposition 5.7, \(S\) is compact. This completes the proof of the theorem.

\[\]

**Proposition 7.123** Let \(X\) and \(Y\) be normed linear spaces over \(K\), \(A \in B(X,Y)\), and \(X_{weak}\) and \(Y_{weak}\) be the topological spaces of \(X\) and \(Y\) endowed with the weak topology. Then, \(A : X_{weak} \rightarrow Y_{weak}\) is continuous.

**Proof** Fix any basis open set \(O_Y \in \mathcal{O}_{weak}(Y)\), we will show that \(A_{inv}(O_Y) \in \mathcal{O}_{weak}(X)\). By Proposition 7.116, \(O_Y = \{y \in Y \mid \langle\langle y_{i}, y \rangle\rangle \in O_i, i = 1, \ldots, n\}\), where \(n \in \mathbb{N}\), \(O_i \in \mathcal{O}_{K}, y_{i} \in Y^*, i = 1, \ldots, n\). Then, \(A_{inv}(O_Y) = \{x \in X \mid \langle\langle y_{i}, Ax \rangle\rangle \in O_i, i = 1, \ldots, n\}\).
Let $X$ be a normed linear space and $D \subseteq X$. $D$ is said to be locally convex at $x_0 \in D$ if, and only if, $D \cap B_X(x_0, \delta)$ is convex. $D$ is said to be locally convex if it is locally convex at any $x \in D$.

Then, $D \subseteq X$ is locally convex at $x_0 \in D$ if, and only if, there exists a basis $B \subseteq O$ at $x_0$ for the subset topological space $(D, O)$ such that each $B \in B$ is convex. Furthermore, $D \subseteq X$ is locally convex if, and only if, there exists a basis $B \subseteq O$ for the subset topological space $(D, O)$ such that each $B \in B$ is convex.

Proposition 7.125 Let $X$ be a Banach space over $K$. Then, $X$ is $\sigma$-compact if, and only if, $X$ is finite dimensional.

Proof “Sufficiency” Assume that $X$ is finite dimensional. Let $K_n := B_X(\vartheta_X, n), \forall n \in \mathbb{N}$. By Proposition 7.42, $K_n$ is compact, $\forall n \in \mathbb{N}$. Clearly, $\bigcup_{n=1}^{\infty} K_n = X$. Hence, $X$ is $\sigma$-compact.

“Necessity” Assume that $X$ is $\sigma$-compact. Then, $X = \bigcup_{n=1}^{\infty} K_n$, where $K_n \subseteq X$ is compact, $\forall n \in \mathbb{N}$. We will show that $\overline{B}_X(\vartheta_X, 1)$ is compact by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\exists n_0 \in \mathbb{N}, \exists x_0 \in K_{n_0}, \exists \delta \in (0, \infty) \subseteq \mathbb{R}$, such that $B_X(x_0, \delta) \subseteq K_{n_0}$; Case 2: $\forall n \in \mathbb{N}, K_n$ does not contain any nonempty open ball as its subset. Case 1: $\exists n_0 \in \mathbb{N}, \exists x_0 \in K_{n_0}, \exists \delta \in (0, \infty) \subseteq \mathbb{R}$, such that $B_X(x_0, \delta) \subseteq K_{n_0}$. Introduce the transformation $F : X \to X$ by $F(x) = \frac{\delta}{2}(x - x_0), \forall x \in X$. Clearly, $F$ is a homeomorphism. Then, $\overline{B}_X(\vartheta_X, 1) \subseteq F(K_{n_0})$. By $K_{n_0}$ being compact, we have $F(K_{n_0})$ is compact since $F$ is a homeomorphism. Then, $\overline{B}_X(\vartheta_X, 1)$ is a closed subset of $F(K_{n_0})$ and therefore compact by Proposition 5.5. This proves our intended conclusion. Case 2: $\forall n \in \mathbb{N}, K_n$ does not contain any nonempty open ball as its subset. By Proposition 5.5, $K_n$ is closed, $\forall n \in \mathbb{N}$. Then, $\forall x \in K_n, x \in K_n, \forall n \in \mathbb{N}$, since $x$ is not an interior point of $K_n$. Then, $X \supseteq \overline{K_n} \supseteq K_n \cup \overline{K_n} = X$. This implies that $X = K_n = \overline{K_n}$. Hence, $K_n$ is a nowhere dense subset in $X$. This then leads to the conclusion that $X = \bigcup_{n=1}^{\infty} K_n$ is of first category. By Baire Category Theorem 4.40, $X$ is second category everywhere. Then, $X \neq \emptyset$ is not of first category. This is a contradiction. Hence, this case is impossible.

We conclude from the previous paragraph that $K := \overline{B}_X(\vartheta_X, 1)$ is compact. Suppose that $X$ is not finite dimensional. Then, there exists $(y_n)_{n=1}^{\infty} \subseteq X$ such that $y_n$’s are distinct and the set $E := \{ y_n \mid n \in \mathbb{N} \} \subseteq X$ is a linearly independent set. Without loss of generality, we may assume that $\|y_n\| = 1, \forall n \in \mathbb{N}$. Let $x_1 := y_1 \in K, z_n \in \text{arg min}_{z \in \text{span}(\{x_1, \ldots, x_{n-1}\})} \|y_n - z\|$, and $x_n := \frac{1}{\|y_n - z_n\|}(y_n - z_n) \in K, n = 2, 3, \ldots$. Clearly, $z_n$ exists since the minimization is over finite dimensional.
subspaces of \( \mathcal{X} \). By \( E \) being a linearly independent set, we have \( \|y_n - z_n\| > 0, \forall n \in \mathbb{N} \). Hence, \( x_n \in \mathcal{X} \) is well-defined, \( \forall n \in \mathbb{N} \). Clearly, \( \|x_n\| = 1 \) and \( x_n \in K, \forall n \in \mathbb{N} \). By \( K \) being compact and Borel-Lebesgue Theorem 5.37, \( \exists x_0 \in K \) and a subsequence \((x_{n_i})_{i=1}^{\infty}\) of \((x_n)_{n=1}^{\infty}\) such that \( \lim_{i \in \mathbb{N}} x_{n_i} = x_0 \). Since \( x_{n_i} \in \text{span}(E), \forall i \in \mathbb{N} \), then \( x_0 \in \text{span}(E) \). This implies that \( x_0 = \sum_{j=1}^{\infty} a_j x_j \), where \( a_j \in \mathbb{K}, \forall j \in \mathbb{N} \). Note that \( x_0 = \lim_{i \in \mathbb{N}} x_{n_i} = \lim_{i \in \mathbb{N}} \sum_{j=1}^{n_i} a_j x_j \) and thus \( \lim_{i \in \mathbb{N}}(x_{n_i} - \sum_{j=1}^{n_i} a_j x_j) = \vartheta_X \). But, \( \forall i \in \mathbb{N} \),
\[
\|x_{n_i} - \sum_{j=1}^{n_i} a_j x_j\| = \|\sum_{j=1}^{n_i} a_j x_j - \sum_{j=1}^{n_i} (\|y_n - z_n\|) a_j x_j - z_n\|.
\]
The vector \( \sum_{j=1}^{n_i-1} (\|y_n - z_n\|) a_j x_j - z_n \in \text{span}\{(x_1, \ldots, x_{n_i-1})\} \). By the definition of \( z_{n_i} \), we have
\[
\|x_{n_i} - \sum_{j=1}^{n_i} a_j x_j\| = \frac{1}{\|y_n - z_n\|} \|y_{n_i} - \sum_{j=1}^{n_i-1} (\|y_n - z_n\|) a_j x_j - z_n\| \geq \frac{1}{\|y_n - z_n\|} \|y_n - z_n\| = 1.
\]
This contradicts with \( \lim_{i \in \mathbb{N}}(x_{n_i} - \sum_{j=1}^{n_i} a_j x_j) = \vartheta_X \). Hence, the hypothesis does not hold, and \( \mathcal{X} \) must be finite dimensional.

This completes the proof of the proposition. \( \square \)

**Proposition 7.126** Let \( \mathcal{X} := (X, \rho) \) be a separable metric space, \( \mathcal{Y} \) be a Banach space, \( D \subseteq \mathcal{X} \), and \( f : D \to \mathcal{Y} \) be continuous. Then, \( \mathcal{W} := \text{span}(f(\mathcal{X})) \subseteq \mathcal{Y} \) is a separable normed linear subspace of \( \mathcal{Y} \), and \( \overline{\mathcal{W}} \subseteq \mathcal{Y} \) is a separable Banach subspace of \( \mathcal{Y} \).

**Proof** \( \mathcal{W} \) is a subspace of \( \mathcal{Y} \) and hence a normed linear subspace of \( \mathcal{Y} \). \( \mathcal{W} \) is a closed subspace of \( \mathcal{Y} \) and hence a complete normed linear subspace of \( \mathcal{Y} \), by Proposition 4.39. Then, \( \overline{\mathcal{W}} \) is a Banach subspace of \( \mathcal{Y} \). All we need to show is that they are separable. By Proposition 4.38, \( D \) is separable as a metric subspace of \( \mathcal{X} \). Let \( D_1 \subseteq D \) be a countable dense subset of \( D \). By the continuity of \( f, f(D_1) \) is a countable dense subset of \( f(D) = f(\mathcal{X}) \). Then, \( f(\mathcal{X}) \) is a separable subset of \( \mathcal{Y} \). By Proposition 7.35, \( \mathcal{W} \) and \( \overline{\mathcal{W}} \) are separable subsets of \( \mathcal{Y} \). This completes the proof of the proposition. \( \square \)
Chapter 8

Global Theory of Optimization

In this chapter, we are going to develop a number of tools for optimization in real normed linear spaces, including the geometric form of Hahn-Banach Theorem, minimum norm duality for convex sets, Fenchel Duality Theorem, and Lagrange multiplier theory for convex programming. In this chapter, we will restrict our attention to real spaces, rather than complex ones.

8.1 Hyperplanes and Convex Sets

Definition 8.1 Let \( X \) be a real vector space. A hyperplane \( H \) is a maximal proper linear variety in \( X \), that is, a linear variety \( H \subset X \), and if \( V \supset H \) is a linear variety, then either \( V = H \) or \( V = X \).

Proposition 8.2 Let \( X \) be a real vector space. \( H \subset X \) is a hyperplane if, and only if, there exist a linear functional \( f : X \to \mathbb{R} \) and \( c \in \mathbb{R} \) with \( f \) being not identically equal to zero, such that \( H = \{ x \in X \mid f(x) = c \} \).

Proof “Necessity” Let \( H \) be a hyperplane. Then, \( H \) is a linear variety. There exists a subspace \( M \) and \( x_0 \in X \) such that \( H = x_0 + M \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_0 \notin M \); Case 2: \( x_0 \in M \). Case 1: \( x_0 \notin M \). Let \( \bar{M} := \text{span} \{ M \cup \{ x_0 \} \} \). Clearly, \( \bar{M} \supset H \) and is a linear variety. Then, \( \bar{M} = X \) by Definition 8.1. \( \forall x \in X \), \( x \) can be uniquely written as \( \alpha x_0 + m \), where \( \alpha \in \mathbb{R} \) and \( m \in M \). Define \( f : X \to \mathbb{R} \) by \( f(x) = f(\alpha x_0 + m) = \alpha \), \( \forall x \in X \). Clearly, \( f \) is a linear functional and is not identically equal to zero. It is straightforward to verify that \( H = \{ x \in X \mid f(x) = 1 \} \). Case 2: \( x_0 \in M \). Then, \( H = M \). By Definition 8.1, \( \exists x_1 \in X \setminus H \). Let \( \bar{M} := \text{span} \{ M \cup \{ x_1 \} \} \). Clearly, \( \bar{M} \supset H \) and is a linear variety. Then, \( \bar{M} = X \) by Definition 8.1. \( \forall x \in X \), \( x \) can be uniquely written as \( \alpha x_1 + m \), where \( \alpha \in \mathbb{R} \) and \( m \in M \). Define \( f : X \to \mathbb{R} \)
by \( f(x) = f(\alpha x_1 + m) = \alpha, \forall x \in X \). Clearly, \( f \) is a linear functional and is not identically equal to zero. It is straightforward to verify that \( H = \{ x \in X \mid f(x) = 0 \} \).

"Sufficiency" Let \( H = \{ x \in X \mid f(x) = c \} \), where \( f : X \to \mathbb{R} \) is a linear functional, \( c \in \mathbb{R} \), and \( f \) is not identically equal to zero. Let \( M := \mathcal{N}(f) \). Clearly, \( M \) is a proper subspace of \( X \). Since \( f \) is not identically equal to zero, then \( \exists x_0 \in X \setminus M \) such that \( f(x_0) = 1 \). \( \forall x \in X, f(x - f(x)x_0) = 0. \)

Then, \( x - f(x)x_0 \in M \) and \( x \in \text{span}(M \cup \{ x_0 \}) \). Hence, \( X = \text{span}(M \cup \{ x_0 \}) \) and \( M \) is a maximal proper subspace. Then, \( H = c x_0 + M \) is a hyperplane.

This completes the proof of the proposition. \( \square \)

Consider a hyperplane \( H \) in a real normed linear space \( X \). By Proposition 7.17, \( \overline{H} \) is a linear variety. By Definition 8.1, \( \overline{H} = X \) or \( \overline{H} = H \). Thus, a hyperplane in a real normed linear space must either be dense or closed.

**Proposition 8.3** Let \( X \) be a real normed linear space and \( H \subseteq X \). \( H \) is a closed hyperplane if, and only if, there exist \( x_0 \in X^* \) and \( c \in \mathbb{R} \) with \( x_0 \neq \varnothing \) such that \( H = \{ x \in X \mid \langle x_0, x \rangle = c \} \).

**Proof** "Necessity" Let \( H \) be a closed hyperplane. By Proposition 8.2, there exists a linear functional \( f : X \to \mathbb{R} \) and \( c \in \mathbb{R} \) with \( f \) being not identically equal to zero such that \( H = \{ x \in X \mid f(x) = c \} \). All we need to show is that \( f \in X^* \). Since \( H \) is a linear variety, then \( H \neq \emptyset \). Fix \( x_0 \in H \).

It is easy to show that \( M := H - x_0 = \mathcal{N}(f) \). By Proposition 7.16, \( M \) is closed. By Proposition 7.72, \( f \in X^* \).

"Sufficiency" By Proposition 8.2, \( H \) is a hyperplane. By the continuity of \( x_0 \) and Proposition 3.10, \( H \) is closed. Hence, \( H \) is a closed hyperplane.

This completes the proof of the proposition. \( \square \)

For a real normed linear space \( X \) and a closed hyperplane \( H \subseteq X \). Then, \( H = \{ x \in X \mid \langle x_0, x \rangle = c \} \), where \( x_0 \neq \varnothing \) and \( c \in \mathbb{R} \). We associate four sets with \( H \): (a) \( \{ x \in X \mid \langle x_0, x \rangle \leq c \} \); (b) \( \{ x \in X \mid \langle x_0, x \rangle < c \} \); (c) \( \{ x \in X \mid \langle x_0, x \rangle \geq c \} \); (d) \( \{ x \in X \mid \langle x_0, x \rangle > c \} \), which are called half-spaces. The first two are negative half-spaces. The last two are positive half-spaces. The first and third are closed. The second and fourth are open.

**Definition 8.4** Let \( X \) be a real normed linear space and \( K \subseteq X \) be convex with \( \varnothing \in K^0 \). The Minkowski functional \( p : X \to \mathbb{R} \) of \( K \) is defined by \( p(x) = \inf \{ r \in \mathbb{R} \mid r^{-1} x \in K, r > 0 \} \), \( \forall x \in X \).

**Proposition 8.5** Let \( X \) be a real normed linear space and \( K \subseteq X \) be convex with \( \varnothing \in K^0 \). Then, the Minkowski functional \( p : X \to \mathbb{R} \) of \( K \) satisfies

(i) \( 0 \leq p(x) < +\infty, \forall x \in X \);

(ii) \( p(\alpha x) = \alpha p(x), \forall x \in X, \forall \alpha \in [0, \infty) \subseteq \mathbb{R} \);
(iii) \( p(x_1 + x_2) \leq p(x_1) + p(x_2), \forall x_1, x_2 \in \mathcal{X}; \)

(iv) \( p \) is uniformly continuous;

(v) \( \overline{K} = \{ x \in \mathcal{X} \mid p(x) \leq 1 \}; \ K^\circ = \{ x \in \mathcal{X} \mid p(x) < 1 \}. \)

Furthermore, (ii) and (iii) implies that \( p \) is a sublinear functional.

**Proof**

(i) Since \( \vartheta \in K^\circ \), then \( \exists \epsilon_0 \in (0, \infty) \subset \mathbb{R} \) such that \( B ( \vartheta, \epsilon_0 ) \subseteq K, \forall x \in \mathcal{X}, \) we have either \( x = \vartheta \), then \( p(\vartheta) = 0 \); or \( x \neq \vartheta \), then \( 0 \leq p(x) \leq \| x \| / \epsilon_0 + \infty. \)

(ii) \( \forall x \in \mathcal{X}, \forall \alpha \in [0, \infty) \subset \mathbb{R} \), we will distinguish three exhaustive and mutually exclusive cases: Case 1: \( x = \vartheta \); Case 2: \( x \neq \vartheta \) and \( \alpha = 0 \); Case 3: \( x \neq \vartheta \) and \( \alpha > 0 \). Case 1: \( x = \vartheta \). We have \( p(\alpha x) = p(\vartheta) = 0 = \alpha p(x). \)

Case 2: \( x \neq \vartheta \) and \( \alpha = 0 \). Then, we have \( p(\alpha x) = p(\vartheta) = 0 = \alpha p(x). \)

Case 3: \( x \neq \vartheta \) and \( \alpha > 0 \). \( \forall \alpha \in (0, \infty) \subset \mathbb{R} \), we have \( \alpha > 0 \) and \( \alpha \epsilon \in (0, \infty) \subset \mathbb{R} \). Hence, \( \alpha p(x) \geq p(\alpha x). \)

On the other hand, \( \forall \alpha \in (0, \infty) \subset \mathbb{R} \), we have \( r(\alpha) \in (r \alpha \epsilon > K, \alpha x \epsilon > K, r \epsilon > 0 \). Hence, \( \alpha = \alpha p(x) \geq p(x). \)

(iii) \( \forall x_1, x_2 \in \mathcal{X}, \forall r_1 \in \{ r \in \mathbb{R} \mid r^{-1} x_1 \in K, r > 0 \}, \forall r_2 \in \{ r \in \mathbb{R} \mid r^{-1} x_2 \in K, r > 0 \}, \) we have \( r_1^{-1} x_1, r_2^{-1} x_2 \in K \). By the convexity of \( K \), we have

\[
(r_1 + r_2)^{-1} (x_1 + x_2) = \frac{r_1}{r_1 + r_2} r_1^{-1} x_1 + \frac{r_2}{r_1 + r_2} r_2^{-1} x_2 \in K
\]

Then, \( r_1 + r_2 \in \{ r \in \mathbb{R} \mid r^{-1} (x_1 + x_2) \in K, r > 0 \} \). Hence, we have \( r_1 + r_2 \leq p(x_1 + x_2) \) and \( p(x_1) + p(x_2) \geq p(x_1 + x_2). \)

(iv) \( \forall x \in \mathcal{X}, 0 \leq p(x) \leq \| x \| / \epsilon_0. \) \( \forall \alpha \in (0, \infty) \subset \mathbb{R} \), let \( \delta = \epsilon_0 \epsilon \in (0, \infty) \subset \mathbb{R} \). \( \forall x_1, x_2 \in \mathcal{X} \) with \( \| x_1 - x_2 \| / \| x_2 \| < \delta \), we have \( p(x_1) \leq p(x_2) + p(x_1 - x_2) \leq p(x_2) + \| x_1 - x_2 \| / \epsilon_0 \leq p(x_2) + \epsilon \) and \( p(x_1) \leq p(x_1) + p(x_2 - x_1) \leq p(x_1) + \| x_2 - x_1 \| / \epsilon_0 < p(x_1) + \epsilon. \) Hence, \( | p(x_1) - p(x_2) | < \epsilon. \) This shows that \( p \) is uniformly continuous.

(v) \( \forall x \in K^\circ, \) we have either \( x = \vartheta \), then \( p(x) = 0 < 1; \) or \( x \neq \vartheta \), then \( \exists \epsilon \in (0, \infty) \subset \mathbb{R} \) such that \( B ( x, \epsilon / \| x \| ) \subseteq K, \) which implies that \( p(x) \leq 1/(1 + \epsilon) < 1. \) Therefore, \( K^\circ \subseteq \{ x \in \mathcal{X} \mid p(x) < 1 \}. \)

\( \forall x \in \mathcal{X} \) with \( p(x) < 1, \) by the continuity of \( p, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( p(y) < 1, \forall y \in B ( x, \delta \), \( y \in K \) by the convexity of \( K, p(y) < 1, \) and the fact that \( \vartheta \in K, \) then \( B ( x, \delta ) \subseteq K \) and \( x \in K^\circ. \) Therefore, we have \( \{ x \in \mathcal{X} \mid p(x) < 1 \} \subseteq K^\circ. \) Thus, \( K^\circ = \{ x \in \mathcal{X} \mid p(x) < 1 \}. \)

\( \forall x \in K, \) we have \( p(x) \leq 1. \) Then, \( K \subseteq \{ x \in \mathcal{X} \mid p(x) \leq 1 \}. \) By the continuity of \( p \) and Proposition 3.10, \( \{ x \in \mathcal{X} \mid p(x) \leq 1 \} \) is closed. Then, \( \overline{K} \subseteq \{ x \in \mathcal{X} \mid p(x) \leq 1 \}. \)

\( \forall x \in \mathcal{X} \) with \( p(x) \leq 1, \forall \rho \in (0, 1) \subset \mathbb{R}, p(\rho x) = \rho p(x) < 1. \) Then, \( \rho x \in K^\circ. \) Then, by Proposition 4.13, \( x \in \overline{K^\circ} \subseteq \overline{K}. \) Hence, \( \{ x \in \mathcal{X} \mid p(x) \leq 1 \} \subseteq \overline{K}. \) Therefore, \( \overline{K} = \{ x \in \mathcal{X} \mid p(x) \leq 1 \}. \)

This completes the proof of the proposition. \( \square \)
Proposition 8.6 Let $X$ be a real normed linear space and $K \subseteq X$ be convex with $K^o \neq \emptyset$. Then, $K = K^o$.

Proof Clearly, $K^o \subseteq K$. Then, $K^o \subseteq K$.

To show that $K \subseteq K^o$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $\emptyset \in K^o$; Case 2: $\emptyset \not\in K^o$. Let $p : X \to [0, \infty) \subseteq \mathbb{R}$ be the Minkowski functional of $K$. By Proposition 8.5, $p$ is a continuous sublinear functional, $K^o = \{ x \in X \mid p(x) < 1 \}$, and $K = \{ x \in X \mid p(x) \leq 1 \}$. Thus, $K^o = K$.

By Proposition 8.6, $\emptyset \not\in K^o$. Let $K_1 = K - x_0$. By Propositions 7.16 and 6.39, $K_1^o = K^o - x_0 = \emptyset$ and $K_1$ is convex. By Case 1, $K_1 \subseteq K_1^o$. By Proposition 7.16, $K = x_0 + K_1 = x_0 + K_1 \subseteq x_0 + K_1^o = x_0 + K_1^o = (x_0 + K_1)^o = K^o$. This case is proved.

Hence, $K = K^o$. This completes the proof of the proposition.

8.2 Geometric Form of Hahn-Banach Theorem

The next theorem is the geometric form of Hahn-Banach Theorem.

Theorem 8.7 (Mazur’s Theorem) Let $X$ be a real normed linear space, $K \subseteq X$ be a convex set with nonempty interior, and $V \subseteq X$ is a linear variety with $V \cap K^o = \emptyset$. Then, there exists a closed hyperplane $H$ containing $V$ but no interior point of $K$ and $K$ is contained in one of the closed half-spaces associated with $H$; that is, $\exists c \in \mathbb{R}$ and $\exists x_* \in X^*$ with $x_* \not\in \emptyset$, such that $\langle x_*, v \rangle = c$, $\forall v \in V$, $\langle x_*, k \rangle < c$, $\forall k \in K^o$, and $\langle x_*, k \rangle \leq c$, $\forall k \in K$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\emptyset \in K^o$; Case 2: $\emptyset \not\in K^o$. Case 1: $\emptyset \in K^o$. Let $p : X \to [0, \infty) \subseteq \mathbb{R}$ be the Minkowski functional of $K$. By Proposition 8.5, $p$ is a continuous sublinear functional. Let $M := \text{span } (V)$. Since $V$ is a linear variety and $\emptyset \not\in V$, then $\exists x_1 \in V$ such that $V - x_1 =: M$ is a subspace in $M$ and $x_1 \not\in M$. $\exists m \in M$, $\exists x \in V$ such that $m = \alpha x_1 + \bar{m}$. Then, we may define a functional $f : M \to \mathbb{R}$ by $f(m) = f(\alpha x_1 + \bar{m}) = \alpha$, $\forall m \in M$. Clearly, $f$ is a linear functional on $M$ and $f(v) = 1$, $\forall v \in V$. $\forall v \in V$, we have $v \not\in K^o$ and $p(v) \geq 1 = f(v)$. $\forall m = \alpha x_1 + \bar{m} \in M$, we have either $\alpha > 0$, then $f(m) = f(x_1 + \alpha^{-1} \bar{m}) = \alpha \leq \alpha p(x_1 + \alpha^{-1} \bar{m}) = p(m)$, where the inequality follows from the fact that $x_1 + \alpha^{-1} \bar{m} \in V$; or $\alpha \leq 0$, then $f(m) = \alpha \leq 0 \leq p(m)$. Thus, $\forall m \in M$, we have $f(m) \leq p(m)$. By the extension form of Hahn-Banach Theorem, there exists $x_* \in X^*$ such
that \( x_*|_{M} = f \) and \( \langle x_*, x \rangle \leq p(x), \forall x \in X \). Clearly, \( x_* \neq \varnothing \), \( \forall v \in V, \langle x_*, v \rangle = f(v) = 1 := c \in \mathbb{R}. \forall k \in K, \) we have \( \langle x_*, k \rangle \leq p(k) \leq 1 \).

Then, \( K \) is a convex set.

Then, \( \forall \in H \), we have \( \langle x_*, k \rangle \leq p(k) < 1 \). Hence, the closed hyperplane \( H := \{ x \in X \mid \langle x_*, x \rangle = c \} \) is the one that we seek.

Case 2: \( \vartheta \notin K^o \). Let \( x_0 \in K^o \). Then, by Proposition 6.39, \( K_1 := K - x_0 \) is a convex set. \( V_1 := V - x_0 \) is a linear variety. By Proposition 7.16, \( K_1^o = K^o - x_0 \). Then, \( \vartheta \in K_1^o \) and \( V_1 \cap K_1^o \neq \emptyset \). By Case 1, there exist \( c_1 \in \mathbb{R} \) and \( x_* \in X^* \) with \( x_\neq \varnothing \), such that \( \langle x_*, v_1 \rangle = c_1, \forall v_1 \in V_1, \langle x_*, k_1 \rangle \leq c_1, \forall k_1 \in K_1 \). Let \( c := c_1 + \langle x_*, x_0 \rangle \in \mathbb{R} \). Then, \( \forall v \in V, \langle x_*, v \rangle \leq c, \forall k \in K^o, \langle x_*, k \rangle < c \). By Proposition 7.16, \( K_1 = K - x_0 \). Then, \( \forall k \in K^o, \langle x_*, k \rangle \leq c \). Thus, the closed hyperplane \( H := \{ x \in X \mid \langle x_*, x \rangle = c \} \) is the one that we seek.

This completes the proof of the theorem. \( \square \)

**Definition 8.8** Let \( X \) be a real normed linear space. A closed hyperplane \( H := \{ x \in X \mid \langle x_*, x \rangle = c \} \) with \( x_* \in X^* \), \( x_\neq \varnothing \), and \( c \in \mathbb{R} \) is called a supporting hyperplane of a convex set \( K \subseteq X \) if either \( \inf_{k \in K} \langle x_*, k \rangle = c \) or \( \sup_{k \in K} \langle x_*, k \rangle = c \).

Clearly, for a convex set \( K \) in a real normed linear space, if \( K \) admits interior points, then, by Mazur’s Theorem, there exists a supporting hyperplane of \( K \) passing through each boundary point of \( K \).

**Theorem 8.9 (Eidelheit Separation Theorem)** Let \( X \) be a real normed linear space and \( K_1, K_2 \subseteq X \) be nonempty convex sets with \( K_1^o \neq \emptyset \) and \( K_1^o \cap K_2 = \emptyset \). Then, there exists a closed hyperplane \( H \) that separates \( K_1 \) and \( K_2 \), that is \( \exists x_* \in X^* \) with \( x_* \neq \varnothing \) and \( \exists c \in \mathbb{R} \) such that

\[
\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle
\]

**Proof** Let \( K := K_1^o - K_2 \). By Propositions 7.15 and 6.39, \( K \) is convex. Clearly, \( K^o \neq \emptyset \) since \( K_1^o \neq \emptyset \) and \( K_2 \neq \emptyset \). Since \( K_1^o \cap K_2 = \emptyset \), then \( \vartheta \notin K \).

Let \( V = \{ \vartheta \} \). Then, \( V \) is a linear variety and \( V \cap K^o = \emptyset \). By Mazur’s Theorem, \( \exists x_* \in X^* \) with \( x_* \neq \varnothing \) such that \( \langle x_*, k \rangle \leq \langle x_*, \vartheta \rangle = 0 \), \( \forall k \in K \). \( \forall k_1 \in K_1^o, \forall k_2 \in K_2, k_1 - k_2 \in K \) and \( \langle x_*, k_1 - k_2 \rangle \leq 0 \). Then, \( \langle x_*, k_1 \rangle \leq \langle x_*, k_2 \rangle \). Hence, \( -\infty < \sup_{k_1 \in K_1^o} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle < +\infty \) for some \( c \in \mathbb{R} \). By Proposition 8.6, \( K_1^o = K_1 \).

By Proposition 4.13, \( \forall k_1 \in K_1 \), \( \langle x_*, k_1 \rangle \leq \sup_{k_1 \in K_1^o} \langle x_*, k_1 \rangle \leq c \). Hence, we have

\[
\sup_{k_1 \in K_1} \langle x_*, k_1 \rangle = \sup_{k_1 \in K_1^o} \langle x_*, k_1 \rangle \leq c \leq \inf_{k_2 \in K_2} \langle x_*, k_2 \rangle
\]

Thus, the closed hyperplane \( H := \{ x \in X \mid \langle x_*, x \rangle = c \} \) is the one we seek. This completes the proof of the theorem. \( \square \)
Proposition 8.10 Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be a closed convex set. Then, $\forall x_0 \in \mathcal{X} \setminus K$. Then, there exists an $x_* \in \mathcal{X}^*$ such that $\langle \langle x_*, x_0 \rangle \rangle < \inf_{k \in K} \langle \langle x_*, k \rangle \rangle$. Hence, $K$ is weakly closed.

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: $K = \emptyset$; Case 2: $K \neq \emptyset$. Case 1: $K = \emptyset$. Let $x_* = \emptyset$. Then, $\langle \langle x_*, x_0 \rangle \rangle = 0 < +\infty = \inf_{k \in K} \langle \langle x_*, k \rangle \rangle$. Clearly, $K$ is weakly closed. This case is proved.

Case 2: $K \neq \emptyset$. $\forall x_0 \in \mathcal{X} \setminus K$, by Proposition 4.10, $\text{dist}(x_0, K) = \inf_{k \in K} \| x_0 - k \| =: d \in (0, \infty) \subset \mathbb{R}$. Then, $K_1 := \mathcal{B}_\mathcal{X}(x_0, d)$ is a convex set with nonempty interior and $K_1 \cap K = \emptyset$. By Eidelheit Separation Theorem, there exists $x_* \in \mathcal{X}^*$ with $x_* \neq \emptyset$ and $\exists c \in \mathbb{R}$ such that $\sup_{k \in K_1} \langle \langle x_*, k \rangle \rangle \leq c \leq \inf_{k \in K} \langle \langle x_*, k \rangle \rangle$. By Lemma 7.75, $\langle \langle x_*, x_0 \rangle \rangle < \sup_{k \in K_1} \langle \langle x_*, k \rangle \rangle \leq c$. Then, $\langle \langle x_*, x_0 \rangle \rangle < \inf_{k \in K} \langle \langle x_*, k \rangle \rangle$.

Let $\bar{K}$ be the closure of $K$ in the weak topology $\mathcal{O}_{\text{weak}}(\mathcal{X})$. $x_0 \in \{ x \in \mathcal{X} \mid \langle \langle x_*, x \rangle \rangle < c \} := O$. Clearly, $O$ is a weakly open set and $O \cap K = \emptyset$. Then, by Proposition 3.3, $x_0 \not\in \bar{K}$. Thus, we have shown that $\forall x_0 \in \bar{K}$, $x_0 \not\in \bar{K}$. Hence, $\bar{K} \subseteq K$. Hence, $K$ is weakly closed by Proposition 3.3. This case is proved.

This completes the proof of the proposition. \(\square\)

Proposition 8.11 Let $\mathcal{X}$ be a reflexive real normed linear space and $K \subseteq \mathcal{X}^*$ be a bounded closed convex set. Then, $K$ is weak$^*$ compact.

**Proof** Since $K$ is bounded, then there exists $n \in \mathbb{Z}_+$ such that $K \subseteq \mathcal{B}_{\mathcal{X}^*}(\emptyset, n) := S$. By Alaoglu Theorem, $S$ is weak$^*$ compact. Since $K$ is closed and convex, then, by Proposition 8.10, $K$ is weakly closed. Since $\mathcal{X}$ is reflexive, then the weak topology and the weak$^*$ topology on $\mathcal{X}^*$ coincide. Then, $K$ is weak$^*$ closed. By Proposition 5.5, $K$ is weak$^*$ compact. This completes the proof of the proposition. \(\square\)

8.3 Duality in Minimum Norm Problems

**Definition 8.12** Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be a nonempty convex set. The support of $K$ is the set $K_{\text{supp}} := \{ x_* \in \mathcal{X}^* \mid \sup_{k \in K} \langle \langle x_*, k \rangle \rangle < +\infty \}$.

**Proposition 8.13** Let $\mathcal{X}$ be a real normed linear space and $K, G \subseteq \mathcal{X}$ be nonempty convex sets. Then, the following statements hold.

(i) $K_{\text{supp}} \subseteq \mathcal{X}^*$ is a convex cone.

(ii) If $K$ is closed, then $K = \bigcap_{x_* \in K_{\text{supp}}} \{ x \in \mathcal{X} \mid \langle \langle x_*, x \rangle \rangle \leq \sup_{k \in K} \langle \langle x_*, k \rangle \rangle \}
= \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \emptyset} \{ x \in \mathcal{X} \mid \langle \langle x_*, x \rangle \rangle \leq \sup_{k \in K} \langle \langle x_*, k \rangle \rangle \}$, that is, $K$ equals to the intersection of all closed half-spaces containing $K$. 


8.3. DUALITY IN MINIMUM NORM PROBLEMS

(iii) \((K + G)\supp = K\supp \cap G\supp\).

(iv) \((K \cap G)\supp \supseteq K\supp + G\supp\).

Proof 

(i) Clearly, \(\vartheta_* \in K\supp\). \(\forall x_1, x_2 \in K\supp\), \(\forall \alpha \in [0, \infty) \subseteq \mathbb{R}\), we have either \(\alpha = 0\), then \(\alpha x_1 = \vartheta_* \in K\supp\), or \(\alpha > 0,\) then 
\[
\sup_{k \in K} \langle \alpha x_*, k \rangle = \sup_{k \in K} \alpha \langle x_*, k \rangle = \alpha \sup_{k \in K} \langle x_*, k \rangle < +\infty,
\]
where we have applied Proposition 3.81, which further implies that \(\alpha x_* \in K\supp\).

This shows that \(K\supp\) is a cone. Note that \(\sup_{k \in K} \langle (x_1 + x_2)_*, k \rangle = \sup_{k \in K} \langle (x_1)_*, k \rangle + \sup_{k \in K} \langle (x_2)_*, k \rangle < +\infty\), where we have applied Proposition 3.81. Then, \(x_1 + x_2 \in K\supp\).

This coupled with the fact that \(K\supp\) is a cone implies that \(K\supp\) is a convex cone.

(ii) Clearly \(K \subseteq \bigcap_{x_*, x \in K\supp} \{ x \in X \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle (x_*, k) \rangle \} =: \hat{K}\). \(\forall x \in X \setminus K\), by Proposition 8.10, \(\exists x_0 \in X^*\) such that \(\langle x_0, x \rangle \in \inf_{k \in K} \langle (x_0)_*, k \rangle\).

Then, \(\sup_{k \in K} \langle (x_0)_*, k \rangle < \langle (x_0)_*, k \rangle \). Then, \(\inf_{k \in K} \langle (x_0)_*, k \rangle \). Hence, \(-x_0 \in K\supp\) and \(x_0 \in X \setminus \hat{K}\). This shows that \(X \setminus K \subseteq X \setminus \hat{K}\) and \(\hat{K} \subseteq K\). Hence, \(K = \hat{K}\). Note that, for \(\vartheta_* \in K\supp\), \(\forall x \in X \mid \langle \vartheta_*, x \rangle \leq \sup_{k \in K} \langle (\vartheta_*, k) \rangle \) \(\subseteq K\supp\). Then, \(K = \bigcap_{x_*, x \in K\supp} \{ x \in X \mid \langle \vartheta_*, x \rangle \leq \sup_{k \in K} \langle (\vartheta_*, k) \rangle \} =: K\). Hence, \(K = \hat{K}\).

(iii) \(\forall x_* \in K\supp \cap G\supp\), then \(\sup_{k \in K} \langle (x_*, k) \rangle =: c_1 < +\infty\) and \(\sup_{g \in G} \langle (x_*, g) \rangle =: c_2 < +\infty\). \(\forall x \in K + G, x = k + g\) for some \(k \in K\) and some \(g \in G\). Then, \(\langle x, x \rangle = \langle (x_*, k) \rangle + \langle (x_*, g) \rangle \leq c_1 + c_2 < +\infty\). This shows that \(x_* \in (K + G)\supp\). Hence, \(K\supp \cap G\supp \subseteq (K + G)\supp\).

On the other hand, \(\forall x_* \in (K + G)\supp\), \(\sup_{x \in K + G} \langle (x_*, x) \rangle =: c < +\infty\). \(\forall k_0 \in K\) and \(g_0 \in G\), since \(K\) and \(G\) are nonempty. Then, 
\[
\sup_{k \in K} \langle (x_*, k) \rangle = \sup_{x \in K + G} \langle (x_*, x) \rangle - \langle (x_*, g_0) \rangle \leq c - \langle (x_*, g_0) \rangle < +\infty
\]
and 
\[
\sup_{g \in G} \langle (x_*, g) \rangle = \sup_{x \in K + G} \langle (x_*, x) \rangle - \langle (x_*, k_0) \rangle \leq c - \langle (x_*, k_0) \rangle < +\infty.
\]
Hence, \(x_* \in K\supp \cap G\supp\). This shows that \((K + G)\supp \subseteq K\supp \cap G\supp\). Therefore, we have \((K + G)\supp = K\supp \cap G\supp\).

(iv) \(\forall x_* \in K\supp + G\supp\), let \(x_* = x_1 + x_2\) with \(x_1 \in K\supp\) and \(x_2 \in G\supp\). \(\forall x \in K \cap G\), we have \(\langle x, x \rangle = \langle (x_1)_*, x \rangle + \langle (x_2)_*, x \rangle \leq \sup_{k \in K} \langle (x_1)_*, k \rangle + \sup_{g \in G} \langle (x_2)_*, g \rangle < +\infty\). Hence, \(x_* \in (K \cap G)\supp\).

This completes the proof of the proposition.

\qed

Definition 8.14 Let \(X\) be a real normed linear space and \(K \subseteq X\) be a nonempty convex set. The support functional of \(K\) is \(h : K\supp \to \mathbb{R}\) given by \(h(x_*) = \sup_{k \in K} \langle (x_*, k) \rangle\), \(\forall x_* \in K\supp\).
In the above definition, $h$ takes value in $\mathbb{R}$ since $K \neq \emptyset$ and $x_* \in K_{\text{supp}}$.

**Proposition 8.15** Let $X$ be a real normed linear space, $K \subseteq X$ be a nonempty convex set, $h : K_{\text{supp}} \to \mathbb{R}$ be the support functional of $K$, and $y \in X$. Then, the following statements hold.

(i) $\delta := \inf_{k \in K} \Vert y - k \Vert = \max_{x_* \in K_{\text{supp}}, \Vert x_* \Vert \leq 1} (\langle x_*, y \rangle - h(x_*))$, where the maximum is achieved at some $x_0 \in K_{\text{supp}}$ with $\Vert x_0 \Vert \leq 1$. If the infimum is achieved at $k_0 \in K$ then $y - k_0$ is aligned with $x_0$ and $\langle (x_0, k_0) \rangle = h(x_0)$.

(ii) If $\exists k_0 \in K$ and $\exists x_0 \in K_{\text{supp}}$ with $\Vert x_0 \Vert = 1$ such that $y - k_0$ is aligned with $x_0$ and $\langle (x_0, k_0) \rangle = h(x_0)$, then the infimum is achieved at $k_0$ and the maximum is achieved at $x_0$, that is $\delta = \Vert y - k_0 \Vert = \langle (x_0, y) \rangle - h(x_0)$.

**Proof**

(i) $\forall x_* \in K_{\text{supp}}$ with $\Vert x_* \Vert \leq 1$, $\forall k \in K$, we have $\Vert y - k \Vert \geq \Vert x_* \Vert \Vert y - k \Vert \geq \langle x_*, y - k \rangle = \langle x_*, y \rangle - \langle x_*, k \rangle = \langle (x_*, y) \rangle - h(x_*)$. Hence, $\delta = \inf_{k \in K} \Vert y - k \Vert \geq \sup_{x_* \in K_{\text{supp}}, \Vert x_* \Vert \leq 1} (\langle (x_*, y) \rangle - h(x_*))$. If the infimum is achieved at $k_0 \in K$, then $y - k_0$ is aligned with $x_0$ and $\langle (x_0, k_0) \rangle = h(x_0)$. This case is proved.

Case 2: $\delta > 0$. Since $K \neq \emptyset$, then $\delta < +\infty$. Take $K_1 = \mathcal{B}_X(y, \delta)$. Then, $K_1$ is a convex set with nonempty interior. $K \cap K_1 = \emptyset$. By Eidelheit Separation Theorem, there exists $x_* \in \mathcal{X}$ with $x_* \notin \partial \mathcal{B}$ and $\exists c \in \mathbb{R}$ such that $\sup_{k \in K_1} \langle (x_*, k) \rangle = c \leq \inf_{k \in K} \langle (x_*, k) \rangle$. Take $x_0 = -\Vert x_* \Vert^{-1} x_*$. Then, by Proposition 3.81, we have $\inf_{k \in K_1} \langle (x_0, k) \rangle \geq -c/\Vert x_* \Vert \geq \sup_{k \in K} \langle (x_0, k) \rangle$. This shows, $x_0 \in K_{\text{supp}}$ and $\Vert x_0 \Vert = 1$. The above inequality is equivalent to $\langle (x_0, y) \rangle + \inf_{x \in \mathcal{B}_X(y, \delta)} \langle (x_0, x) \rangle \geq h(x_0)$. By Lemma 7.75, we have $\langle (x_0, y) \rangle - h(x_0) \geq \delta$. Hence, we have $\delta = \langle (x_0, y) \rangle - h(x_0) = \max_{x_* \in K_{\text{supp}}, \Vert x_* \Vert \leq 1} (\langle (x_*, y) \rangle - h(x_*))$. If the infimum is achieved at some $k_0 \in K$, then $\delta = \Vert y - k_0 \Vert = \langle (x_0, y) \rangle - h(x_0) \leq \langle (x_0, y) \rangle - \langle (x_0, k_0) \rangle = \langle (x_0, y - k_0) \rangle \leq \Vert x_0 \Vert \Vert y - k_0 \Vert \leq \Vert y - k_0 \Vert$, where the second inequality follows from Proposition 7.72. Hence, $\langle (x_0, y - k_0) \rangle = \Vert x_0 \Vert \Vert y - k_0 \Vert$, $x_0$ is aligned with $y - k_0$, and $\langle (x_0, k_0) \rangle = h(x_0)$. This case is proved.

(ii) Note that $\delta \leq \Vert y - k_0 \Vert = \Vert x_0 \Vert \Vert y - k_0 \Vert = \langle (x_0, y - k_0) \rangle = \langle (x_0, y) \rangle - h(x_0) \leq \delta$. Then, the result follows.

This completes the proof of the proposition.

Now, we will state a proposition that guarantees the existence of a minimizing solution to a minimum norm problem. This proposition is based on the following result.
Proposition 8.16 Let $\mathcal{X}$ be a real normed linear space. Then, $\| \cdot \| : \mathcal{X} \to (0, \infty) \subset \mathbb{R}$ is weakly lower semicontinuous.

Proof By Definition 3.14, all we need to show is that, $\forall a \in \mathbb{R}$, the set $S_a := \{ x \in \mathcal{X} \mid \| x \| > a \}$ is weakly open. Note that $S_a$ is strongly closed and convex. Then, by Proposition 8.10, $\bar{S}_a$ is weakly closed. Hence, the norm is a weakly lower semicontinuous functional on $\mathcal{X}$. This completes the proof of the proposition. □

Proposition 8.17 Let $\mathcal{X}$ be a reflexive real normed linear space, $x_0 \in \mathcal{X}^*$, and $K \subseteq \mathcal{X}^*$ be a nonempty closed convex set. Then, $\delta = \min_{k \in K} \| x_0 - k \|$ and the minimum is achieved at some $k_0 \in K$.

Proof Fix a $k_1 \in K \neq \emptyset$. Let $\mu := \| x_0 - k_1 \| \in (0, \infty) \subset \mathbb{R}$ and $d = \mu + \| x_0 \| + 1 \in (0, \infty) \subset \mathbb{R}$. Then, $\forall k \in K$ with $\| k \| > d$, we have $\| x_0 - k \| \geq \| k \| - \| x_0 \| > \mu + 1$. Thus, $\delta = \inf_{k \in K} \| x_0 - k \|$ and any $k_0$ achieving the infimum for the original problem must be in the set $K \cap \overline{B}_{\mathcal{X}^*}(\vartheta_* , d ) =: K_1$. By Proposition 6.40, $K_1$ is bounded closed and convex. Note that $k_1 \in K_1 \neq \emptyset$. Then, by Proposition 8.11, $K_1$ is weak* compact. By Propositions 7.121, 8.16, and 3.16, we have $f : \mathcal{X}^* \to \mathbb{R}$ given by $f(k_* ) = \| x_0 - k_* \|$, $\forall k_* \in \mathcal{X}^*$, is lower semicontinuous. By Proposition 5.30 and Definition 3.14, there exists $k_{*0} \in K_1$ that achieves the infimum on $K_1$. Such $k_{*0}$ achieves the infimum on $K$. This completes the proof of the proposition. □

### 8.4 Convex and Concave Functionals

**Definition 8.18** Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$ be a convex set, $f : C \to \mathbb{R}$ and $g : C \to \mathbb{R}$. $f$ is said to be convex, if $\forall x_1 , x_2 \in C$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$$

$f$ is said to be strictly convex if $\forall x_1 , x_2 \in C$ with $x_1 \neq x_2$ and $\forall \alpha \in (0, 1) \subset \mathbb{R}$, we have that the strict inequality holds in the above. $g$ is said to be (strictly) concave if $-g$ is (strictly) convex.

**Definition 8.19** Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$, and $f : C \to \mathbb{R}$. The epigraph of $f$ over $C$ is the set

$$[f,C] := \{ (r,x) \in \mathbb{R} \times \mathcal{X} \mid x \in C, f(x) \leq r \}$$

**Proposition 8.20** Let $\mathcal{X}$ be a real vector space, $C \subseteq \mathcal{X}$ be convex, and $f : C \to \mathbb{R}$. Then, $f$ is convex if, and only if, the epigraph $[f,C]$ is convex.
Proof  "Sufficiency" $\forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $(f(x_1), x_1)$, $(f(x_2), x_2) \in \{f, C\}$. By the convexity of the epigraph, then $\alpha (f(x_1), x_1) + (1 - \alpha) (f(x_2), x_2) = (\alpha f(x_1) + (1 - \alpha) f(x_2), \alpha x_1 + (1 - \alpha) x_2) \in \{f, C\}$. Then, $\alpha f(x_1) + (1 - \alpha) f(x_2) \geq f(\alpha x_1 + (1 - \alpha) x_2)$. Hence, $f$ is convex.

"Necessity" Let $f$ be convex. $\forall (r_1, x_1), (r_2, x_2) \in \{f, C\}, \forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $r_1 \geq f(x_1)$ and $r_2 \geq f(x_2)$. By the convexity of $f$ and $C$, we have $\alpha f(x_1) + (1 - \alpha) f(x_2) \leq f(\alpha x_1 + (1 - \alpha) x_2)$. Then, we have $\alpha (r_1, x_1) + (1 - \alpha) (r_2, x_2) = (\alpha r_1 + (1 - \alpha)r_2, \alpha x_1 + (1 - \alpha) x_2) \in \{f, C\}$. Hence, the epigraph is convex.

This completes the proof of the proposition. □

Proposition 8.21 Let $X$ be a real normed linear space, $C \subseteq X$ be nonempty, and $f : C \to \mathbb{R}$. Then, $V(\{f, C\}) = \mathbb{R} \times V(C)$.

Proof  We will first show that $v(\{f, C\}) = \mathbb{R} \times v(C)$. Fix $x_0 \in C$. Then, $(f(x_0), x_0) \in \{f, C\}$. $\forall (r, x) \in v(\{f, C\}) = (f(x_0), x_0) + \text{span}(\{f, C\} - (f(x_0), x_0))$, where the equality follows from Proposition 6.37, then $\exists n \in \mathbb{Z}_+, \exists (1, x_1), \ldots, (r_n, x_n) \in \{f, C\}$, and $\exists \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that $(r - f(x_0), x - x_0) = \sum_{i=1}^{n} \alpha_i (r_i - f(x_0), x_i - x_0)$. Then, $x - x_0 = \sum_{i=1}^{n} \alpha_i (x_i - x_0)$. $\forall \alpha \in (0, \infty) \subset \mathbb{R}$ such that $r_0 > f(x_0)$. Then, $(r_0, x_0) \in \{f, C\}$. Let $\alpha_0 = (r - f(x_0) - \sum_{i=1}^{n} \alpha_i (f(x_i) - f(x_0))/r_0 - f(x_0))$. Now it is easy to check that $(r - f(x_0), x - x_0) = \sum_{i=1}^{n} \alpha_i (f(x_i) - f(x_0), x_i - x_0) + r_0 (r_0 - f(x_0), x_0 - x_0)$. Note that $(f(x_i), x_i) \in \{f, C\}, \forall i = 1, \ldots, n$. Then, we have $(r, x) \in (f(x_0), x_0) + \text{span}(\{f, C\} - (f(x_0), x_0)) = v(\{f, C\})$. Hence, $\mathbb{R} \times v(C) \subseteq v(\{f, C\})$. Therefore, we have $v(\{f, C\}) = \mathbb{R} \times v(C)$.

By Proposition 7.18, $V(\{f, C\}) = v(\{f, C\}) = \mathbb{R} \times v(C)$. By Proposition 4.13 and 3.67, we have $\mathbb{R} \times v(C) = \mathbb{R} \times V(C)$. Therefore, we have $V(\{f, C\}) = \mathbb{R} \times V(C)$. This completes the proof of the proposition. □

Proposition 8.22 Let $X$ be a real normed linear space, $C \subseteq X$ be convex, $f : C \to \mathbb{R}$ be convex, and $\partial C \neq \emptyset$. Then, $\{f, C\}$ has a relative interior point $(r_0, x_0)$ if, and only if, $f$ is continuous at $x_0$, $x_0 \in \partial C$, and $r_0 \in (f(x_0), +\infty) \subset \mathbb{R}$.

Proof  "Sufficiency" Fix $x_0 \in \partial C$ and $r_0 \in (f(x_0), +\infty) \subset \mathbb{R}$. Let $f$ be continuous at $x_0$. Then, for $\epsilon = (r_0 - f(x_0))/2 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, r_0 - f(x_0) - \epsilon) \subset \mathbb{R}$ such that $\forall x \in B_X(x_0, \delta) \cap V(C)$, we have $x \in C$ and $|f(x) - f(x_0)| < \epsilon$. Clearly, $(r_0, x_0) \in \{f, C\}$. $\forall (r, x) \in B_{\mathbb{R} \times X}(\{r_0, x_0\}, \delta) \cap V(\{f, C\}) = B_{\mathbb{R} \times X}(\{r_0, x_0\}, \delta) \cap (\mathbb{R} \times V(C))$, we have $x \in B_X(x_0, \delta) \cap V(C)$ and $r \in B(r_0, \delta)$. Then, $x \in C$ and $r > r_0 - \delta \geq f(x_0) + \epsilon > f(x)$. Hence, $(r, x) \in \{f, C\}$. This shows that $(r_0, x_0) \in \partial \{f, C\}$. 

“Necessity” Let \((r_0, x_0) \in \overset{\circ}{[f, C]}\). Then, \(\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}\) such that \((B_{\mathbb{R}}(r_0, \epsilon_0) \times B_X(x_0, \epsilon_0)) \cap V([f, C]) \subseteq [f, C]\). By Proposition 8.21, we have \(B_{\mathbb{R}}(r_0, \epsilon_0) \times (B_X(x_0, \epsilon_0) \cap V(C)) \subseteq [f, C]\). Therefore, \(B_X(x_0, \epsilon_0) \cap V(C) \subseteq C\), \(x_0 \in \overset{\circ}{C}\), \(f(x) \leq r_0 - \epsilon_0\), \(\forall x \in B_X(x_0, \epsilon_0) \cap V(C)\), and \(f(x_0) \leq r_0 - \epsilon_0 < r_0\).

\(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), let \(\delta = \min \{\epsilon/(r_0 - f(x_0)), 1\} \in (0, 1] \subset \mathbb{R}\). \(\forall x \in B_X(x_0, \delta \epsilon_0) \cap C\), we have \(x, x_0, x_0 + \delta^{-1}(x - x_0), x_0 - \delta^{-1}(x - x_0) \in B_X(x_0, \epsilon_0) \cap V(C)\). By the convexity of \(f\), we have

\[
\begin{align*}
f(x) &= f((1 - \delta)x_0 + \delta(x_0 + \delta^{-1}(x - x_0))) \\
&\leq (1 - \delta)f(x_0) + \delta f(x_0 + \delta^{-1}(x - x_0)) \\
&\leq f(x_0) + (r_0 - \epsilon_0 - f(x_0)) \delta < f(x_0) + \epsilon \\
f(x_0) &= f(\frac{1 - \delta}{1 + \delta}x + \frac{\delta}{1 + \delta}(x_0 - \delta^{-1}(x - x_0))) \\
&\leq \frac{1}{1 + \delta}f(x) + \frac{\delta}{1 + \delta}f(x_0 - \delta^{-1}(x - x_0)) \\
&\leq \frac{1}{1 + \delta}f(x) + \frac{\delta}{1 + \delta}(r_0 - \epsilon_0) \Rightarrow (1 + \delta)f(x_0) \leq f(x) + (r_0 - \epsilon_0) \delta \\
&\Rightarrow f(x) \geq f(x_0) - (r_0 - f(x_0) - \epsilon_0) \delta > f(x_0) - \epsilon
\end{align*}
\]

Hence, \(|f(x) - f(x_0)| < \epsilon\). This shows that \(f\) is continuous at \(x_0\).

This completes the proof of the proposition.

\(\square\)

**Proposition 8.23** Let \(X\) be a real normed linear space, \(C \subseteq X\) be convex, \(x_0 \in \overset{\circ}{C} \neq \emptyset\), and \(f : C \to \mathbb{R}\) be convex. If \(f\) is continuous at \(x_0\), then \(f\) is continuous at \(x\), \(\forall x \in \overset{\circ}{C}\).

**Proof** Fix any \(x \in \overset{\circ}{C}\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(x = x_0\); Case 2: \(x \neq x_0\). Case 1: \(x = x_0\). Then \(f\) is continuous at \(x\). Case 2: \(x \neq x_0\). \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), by the continuity of \(f\) at \(x_0\) and the fact that \(x_0 \in \overset{\circ}{C}\), \(\exists \delta \in (0, \infty) \subset \mathbb{R}\) such that \(|f(y) - f(x_0)| < \epsilon\), \(\forall y \in B_X(x_0, \delta) \cap V(C) \subseteq C\). Since \(x \in \overset{\circ}{C}\), then \(\exists \delta_0 \in (0, \delta] \subset \mathbb{R}\) such that \(B_X(x_0, \delta_0) \cap V(C) \subseteq C\). Take \(\beta = \frac{\|x - x_0\| + \delta_0}{\|x - x_0\|} \in (1, \infty) \subset \mathbb{R}\). Then, \(\beta(x - x_0) + x_0 \in B_X(x, \delta_1) \cap V(C) \subseteq C\). \(\forall y \in B_X(x, (\beta - 1)\delta_1/\beta) \cap C\), \(y = (1 - 1/\beta) (\beta/(\beta - 1)(y - x) + x_0) + (1/\beta) (\beta(x - x_0) + x_0)\). Note that \((\beta/\beta - 1) (y - x) + x_0 \in B_X(x_0, \delta) \cap V(C) \subseteq C\). By the convexity of \(f\), we have

\[
\begin{align*}
f(y) &\leq (1 - 1/\beta)f(\beta/(\beta - 1)(y - x) + x_0) + (1/\beta)f(\beta(x - x_0) + x_0) \\
&< (1 - 1/\beta)(f(x_0) + \epsilon) + (1/\beta)f(\beta(x - x_0) + x_0) =: r_1
\end{align*}
\]

Define \(r := r_1 + (\beta - 1)\delta_1/\beta\). Then, by Proposition 8.21, \(B_{\mathbb{R} \times X}((r, x), (\beta - 1)\delta_1/\beta) \cap V([f, C]) = B_{\mathbb{R} \times X}((r, x), (\beta - 1)\delta_1/\beta) \cap (\mathbb{R} \times V(C)) \subseteq [f, C]\). Hence, \((r, x) \in \overset{\circ}{[f, C]}\). By Proposition 8.22, \(f\) is continuous at \(x\). This completes the proof of the proposition.
Proposition 8.24  Let $X$ be a finite-dimensional real normed linear space, $C \subseteq X$ be a convex set, and $f : C \to \mathbb{R}$ be convex. Then, $\forall x_0 \in C$, $f$ is continuous at $x_0$.

Proof  We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \delta C = 0 \); Case 2: \( \delta C \neq 0 \). This is trivial.

Case 2: \( \delta C \neq 0 \). Fix any $x_0 \in \delta C$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $K := \overline{B}_X (x_0, \delta) \cap V(C) \subseteq C$. Let $M := V(C) - x_0$, which is a closed subspace. Let $n \in \mathbb{Z}_+$ be the dimension of $M$, which is well-defined by Theorem 6.51. We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: $n = 0$; Case 2b: $n \in \mathbb{N}$. Case 2a: $n = 0$. Then, $C = \{ x_0 \}$. Clearly, $f$ is continuous at $x_0$.

Case 2b: $n \in \mathbb{N}$. Let $\{ e_1, \ldots, e_n \} \subseteq M$ be a basis in $M$ such that $\| e_i \| = 1$, $\forall i = 1, \ldots, n$. Then, $K := \text{co} \{ \{ x_0 \pm \delta e_i \mid i = 1, \ldots, n \} \} \subseteq \text{co} (K) \subseteq C$ since $C$ is convex. $\forall x \in K$, by Proposition 6.43, $x = \sum_{i=1}^n \alpha_i (x_0 + \delta e_i) + \sum_{i=1}^n \beta_i (x_0 - \delta e_i)$ for some $\alpha_i, \beta_i \in [0,1] \subset \mathbb{R}$, $i = 1, \ldots, n$, with $\sum_{i=1}^n (\alpha_i + \beta_i) = 1$. By the convexity of $f$, we have

$$f(x) \leq \sum_{i=1}^n \alpha_i f(x_0 + \delta e_i) + \sum_{i=1}^n \beta_i f(x_0 - \delta e_i)$$

$$\leq \sum_{i=1}^n |f(x_0 + \delta e_i)| + \sum_{i=1}^n |f(x_0 - \delta e_i)| =: r_1$$

It is easy to see that $p(m) := \sum_{i=1}^n |\alpha_i|$, $\forall m = \sum_{i=1}^n \alpha_i e_i \in M$ defines a norm on $M$. By Theorem 7.38, $\exists \xi \in [1, \infty) \subset \mathbb{R}$ such that $p(m)/\xi \leq \| m \| \leq \xi p(m)$, $\forall m \in M$. $\forall x \in \mathcal{B}_X (x_0, \delta/\xi) \cap V(C)$, we have $x = x_0 + \sum_{i=1}^n \alpha_i e_i$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Then, $\delta/\xi > \| x - x_0 \| \geq p(\sum_{i=1}^n \alpha_i e_i)/\xi = \delta/\xi \sum_{i=1}^n |\alpha_i|$. Then, we have $\sum_{i=1}^n |\alpha_i| < 1$. Then, $x$ can be expressed as a convex combination of vectors in $\{ x_0 \pm \delta e_i \mid i = 1, \ldots, n \}$ and $x \in K$. Hence, $\mathcal{B}_X (x_0, \delta/\xi) \cap V(C) \subseteq K \subseteq C$. Take $r = \delta/\xi + r_1$. It is easy to show that $\mathcal{B}_{\mathbb{R} \times X} ((r, x_0), \delta/\xi) \cap V(\{ f, C \}) \subseteq [f, C]$, by Proposition 8.21. Then, $(r, x_0) \in \delta [f, C]$. By Proposition 8.22, $f$ is continuous at $x_0$.

This completes the proof of the proposition. \qed

Proposition 8.25  Let $X$ be a real normed linear space, $C \subseteq X$, and $f : C \to \mathbb{R}$. Then, the following statements hold.

(i) If $[f, C]$ is closed, then, $f$ is lower semicontinuous.

(ii) If $C$ is closed and $f$ is lower semicontinuous, then $[f, C]$ is closed.

Proof  (i) $\forall a \in \mathbb{R}$, $V_a := \{ (a, x) \in \mathbb{R} \times X \mid x \in X \}$ is closed. Then, $[f, C] \cap V_a = \{ (a, x) \in \mathbb{R} \times X \mid x \in C, f(x) \leq a \}$ is closed. Hence, by Proposition 4.13, $T_a := \{ x \in C \mid f(x) \leq a \}$ is closed. Note that

$$\{ x \in C \mid -f(x) < a \} = C \setminus \{ x \in C \mid f(x) \leq -a \} = C \setminus T_{-a}$$
Clearly, $C \setminus T_{-a}$ is open in the subset topology of $C$. Then, $-f$ is upper semicontinuous and $f$ is lower semicontinuous.

(ii) $\forall (r_0, x_0) \in [f, C]$, by Proposition 4.13, $\exists ((r_n, x_n))_{n=1}^{\infty} \subseteq [f, C]$ such that $\lim_{n \in \mathbb{N}} (r_n, x_n) = (r_0, x_0)$. By Proposition 3.67, we have $\lim_{n \in \mathbb{N}} r_n = r_0$ and $\lim_{n \in \mathbb{N}} x_n = x_0$. By Definition 8.19, we have $(x_n)_{n=1}^{\infty} \subseteq C$ and $r_n \geq f(x_n)$, $\forall n \in \mathbb{N}$. By Proposition 4.13, $x_0 \in \overline{C} = C$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: there exists infinitely many $n \in \mathbb{N}$ such that $x_n = x_0$; Case 2: there exists only finitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Case 1: there exists infinitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Then, without loss of generality, assume $x_n = x_0$, $\forall n \in \mathbb{N}$. Then, $r_n \geq f(x_0)$, $\forall n \in \mathbb{N}$. Hence, $f(x_0) \leq \lim_{n \in \mathbb{N}} r_n = r_0$. Then, $(r_0, x_0) \in [f, C]$.

Case 2: there exists only finitely many $n \in \mathbb{N}$ such that $x_n = x_0$. Then, without loss of generality, assume $x_n \neq x_0$, $\forall n \in \mathbb{N}$. Then, $x_0$ is an accumulation point of $C$. By Propositions 3.86 and 3.85 and Definition 3.14, $-f(x_0) \geq \limsup_{x \to x_0} (-f(x)) = -\liminf_{x \to x_0} f(x)$. By Proposition 3.87, we have $f(x_0) \leq \liminf_{x \to x_0} f(x) \leq \liminf_{n \in \mathbb{N}} f(x_n) \leq \liminf_{n \in \mathbb{N}} r_n = r_0$. Then, $(r_0, x_0) \in [f, C]$.

In both cases, we have $(r_0, x_0) \in [f, C]$. Then, $[f, C] \subseteq [f, C]$. By Proposition 3.3, $[f, C]$ is closed.

This completes the proof of the proposition. \hfill $\square$

**Proposition 8.26** Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be convex, and $f : C \rightarrow \mathbb{R}$ be convex. Assume that $[f, C]$ is closed. Then, $f$ is weakly lower semicontinuous.

**Proof** $\forall a \in \mathbb{R}$, $V_a := \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in \mathcal{X}\}$ is closed. Then, $[f, C] \cap V_a = \{(a, x) \in \mathbb{R} \times \mathcal{X} \mid x \in C, f(x) \leq a\}$ is closed. Hence, by Proposition 4.13, $T_a := \{x \in C \mid f(x) \leq a\}$ is closed. Since $f$ is convex, then $T_a$ is convex. By Proposition 8.10, $T_a$ is weakly closed. Note that 

$$\{x \in C \mid -f(x) < a\} = C \setminus \{x \in C \mid f(x) \leq -a\} = C \setminus T_{-a}$$

Clearly, $C \setminus T_{-a}$ is weakly upper semicontinuous and $f$ is weakly lower semicontinuous. This completes the proof of the proposition. \hfill $\square$

### 8.5 Conjugate Convex Functionals

**Definition 8.27** Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, and $f : C \rightarrow \mathbb{R}$ be convex. The conjugate set $C_{\text{conj}}$ is defined as

$$C_{\text{conj}} := \{x_* \in \mathcal{X}^* \mid \sup_{x \in C} (\langle x_*, x \rangle - f(x)) < +\infty\}$$

and the functional $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ conjugate to $f$ is defined by

$$f_{\text{conj}}(x_*) = \sup_{x \in C} (\langle x_*, x \rangle - f(x)), \quad \forall x_* \in C_{\text{conj}}$$
We will use a compact notation \([f, C]_{\text{conj}}\) for \([f_{\text{conj}}, C_{\text{conj}}]\).

In the above definition, \(f_{\text{conj}}\) takes value in \(\mathbb{R}\) since \(x_* \in C_{\text{conj}}\) and \(C \neq \emptyset\).

**Proposition 8.28** Let \(X\) be a real normed linear space, \(C \subseteq X\) be a nonempty convex set, \(f : C \to \mathbb{R}\) be convex. Then, \(C_{\text{conj}} \subseteq X^*\) is convex, \(f_{\text{conj}} : C_{\text{conj}} \to \mathbb{R}\) is convex, and \([f_{\text{conj}}, C_{\text{conj}}] \subseteq \mathbb{R} \times X^*\) is a closed convex set.

**Proof** \(\forall x_1, x_2 \in C_{\text{conj}}, \forall \alpha \in [0, 1] \subset \mathbb{R}, \) let \(M_i := f_{\text{conj}}(x_i) = \sup_{x \in C} (\langle \langle x_1, x \rangle \rangle - f(x)) \in \mathbb{R}, i = 1, 2. \forall x \in C, \) we have \(\langle \langle \alpha x_1 + (1 - \alpha) x_2, x \rangle \rangle - f(x) = \alpha (\langle \langle x_1, x \rangle \rangle - f(x)) + (1 - \alpha) (\langle \langle x_2, x \rangle \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha) M_2 < +\infty. \) Then, \(\sup_{x \in C} (\langle \langle \alpha x_1 + (1 - \alpha) x_2, x \rangle \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha) M_2. \) Hence, \(\alpha x_1 + (1 - \alpha) x_2 \in C_{\text{conj}}\) and \(C_{\text{conj}}\) is convex.

The above implies that \(f_{\text{conj}}(\alpha x_1 + (1 - \alpha) x_2) \leq \alpha f_{\text{conj}}(x_1) + (1 - \alpha) f_{\text{conj}}(x_2). \) Hence, \(f_{\text{conj}}\) is convex.

By Proposition 8.20, \([f_{\text{conj}}, C_{\text{conj}}]\) is convex. \(\forall (s_0, x_0) \in [f_{\text{conj}}, C_{\text{conj}}],\) by Proposition 4.13, there exists \((s_k, x_k)\) such that \(\lim_{k \to \infty} (s_k, x_k) = (s_0, x_0).\) By Proposition 3.67, we have \(\lim_{k \to \infty} s_k = s_0\) and \(\lim_{k \to \infty} x_k = x_0. \forall x \in C, \) \(\forall k \in \mathbb{N}, \) since \((s_k, x_k) \in [f_{\text{conj}}, C_{\text{conj}}]\), then, \(\langle \langle x_k, x \rangle \rangle - f(x) \leq f_{\text{conj}}(x_k) \leq s_k. \) By Propositions 7.72 and 3.66, we have \(\langle \langle x_0, x \rangle \rangle - f(x) = \lim_{k \to \infty} (\langle \langle x_k, x \rangle \rangle - f(x)) \leq \lim_{k \to \infty} s_k = s_0 < +\infty. \) Hence, \(x_0 \in C_{\text{conj}}\) and \(f_{\text{conj}}(x_0) \leq s_0.\) Then, \((s_0, x_0) \in [f_{\text{conj}}, C_{\text{conj}}]\). This shows that \([f_{\text{conj}}, C_{\text{conj}}] \subseteq [f_{\text{conj}}, C_{\text{conj}}].\) By Proposition 3.3, \([f_{\text{conj}}, C_{\text{conj}}]\) is closed.

This completes the proof of the proposition. \(\square\)

Geometric interpretation of \(f_{\text{conj}} : C_{\text{conj}} \to \mathbb{R}. \) Fix \(x_* \in C_{\text{conj}}, \forall (r, x) \in [f, C],\) we have \(f(x) \leq r. \) Then, \(\langle \langle (-1, x_*), (r, x) \rangle \rangle = \langle \langle x_*, x \rangle \rangle - r \leq f_{\text{conj}}(x_*).\) It is easy to recognize that \(\sup_{(r, x) \in [f, C]} (\langle \langle (-1, x_*), (r, x) \rangle \rangle = f_{\text{conj}}(x_*).\) Hence, the closed hyperplane \(H := \{(r, x) \in \mathbb{R} \times X \mid \langle \langle (-1, x_*), (r, x) \rangle \rangle = f_{\text{conj}}(x_*)\}\) is a supporting hyperplane of \([f, C].\)

**Proposition 8.29** Let \(X\) be a real normed linear space, \(C \subseteq X\) be a nonempty convex set, and \(f : C \to \mathbb{R}\) be a convex functional. Assume that \([f, C] := K \subseteq \mathbb{R} \times X\) is closed. Then, \(\exists x_0 \in X^*\) such that

\[
\sup_{(r, x) \in K} \langle \langle (-1, x_0), (r, x) \rangle \rangle < +\infty
\]

Then, \(x_0 \in C_{\text{conj}} \neq \emptyset.\)

**Proof** Fix \(x_0 \in C \neq \emptyset.\) By Proposition 8.20, \(K\) is convex. Clearly, \(K \neq \emptyset.\) By the assumption of the proposition, \(K\) is closed. Let \(r_0 := f(x_0) - 1 \in \mathbb{R}.\) Then, \((r_0, x_0) \notin K.\) By Proposition 8.10 and Example 7.76, \(\exists (s_0, x_0) \in \mathbb{R} \times X^*\) such that

\[
\langle \langle (s_0, x_0), (r_0, x_0) \rangle \rangle < \inf_{(r, x) \in K} \langle \langle (s_0, x_0), (r, x) \rangle \rangle \tag{8.1}
\]
8.5. CONJUGATE CONVEX FUNCTIONALS

Since \((f(x_0), x_0) \in K\), then \(\bar{s}_0r_0 + \langle \bar{x}_0, x_0 \rangle < \bar{s}_0f(x_0) + \langle \bar{x}_0, x_0 \rangle\), which implies that \(\bar{s}_0 > 0\). Let \(x_0 = -\bar{s}_0^{-1}\bar{x}_0 \in X^*\). Then, (8.1) is equivalent to

\[
\langle \langle -1, x_0 \rangle, (r_0, x_0) \rangle > \sup_{(r, x) \in K} \langle \langle -1, x_0 \rangle, (r, x) \rangle
\]

The above implies that \(\langle \langle -1, x_0 \rangle, (r_0, x_0) \rangle > \sup_{x \in C} \langle \langle x_0, x \rangle - f(x) \rangle\).
Hence, \(x_0 \in C_{\text{conj}} \neq \emptyset\).

This completes the proof of the proposition.

\[\Box\]

**Proposition 8.30** Let \(X\) be a real normed linear space and \(K \subseteq \mathbb{R} \times X =: \mathcal{W}\) be a closed convex set.

Assume that there exists a nonvertical hyperplane such that \(K\) is contained in one of the half-spaces associated with the hyperplane, that is, \(\exists (s_1, x_1) \in \mathcal{W}^* = \mathbb{R} \times X^*\) with \(s_1 \neq 0\) such that \(\sup_{(r, x) \in K} \langle \langle s_1, x_1 \rangle, (r, x) \rangle < +\infty\).

Then, \(\forall (r_0, x_0) \in \mathcal{W} \setminus K\), there exists a nonvertical hyperplane separating \((r_0, x_0)\) and \(K\), that is, \(\exists x_0 \in X^*\) such that either

\[
\sup_{(r, x) \in K} \langle \langle -1, x_0 \rangle, (r, x) \rangle < \langle \langle -1, x_0 \rangle, (r_0, x_0) \rangle \tag{8.2a}
\]

or

\[
\inf_{(r, x) \in K} \langle \langle -1, x_0 \rangle, (r, x) \rangle > \langle \langle -1, x_0 \rangle, (r_0, x_0) \rangle \tag{8.2b}
\]

**Proof** Fix \((r_0, x_0) \in \mathcal{W} \setminus K\). By Proposition 8.10, \(\exists (\bar{s}_0, \bar{x}_0) \in \mathbb{R} \times X^*\) such that

\[
\langle \langle \bar{s}_0, \bar{x}_0 \rangle, (r_0, x_0) \rangle < \inf_{(r, x) \in K} \langle \langle \bar{s}_0, \bar{x}_0 \rangle, (r, x) \rangle \tag{8.3}
\]

We will distinguish three exhaustive and mutually exclusive cases: Case 1: \(\bar{s}_0 > 0\); Case 2: \(\bar{s}_0 < 0\); Case 3: \(\bar{s}_0 = 0\). Case 1: \(\bar{s}_0 > 0\). Let \(x_0 = -\bar{s}_0^{-1}\bar{x}_0 \in X^*\). Then, (8.3) is equivalent to

\[
-r_0 + \langle \langle x_0, x_0 \rangle \rangle > \sup_{(r, x) \in K} \langle -r + \langle \langle x_0, x \rangle \rangle \rangle
\]

Hence, (8.2a) holds.

Case 2: \(\bar{s}_0 < 0\). Let \(x_0 = -\bar{s}_0^{-1}\bar{x}_0 \in X^*\). Then, (8.3) is equivalent to

\[
-r_0 + \langle \langle x_0, x_0 \rangle \rangle < \inf_{(r, x) \in K} \langle -r + \langle \langle x_0, x \rangle \rangle \rangle
\]

Hence, (8.2b) holds.

Case 3: \(\bar{s}_0 = 0\). We will further distinguish two exhaustive and mutually exclusive cases: Case 3a: \(K = \emptyset\); Case 3b: \(K \neq \emptyset\). Case 3a: \(K = \emptyset\). Take \(x_0 = 0\). Then,

\[
-r_0 = -\infty = \sup_{(r, x) \in K} \langle \langle -1, x_0 \rangle, (r, x) \rangle
\]
Hence, (8.2a) holds.

Case 3b: $K \neq \emptyset$. Define

\[
M_1 := \sup_{(r,x) \in K} \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle - \langle\langle (s_1, x_{s1}), (r_0, x_0) \rangle\rangle \in \mathbb{R}
\]

\[
M_2 := \inf_{(r,x) \in K} \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r, x) \rangle\rangle - \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
\in (0, \infty) \subset \mathbb{R}
\]

Let $\delta := M_2/(1 + |M_1|) \in (0, \infty) \subset \mathbb{R}$ and $(\bar{s}_0, \bar{x}_{s0}) = (\bar{s}_0, \bar{x}_{s0}) - \delta(s_1, x_{s1}) \in \mathbb{R} \times \mathbb{X}^\ast$. Then, we have

\[
\inf_{(r,x) \in K} \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r, x) \rangle\rangle - \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
= \inf_{(r,x) \in K} \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r, x) \rangle\rangle - \delta \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
\geq \inf_{(r,x) \in K} \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r, x) \rangle\rangle + \inf_{(r,x) \in K} \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \delta \sup_{(r,x) \in K} \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \langle\langle (\bar{s}_0, \bar{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
= M_2 - \delta M_1 > 0
\]

where we have applied Proposition 3.81 in the above. Hence, we have obtained an alternative pair $(\bar{s}_0, \bar{x}_{s0}) := (\bar{s}_0, \bar{x}_{s0})$ such that (8.3) holds. For this alternative pair, $\bar{s}_0 = -\delta s_1 \neq 0$. Hence, this case can be solved by Case 1 or Case 2 with the alternative pair.

This completes the proof of the proposition. \(\square\)

**Definition 8.31** Let $\mathbb{X}$ be a real normed linear space, $\Gamma \subseteq \mathbb{X}^\ast$ be a nonempty convex set, and $\varphi : \Gamma \to \mathbb{R}$ be convex. The pre-conjugate set $\text{conj} \Gamma$ is defined as

\[
\text{conj} \Gamma := \{ x \in \mathbb{X} \mid \sup_{x_\ast \in \Gamma} \langle\langle x_\ast, x \rangle\rangle - \varphi(x_\ast) < +\infty \}
\]

and the functional $\text{conj} \varphi : \text{conj} \Gamma \to \mathbb{R}$ pre-conjugate to $\varphi$ is defined by

\[
\text{conj} \varphi(x) = \sup_{x_\ast \in \Gamma} \langle\langle x_\ast, x \rangle\rangle - \varphi(x_\ast), \quad \forall x \in \text{conj} \Gamma
\]

We will use a compact notation $\text{conj} \varphi, \Gamma$ for $[\text{conj} \varphi, \text{conj} \Gamma]$.

In the above definition, $\text{conj} \varphi$ takes value in $\mathbb{R}$ since $x \in \text{conj} \Gamma$ and $\Gamma \neq \emptyset$.

Next, we state two duality results regarding conjugate convex functionals.
8.5. CONJUGATE CONVEX FUNCTIONALS

Proposition 8.32 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, and $f : C \rightarrow \mathbb{R}$ be a convex functional. Assume that $[f, C]$ is closed. Then, $[f, C] = \text{conj}[[f, C]]$. Therefore, $\forall x_0 \in C$, $f(x_0) = \sup_{x, \in C_{\text{conj}}}((\langle x_*, x \rangle) - f_{\text{conj}}(x_*))$.

**Proof** Since $C \neq \emptyset$, then $C_{\text{conj}}$ and $f_{\text{conj}}$ are well-defined. By Proposition 8.29, $C_{\text{conj}} \neq \emptyset$. By Proposition 8.28, $C_{\text{conj}}$ is convex and $f_{\text{conj}}$ is convex. Then, $\text{conj}(f_{\text{conj}})$ and $\text{conj}(f_{\text{conj}})$ are well defined. Thus, $\text{conj}[[f, C]]$ is a well-defined set.

We will first show that $[f, C] \subseteq \text{conj}[[f_{\text{conj}}, C_{\text{conj}}]] = \text{conj}[[f, C]]$. Fix any $(r, x) \in [f, C]$; $\forall x_* \in C_{\text{conj}}$, we have $f_{\text{conj}}(x_*) \geq \langle x_*, x \rangle - f(x)$. Hence, we have $r \geq f(x) \geq \langle x_*, x \rangle - f_{\text{conj}}(x_*)$. Thus, we have

$$r \geq \sup_{x_* \in C_{\text{conj}}} \langle x_*, x \rangle - f_{\text{conj}}(x_*) = \text{conj}(f_{\text{conj}})(x)$$

Hence, $(r, x) \in \text{conj}[[f_{\text{conj}}, C_{\text{conj}}]]$ and $[f, C] \subseteq \text{conj}[[f_{\text{conj}}, C_{\text{conj}}]]$.

On the other hand, fix any $(r_0, x_0) \in (\mathbb{R} \times \mathcal{X}) \setminus [f, C]$. By Proposition 8.29, $\exists x_0 \in \mathcal{X}$ such that

$$\sup_{(r, x) \in [f, C]} \langle (-1, x_0), (r, x) \rangle < +\infty$$

Note that $[f, C]$ is convex, by Proposition 8.20. By Proposition 8.30, there exists a nonvertical hyperplane separating $(r_0, x_0)$ and $K := [f, C]$, that is, $\exists x_0 \in \mathcal{X}$ such that either

$$\sup_{(r, x) \in K} \langle (-1, x_0), (r, x) \rangle < \langle (-1, x_0), (r_0, x_0) \rangle$$

or

$$\inf_{(r, x) \in K} \langle (-1, x_0), (r, x) \rangle > \langle (-1, x_0), (r_0, x_0) \rangle$$

Since $K$ is the epigraph of $f$, then (8.4b) is impossible since the left-hand-side equals to $-\infty$. Therefore, (8.4a) must hold. Then, $c := \sup_{(r, x) \in K} \langle (-1, x_0), (r, x) \rangle = \sup_{x \in C} \langle (x_0, x) - f(x) \rangle = f_{\text{conj}}(x_0) \in \mathbb{R}$ and $x_0 \in C_{\text{conj}}$. Then, $(c, x_0) \in [f_{\text{conj}}, C_{\text{conj}}]$. But, $c < -r_0 + \langle (x_0, x_0) \rangle$ implies that $r_0 < \langle x_0, x_0 \rangle - f_{\text{conj}}(x_0) \leq \sup_{x_* \in C_{\text{conj}}} \langle x_*, x \rangle - f_{\text{conj}}(x_*)$. Therefore, $(r_0, x_0) \not\in \text{conj}[[f_{\text{conj}}, C_{\text{conj}}]]$. Hence, $\text{conj}[[f_{\text{conj}}, C_{\text{conj}}]] \subseteq [f, C]$.

This yields $[f, C] = \text{conj}[[f, C]]$, which implies that $f$ and $\text{conj}(f_{\text{conj}})$ admit the same domain of definition and equal to each other on $C$. Therefore, $\forall x_0 \in C$, $f(x_0) = \text{conj}(f_{\text{conj}})(x_0) = \sup_{x_* \in C_{\text{conj}}} \langle (x_*, x_0) \rangle - f_{\text{conj}}(x_*)$.

\[ \square \]

Proposition 8.33 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ be a nonempty convex set, $f : C \rightarrow \mathbb{R}$ be a convex functional, $f$ be lower semicontinuous at $x_0 \in C$, $C_{\text{conj}} \subseteq \mathcal{X}^*$ be the conjugate set, and $f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}$ be the conjugate functional. Then, $f(x_0) = \sup_{x_* \in C_{\text{conj}}} \langle (x_*, x_0) \rangle - f_{\text{conj}}(x_*)$.
Proof. For all $x^* \in C_{\text{conj}}$, we have $\langle \langle x^*, x_0 \rangle\rangle - f_{\text{conj}}(x^*) = \langle \langle x^*, x_0 \rangle\rangle + \inf_{x \in C}(-\langle \langle x^*, x \rangle\rangle + f(x)) \leq f(x_0)$. Then,

$$f(x_0) \geq \sup_{x^* \in C_{\text{conj}}} (\langle \langle x^*, x_0 \rangle\rangle - f_{\text{conj}}(x^*))$$

For all $\epsilon \in (0, \infty) \subseteq \mathbb{R}$, by the lower semicontinuity of $f$ at $x_0$, there exists $\delta \in (0, \epsilon/2] \subseteq \mathbb{R}$ such that $f(x) > f(x_0) - \epsilon/2$, $\forall x \in B_X(x_0, \delta) \cap C$. Consider the real normed linear space $\mathcal{W} := \mathbb{R} \times X$. Define $K_2 := [f, C]$, which is clearly nonempty. By Proposition 8.20, $K_2$ is convex. Define $K_1 := \mathcal{B}_W((f(x_0) - \epsilon, x_0), \delta)$. Then, $K_1$ is convex and $K^o_1 = K_1 \neq \emptyset$.

Claim 8.33.1 $K^o_1 \cap K_2 = \emptyset$.

Proof of claim: Suppose $\exists (r, x) \in K^o_1 \cap K_2$. Then, $x \in B_X(x_0, \delta) \cap C$, $r < f(x_0) - \epsilon + \delta \leq f(x_0) - \epsilon/2$ and $r \geq f(x)$. This leads to $f(x) < f(x_0) - \epsilon/2$, which is a contradiction. Hence, $K^o_1 \cap K_2 = \emptyset$. This completes the proof of the claim.

By Eidelheit Separation Theorem and Example 7.76, $\exists c \in \mathbb{R}$ and $\exists (\bar{s}_0, \bar{x}_{\ast 0}) \in \mathcal{W}^* = \mathbb{R} \times X^*$ with $(\bar{s}_0, \bar{x}_{\ast 0}) \neq (0, \partial_X x_0)$ such that

$$\sup_{(r, x) \in K_1} \langle \langle \bar{s}_0, \bar{x}_{\ast 0} \rangle, (r, x) \rangle \leq c \leq \inf_{(r, x) \in K_2} \langle \langle \bar{s}_0, \bar{x}_{\ast 0} \rangle, (r, x) \rangle \quad (8.5)$$

The second inequality in (8.5) implies that $-\infty < c \leq \inf_{(r, x) \in K_2} \langle \langle \bar{s}_0, \bar{x}_{\ast 0} \rangle, (r, x) \rangle$. Since $K_2 = [f, C]$, then $\bar{s}_0 \geq 0$.

Claim 8.33.2 $\bar{s}_0 > 0$.

Proof of claim: Suppose $\bar{s}_0 = 0$. By the fact that $(\bar{s}_0, \bar{x}_{\ast 0}) \neq (0, \partial_X x_0)$, we have $\bar{x}_{\ast 0} \neq \partial_X x_0$. Then, (8.5) implies that $\langle \langle \bar{x}_{\ast 0}, x_0 \rangle\rangle \geq c \geq \sup_{(r, x) \in K_1} \langle \langle \bar{s}_0, \bar{x}_{\ast 0} \rangle, (r, x) \rangle = \sup_{x \in B_{\bar{x}_{\ast 0}}(x_0, \delta)} \langle \langle \bar{s}_0, x \rangle\rangle$. This is impossible since $\bar{x}_{\ast 0} \neq \partial_X x_0$. Therefore, $\bar{s}_0 < 0$. This completes the proof of the claim.

Let $x^*_0 := -\bar{s}_0^{-1} \bar{x}_{\ast 0} \in X^*$. Then, (8.5) is equivalent to

$$\inf_{(r, x) \in K_1} \langle \langle \bar{x}_{\ast 0}, x_0 \rangle\rangle - r) \geq -\bar{s}_0^{-1} c \geq \sup_{(r, x) \in K_2} (\langle \langle \bar{x}_{\ast 0}, x \rangle\rangle - r)$$

$$= \sup_{x \in C} (\langle \langle \bar{x}_{\ast 0}, x \rangle\rangle - f(x))$$

Hence, $x_{\ast 0} \in C_{\text{conj}}$ and $f_{\text{conj}}(x_{\ast 0}) \leq \langle \langle \bar{x}_{\ast 0}, x_0 \rangle\rangle - f(x_0) + \epsilon$. Therefore, we have

$$\langle \langle \bar{x}_{\ast 0}, x_0 \rangle\rangle - f_{\text{conj}}(x_{\ast 0}) \geq f(x_0) - \epsilon$$

By the arbitrariness of $\epsilon$, $f(x_0) = \sup_{x_{\ast 0} \in C_{\text{conj}}} (\langle \langle \bar{x}_{\ast 0}, x \rangle\rangle - f_{\text{conj}}(x_{\ast 0}))$. This completes the proof of the proposition.

Proposition 8.34 Let $X$ be a real normed linear space, $C \subseteq X$ be a nonempty convex set, $C^o \neq \emptyset$, $f : C \to \mathbb{R}$ be a convex functional, $f$ be continuous at $\bar{x} \in C^o$, $C_{\text{conj}} \subseteq X^*$ be the conjugate set, and $f_{\text{conj}} : C_{\text{conj}} \to \mathbb{R}$ be the conjugate functional. Then, $\forall x_0 \in C^o$, $f(x_0) = \max_{x_{\ast 0} \in C_{\text{conj}}} (\langle \langle \bar{x}_{\ast 0}, x_0 \rangle\rangle - f_{\text{conj}}(x_{\ast 0}))$. 


Proof \(\forall x_0 \in C_{\text{conj}}, \text{we have } \langle\langle x_0, x_0 \rangle\rangle - f_{\text{conj}}(x_0) = \langle\langle x_0, x_0 \rangle\rangle + \inf_{x \in C} (- \langle\langle x, x \rangle\rangle + f(x)) \leq f(x_0). \) Then,

\[
f(x_0) \geq \sup_{x_0 \in C_{\text{conj}}} (\langle\langle x_0, x_0 \rangle\rangle - f_{\text{conj}}(x_0))
\]

Consider the real normed linear space \(W := \mathbb{R} \times X.\) Let \(V = \{(f(x_0), x_0)\} \subseteq W.\) Clearly, \(V\) is a linear variety. Define \(K := \{f, C\}.\) By Proposition 8.20, \(K\) is convex. By Proposition 8.22, \(K\) admits relative interior point \((\tilde{r}, \tilde{x})\) for some \(\tilde{r} > f(\tilde{x}).\) By Proposition 8.21, \(V(\{f, C\}) = \mathbb{R} \times V(C) = \mathbb{R} \times X = W\) since \(C^\circ \neq \emptyset.\) Hence, \((\tilde{r}, \tilde{x}) \in K^\circ \neq \emptyset.\) Note that \((f(x_0) - \delta, x_0) \notin \{f, C\}, \forall \delta \in (0, \infty) \subseteq \mathbb{R}.\) Then, \((f(x_0), x_0) \notin K^\circ.\) Hence, \(V \cap K^\circ = \emptyset.\) By Mazur’s Theorem and Example 7.76, \(\exists \in \mathbb{R}\) and \(\exists(s_0, \tilde{x}_0) \in W^* = \mathbb{R} \times X^*\) with \((s_0, \tilde{x}_0) \neq (0, 0)\) such that

\[
\langle\langle s_0, \tilde{x}_0 \rangle\rangle, (f(x_0), x_0) \rangle \rangle = c \geq \sup_{(r, x) \in K} \langle\langle s_0, \tilde{x}_0 \rangle\rangle, (r, x) \rangle
\]

which is equivalent to

\[
s_0 f(x_0) + \langle\langle \tilde{x}_0, x_0 \rangle\rangle \geq \sup_{(r, x) \in K} (s_0 r + \langle\langle \tilde{x}_0, x \rangle\rangle)
\] (8.6)

Since \(K = \{f, C\},\) then \(s_0 \geq 0,\) otherwise the right-hand-side of (8.6) equals to \(+\infty.\)

Claim 8.34.1 \(s_0 \leq 0.\)

Proof of claim: Suppose \(s_0 = 0.\) By the fact that \((s_0, \tilde{x}_0) \neq (0, 0)\) \(\notin X^*,\) we have \(\tilde{x}_0 \neq \tilde{x}_0.\) Then, (8.6) implies that \(\langle\langle \tilde{x}_0, x_0 \rangle\rangle \geq \sup_{x \in C} \langle\langle \tilde{x}_0, x \rangle\rangle.\) Note that \(x_0 \in C^\circ\) and \(\tilde{x}_0 \neq \tilde{x}_0\) implies that \(\sup_{x \in C} \langle\langle \tilde{x}_0, x \rangle\rangle > \langle\langle \tilde{x}_0, x_0 \rangle\rangle.\) This is a contradiction. Therefore, \(s_0 < 0.\)

This completes the proof of the claim. \(\square\)

Let \(x_0 := |s_0|^{-1} \tilde{x}_0 \in X^*.\) Then, (8.6) is equivalent to

\[
-f(x_0) + \langle\langle x_0, x_0 \rangle\rangle \geq \sup_{(r, x) \in K} (-r + \langle\langle x_0, x \rangle\rangle) = \sup_{x \in C} \langle\langle x_0, x \rangle\rangle - f(x)
\]

Hence, \(x_0 \in C_{\text{conj}}\) and \(-f(x_0) + \langle\langle x_0, x_0 \rangle\rangle \geq f_{\text{conj}}(x_0).\)

Therefore, we have

\[
f(x_0) \leq \langle\langle x_0, x_0 \rangle\rangle - f_{\text{conj}}(x_0) \leq \sup_{x_0 \in C_{\text{conj}}} (\langle\langle x_0, x_0 \rangle\rangle - f_{\text{conj}}(x_0)) \leq f(x_0)
\]

Hence, \(f(x_0) = \max_{x_0 \in C_{\text{conj}}} (\langle\langle x_0, x_0 \rangle\rangle - f_{\text{conj}}(x_0)),\) where the maximum is achieved at \(x_0.\) \(\square\)
8.6 Fenchel Duality Theorem

Let $X$ be a real normed linear space, $C_1, C_2 \subseteq X$ be nonempty convex sets, and $f_1 : C_1 \to \mathbb{R}$ and $f_2 : C_2 \to \mathbb{R}$ be convex functionals. We consider the problem of

$$\mu := \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x))$$

We assume that $f_{1\text{conj}}, C_{1\text{conj}}, f_{2\text{conj}}, C_{2\text{conj}}$ are easily characterized yet $(f_1 + f_2)_{\text{conj}}$ and $(C_1 \cap C_2)_{\text{conj}}$ are difficult to determine. Then, the above infimum can be equivalently calculated by

$$\mu = \inf_{(r_1, r_2) \in [f_1, C_1], (r_2, x) \in [f_2, C_2]} r_1 + r_2$$

The idea of Fenchel Duality Theorem can be illustrated in Figure 8.1.

**Theorem 8.35 (Fenchel Duality Theorem)** Let $X$ be a real normed linear space, $C_1, C_2 \subseteq X$ be nonempty convex sets, $f_1 : C_1 \to \mathbb{R}$ and $f_2 : C_2 \to \mathbb{R}$ be convex functionals, and $C_2 \cap C_1^\circ \neq \emptyset$. Assume that that $f_1$ is continuous at $\bar{x} \in C_1^\circ$ and $\mu := \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x))$ is finite. Let $f_{1\text{conj}} : C_{1\text{conj}} \to \mathbb{R}$ and $f_{2\text{conj}} : C_{2\text{conj}} \to \mathbb{R}$ be conjugate functionals of $f_1$ and $f_2$, respectively. Then, the following statements hold.

(i) $\mu = \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \max_{x_0 \in C_{1\text{conj}} \cap (-C_{2\text{conj}})} (-f_{1\text{conj}}(x_0) - f_{2\text{conj}}(-x_0))$, where the maximum is achieved at some $x_0 \in C_{1\text{conj}} \cap \ldots$
We will show that $s$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: with the definition of $V(r, x)$ admits relative interior point (8.6). Fenchel Duality Theorem

(ii) If there exists $x_0 \in C_1 \cap (-C_2)$ and $x_0 \in C_1 \cap C_2$ such that (8.7) holds, then the infimum is achieved at $x_0$ and the maximum is achieved at $x_0$.

Proof. $\forall x \in C_1 \cap (-C_2)$, we have

$$f_{1\text{conj}}(x) = -\sup_{x \in C_1} (\langle x, x \rangle - f_1(x)) = \inf_{x \in C_1} (-\langle x, x \rangle + f_1(x))$$

$$-f_{2\text{conj}}(x) = \inf_{x \in C_2} (\langle x, x \rangle + f_2(x))$$

Then, we have

$$-f_{1\text{conj}}(x) - f_{2\text{conj}}(x) \leq \inf_{x \in C_1 \cap C_2} (-\langle x, x \rangle + f_1(x)) + \inf_{x \in C_1 \cap C_2} (\langle x, x \rangle + f_2(x)) = \mu (8.8)$$

(i) Consider the sets $K_1 := [f_1 - \mu, C_1]$ and $K_2 := \{(r, x) \in \mathbb{R} \times \mathbb{X} | x \in C_2, r \leq -f_2(x)\} = \{(r, x) \in \mathbb{R} \times \mathbb{X} | (r, x) \in [f_2, C_2]\}$. By Proposition 8.20, $K_1$ and $K_2$ are nonempty convex sets. By Proposition 8.22, $K_1$ admits relative interior point $(\bar{r}, \bar{x})$ for some $\bar{r} \in \mathbb{R}$. By Proposition 8.21, $V([f_1 - \mu, C_1]) = \mathbb{R} \times V(C_1) = \mathbb{R} \times \mathbb{X}$. Then, $(\bar{r}, \bar{x}) \in K_1 \neq \emptyset$.

We will show that $K_1^o \cap K_2 = \emptyset$ by an argument of contradiction. Suppose $(r, x) \in K_1^o \cap K_2 \neq \emptyset$. Then, we have $(r, x) \in K_1^o$, which implies that $x \in C_1$ and $r > f_1(x) - \mu$; and $(r, x) \in K_2$, which implies that $x \in C_2$ and $r \leq -f_2(x)$. Then, $x \in C_1 \cap C_2$ and $f_1(x) + f_2(x) < \mu$. This contradicts with the definition of $\mu$. Hence, $K_1^o \cap K_2 = \emptyset$.

By Eidelheit Separation Theorem and Example 7.76, $\exists (\tilde{s}_0, \tilde{x}_0) \in \mathbb{R} \times \mathbb{X}^*$ with $(\tilde{s}_0, \tilde{x}_0) \neq \emptyset$, such that

$$-\infty < \sup_{(r, x) \in K_1^o} \langle (\tilde{x}_0, x) + rs_0 \rangle \leq \inf_{(r, x) \in K_2} \langle (\tilde{x}_0, x) + rs_0 \rangle < +\infty (8.9)$$

We will show that $\bar{s}_0 < 0$ by an argument of contradiction. Suppose $\bar{s}_0 \geq 0$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\bar{s}_0 > 0$; Case 2: $\bar{s}_0 = 0$. Case 1: $\bar{s}_0 > 0$. Then, $\sup_{(r, x) \in K_1} \langle (\tilde{x}_0, x) + rs_0 \rangle = +\infty$, which contradicts (8.9). Case 2: $\bar{s}_0 = 0$. Then, $\tilde{x}_0 \neq \emptyset$. By (8.9), we have

$$-\infty < \sup_{x \in C_1} \langle (\tilde{x}_0, x) \rangle \leq \inf_{x \in C_2} \langle (\tilde{x}_0, x) \rangle < +\infty$$
Let $\hat{x} \in C_1^{\circ} \cap C_2 \neq \emptyset$. Then, $\sup_{x \in C_1} \langle \langle \hat{x}, x \rangle \rangle \leq \langle \langle \hat{x}, \hat{x} \rangle \rangle$. This is not possible since $\hat{x} \in C_1^{\circ}$ and $\hat{x} \neq \partial_XX$. Hence, we have a contradiction in both cases. Therefore, $\bar{s}_0 < 0$.

Let $x^* = |\bar{s}_0|^{-1}\bar{x}^*$. Then, (8.9) is equivalent to

$$-\infty < \sup_{(r,x) \in K_1} \langle \langle (x^*, x) \rangle \rangle - r \leq \inf_{(r,x) \in K_2} \langle \langle (x^*, x) \rangle \rangle - r < +\infty \quad (8.10)$$

Note that

$$+\infty > \sup_{(r,x) \in K_1} \langle \langle (x^*, x) \rangle \rangle - r = \sup_{x \in C_1} \langle \langle (x^*, x) \rangle \rangle - f_1(x) + \mu$$

$$\geq \sup_{x \in C_1 \cap C_2} \langle \langle (x^*, x) \rangle \rangle - f_1(x) + \mu =: d_1 > -\infty$$

Hence, $x^* \in C_{1\text{conj}}$. Note also that

$$-\infty < \inf_{(r,x) \in K_2} \langle \langle (x^*, x) \rangle \rangle - r = \inf_{x \in C_2} \langle \langle (x^*, x) \rangle \rangle + f_2(x)$$

$$= - \sup_{x \in C_2} \langle \langle -x^*, x \rangle \rangle - f_2(x) \leq \inf_{x \in C_1 \cap C_2} \langle \langle (x^*, x) \rangle \rangle + f_2(x)$$

$$=: d_2 < +\infty$$

Hence, $x^* \in -C_{2\text{conj}}$. By (8.10), $d_1 \leq d_2$. On the other hand, we have

$$d_2 - d_1$$

$$= \inf_{x \in C_1 \cap C_2} \langle \langle (x^*, x) \rangle \rangle + f_2(x) + \inf_{x \in C_1 \cap C_2} \langle \langle -x^*, x \rangle \rangle + f_1(x) - \mu$$

$$\leq \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) - \mu = 0$$

Hence, we have

$$d_1 = \sup_{x \in C_1} \langle \langle (x^*, x) \rangle \rangle - f_1(x) + \mu = - \sup_{x \in C_2} \langle \langle -x^*, x \rangle \rangle - f_2(x) = d_2$$

This implies that $d_1 = f_{1\text{conj}}(x^*) + \mu = -f_{2\text{conj}}(-x^*) = d_2$. Therefore, we have $\mu = -f_{1\text{conj}}(x^*) - f_{2\text{conj}}(-x^*)$ and $x^* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$. This coupled with (8.8) implies

$$\mu = \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \max_{x \in C_{1\text{conj}} \cap (-C_{2\text{conj}})} (-f_{1\text{conj}}(x^*) - f_{2\text{conj}}(-x^*))$$

where the maximum is achieved at some $x^* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})$. If the infimum is achieved at $x_0 \in C_1 \cap C_2$, then we have

$$-f_{1\text{conj}}(x^*) \leq \inf_{x \in C_1 \cap C_2} (\langle \langle x^*, x \rangle \rangle + f_1(x)) \leq \langle \langle x^*, x_0 \rangle \rangle$$

$$+ f_1(x_0)$$

$$-f_{2\text{conj}}(-x^*) \leq \inf_{x \in C_1 \cap C_2} (\langle \langle x^*, x \rangle \rangle + f_2(x)) \leq \langle \langle x^*, x_0 \rangle \rangle + f_2(x_0)$$
8.6. FENCHEL DUALITY THEOREM

\[ \mu = -f_{1\text{conj}}(x_{0}) - f_{2\text{conj}}(x_{0}) \leq f_{1}(x_{0}) + f_{2}(x_{0}) = \mu \]

Hence, (8.7) holds.

(ii) If there exists \( x_{0} \in C_{1\text{conj}} \cap (-C_{2\text{conj}}) \) and \( x_{0} \in C_{1} \cap C_{2} \) such that (8.7) holds, then, by (8.8), we have

\[
\mu \geq -f_{1\text{conj}}(x_{0}) - f_{2\text{conj}}(-x_{0}) = -\langle\langle x_{0}, x_{0} \rangle\rangle + f_{1}(x_{0}) \\
+ \langle\langle x_{0}, x_{0} \rangle\rangle + f_{2}(x_{0}) = f_{1}(x_{0}) + f_{2}(x_{0}) = \mu
\]

Hence, the infimum is achieved at \( x_{0} \) and the maximum is achieved at \( x_{0} \).

This completes the proof of the theorem.

Proposition 8.36

Let \( \Omega \) be a set, \( J : \Omega \to \mathbb{R} \), \( \Lambda \) be an index set. \( A : \Lambda \to \Omega_{2} \). Assume \( \bigcup_{\lambda \in \Lambda} A(\lambda) = \Omega \). Then,

\[
\inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega) \quad \text{sup} \quad \sup_{\lambda \in \Lambda} \sup_{\omega \in A(\lambda)} J(\omega)
\]

**Proof**

\[
\forall \omega_{0} \in \Omega, \text{there exists } \lambda_{0} \in \Lambda, \text{such that } \omega_{0} \in A(\lambda_{0}). \text{Then,}
\]

\[
J(\omega_{0}) \geq \inf_{\omega \in A(\lambda_{0})} J(\omega) \geq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)
\]

Hence, by Proposition 3.81, we have

\[
\inf_{\omega \in \Omega} J(\omega) \geq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)
\]

On the other hand, \( \forall \lambda \in \Lambda, A(\lambda) \subseteq \Omega \) implies that

\[
\inf_{\omega \in \Omega} J(\omega) \leq \inf_{\omega \in A(\lambda)} J(\omega)
\]

Then,

\[
\inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)
\]

Combining the above arguments, we have

\[
\inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega)
\]

Substituting \(-J\) into this inequality, by Proposition 3.81, yields the other equality.

This completes the proof of the proposition.

Proposition 8.37

Let \( X \) and \( Y \) be sets, \( g_{1} : X \to \mathbb{R} \), \( g_{2} : Y \to \mathbb{R} \). Assume that, \( \forall x \in X, \forall y \in Y, g_{1}(x) + g_{2}(y) \in \mathbb{R} \) is well-defined. Then,

\[
\inf_{(x, y) \in X \times Y} (g_{1}(x) + g_{2}(y)) = \inf_{x \in X} g_{1}(x) + \inf_{y \in Y} g_{2}(y)
\]

when the right-hand-side makes sense.
**Proof** Define \( \mu := \inf_{(x,y) \in X \times Y} (g_1(x) + g_2(y)) \) and \( \mu_r := \inf_{y \in Y} g_2(y) \). We will distinguish two exhaustive and mutually exclusive cases:

Case 1: \( X \times Y = \emptyset \); Case 2: \( X \times Y \neq \emptyset \). Without loss of generality, assume \( X = \emptyset \). Then, \( \mu_l = +\infty = \inf_{x \in X} g_1(x) \). Since \( \mu_r \in \mathbb{R} \) is well-defined, then \( \inf_{y \in Y} g_2(y) > -\infty \). Hence, the desired equality holds.

Case 2: \( X \times Y \neq \emptyset \). Then, \( X \neq \emptyset \) and \( Y \neq \emptyset \). This implies that \( \inf_{x \in X} g_1(x) = \inf_{(x,y) \in X \times Y} g_1(x) \) and \( \inf_{y \in Y} g_2(y) = \inf_{(x,y) \in X \times Y} g_2(y) \).

By Proposition 3.81, we have

\[
\mu_l \geq \inf_{(x,y) \in X \times Y} g_1(x) + \inf_{(x,y) \in X \times Y} g_2(y) = \mu_r
\]

We will further distinguish two exhaustive and mutually exclusive cases:

Case 2a: \( \mu_r > +\infty \); Case 2b: \( \mu_r < +\infty \). Then, \( +\infty \geq \mu_l \geq \mu_r = +\infty \). Hence, the desired equality holds.

Case 2b: \( \mu_r < +\infty \). \( \forall m \in \mathbb{R} \) with \( m > \mu_r \), by Proposition 3.81, \( \exists x_0 \in X \) and \( \exists y_0 \in Y \) such that \( g_1(x_0) + g_2(y_0) < m \). Then, again by Proposition 3.81, we have \( \mu_l < m \). This implies that \( \mu_l \leq \mu_r \). Hence, the desired equality holds.

This completes the proof of the proposition. \( \square \)

**Proposition 8.38** Let \( \mathcal{X} \) be a topological space, \( Y \) be a set, and \( V : \mathcal{X} \times Y \to \mathbb{R} \). Assume that \( V \) satisfies the following two conditions.

(i) \( \forall x_1 \in \mathcal{X} \), \( W(x_1) := \inf_{y \in Y} V(x_1, y) \in \mathbb{R} \). This defines the function \( W : \mathcal{X} \to \mathbb{R} \).

(ii) \( \forall y \in Y \), define the function \( f_y : \mathcal{X} \to \mathbb{R} \) by \( f_y(x) = V(x, y) \), \( \forall x \in \mathcal{X} \).

The collection of functions \( \{ f_y \mid y \in Y \} \) is equicontinuous.

Then, \( W \) is continuous.

**Proof** \( \forall x_0 \in \mathcal{X} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by (ii), \( \exists U \in \mathcal{O}_{\mathcal{X}} \) with \( x_0 \in U \) such that, \( \forall x \in U \), \( \forall y \in Y \), we have \( |f_y(x) - f_y(x_0)| = |V(x, y) - V(x_0, y)| < \epsilon \).

By (i), \( \exists y_0 \in Y \) such that \( |W(x) - V(x, y)| < \epsilon \) and \( |W(x_0) - V(x_0, y_0)| < \epsilon \). Note that

\[
\begin{align*}
-2\epsilon &< -|W(\bar{x}) - V(\bar{x}, \bar{y})| - |V(\bar{x}, \bar{y}) - V(x_0, \bar{y})| \leq W(\bar{x}) - V(x_0, \bar{y}) \\
&\leq W(\bar{x}) - V(x_0, \bar{y}) + V(x_0, \bar{y}) - W(x_0) = W(\bar{x}) - W(x_0) \\
&= W(\bar{x}) - V(\bar{x}, y_0) + V(\bar{x}, y_0) - W(x_0) \leq V(\bar{x}, y_0) - W(x_0) \\
&\leq |V(\bar{x}, y_0) - V(x_0, y_0)| + |V(x_0, y_0) - W(x_0)| < 2\epsilon
\end{align*}
\]

Hence, we have \( |W(\bar{x}) - W(x_0)| < 2\epsilon \) and \( W \) is continuous at \( x_0 \). By Proposition 3.9, \( W \) is continuous.

This completes the proof of the proposition. \( \square \)

Next, we present a result on game theory.
Proposition 8.39 Let $X$ be a reflexive real normed linear space and $A \subseteq X$ and $B \subseteq X^*$ be nonempty bounded closed convex sets. Then,
\[
\min_{x \in A} \max_{x^* \in B} \langle x^*, x \rangle = \max_{x^* \in B} \min_{x \in A} \langle x^*, x \rangle
\]

Proof Let $\mu := \inf_{x \in A} \sup_{x^* \in B} \langle x^*, x \rangle$. Let $h : X \times B \to \mathbb{R}$ be given by $h(x, x^*) = \langle x^*, x \rangle$, $\forall x \in X$, $\forall x^* \in B$. By Proposition 7.72, $h$ is continuous. Define $f_1 : X \to \mathbb{R}$ by $f_1(x) = \sup_{x^* \in B} h(x, x^*)$, $\forall x \in X$. $\forall x \in X$, define $h : X^* \to \mathbb{R}$ by $h_x(x^*) = h(x, x^*) = \langle x^*, x \rangle$, $\forall x^* \in X^*$. Then, $h_x$ is weak* continuous. By Proposition 8.11, $B$ is weak* compact. Then, by Proposition 5.29, $\exists x^* \in B$ such that $f_1(x) = h_x(x^*) = \max_{x^* \in B} h(x, x^*) = \max_{x^* \in B} h(x, x^*) \in \mathbb{R}$. Hence, $f_1 : X \to \mathbb{R}$ takes value in $\mathbb{R}$.

Since $B$ is bounded, then $\exists M_B \in [0, \infty) \subset \mathbb{R}$ such that $\|x^*\| \leq M_B$, $\forall x^* \in B$. $\forall x \in X$, $\forall x^* \in B$, define $h_{x^*} : X \to \mathbb{R}$ by $h_{x^*}(x) = h(x, x^*)$, $\forall x \in X$. $\forall x_0 \in X$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon/(1 + M_B) \in (0, \infty) \subset \mathbb{R}$. $\forall x \in B_x(x_0, \delta)$, $h_{x^*} \in B$, $\forall x^* \in B$, we have $|h_{x^*}(x) - h_{x^*}(x_0)| = |h(x, x^*) - h(x_0, x^*)| = |\langle x^*, x-x_0 \rangle| \leq \|x^*\| \|x-x_0\| \leq M_B \delta < \epsilon$, where we have applied Proposition 7.72. Hence, $\{h_{x^*} \mid x^* \in B\}$ is equicontinuous. By Proposition 8.38, $f_1 : X \to \mathbb{R}$ is continuous.

$\forall x_1, x_2 \in X$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have
\[
f_1(\alpha x_1 + (1-\alpha)x_2) = \max_{x^* \in B} \langle x^*, \alpha x_1 + (1-\alpha)x_2 \rangle
\]
\[
\leq \sup_{x^* \in B} \alpha \langle x^*, x_1 \rangle + (1-\alpha) \langle x^*, x_2 \rangle
\]
\[
= \alpha \max_{x^* \in B} \langle x^*, x_1 \rangle + (1-\alpha) \max_{x^* \in B} \langle x^*, x_2 \rangle
\]
where we have applied Proposition 3.81 in the second equality. Hence, $f_1$ is convex.

Since $A$ is bounded, then, $\exists M_A \in [0, \infty) \subset \mathbb{R}$ such that $\|x\| \leq M_A$, $\forall x \in A$. $\forall x \in A$, $\forall x \in X$, $\forall x^* \in B$, by Proposition 7.72, we have $\langle x^*, x \rangle \geq -\|x^*\| \|x\| \geq -M_A M_B$. Then, $f_1(x) \geq -M_A M_B$ since $B \neq \emptyset$. Then, $\mu \geq -M_A M_B$. Since $A \neq \emptyset$, then $\mu < +\infty$. Hence, $\mu$ is finite.

Define $f_2 : A \to \mathbb{R}$ by $f_2(x) = 0$, $\forall x \in A$. Clearly, $f_2$ is convex. Note $\mu = \inf_{x \in X \cap A} (f_1(x) + f_2(x))$. Now, it is easy to check that all assumptions for the Fenchel Duality Theorem are satisfied. Then,
\[
\mu = \max_{x^* \in X^*} (-f_1(x^*) - f_2(-x^*))
\]

Claim 8.39.1 $X^* = B$ and $f_1(x^*) = 0$, $\forall x^* \in B$.

Proof of claim: $X^* = \{x^* \in X^* \mid \sup_{x \in X} \langle x^*, x \rangle - f_1(x) < +\infty\}$. $\forall x_0 \in B$, $\forall x \in X$, we have $f_1(x) = \max_{x^* \in B} \langle x^*, x \rangle \geq \langle x_0, x \rangle$. Then, $\langle x_0, x \rangle - f_1(x) \leq 0$. Hence, $x_0 \in X^*$. Hence, $B \subseteq X^*$. The above also implies that $f_1(x^*) = \sup_{x \in X} \langle x_0, x \rangle - f_1(x) = 0$. 


On the other hand, \( \forall x_0 \in X^* \setminus B \), by Proposition 8.10, \( \exists x_\ast \in X^{**} \) such that \( \langle x_\ast, 0 \rangle \rangle < \inf_{x \in B} \langle x_\ast, x \rangle \rangle \). Since \( X \) is reflexive, then \( x_\ast \in \phi(x_0) \) for some \( x_0 \in X \), where \( \phi : X \to X^{**} \) is the natural mapping. Then, \( \langle x_\ast, 0 \rangle \rangle < \inf_{x \in B} \langle x_\ast, x \rangle \rangle \) and \( \langle x_\ast, x_0 \rangle \rangle > \sup_{x \in B} \langle x_\ast, x \rangle \rangle = f_1(-x_0) \). Then, \( \langle x_\ast, -x_0 \rangle \rangle = f_1(-x_0) \). Therefore, we have \( x_\ast \in X^* \setminus X_{\text{conj}} \). This implies that \( X_{\text{conj}} \subseteq B \). Therefore, we have \( X_{\text{conj}} = B \).

This completes the proof of the claim. \( \square \)

Claim 8.39.2 \( A_{\text{conj}} = X^* \) and \( f_{2,\text{conj}}(x_\ast) = \max_{x \in A} \langle x_\ast, x \rangle \rangle, \forall x_\ast \in X^* \).

Proof of claim: \( A_{\text{conj}} = \{ x_\ast \in X^* \mid \sup_{x \in A} \langle x_\ast, x \rangle \rangle < +\infty \} \).

Since \( X \) is reflexive, then, by Definition 7.89, \( \phi : X \to X^{**} \) is an isometrical isomorphism. Then, \( \phi(A) \subseteq X^{**} \) is a nonempty bounded closed convex set. By Proposition 7.90, \( Y := X^* \) is a reflexive real normed linear space. By Proposition 8.11, \( \phi(A) \) is weak* compact. \( \forall x_\ast \in X^* \), \( h_{x_\ast} \circ \phi_{\text{inv}} : X^{**} \to \mathbb{R} \) is weak* continuous. By Proposition 5.29, we have

\[
\sup_{x \in A} \langle x_\ast, x \rangle \rangle = \sup_{x_\ast \in \phi(A)} \langle x_\ast, x \rangle \rangle = \langle x_\ast, \phi_{\text{inv}}(x_\ast(0)) \rangle \rangle
\]

for some \( x_\ast(0) \in \phi(A) \) and \( x_\ast(0) = \phi_{\text{inv}}(x_\ast(0)) \). Then, \( x_\ast \in A_{\text{conj}} \) and \( A_{\text{conj}} = X^* \). Clearly, \( f_{2,\text{conj}}(x_\ast) = \langle x_\ast, x_0 \rangle \rangle = \max_{x \in A} \langle x_\ast, x \rangle \rangle \). This completes the proof of the claim. \( \square \)

Then, \( \mu = \inf_{x \in A} \max_{x \ast \in B} \langle x_\ast, x \rangle \rangle = \max_{x \ast \in B \setminus X^*} \langle x_\ast, x \rangle \rangle = \max_{x \ast \in B \setminus X^*} \langle x_\ast, x \rangle \rangle \).

To show that the infimum in the above equation is actually achieved, we note that

\[ -\mu = \min_{x \ast \in B \setminus x \in (-A)} \langle x_\ast, x \rangle \rangle = \max_{x \ast \in B \setminus \phi(-A)} \langle x_\ast, x \rangle \rangle \]

By Proposition 7.90, \( Y := X^* \) is a reflexive real normed linear space and \( \phi(X) = X^{**} \). Since \( \phi \) is an isometrical isomorphism, then \( \phi(-A) \) is a nonempty bounded closed convex set. Applying the result that we have obtained in this proof to the above, we have

\[ -\mu = \max_{x \ast \in \phi(-A)} \min_{x \ast \in B \setminus x \in (-A)} \langle x_\ast, x \rangle \rangle = \max_{x \ast \in B \setminus x \in (-A)} \langle x_\ast, x \rangle \rangle \]

which is equivalent to \( \mu = \min_{x \in A} \max_{x \ast \in B} \langle x_\ast, x \rangle \rangle \).

This completes the proof of the proposition. \( \square \)
8.7 Positive Cones and Convex Mappings

Definition 8.40 Let \( X \) be a normed linear space and \( P \subseteq X \) be a closed convex cone. For \( x, y \in X \), we will write \( x \succeq y \) (with respect to \( P \)) if \( x - y \in P \). The cone \( P \) defining this relation is called the positive cone in \( X \). The cone \( N = -P \) is called the negative cone in \( X \) and we write \( y \preceq x \) if \( y - x \in N \). We will write \( x \succ y \) \((y \prec x)\) if \( x - y \in P^\circ \) \((y - x \in N^\circ = -P^\circ)\).

It is easy to check that relations \( \preceq \) and \( \succeq \) are reflexive and transitive.

Proposition 8.41 Let \( X \) be a normed linear space with positive cone \( P \). \( \forall x_1, x_2, x_3, x_4 \in X, \forall \alpha \in [0, \infty) \subset \mathbb{R} \), we have

(i) \( x_1, x_2 \in P \implies x_1 + x_2 \in P, \alpha x_1 \in P, \) and \( -\alpha x_1 \preceq \emptyset \);
(ii) \( x_1 \preceq x_1 \);
(iii) \( x_1 \preceq x_2 \) and \( x_2 \preceq x_3 \) implies \( x_1 \preceq x_3 \);
(iv) \( x_1 \preceq x_2 \) and \( x_3 \preceq x_4 \) implies \( x_1 + x_3 \preceq x_2 + x_4 \);
(v) \( x_1 \preceq x_2 \) implies \( \alpha x_1 \preceq \alpha x_2 \);
(vi) \( x_1 \succ \emptyset \) and \( x_2 \succeq \emptyset \) implies \( x_1 + x_2 \succ \emptyset \);
(vii) \( x_1 < x_2 \) and \( \alpha > 0 \) implies \( \alpha x_1 < \alpha x_2 \).

Proof This is straightforward.

Definition 8.42 Let \( X \) be a real normed linear space and \( S \subseteq X \). The set \( S^\oplus := \{ x \in X^* \mid \langle x, x \rangle \geq 0, \forall x \in S \} \) is called the positive conjugate cone of \( S \). The set \( S^\ominus := \{ x \in X^* \mid \langle x, x \rangle \leq 0, \forall x \in S \} \) is called the negative conjugate cone of \( S \). Clearly, \( S^\ominus = -S^\oplus \).

Proposition 8.43 Let \( X \) be a real normed linear space and \( S, T \subseteq X \). Then,

(i) \( S^\oplus \subseteq X^* \) is a closed convex cone;
(ii) if \( S \subseteq T \), then \( T^\oplus \subseteq S^\oplus \).

Proof

(i) Clearly, \( \emptyset \in S^\oplus \), \( \forall x \in S^\oplus, \forall \alpha \in [0, \infty) \subset \mathbb{R} \), \( \forall x \in S \), we have \( \langle \alpha x, x \rangle \rangle = \alpha \langle x, x \rangle \rangle \geq 0 \). Hence, \( \alpha x \in S^\oplus \). Therefore, \( S^\oplus \) is a cone with vertex at origin. \( \forall x_{1}, x_{2} \in S^\oplus, \forall x \in S \), we have \( \langle x_{1} + x_{2}, x \rangle \rangle = \langle x_{1}, x \rangle \rangle + \langle x_{2}, x \rangle \rangle \geq 0 \). Hence, \( x_{1} + x_{2} \in S^\oplus \). Therefore, \( S^\oplus \) is a convex cone.
\( \forall x \in S^\ominus \), by Proposition 4.13, \( \exists \{ x_{n} \}_{n=1}^{\infty} \subseteq S^\oplus \) such that \( \lim_{n \to \infty} x_{n} = x \). \( \forall x \in S \), by Propositions 7.72 and 3.66, \( \langle x, x \rangle \rangle = \lim_{n \to \infty} \langle x_{n}, x \rangle \rangle \geq 0 \). Therefore, \( x \in S^\oplus \) and \( S^\ominus \subseteq S^\oplus \). By Proposition 3.3, \( S^\oplus \) is closed.
(ii) This is straightforward.
This completes the proof of the proposition.
Proposition 8.44 Let $X$ and $Y$ be real normed linear spaces, $A \in B(X,Y)$, and $S \subseteq X$. Then, $(A(S))^\oplus = A'_\text{inv}(S^\oplus)$.

Proof $\forall y_* \in (A(S))^\ominus \subseteq Y^*$, $\forall s \in S$, we have $\langle \langle A'y_*, s \rangle \rangle = \langle \langle y_*, As \rangle \rangle \geq 0$. Then, $A'y_* \in S^\ominus$ and $y_* \in A'_\text{inv}(S^\ominus)$. Hence, $(A(S))^\ominus \subseteq A'_\text{inv}(S^\ominus)$.

On the other hand, $\forall y_* \in A'_\text{inv}(S^\ominus)$, $A'y_* \in S^\oplus$. $\forall s \in S$, we have $\langle \langle y_*, As \rangle \rangle = \langle \langle A'y_*, s \rangle \rangle \geq 0$. Therefore, we have $y_* \in (A(S))^\ominus$ and $A'_\text{inv}(S^\ominus) \subseteq (A(S))^\ominus$.

Hence, $(A(S))^\ominus = A'_\text{inv}(S^\ominus)$. This completes the proof of the proposition. $\square$

Definition 8.45 Let $X$ be a real normed linear space and $P \subseteq X$ be the positive cone. We will defined $P^\ominus \subseteq X^*$ to be the positive cone in the dual.

Proposition 8.46 Let $X$ be a real normed linear space with the positive cone $P$. If $x_0 \in X$ satisfies that $\langle \langle x_*, x_0 \rangle \rangle \geq 0$, $\forall x_* \in P^\ominus$ (or $x_* \triangleright \vartheta$), then $x_0 \in P$.

Proof Suppose $x_0 \notin P$, by Proposition 8.10, $\exists x_* \in X^*$, such that $-\infty < \langle \langle x_*, x_0 \rangle \rangle < \inf_{x \in P} \langle \langle x_*, x \rangle \rangle$. Since $P$ is a cone, then $\inf_{x \in P} \langle \langle x_*, x \rangle \rangle = 0$ (it must be greater than or equal to 0 since otherwise the infimum must be $-\infty$; and it must be less than or equal to 0 since $\vartheta x \in P$). Hence, $x_0 \in P^\ominus$ and $\langle \langle x_*, x_0 \rangle \rangle < 0$. This contradicts the assumption. Therefore, we must have $x_0 \in P$. This completes the proof of the proposition. $\square$

Proposition 8.47 Let $X$ be a real normed linear space with positive cone $P$, $\forall x_* \in X^*$, $\forall x \in X$, we have

(i) $x \triangleright \vartheta$ and $x_* \triangleright \vartheta_*$ implies $\langle \langle x_*, x \rangle \rangle \geq 0$;

(ii) $x \triangleleft \vartheta$ and $x_* \triangleright \vartheta_*$ implies $\langle \langle x_*, x \rangle \rangle \leq 0$;

(iii) $x \triangleright \vartheta$, $x_* \triangleright \vartheta_*$, and $x_* \neq \vartheta_*$ implies $\langle \langle x_*, x \rangle \rangle > 0$;

(iv) $x \triangleright \vartheta$, $x \neq \vartheta$, and $x_* \triangleright \vartheta_*$ implies $\langle \langle x_*, x \rangle \rangle > 0$.

Proof This is straightforward. $\square$

Definition 8.48 Let $X$ be a real vector space, $\Omega \subseteq X$, and $Z$ be a real normed linear space with the positive cone $P \subseteq Z$. A mapping $G : \Omega \rightarrow Z$ is said to be convex if $\Omega$ is convex and $\forall x_1, x_2 \in \Omega$, $\forall \alpha \in [0,1] \subset \mathbb{R}$, we have $G(\alpha x_1 + (1 - \alpha) x_2) \leq \alpha G(x_1) + (1 - \alpha) G(x_2)$.

We note that the convexity of a mapping depends on the definition of the positive cone $P$. 
Proposition 8.49 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be convex, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, and $G : \Omega \to \mathcal{Z}$ be a convex mapping. Then, $\forall z \in \mathcal{Z}$, the set $\Omega_z := \{ x \in \Omega \mid G(x) \preceq z \}$ is convex.

Proof Fix any $z \in \mathcal{Z}$. $\forall x_1, x_2 \in \Omega_Z$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have $G(x_1) \preceq z$ and $G(x_2) \preceq z$. Since $P$ is a convex cone, then, $\alpha G(x_1) + (1 - \alpha)G(x_2) \preceq \alpha z + (1 - \alpha)z = z$. By the convexity of $G$, we have $G(\alpha x_1 + (1 - \alpha) x_2) \preceq \alpha G(x_1) + (1 - \alpha)G(x_2)$. By Proposition 8.41, we have $G(\alpha x_1 + (1 - \alpha) x_2) \preceq z$. Hence, $\alpha x_1 + (1 - \alpha) x_2 \in \Omega_z$. Hence, $\Omega_z$ is convex. This completes the proof of the proposition. \qed

8.8 Lagrange Multipliers

The basic problem to be considered in this section is

$$\mu_0 := \inf_{x \in \Omega, \ G(x) \preceq \vartheta} f(x) \quad (8.11)$$

where $\mathcal{X}$ is a real vector space, $\Omega \subseteq \mathcal{X}$ is a nonempty convex set, $f : \Omega \to \mathbb{R}$ is a convex functional, $\mathcal{Z}$ is a real vector space, $\vartheta \preceq \mathcal{Z}$, $G : \Omega \to \mathcal{Z}$ is a convex mapping.

Toward a solution to the above problem, we consider a class of problems:

$$\Gamma := \{ z \in \mathcal{Z} \mid \exists x \in \Omega \exists \vartheta : G(x) \preceq z \} \quad (8.12a)$$

$$\omega(z) := \inf_{x \in \Omega, \ G(x) \preceq z} f(x); \quad \forall z \in \Gamma \quad (8.12b)$$

where $\omega : \Gamma \to \mathbb{R}_+$ is the primal functional. To guarantee that $\omega$ is real-valued, we make the following assumption.

Assumption 8.50 $\vartheta \preceq \mathcal{Z}$ and (i) $\exists \bar{z} \in \Gamma$ such that $\omega(\bar{z}) \in \mathbb{R}$ or (ii) $\mu := \inf_{x \in \Omega} f(x) > -\infty$ holds.

Fact 8.51 $\Gamma \subseteq \mathcal{Z}$ is convex.

Proof $\forall z_1, z_2 \in \Gamma$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, there exist $x_i \in \Omega$ such that $G(x_i) \preceq z_i$, $i = 1, 2$. By the convexity of $\Omega$, we have $\alpha x_1 + (1 - \alpha) x_2 \in \Omega$. Then, by the convexity of $G$ and Proposition 8.41, $G(\alpha x_1 + (1 - \alpha) x_2) \preceq \alpha G(x_1) + (1 - \alpha)G(x_2) \preceq \alpha z_1 + (1 - \alpha)z_2$. Hence, $\alpha z_1 + (1 - \alpha)z_2 \in \Gamma$. This completes the proof of the fact. \qed

Fact 8.52 Under Assumption 8.50, $\omega : \Gamma \to \mathbb{R}$ is real-valued, convex, and nonincreasing, that is, $\forall z_1, z_2 \in \Gamma$ with $z_1 \preceq z_2$, we have $\omega(z_1) \geq \omega(z_2)$.
We will first show that \( \omega(z) \in \mathbb{R}, \forall z \in \Gamma \). Let (i) in Assumption 8.50 hold. Fix any \( z \in \Gamma \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( z = \bar{z} \); Case 2: \( z \neq \bar{z} \). Case 1: \( z = \bar{z} \). Then, \( \omega(z) = \omega(\bar{z}) \in \mathbb{R} \).

Case 2: \( z \neq \bar{z} \). By the definition of \( \Gamma \) and the fact that \( f \) is real-valued, we have \( \omega(z) < +\infty \). By (i) of Assumption 8.50, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( B_2(\bar{z}, \delta) \cap \mathcal{V}(\Gamma) \subseteq \Gamma \). Let \( \bar{z}_1 := \bar{z} + \frac{\delta}{\|z - z\|}(\bar{z} - z) \in B_2(\bar{z}, \delta) \cap \mathcal{V}(\Gamma) \subseteq \Gamma \) and \( \bar{\alpha} = \frac{\|z - z\|}{\|\bar{z} - z\|} \in (0, 1) \subset \mathbb{R} \). It is easy to verify that \( \bar{z} = \bar{\alpha}\bar{z}_1 + (1 - \bar{\alpha})z \). Then, we have

\[
-\infty < \omega(\bar{z}) = \inf_{\bar{x} \in \Omega, \ G(\bar{x}) \bar{\alpha}\bar{z}_1 + (1 - \bar{\alpha})z} f(\bar{x})
\]

\[
\leq \inf_{\bar{x} = \bar{\alpha}\bar{z}_1 + (1 - \bar{\alpha})x, \bar{x}_1 \in \Omega, \ G(\bar{x}_1) \bar{\alpha}\bar{z}_1, \ x \in \Omega, \ G(x)z} f(\bar{x})
\]

\[
\leq \inf_{\bar{x}_1 \in \Omega, \ G(\bar{x}_1) \bar{\alpha}\bar{z}_1, \ x \in \Omega, \ G(x)z} (\bar{\alpha}f(\bar{x}_1) + (1 - \bar{\alpha})f(x))
\]

\[
= \inf_{\bar{x}_1 \in \Omega, \ G(\bar{x}_1) \bar{\alpha}\bar{z}_1} \bar{\alpha}f(\bar{x}_1) + \inf_{x \in \Omega, \ G(x)z} (1 - \bar{\alpha})f(x)
\]

\[
= \bar{\alpha} \omega(\bar{z}_1) + (1 - \bar{\alpha})\omega(z)
\]

where the second equality follows from Proposition 8.37, and the third equality follows from Proposition 3.81. Then, \( \omega(z) > -\infty \). Hence, \( \omega(z) \in \mathbb{R} \).

In both cases, we have \( \omega(z) \in \mathbb{R} \). Therefore, \( \omega : \Gamma \rightarrow \mathbb{R} \) is real-valued.

Let (ii) in Assumption 8.50 hold. Fix any \( z \in \Gamma \). Then, \( \omega(z) \geq \mu > -\infty \). By the definition of \( \mu \), we have \( \omega(z) < +\infty \). Hence, \( \omega(z) \in \mathbb{R} \).

Thus, under Assumption 8.50, \( \omega : \Gamma \rightarrow \mathbb{R} \) is real-valued.

\( \forall z_1, z_2 \in \Gamma, \forall \alpha \in [0, 1] \subset \mathbb{R} \), we have

\[
\omega(\alpha z_1 + (1 - \alpha)z_2) = \inf_{x \in \Omega, \ G(x) \bar{\alpha}z_1 + (1 - \bar{\alpha})z_2} f(x)
\]

\[
\leq \inf_{x = \alpha x_1 + (1 - \alpha)x_2, \ x_1 \in \Omega, \ G(x_1) \bar{\alpha}z_1, \ x_2 \in \Omega, \ G(x_2) \bar{\alpha}z_2} f(x)
\]

\[
\leq \inf_{x_1 \in \Omega, \ G(x_1) \bar{\alpha}z_1, \ x_2 \in \Omega, \ G(x_2) \bar{\alpha}z_2} (\alpha f(x_1) + (1 - \alpha)f(x_2))
\]

\[
= \inf_{x_1 \in \Omega, \ G(x_1) \bar{\alpha}z_1} \alpha f(x_1) + \inf_{x_2 \in \Omega, \ G(x_2) \bar{\alpha}z_2} (1 - \alpha)f(x_2)
\]

\[
= \alpha \inf_{x_1 \in \Omega, \ G(x_1) \bar{\alpha}z_1} f(x_1) + (1 - \alpha) \inf_{x_2 \in \Omega, \ G(x_2) \bar{\alpha}z_2} f(x_2)
\]

\[
= \alpha \omega(z_1) + (1 - \alpha)\omega(z_2)
\]

where the second equality follows from Proposition 8.37, and the third equality follows from Proposition 3.81. Hence, \( \omega \) is convex.
8.8. LAGRANGE MULTIPLIERS

It is obvious that \( \omega \) is nonincreasing. This completes the proof of the fact. \( \square \)

**Fact 8.53** Under Assumption 8.50, we have \( \Gamma_{\text{conj}} \subseteq \mathcal{P}^\ominus \) and \( -\Gamma_{\text{conj}} \subseteq \mathcal{P}^\oplus \).

**Proof** Since \( \vartheta_{\mathcal{G}} \subseteq \Gamma \) implies that \( \exists x_1 \in \Omega \) such that \( G(x_1) \leq \vartheta_{\mathcal{G}} \), then \( \Gamma \supseteq G(x_1) + P \). \( \forall z_* \in \Gamma_{\text{conj}} \), by Definition 8.27, we have

\[
+\infty > \sup_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle - \omega(z)) \geq \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle - \omega(G(x_1))) \\
\geq \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle - \omega(G(x_1)))
\]

where the third inequality follows from Fact 8.52, and the equality follows from Proposition 8.37. Hence, \( \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle - \omega(G(x_1))) = \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle) \). Since \( P \) is a cone, \( \sup_{z \in G(x_1) + P} (\langle\langle z_*, z \rangle\rangle) = 0 \). This implies that \( z_* \in \mathcal{P}^\ominus \), \( \Gamma_{\text{conj}} \subseteq \mathcal{P}^\ominus \), and \( -\Gamma_{\text{conj}} \subseteq \mathcal{P}^\oplus \). This completes the proof of the fact. \( \square \)

**Fact 8.54** Let Assumption 8.50 hold. Define \( \bar{\omega} : \mathcal{P}^\ominus \to \mathbb{R}_e \) by

\[
\bar{\omega}(z_*) = \sup_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle - \omega(z)); \quad \forall z_* \in \mathcal{P}^\ominus
\]

Then, \( \sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in \mathcal{P}^\ominus} -\bar{\omega}(z_*) \). \( \forall z_* \in \mathcal{P}^\ominus \), we have that

\[
-\bar{\omega}(-z_*) = \inf_{x \in \Omega} (f(x) + \langle\langle z_*, G(x) \rangle\rangle) =: \varphi(z_*)
\]

where \( \varphi : \mathcal{P}^\ominus \to \mathbb{R}_e \) is called the dual functional.

**Proof** Clearly, \( \forall z_* \in \Gamma_{\text{conj}} \), \( \bar{\omega}(z_*) = \omega_{\text{conj}}(z_*) \) and, by the definition of \( \Gamma_{\text{conj}} \), \( \forall z_* \in \mathcal{P}^\ominus \setminus \Gamma_{\text{conj}} \), \( \bar{\omega}(z_*) = +\infty \). Hence, \( \sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in \mathcal{P}^\ominus} -\bar{\omega}(z_*) \) ,

\[
\forall z_* \in \mathcal{P}^\ominus, -\bar{\omega}(z_*) = -\sup_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle - \omega(z)) = \inf_{z \in \Gamma} (\langle\langle z_*, z \rangle\rangle + \omega(z))
\]

where the second equality follows from Proposition 3.81; the fourth equality follows from Proposition 8.37; the fifth and sixth equality follows from
Proposition 8.36; and the eighth equality follows from Proposition 8.37. This completes the proof of the fact.

The desired theory follows by applying either duality results for convex functionals, which are Propositions 8.33 and 8.34. This leads to two different regularity conditions. To apply Proposition 8.33, we assume that

**Assumption 8.55** \( \omega \) is lower semicontinuous at \( \vartheta_Z \).

To apply Proposition 8.34, we assume that

**Assumption 8.56** \( \exists x_1 \in \Omega \) such that \( G(x_1) \not\geq \vartheta_Z \).

#### Proposition 8.57

Let \( \mathcal{X} \) be a real vector space, \( \Omega \subseteq \mathcal{X} \) be nonempty and convex, \( f : \Omega \rightarrow \mathbb{R} \) be a convex functional, \( Z \) be a real normed linear space with the positive cone \( P \subseteq Z \), \( G : \Omega \rightarrow Z \) be a convex mapping, and \( \mu_0 \) be defined as in (8.11). Let Assumptions 8.50 and 8.55 hold. Then,

\[
\mu_0 = \sup_{z_* \in \theta_Z} \inf_{x \in \Omega} (f(x) + \langle \langle z_*, G(x) \rangle \rangle) = \sup_{z_* \in \theta_Z} \varphi(z_*) \tag{8.13}
\]

Furthermore, if the supremum in (8.13) is achieved at \( z_{*0} \in P^\oplus \), that is,

\[
\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) \tag{8.14}
\]

then the following statement holds: the infimum in (8.11) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \not\leq \vartheta_Z \) if, and only if, the infimum in (8.14) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \not\leq \vartheta_Z \) and \( \langle \langle z_{*0}, G(x_0) \rangle \rangle = 0 \).

#### Proof

By Facts 8.51 and 8.52, Assumption 8.55, and Proposition 8.33, we have

\[
\mu_0 = \omega(\vartheta_Z) = \sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*)
\]

Then, by Facts 8.53 and 8.54, we have

\[
\mu_0 = \sup_{z_* \in P^\oplus} -\omega(z_*) = \sup_{z_* \in \theta_Z} -\omega(-z_*) = \sup_{z_* \in \theta_Z} \varphi(z_*)
\]

Therefore, (8.13) holds.

Let the supremum in (8.13) be achieved at \( z_{*0} \in P^\oplus \), that is, (8.14) holds.

If the infimum in (8.11) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \not\leq \vartheta_Z \), then, we have

\[
\mu_0 = f(x_0) \geq f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle \geq \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) = \mu_0.
\]

Hence, the infimum in (8.14) is achieved at \( x_0 \) and \( \langle \langle z_{*0}, G(x_0) \rangle \rangle = 0 \).

On the other hand, if the infimum in (8.14) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \not\leq \vartheta_Z \) and \( \langle \langle z_{*0}, G(x_0) \rangle \rangle = 0 \), then \( \mu_0 = f(x_0) + \langle \langle z_{*0}, G(x_0) \rangle \rangle = f(x_0) \). Hence, the infimum in (8.11) is achieved at \( x_0 \).

This completes the proof of the proposition. \[\square\]
8.8. LAGRANGE MULTIPLIERS

Proposition 8.58 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty and convex, $f : \Omega \to \mathbb{R}$ be a convex functional, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $P^\circ \neq \emptyset$, $G : \Omega \to \mathcal{Z}$ be a convex mapping. Let Assumptions 8.50 and 8.56 hold. Then,

$$
\mu_0 = \max_{z_* \in \bar{\theta}_Z} \inf_{x \in \Omega} (f(x) + \langle \langle z_*, G(x) \rangle \rangle) = \max_{z_* \in \bar{\theta}_Z} \varphi(z_*)
$$

(8.15)

where the maximum is achieved at $z_0 \in \mathcal{P}^\circ$, that is,

$$
\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_0, G(x) \rangle \rangle)
$$

(8.16)

Furthermore, the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \theta_\mathcal{Z}$ if, and only if, the infimum in (8.16) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \theta_\mathcal{Z}$ and $\langle \langle z_0, G(x_0) \rangle \rangle = 0$.

Proof By Assumption 8.56, $G(x_1) \leq \theta_\mathcal{Z}$. Then, $-G(x_1) \in P^\circ$ and $G(x_1) + P \subseteq \Gamma$. By Proposition 7.16, $\Gamma^\circ \supseteq G(x_1) + P^\circ$ and $\theta_\mathcal{Z} \in \Gamma^\circ$. $\forall z \in G(x_1) + P$, we have $z \leq G(x_1)$ and, by Fact 8.52, $\omega(z) \leq \omega(G(x_1))$. Since $G(x_1) \leq \theta_\mathcal{Z}$, then $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $B_\mathcal{Z}(\theta_\mathcal{Z}, \delta) \subseteq G(x_1) + P$. Take $r_0 := \omega(G(x_1)) + \delta \in \mathbb{R}$. It is easy to check that $B_{\mathbb{R}^\mathcal{Z}}((r_0, \theta_\mathcal{Z}), \delta) \subseteq [\omega, \Gamma]$. Hence, $(r_0, \theta_\mathcal{Z}) \in [\omega, \Gamma]^\circ$. By Proposition 8.22, $\omega$ is continuous at $\theta_\mathcal{Z}$. By Facts 8.51 and 8.52 and Proposition 8.34, we have

$$
\mu_0 = \omega(\theta_\mathcal{Z}) = \max_{z_* \in \Gamma^\circ} -\omega(\phi^\circ(z_*)
$$

Then, by Facts 8.53 and 8.54, we have

$$
\mu_0 = \max_{z_* \in \mathcal{P}^\circ} -\phi^\circ(z_*) = \max_{z_* \in \bar{\theta}_Z} -\phi^\circ(z_*) = \max_{z_* \in \bar{\theta}_Z} \varphi(z_*)
$$

Therefore, (8.15) holds and the maximum is achieved at $z_0 \in \mathcal{P}^\circ$.

If the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \theta_\mathcal{Z}$, then, we have $\mu_0 = f(x_0) \geq f(x_0) + \langle \langle z_0, G(x_0) \rangle \rangle = \inf_{x \in \Omega} (f(x) + \langle \langle z_0, G(x) \rangle \rangle) = \mu_0$. Hence, the infimum in (8.16) is achieved at $x_0$ and $\langle \langle z_0, G(x_0) \rangle \rangle = 0$.

On the other hand, if the infimum in (8.16) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \theta_\mathcal{Z}$ and $\langle \langle z_0, G(x_0) \rangle \rangle = 0$, then $\mu_0 = f(x_0) + \langle \langle z_0, G(x_0) \rangle \rangle = f(x_0)$. Hence, the infimum in (8.11) is achieved at $x_0$.

This completes the proof of the proposition. □

In the Propositions 8.57 and 8.58, $z_0$ is called the Lagrange multiplier. Assumption 8.50 guarantees that the primal functional is real-valued and convex. Assumption 8.56 guarantees the existence of a Lagrange multiplier. This assumption is restrictive. On the other hand, Assumption 8.55 guarantees the duality but not the existence of a Lagrange multiplier. This condition is more relaxed.
Corollary 8.59 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty and convex, $f : \Omega \to \mathbb{R}$ be a convex functional, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $P^\circ \neq \emptyset$, $G : \Omega \to \mathcal{Z}$ be a convex mapping. Let Assumptions 8.50 and 8.56 hold and $x_0 \in \Omega$ with $G(x_0) \overset{\circ}{=} \partial \mathcal{Z}$ achieves the infimum in (8.11). Then, there exists $z_{a_0} \in \mathcal{Z}^*$ with $z_{a_0} \overset{\circ}{=} \partial \mathcal{Z}^*$ such that the Lagrangian $L : \Omega \times P^\circ \to \mathbb{R}$ defined by

$$L(x, z_*) := f(x) + \langle \langle z_*, G(x) \rangle \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

admits a saddle point at $(x_0, z_{a_0})$, i.e.

$$L(x_0, z_*) \leq L(x, z_{a_0}) \leq L(x, z_0); \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

Proof By Proposition 8.58, there exists $z_{a_0} \in \mathcal{Z}^*$ with $z_{a_0} \overset{\circ}{=} \partial \mathcal{Z}^*$ such that $L(x_0, z_{a_0}) \leq L(x, z_{a_0})$, $\forall x \in \Omega$ and $\langle \langle z_{a_0}, G(x_0) \rangle \rangle = 0$. Then, $\forall z_* \in P^\circ$, we have $L(x_0, z_*) - L(x, z_{a_0}) = \langle \langle z_*, G(x_0) \rangle \rangle - \langle \langle z_{a_0}, G(x_0) \rangle \rangle = \langle \langle z_*, G(x_0) \rangle \rangle \leq 0$. Hence, the saddle point condition holds. This completes the proof of the corollary.

Next, we present two sufficiency results on Lagrange multipliers.

Proposition 8.60 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty, $f : \Omega \to \mathbb{R}$, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \to \mathcal{Z}$.

Assume that there exist $x_0 \in \Omega$ and $z_{a_0} \in \mathcal{Z}^*$ with $z_{a_0} \overset{\circ}{=} \partial \mathcal{Z}^*$ such that

$$f(x_0) + \langle \langle z_{a_0}, G(x_0) \rangle \rangle \leq f(x) + \langle \langle z_{a_0}, G(x) \rangle \rangle; \quad \forall x \in \Omega$$

Then,

$$f(x_0) = \inf_{x \in \Omega, G(x) \overset{\circ}{=} G(x_0)} f(x)$$

Proof $\forall x \in \Omega$ with $G(x) \overset{\circ}{=} G(x_0)$, we have $\langle \langle z_{a_0}, G(x) \rangle \rangle \leq \langle \langle z_{a_0}, G(x_0) \rangle \rangle$, since $z_{a_0} \in P^\circ$. By the assumption of the proposition, $f(x_0) + \langle \langle z_{a_0}, G(x_0) \rangle \rangle \leq f(x) + \langle \langle z_{a_0}, G(x) \rangle \rangle$. Then, $f(x_0) \leq f(x)$. This completes the proof of the proposition.

Proposition 8.61 Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be nonempty, $f : \Omega \to \mathbb{R}$, $\mathcal{Z}$ be a real normed linear space with the positive cone $P \subseteq \mathcal{Z}$, $G : \Omega \to \mathcal{Z}$.

Assume that there exist $x_0 \in \Omega$ and $z_{a_0} \in \mathcal{Z}^*$ with $z_{a_0} \overset{\circ}{=} \partial \mathcal{Z}^*$ such that the Lagrangian $L : \Omega \times P^\circ \to \mathbb{R}$ given by

$$L(x, z_*) = f(x) + \langle \langle z_*, G(x) \rangle \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

admits a saddle point at $(x_0, z_{a_0})$, i.e.

$$L(x_0, z_*) \leq L(x, z_{a_0}) \leq L(x, z_0); \quad \forall x \in \Omega, \forall z_* \in P^\circ$$
Then, \( G(x_0) \leq \partial Z \) and
\[
f(x_0) = L(x_0, z_{*0}) = \inf_{x \in \Omega, G(x) \geq \partial z} f(x)
\]

**Proof**  
By the first inequality in the saddle-point condition, we have
\[
\langle z_{*0}, G(x_0) \rangle \leq \langle z_{*0}, G(x_0) \rangle \leq 0.
\]
\[
\forall \forall \in P^\oplus, \text{ we have } z_{*0} \leq \partial z \text{ and } z_{*0} + z_{*0} \geq \partial z.
\]

Next, we present a result on the sensitivity of the infimization problem.

**Proposition 8.62**  
Let \( X \) be a real vector space, \( \Omega \subseteq X \) be nonempty, \( f : \Omega \rightarrow \mathbb{R}, Z \) be a real normed linear space with the positive cone \( P \subseteq Z \), \( G : \Omega \rightarrow Z \).

Let \( z_i \in Z, \mu_i = \inf_{x \in \Omega, G(x) \leq z_i} f(x), i = 0, 1 \). Let \( z_{*0} \in P^\oplus \subseteq Z^* \) be the Lagrange multiplier associated with \( \mu_0 \), that is
\[
\mu_0 = \inf_{x \in \Omega} (f(x) + \langle z_{*0}, G(x) - z_0 \rangle)
\]

Assume that \( \mu_0 \in \mathbb{R} \). Then, we have
\[
\mu_1 - \mu_0 \geq \langle -z_{*0}, z_1 - z_0 \rangle
\]

**Proof**  
\( \forall \in \Omega \) with \( G(x) \leq z_1 \), we have \( \mu_0 \leq f(x) + \langle z_{*0}, G(x) - z_0 \rangle \), which implies that
\[
f(x) \geq \mu_0 - \langle z_{*0}, G(x) - z_0 \rangle \geq \mu_0 - \langle z_{*0}, z_1 - z_0 \rangle
\]
Hence, \( \mu_1 \geq \mu_0 + \langle -z_{*0}, z_1 - z_0 \rangle \). This completes the proof of the proposition.  \( \square \)
Chapter 9

Differentiation in Banach Spaces

In this chapter, we are going to develop the concept of derivative in normed linear spaces.

9.1 Fundamental Notion

Definition 9.1 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $D \subseteq \mathcal{X}$, and $x_0 \in D$. $u \in \mathcal{X}$ is said to be an admissible deviation in $D$ at $x_0$ if

$$\forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \text{ we have } \{x_0 + r\bar{u} \mid r \in (0, \varepsilon) \subset \mathbb{R}, \bar{u} \in B(u, \varepsilon)\} \cap D \neq \emptyset.$$  

Let $A_D(x_0)$ be the set of admissible deviations in $D$ at $x_0$.

Clearly, if $x_0 \in D^\circ$, then $A_D(x_0) = \mathcal{X}$. Another fact is that when $D_1 \subseteq D_2 \subseteq \mathcal{X}$ and $x_0 \in D_1$, then $A_{D_1}(x_0) \subseteq A_{D_2}(x_0)$. Yet another fact is that when $x_0 \in D_1 \cap D_2 \subset \mathcal{X}$ and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap B(x_0, \delta) = D_2 \cap B(x_0, \delta)$, then $A_{D_1}(x_0) = A_{D_2}(x_0)$.

Proposition 9.2 Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$, $D \subseteq \mathcal{X}$, and $x_0 \in \overline{D}$. Then, $A_D(x_0) \subseteq \mathcal{X}$ is a closed cone.

Proof Clearly, if $x_0 \in D$, then $\vartheta \in A_D(x_0)$. On the other hand, if $x_0 \in \overline{D} \setminus D$, then, by Proposition 4.13, $x_0$ is an accumulation point of $D$. $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \bar{x} \in (B(x_0, \varepsilon^2) \cap D) \setminus \{x_0\}$. Let $v := \bar{x} - x_0$. Then, $\varepsilon := \sqrt{\|v\|} \in (0, \varepsilon) \subset \mathbb{R}$. Let $\bar{v} := \varepsilon^{-1}v$. Then, $\bar{v} \in B(\vartheta, \varepsilon)$ and $\bar{x} = x_0 + \delta \bar{v} \in \{x_0 + r\bar{u} \mid r \in (0, \varepsilon) \subset \mathbb{R}, \bar{u} \in B(\vartheta, \varepsilon)\} \cap D \neq \emptyset$. Hence, $\vartheta \in A_D(x_0)$. Therefore, $\vartheta \in A_D(x_0)$ if $x_0 \in \overline{D}$.

$\forall u \in A_D(x_0)$, $\forall \alpha \in [0, \infty) \subset \mathbb{R}$, we will show that $\alpha u \in A_D(x_0)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\alpha = 0$; Case 2: $\alpha > 0$. Case 1: $\alpha = 0$. Then, $\alpha u = \vartheta \in A_D(x_0)$. Case 2: $\alpha > 0$. $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, let $\bar{\varepsilon} = \min \{\alpha \varepsilon, \varepsilon/\alpha\} \in (0, \infty) \subset \mathbb{R}$.
By \( u \in A_D(x_0) \), \( \exists \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \; \bar{u} \in B(u, \epsilon) \} \cap D \). Hence, \( \bar{x} \in D \) and \( \bar{x} = x_0 + r\bar{u} = x_0 + (r/\alpha)(\alpha\bar{u}) \) with \( r \in (0, \epsilon) \subset \mathbb{R} \) and \( \bar{u} \in B(u, \epsilon) \). Then, \( r/\alpha \in (0, \epsilon) \subset \mathbb{R} \) and \( \alpha\bar{u} \in B(\alpha u, \epsilon) \). Hence, \( \bar{x} \in \{ x_0 + r\bar{u} \mid \bar{r} \in (0, \epsilon) \subset \mathbb{R}, \; \bar{u} \in B(\alpha u, \epsilon) \} \cap D \neq \emptyset \). This implies that \( \alpha u \in A_D(x_0) \). Hence, \( A_D(x_0) \) is a cone.

\( \forall u \in A_D(x_0), \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by Proposition 3.3, \( \exists u_1 \in A_D(x_0) \cap B(u, \epsilon/2) \). By \( u_1 \in A_D(x_0) \), \( \exists \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \epsilon/2) \subset \mathbb{R}, \; \bar{u} \in B(u_1, \epsilon/2) \} \cap D \). Then, \( \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \; \bar{u} \in B(u, \epsilon) \} \cap D \neq \emptyset \). Hence, \( u \in A_D(x_0) \). Then, \( A_D(x_0) \subset A_D(x_0) \) and \( A_D(x_0) \) is closed. This completes the proof of the proposition.

**Definition 9.3** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), and \( x_0 \in D \). Assume that \( \text{span}(A_D(x_0)) = X \). Let \( L \in B(X, Y) \) be such that \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall x \in D \cap B_X(x_0, \delta) \), we have

\[ \| f(x) - f(x_0) - L(x - x_0) \| \leq \epsilon \| x - x_0 \| \]

Then, \( L \) is called the (Fréchet) derivative of \( f \) at \( x_0 \) and denoted by \( f^{(1)}(x_0) \) or \( Df(x_0) \). When \( L \) exists, we will say that \( f \) is (Fréchet) differentiable at \( x_0 \). \( Df \) or \( f^{(1)} \) will denote the \( B(X, Y) \)-valued function whose domain of definition is \( \text{dom}(f^{(1)}) := \{ x \in D \mid Df(x) \in B(X, Y) \text{ exists} \} \). If \( f \) is differentiable at \( x_0 \), \( \forall x_0 \in D \), we say \( f \) is (Fréchet) differentiable. In this case, \( Df : D \to B(X, Y) \) or \( f^{(1)} : D \to B(X, Y) \).

Clearly, when \( X = Y = \mathbb{R} \) and \( D = [a, b] \subseteq X \) with \( a < b \), then \( Df(t) \) is simply the derivative of \( f \) at \( t \in [a, b] \), as we know before.

**Definition 9.4** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), \( x_0 \in D \), and \( u \in A_D(x_0) \). Let \( M_D(x_0) := \{ (r, \bar{u}) \in \mathbb{R} \times X \mid r \in (0, +\infty) \subset \mathbb{R}, \; \bar{u} \in X, \; x_0 + r\bar{u} \in D \} \) and \( \hat{g} : M_D(x_0) \to Y \) be given by \( \hat{g}(r, \bar{u}) = r^{-1}(f(x_0 + r\bar{u}) - f(x_0)) \), \( \forall (r, \bar{u}) \in M_D(x_0) \). Clearly, \( (0, u) \) is an accumulation point of \( M_D(x_0) \) since \( u \in A_D(x_0) \). The directional derivative of \( f \) at \( x_0 \) along \( u \), denoted by \( Df(x_0; u) \), is the limit \( \lim_{(r, \bar{u}) \to (0, u)} \hat{g}(r, \bar{u}) \), when it exists.

Clearly, the directional derivative is unique when it exists, since \( Y \) is Hausdorff.

**Proposition 9.5** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), \( x_0 \in D \), \( \text{span}(A_D(x_0)) = X \), \( L \in B(X, Y) \) be the Fréchet derivative of \( f \) at \( x_0 \), and \( u \in A_D(x_0) \). Then, \( Df(x_0; u) = Lu \).

**Proof** \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), let \( \bar{\epsilon} = \epsilon/(2 + 2\|u\|) \in (0, \infty) \subset \mathbb{R} \). By \( Df(x_0) = L \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \| f(x) - f(x_0) - L(x - x_0) \| \leq \bar{\epsilon} \| x - x_0 \| \), \( \forall x \in B_X(x_0, \delta) \cap D \). Let \( M_D(x_0) \) and \( \hat{g} : M_D(x_0) \to Y \) be as defined in Definition 9.4. Let \( \delta = \min \{ \epsilon/(2 + 2\|u\|), \delta/(1 + \|u\|), 1 \} \in \mathbb{R} \).
\( (0, \infty) \subset \mathbb{R} \). \( \forall (r, \bar{u}) \in \left( M_D(x_0) \cap B_{X,Y} \right) \setminus \{(0,u)\} \), we have \( \bar{x} := x_0 + r\bar{u} \in D \cap B_{X,Y} \). This implies that
\[
\| \bar{g}(r, \bar{u}) - L\bar{u} \| = \| f(x_0 + r\bar{u}) - f(x_0) - rL\bar{u} \| / r
\]
\[
= \| f(\bar{x}) - f(x_0) - L(\bar{x} - x_0) + rL(\bar{u} - u) \| / r
\]
\[
\leq \| f(\bar{x}) - f(x_0) - L(\bar{x} - x_0) \| / r + \| L(\bar{u} - u) \|
\]
\[
\leq \varepsilon\| \bar{x} - x_0 \| / r + \| L \| \| (\bar{u} - u) \| = \varepsilon\| \bar{u} \| + \| L \| \| \bar{u} - u \|
\]
\[
\leq \varepsilon(\| u \| + \| \bar{u} - u \|) + \| L \| \| \bar{u} - u \| \leq \varepsilon(\| u \| + \delta) + \| L \| \delta
\]
\[
< \varepsilon
\]
where the second inequality follows from Proposition 7.64. Hence, we have \( Df(x_0;u) = \lim_{(r,\bar{u}) \to (0,u)} \bar{g}(r, \bar{u}) = Lu \). This completes the proof of the proposition. \( \Box \)

**Proposition 9.6** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), \( x_0 \in D \), and \( \text{span}(A_D(x_0)) = X \). Then, \( f \) has at most one derivative at \( x_0 \).

**Proof** We will prove the result by an argument of contradiction. Suppose, \( \exists L_1, L_2 \in B(X,Y) \) with \( L_1 \neq L_2 \) such that \( L_1 \) and \( L_2 \) are derivatives of \( f \) at \( x_0 \). By \( \text{span}(A_D(x_0)) = X \) and Proposition 3.56, \( \exists \bar{u} \in \text{span}(A_D(x_0)) \) such that \( L_1\bar{u} \neq L_2\bar{u} \). Since \( L_1 \) and \( L_2 \) are linear, then \( \exists u_0 \in A_D(x_0) \) such that \( L_1u_0 \neq L_2u_0 \). By Proposition 9.5, \( Df(x_0;u_0) = L_1u_0 = L_2u_0 \). This is a contradiction. Hence, the result holds. This completes the proof of the proposition. \( \Box \)

**Proposition 9.7** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), \( x_0 \in D \), and \( \text{span}(A_D(x_0)) = X \). Assume that \( Df(x_0) \in B(X,Y) \) exists, then \( f \) is continuous at \( x_0 \).

**Proof** This is straightforward, and is therefore omitted. \( \Box \)

**Definition 9.8** Let \( X \) be a set, \( Y \) and \( Z \) be normed linear spaces over \( K \), \( D \subseteq X \times Y \), \( f : D \to Z \), and \( (x_0,y_0) \in D \). \( f \) is said to be partial (Fréchet) differentiable with respect to \( y \) at \( (x_0, y_0) \) if \( g : D_1 \to Z \) given by \( g(y) = f(x_0, y) \), \( \forall y \in D_1 \), where \( D_1 := \{ y \in Y \mid (x_0, y) \in D \} \), is differentiable at \( y_0 \). \( Dg(y_0) \in B(Y,Z) \) is called the partial derivative of \( f \) with respect to \( y \) at \( (x_0, y_0) \) and is denoted by \( \frac{\partial f}{\partial y}(x_0, y_0) \). \( \frac{\partial f}{\partial y} \) will denote the \( B(Y,Z) \)-valued function whose domain of definition is \( \text{dom} \left( \frac{\partial f}{\partial y} \right) := \{(x,y) \in D \mid \frac{\partial f}{\partial y}(x,y) \in B(Y,Z) \text{ exists} \} \). When \( \text{dom} \left( \frac{\partial f}{\partial y} \right) = D \), we say that \( f \) is partial differentiable with respect to \( y \).

For notational simplicity, we will adopt the “matrix” notation for our later developments. Let \( X_1, \ldots, X_n \) be normed linear spaces over \( K \), \( n \in \mathbb{N} \).
For a vector \((x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n\), we will use the notation of a column vector
\[
\begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix}
\]
to alternatively denote it. Let \(Y_1, \ldots, Y_m\) be normed linear spaces over \(K\), \(m \in \mathbb{N}\). \(\forall L \in B (X_1 \times \cdots \times X_n, Y_1 \times \cdots \times Y_m)\) can be equivalently written as
\[
\begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix}
\]
where \(L_{ij} \in B (X_j, Y_i)\). Then, \((y_1, \ldots, y_m) = L(x_1, \ldots, x_n)\) can be equivalently expressed as
\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_m
\end{bmatrix} = \begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix} \begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} = \begin{bmatrix}
\sum_{j=1}^{n} L_{1j} x_j \\
\vdots \\
\sum_{j=1}^{n} L_{nj} x_j
\end{bmatrix}
\]
Let \(Z_1, \ldots, Z_p\) be normed linear spaces over \(K\), \(p \in \mathbb{N}\), and \(\bar{L} \in B (Y_1 \times \cdots \times Y_m, Z_1 \times \cdots \times Z_p)\) be given by
\[
\begin{bmatrix}
\bar{L}_{11} & \cdots & \bar{L}_{1m} \\
\vdots & & \vdots \\
\bar{L}_{p1} & \cdots & \bar{L}_{pm}
\end{bmatrix}
\]
Then, \(\bar{L}L \in B (X_1 \times \cdots \times X_n, Z_1 \times \cdots \times Z_p)\) and
\[
\bar{L}L = \begin{bmatrix}
\bar{L}_{11} & \cdots & \bar{L}_{1m} \\
\vdots & & \vdots \\
\bar{L}_{p1} & \cdots & \bar{L}_{pm}
\end{bmatrix} \begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix}
\begin{bmatrix}
\sum_{j=1}^{m} \bar{L}_{1j} L_{j1} & \cdots & \sum_{j=1}^{m} \bar{L}_{1j} L_{jn} \\
\vdots & & \vdots \\
\sum_{j=1}^{m} \bar{L}_{pj} L_{j1} & \cdots & \sum_{j=1}^{m} \bar{L}_{pj} L_{jn}
\end{bmatrix}
\]
It is easy to show that the adjoint of \(L\) is
\[
L^* = \begin{bmatrix}
L'_{11} & \cdots & L'_{1m} \\
\vdots & & \vdots \\
L'_{mn} & \cdots & L'_{mn}
\end{bmatrix}
\]

**Proposition 9.9** Let \(X, Y,\) and \(Z\) be normed linear spaces over \(K\), \(D \subseteq X \times Y, f : D \to Z\), and \((x_0, y_0) \in D\). Assume \(Df(x_0, y_0) \in B (X \times Y, Z)\) exists. Let \(D_{x_0} := \{ y \in Y \mid (x_0, y) \in D \}\). Assume that \(\text{span} \left( A_{D_{x_0}} (y_0) \right) =\)
9.2. THE DERIVATIVES OF SOME COMMON FUNCTIONS

Then, $f$ is partial differentiable with respect to $y$ at $(x_0,y_0)$ and $\partial_f(x_0,y_0) \in B(\mathbb{Y},\mathbb{Z})$ is given by $\partial_f(x_0,y_0)(k) = Df(x_0,y_0)(\partial_x,k)$, $\forall k \in \mathbb{Y}$.

Let $D_{y_0} := \{ x \in \mathbb{X} \mid (x,y_0) \in D \}$. Assume that $\text{span} \{ A_{D_{y_0}}(x_0) \} = \mathbb{X}$. By symmetry, $f$ is partial differentiable with respect to $x$ at $(x_0,y_0)$ and $\partial_f(x_0,y_0) \in B(\mathbb{X},\mathbb{Z})$ is given by $\partial_f(x_0,y_0)(h) = Df(x_0,y_0)(h,\partial_y)$, $\forall h \in \mathbb{X}$.

Hence, we have $Df(x_0,y_0) = \left[ \begin{array}{c} \partial_f(x_0,y_0) \\ \partial_f(x_0,y_0) \end{array} \right]$ in “matrix” notation.

**Proof** Let $\tilde{g} : D_{x_0} \to \mathbb{Z}$ be given by $\tilde{g}(y) = f(x_0,y)$, $\forall y \in D_{x_0}$. Clearly $y_0 \in D_{x_0}$. By the fact that $f$ is differentiable at $(x_0,y_0)$, then $\forall \epsilon \in (0,\infty) \subset \mathbb{R}$, $\exists \delta \in (0,\infty) \subset \mathbb{R}$, $\forall (x,y) \in D \cap \mathcal{B}_{x,y}(\mathbb{X},\mathbb{Z})$, we have $\lVert f(x,y) - f(x_0,y_0) - Df(x_0,y_0)(x-x_0,y-y_0) \rVert \leq \epsilon \lVert (x-x_0,y-y_0) \rVert$.

Then, $\forall y \in D_{y_0} \cap \mathcal{B}_y(y_0,\delta)$, we have $(x,y) \in D \cap \mathcal{B}_{x,y}(\mathbb{X},\mathbb{Z})$ and $\epsilon \lVert y - y_0 \rVert \geq \lVert \tilde{g}(y) - \tilde{g}(y_0) \rVert - \lVert f(x_0,y_0) - Df(x_0,y_0)(\partial_x,y-y_0) \rVert = \lVert \tilde{g}(y) - \tilde{g}(y_0) - L(y - y_0) \rVert$. This implies that $\text{span} \{ A_{D_{y_0}}(x_0) \} = \mathbb{X}$.

By symmetry, $f$ is partial differentiable with respect to $x$ at $(x_0,y_0)$ and $\partial_f(x_0,y_0) \in B(\mathbb{X},\mathbb{Z})$ is given by $\partial_f(x_0,y_0)(h) = Df(x_0,y_0)(h,\partial_y)$, $\forall h \in \mathbb{X}$.

Note that, $\forall h \in \mathbb{X}$ and $\forall k \in \mathbb{Y}$,

$$Df(x_0,y_0)(h,k) = \frac{\partial f}{\partial x}(x_0,y_0)(h) + \frac{\partial f}{\partial y}(x_0,y_0)(k)$$

Then,

$$Df(x_0,y_0) \left[ \begin{array}{c} h \\ k \end{array} \right] = \left[ \begin{array}{c} \frac{\partial f}{\partial x}(x_0,y_0) \\ \frac{\partial f}{\partial y}(x_0,y_0) \end{array} \right] \left[ \begin{array}{c} h \\ k \end{array} \right]$$

Hence, the desired “matrix” notation follows. This completes the proof of the proposition. \[
\square
\]

9.2 The Derivatives of Some Common Functions

**Proposition 9.10** Let $\mathbb{X}$ and $\mathbb{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathbb{X}$, $f : D \to \mathbb{Y}$, $x_0 \in D$, and $\text{span} \{ A_D(x_0) \} = \mathbb{X}$. Assume that $\exists \delta_0 \in (0,\infty) \subset \mathbb{R}$,
Proposition 9.11 Let $X$ and $Y$ be normed linear spaces over $K$, $x_0 \in D_2 \subseteq D_1 \subseteq X$, and $f : D_1 \rightarrow Y$. If $f$ is Fréchet differentiable at $x_0$ and $Dg(x_0) = Df(x_0)$. On the other hand, if $g$ is Fréchet differentiable at $x_0$ and $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D_1 \cap B_X(x_0, \delta_0) = D_2 \cap B_X(x_0, \delta_0)$, then $f$ is Fréchet differentiable at $x_0$ and $Df(x_0) = Dg(x_0)$.

Proof This is straightforward, and is therefore omitted.

Proposition 9.12 Let $X$ be a normed linear space over $K$, and $f : X \rightarrow X$ be given by $f = \text{id}_X$, that is $f(x) = x$, $\forall x \in X$. Then, $f$ is Fréchet differentiable and $Df(x)(h) = h$, $\forall x \in X$ and $\forall h \in X$. Then, $Df(x) = \text{id}_X$, $\forall x \in X$.

Proof This is straightforward, and is therefore omitted.

Proposition 9.13 Let $X$ and $Y$ be normed linear spaces over $K$, and $f : X \times Y \rightarrow X$ be given by $f = \pi_X$, that is $f(x, y) = x$, $\forall (x, y) \in X \times Y$. Then, $f$ is Fréchet differentiable and $Df(x, y)(h, k) = h$, $\forall (x, y) \in X \times Y$ and $\forall (h, k) \in X \times Y$. In “matrix” notation, we have $Df(x, y) = \begin{bmatrix} \text{id}_X & \vartheta_B(y, x) \end{bmatrix}$, $\forall (x, y) \in X \times Y$.

Proof This is straightforward, and is therefore omitted.

Proposition 9.14 Let $X$ be a normed linear space over $K$, $f : X \times X \rightarrow X$ be given by $f(x_1, x_2) = x_1 + x_2$, $\forall (x_1, x_2) \in X \times X$. Then, $f$ is Fréchet differentiable and $Df : X \times X \rightarrow B(X \times X)$ is given by $\forall (x_1, x_2) \in X \times X$, $\forall (h_1, h_2) \in X \times X$, $Df(x_1, x_2)(h_1, h_2) = h_1 + h_2$. In “matrix” notation, we have $Df(x_1, x_2) = \begin{bmatrix} \text{id}_X & \text{id}_X \end{bmatrix}$.

Proof $\forall (x_1, x_2) \in X \times X$, let $L : X \times X \rightarrow X$ be given by $L(h_1, h_2) = h_1 + h_2$; $\forall (h_1, h_2) \in X \times X$. Clearly, $L$ is a linear operator. Note that

$$
\| L \| = \sup_{(h_1, h_2) \in X \times X, \| (h_1, h_2) \| \leq 1} \| L(h_1, h_2) \|
\leq \sup_{(h_1, h_2) \in X \times X, \| (h_1, h_2) \| \leq 1} \| h_1 \| + \| h_2 \|
\leq \sup_{(h_1, h_2) \in X \times X, \| (h_1, h_2) \| \leq 1} \sqrt{2} \left( \| h_1 \|^2 + \| h_2 \|^2 \right)^{1/2} \leq \sqrt{2}
$$

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(X \times X)$.
Clearly, $A_{X 	imes X} (x_1, x_2) = X 	imes X$ since $(x_1, x_2) \in (X \times X)^\circ$. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), set $\delta = 1 \in \mathbb{R}$, \( \forall (\bar{x}_1, \bar{x}_2) \in (X \times X) \cap B_{X 	imes X}((x_1, x_2), \delta) \), we have 
\[ \| f(\bar{x}_1, \bar{x}_2) - f(x_1, x_2) - L(\bar{x}_1 - x_1, \bar{x}_2 - x_2) \| = 0 \leq \epsilon \| (\bar{x}_1 - x_1, \bar{x}_2 - x_2) \|. \] 
By Definition 9.3, $Df(x_1, x_2) = L$. This completes the proof of the proposition.

**Proposition 9.15** Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $f_1 : D \to Y$, $f_2 : D \to Y$, $x_0 \in D$, $\alpha_1, \alpha_2 \in K$, and $g : D \to Y$ be given by $g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$, $\forall x \in D$. Assume that $f_1^{(1)}(x_0)$ and $f_2^{(1)}(x_0)$ exist. Then, $g$ is Fréchet differentiable at $x_0$ and $g^{(1)}(x_0) = \alpha_1 f_1^{(1)}(x_0) + \alpha_2 f_2^{(1)}(x_0)$.

**Proof** Define $L := \alpha_1 f_1^{(1)}(x_0) + \alpha_2 f_2^{(1)}(x_0) \in B(X, Y)$. By assumption, $\text{span}(A_D(x_0)) = X$. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the differentiability of $f_1$ at $x_0$, \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that \( \forall x \in D \cap B_X(x_0, \delta_1) \), we have 
\[ \| f_1(x) - f_1(x_0) - f_1^{(1)}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|. \] 
By the differentiability of $f_2$ at $x_0$, \( \exists \delta_2 \in (0, \infty) \subset \mathbb{R} \) such that \( \forall x \in D \cap B_X(x_0, \delta_2) \), we have 
\[ \| f_2(x) - f_2(x_0) - f_2^{(1)}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|. \] 
Let $\delta := \min \{ \delta_1, \delta_2 \} > 0$. \( \forall x \in D \cap B_X(x_0, \delta) \), we have 
\[ \| g(x) - g(x_0) - L(x - x_0) \| = \| \alpha_1 f_1(x) - f_1(x_0) - f_1^{(1)}(x_0)(x - x_0) \| + \| \alpha_2 f_2(x) - f_2(x_0) - f_2^{(1)}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|. \]

Hence, $g^{(1)}(x_0) = L$. This completes the proof of the proposition.

**Proposition 9.16** Let $X$ be a normed linear space over $K$, $f : K \times X \to X$ be given by $f(\alpha, x) = \alpha x$, $\forall (\alpha, x) \in K \times X$. Then, $f$ is Fréchet differentiable and $Df : K \times X \to B(K \times X, X)$ is given by $\forall (\alpha, x) \in K \times X$, $\forall (d, h) \in K \times X$, $Df(\alpha, x)(d, h) = \alpha h + dx$. Thus, $Df(\alpha, x) = \begin{bmatrix} x & \alpha \text{id}_X \end{bmatrix}$ in “matrix” notation.

**Proof** \( \forall (\alpha, x) \in K \times X \), let $L : K \times X \to X$ be given by $L(d, h) = \alpha h + dx$, $\forall (d, h) \in K \times X$. Clearly, $L$ is a linear operator. Note that 
\[ \| L \| = \sup_{(d, h) \in K \times X, \|(d, h)\| \leq 1} \| L(d, h) \| \leq \sup_{(d, h) \in K \times X, \|(d, h)\| \leq 1} \| \alpha \| \| h \| + |d| \| x \| \leq \sup_{(\alpha, x) \in K \times X, \|(\alpha, x)\| \leq 1} \left( |\alpha|^2 + \| x \|^2 \right)^{1/2} \left( |d|^2 + \| h \|^2 \right)^{1/2} \leq \| (\alpha, x) \| < +\infty \]

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(K \times X, X)$. 
Clearly, \( A_{\mathbb{K} \times \mathcal{X}} (\alpha, x) = \mathbb{K} \times \mathcal{X} \) since \((\alpha, x) \in (\mathbb{K} \times \mathcal{X})^0\). ∀\( \epsilon \in (0, \infty) \subseteq \mathbb{R} \), set \( \delta = 2 \epsilon \in (0, \infty) \subseteq \mathbb{R} \), ∀\((\tilde{\alpha}, \tilde{x}) \in (\mathbb{K} \times \mathcal{X}) \cap B_{\mathbb{K} \times \mathcal{X}} ((\alpha, x), \delta)\), we have

\[
\| f(\tilde{\alpha}, \tilde{x}) - f(\alpha, x) - L(\tilde{\alpha} - \alpha, \tilde{x} - x) \| \\
= \| \tilde{\alpha} \tilde{x} - \alpha x - \alpha(\tilde{x} - x) - (\tilde{\alpha} - \alpha)x \| = \| (\tilde{\alpha} - \alpha)(\tilde{x} - x) \| \\
\leq \frac{1}{2} (|\tilde{\alpha} - \alpha|^2 + |\tilde{x} - x|^2) = \frac{1}{2} \| (\tilde{\alpha} - \alpha, \tilde{x} - x) \|^2 \\
\leq \epsilon \| (\tilde{\alpha} - \alpha, \tilde{x} - x) \|
\]

By Definition 9.3, \( Df(\alpha, x) = L \). This completes the proof of the proposition.

To state the next proposition, we will introduce a new notation. Let \( A \in B(\mathcal{X}, \mathcal{Y}) \) and \( B \in B(\mathcal{Y}, \mathcal{Z}) \), where \( \mathcal{X}, \mathcal{Y}, \) and \( \mathcal{Z} \) are normed linear spaces over \( \mathbb{K} \). Clearly, \( f(A, B) := BA \in B(\mathcal{X}, \mathcal{Z}) \). Let \( g(A) : B(\mathcal{Y}, \mathcal{Z}) \rightarrow B(\mathcal{X}, \mathcal{Z}) \) be given by \( g(A)(B) = BA \), \( \forall B \in B(\mathcal{Y}, \mathcal{Z}) \). Clearly, \( g(A) \) is a bounded linear operator. It is easy to see that \( g : B(\mathcal{X}, \mathcal{Y}) \rightarrow B(B(\mathcal{Y}, \mathcal{Z}), B(\mathcal{X}, \mathcal{Z})) \) is a bounded linear operator with \( \| g \| \leq 1 \). This operator \( g \) is needed in compact “matrix” notation for many linear operators. We will use the \( \text{ro} \) to denote \( g \) for any normed linear spaces \( \mathcal{X}, \mathcal{Y}, \) and \( \mathcal{Z} \). The meaning of \( \text{ro} \) is “right operate”. For \( x \in \mathcal{X} \), we will identify \( \mathcal{X} \) with \( B(\mathbb{R}, \mathcal{X}) \) and write \( \text{ro}(x)(A) = Ax \).

This brings us to the next proposition.

**Proposition 9.17** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), \( f : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y} \) be given by \( f(A, x) = Ax \), \( \forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \). Then, \( f \) is Fréchet differentiable and \( Df : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \) is given by \( \forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \), \( \forall (\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \), \( Df(A, x)(\Delta, h) = Ah + \Delta x \). In “matrix” notation, \( Df(A, x) = [\text{ro}(x) \ A] \).

**Proof** \( \forall (A, x) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \), let \( L : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \rightarrow \mathcal{Y} \) be given by \( L(\Delta, h) = Ah + \Delta x \), \( \forall (\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \). Clearly, \( L \) is a linear operator. Note that

\[
\| L \| = \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \| (\Delta, h) \| \leq 1} \| L(\Delta, h) \| \\
\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \| (\Delta, h) \| \leq 1} \| A \| \| h \| + \| \Delta \| \| x \| \\
\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \| (\Delta, h) \| \leq 1} \left( \| A \|^2 + \| x \|^2 \right)^{1/2} \left( \| \Delta \|^2 + \| h \|^2 \right)^{1/2} \\
\leq \| (A, x) \| < +\infty
\]

where the first inequality follows from Proposition 7.64, and the second inequality follows from Cauchy-Schwarz Inequality. Hence, \( L \in B(B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \mathcal{Y}) \).
Clearly, $A_{B(X,Y)\times X}(A,x) = B(X,Y)\times X$ since $(A,x) \in (B(X,Y)\times X)^\circ = B(X,Y)\times X$. \(\forall \varepsilon \in (0,\infty) \subset \mathbb{R}, \) set $\delta = 2\varepsilon \in (0,\infty) \subset \mathbb{R}$, \(\forall (A,\bar{x}) \in (B(X,Y)\times X) \cap B_{B(X,Y)\times X}((A,x),\delta),\) we have

$$
\|f(\bar{A},\bar{x}) - f(A,x) - L(\bar{A} - A,\bar{x} - x)\|
= \|\bar{A}x - Ax - A(\bar{x} - x) - (\bar{A} - A)x\|
= \|(\bar{A} - A)(\bar{x} - x)\|
\leq \frac{1}{2}\|\bar{A} - A\|^2 + \|\bar{x} - x\|^2
= \frac{1}{2}\|(\bar{A} - A,\bar{x} - x)\|^2
\leq \varepsilon \|(\bar{A} - A,\bar{x} - x)\|
$$

where the first inequality follows from Proposition 7.64. By Definition 9.3, $Df(A,x) = L$. This completes the proof of the proposition. \(\square\)

### 9.3 Chain Rule and Mean Value Theorem

**Theorem 9.18 (Chain Rule)** Let $X$, $Y$, and $Z$ be normed linear spaces over $\mathbb{K}$, $D_x \subseteq X$, $D_y \subseteq Y$, $f : D_x \to D_y$, $g : D_y \to Z$, $x_0 \in D_x$, $y_0 := f(x_0) \in D_y$, and $h := g \circ f : D_x \to Z$. Assume that $f$ is Fréchet differentiable at $x_0$ with $Df(x_0) \in B(X,Y)$ and $g$ is Fréchet differentiable at $y_0$ with $Dg(y_0) \in B(Y,Z)$. Then, $h$ is differentiable at $x_0$ and $Dh(x_0) \in B(X,Z)$ is given by

$$
Dh(x_0) = Dg(f(x_0)) \circ Df(x_0) = Dg(y_0)Df(x_0)
$$

**Proof**

Define $L : X \to Z$ by, \(\forall \bar{x} \in X, \) $L(\bar{x}) = Dg(y_0)Df(x_0)\bar{x} = Dg(y_0)(Df(x_0)(\bar{x}))$. Clearly, $L$ is a linear operator and, by Proposition 7.64, $L \in B(X,Z)$.

\(\forall \varepsilon \in (0,\infty) \subset \mathbb{R}, \) let $\bar{\varepsilon} = (\sqrt{4\varepsilon + (\|Df(x_0)\| + \|Dg(y_0)\|)^2} - \|Df(x_0)\| - \|Dg(y_0)\|)/2 \in (0,\infty) \subset \mathbb{R}$. By the fact that $Dg(y_0) \in B(Y,Z)$, $\exists \delta_1 \in (0,\infty) \subset \mathbb{R}$ such that $\forall y \in D_y \cap B_y(y_0,\delta_1)$, we have

$$
\|g(y) - g(y_0) - Dg(y_0)(y - y_0)\| \leq \bar{\varepsilon} \|y - y_0\|
$$

By the fact that $Df(x_0) \in B(X,Y)$, we have $\text{span}(A_{D_x}((x_0))) = X$ and $\exists \delta_2 \in (0,\delta_1/(\bar{\varepsilon} + \|Df(x_0)\|)) \subset \mathbb{R}$ such that $\forall x \in D_x \cap B_X(x_0,\delta_2)$, we have

$$
\|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \leq \bar{\varepsilon} \|x - x_0\|
$$

Then, by Proposition 7.64, $\|f(x) - y_0\| = \|f(x) - f(x_0)\| \leq \bar{\varepsilon} \|x - x_0\| + \|Df(x_0)\| \|x - x_0\| \leq (\bar{\varepsilon} + \|Df(x_0)\|) \|x - x_0\| < \delta_1$. This implies that $f(x) \in D_y \cap B_y(y_0,\delta_1)$. Then, we have

$$
\|h(x) - h(x_0) - L(x - x_0)\| = \|g(f(x)) - g(y_0) - L(x - x_0)\|
\leq \|g(f(x)) - g(y_0) - Dg(y_0)(f(x) - y_0)\|
+ \|Dg(y_0)(f(x) - f(x_0) - Df(x_0)(x - x_0))\|
$$
Let $D \subset X$, $f_1 : D \to Y$, $f_2 : D \to Z$, $x_0 \in D$, and $g : D \to Y \times Z$ be given by $g(x) = (f_1(x), f_2(x))$, $\forall x \in D$. Then, the following statement holds. $g$ is Fréchet differentiable at $x_0$ if, and only if, $f_1$ and $f_2$ are Fréchet differentiable at $x_0$. In this case, $Dg(x_0)(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$, $\forall h \in X$. In “matrix” notation, $Dg(x_0) = \begin{bmatrix} Df_1(x_0) \\ Df_2(x_0) \end{bmatrix}$.

Proof “Sufficiency” By the differentiability of $f_1$ and $f_2$ at $x_0$, we have span $(A_D(x_0)) = X$. Define $L : X \to Y \times Z$ by, $\forall h \in X$, $L(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$. Clearly, $L$ is a linear operator. Note that

$$
\| L \| = \sup_{h \in X, \| h \| \leq 1} \| L(h) \|
= \sup_{h \in X, \| h \| \leq 1} \sqrt{\| Df_1(x_0)(h) \|^2 + \| Df_2(x_0)(h) \|^2}
\leq \sup_{h \in X, \| h \| \leq 1} \sqrt{\| Df_1(x_0) \|^2 \| h \|^2 + \| Df_2(x_0) \|^2 \| h \|^2}
\leq \sqrt{\| Df_1(x_0) \|^2 + \| Df_2(x_0) \|^2} < +\infty
$$

where the first inequality follows from Proposition 7.64 and the last inequality follows from the fact $Df_1(x_0) \in B(X, Y)$ and $Df_2(x_0) \in B(X, Z)$. Hence, $L \in B(X, Y \times Z)$.

Since $f_1$ and $f_2$ are differentiable at $x_0$, then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap B_X(x_0, \delta_1)$, we have $\| f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0) \| \leq \epsilon \sqrt{2} \| x - x_0 \|$. $\exists \delta_2 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in D \cap B_X(x_0, \delta_2)$, we have $\| f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0) \| \leq \epsilon \sqrt{2} \| x - x_0 \|$. Let $\delta = \min \{ \delta_1, \delta_2 \} \in (0, \infty) \subset \mathbb{R}$, $\forall x \in D \cap B_X(x_0, \delta)$, we have

$$
\| g(x) - g(x_0) - L(x - x_0) \|
= \| (f_1(x), f_2(x)) - (f_1(x_0), f_2(x_0)) - (Df_1(x_0)(x - x_0), Df_2(x_0)(x - x_0)) \|
= \| (f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0), f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0)) \|
= \left( \| f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0) \|^2 + \| f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0) \|^2 \right)^{1/2}
$$
9.3. **CHAIN RULE AND MEAN VALUE THEOREM**

\[ \leq \epsilon \| x - x_0 \| \]

Hence, \( g \) is differentiable at \( x_0 \) and \( Dg(x_0) = L \).

"Necessity" By the differentiability of \( g \) at \( x_0 \), we have \( \text{span}( A_D(x_0) ) = \mathcal{X} \). Note that \( f_1 = \pi_y \circ g \). By Chain Rule and Proposition 9.13, \( f_1 \) is Fréchet differentiable at \( x_0 \) and

\[
Df_1(x_0) = \left[ \begin{array}{c c}
\text{id}_Y & \partial_{B(z,y)}
\end{array} \right] Dg(x_0)
\]

By symmetry, \( f_2 \) is Fréchet differentiable at \( x_0 \) and

\[
Df_2(x_0) = \left[ \begin{array}{c c}
\partial_{B(y,z)} & \text{id}_Z
\end{array} \right] Dg(x_0)
\]

Then, \( Dg(x_0) = \left[ \begin{array}{c c} Df_1(x_0) \\ Df_2(x_0) \end{array} \right] \).

This completes the proof of the proposition. \( \square \)

**Theorem 9.20 (Mean Value Theorem)** Let \( \mathcal{X} \) be a real normed linear space, \( D \subseteq \mathcal{X} \), \( f : D \to \mathbb{R} \), \( x_1, x_2 \in D \), and \( \varphi : [0,1] \to D \) be given by \( \varphi(t) = tx_1 + (1-t)x_2 \), \( \forall t \in I := [0,1] \subseteq \mathbb{R} \). Assume that \( f \) is continuous at \( \varphi(t) \), \( \forall t \in I \) and \( f \) is Fréchet differentiable at \( \varphi(t) \), \( \forall t \in I^0 \). Then, there exists \( t_0 \in I^0 \) such that

\[
f(x_1) - f(x_2) = Df(\varphi(t_0))(x_1 - x_2)
\]

**Proof** By Propositions 9.19, 9.16, and 9.15 and Chain Rule, \( \varphi \) is Fréchet differentiable and \( D\varphi(t)(h) = h(x_1 - x_2) \), \( \forall h \in \mathbb{R}, \forall t \in I \). Define \( g : I \to \mathbb{R} \) by \( g(t) = f(\varphi(t)) \), \( \forall t \in I \). By Chain Rule, \( g \) is Fréchet differentiable at \( t \), \( \forall t \in I^0 \) and \( Dg(t)(h) = Df(\varphi(t))(D\varphi(t)(h)) = hDf(\varphi(t))(x_1 - x_2) \), \( \forall h \in \mathbb{R}, \forall t \in I^0 \). Then, \( Dg(t) = Df(\varphi(t))(x_1 - x_2) \), \( \forall t \in I^0 \). By Proposition 3.12, \( g \) is continuous. By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), we have

\[
f(x_1) - f(x_2) = g(1) - g(0) = Dg(t_0)(1 - 0) = Df(t_0x_1 + (1-t_0)x_2)(x_1 - x_2)
\]

for some \( t_0 \in I^0 \). This completes the proof of the theorem. \( \square \)

Toward a general Mean Value Theorem for vector-valued functions on possibly complex normed linear spaces, we present the following two lemmas.

**Lemma 9.21** Let \( D := \{ a + i0 \mid a \in I := [0,1] \subset \mathbb{R} \} \subset \mathbb{C} \), \( x_0 := x_r + ix_i \in D \), where \( x_r \in I \) and \( x_i = 0 \), \( f : D \to \mathbb{C} \) be Fréchet differentiable at \( x_0 \), \( Df(x_0) = d_r + id_i \in \mathbb{C} \), where \( d_r, d_i \in \mathbb{R} \), and \( g : I \to \mathbb{R} \) be given by \( g(t) = \text{Re}(f(t + i0)) \), \( \forall t \in I \). Then, \( g \) is Fréchet differentiable at \( x_r \) and \( Dg(x_r) = d_r = \text{Re}(Df(x_0)) \in \mathbb{R} \).
Proof. Note that \( \text{span}(A_I(x_r)) = \mathbb{R} \). \( \forall t \in (0, \infty) \subset \mathbb{R} \), by the differentiability of \( f \) at \( x_0 \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \forall \bar{x} = x_r + i\bar{r}, \bar{r} \in D \cap B_{\mathbb{C}}(x_0, \delta) \), we have \( |f(\bar{x}) - f(x_0) - Df(x_0)(\bar{x} - x_0)| \leq \epsilon |\bar{x} - x_0| \). Note that \( \bar{x} \in I \) and \( \bar{x} = 0 \). Then, the above implies that \( |(Re(f(\bar{x})) - Re(f(x_r))) - (\bar{x} - x_r)| \leq \epsilon |\bar{x} - x_r| \). This further implies that \( |Re(f(\bar{x})) - Re(f(x_r))| \leq \epsilon |\bar{x} - x_r| \). Then, \( \forall t \in I \cap B_{\mathbb{R}}(x_r, \delta), \bar{x} = t + i\bar{r} \in D \cap B_{\mathbb{C}}(x_0, \delta) \) and \( |g(t) - g(x_r) - (t - x_r)| \leq \epsilon |t - x_r| \). Hence, \( Dg(x_r) = d_r \). This completes the proof of the lemma.

Lemma 9.22 Let \( D := \{ a + i\bar{o} \mid a \in I := [0, 1] \subset \mathbb{R} \} \subset \mathbb{C}, f : D \to \mathbb{C} \) be continuous, \( f \) be Fréchet differentiable at \( a + i\bar{o}, \forall a \in I^o \). Then, \( \exists t_0 \in I^o \) such that \( \Re(f(t) - f(0)) = \Re(Df(t_0)) \).

Proof. Let \( g : I \to \mathbb{R} \) be given by \( g(t) = \Re(f(t + i\bar{o})), \forall t \in I \). Clearly, \( g \) is continuous since \( f \) is continuous. By Lemma 9.21, \( g \) is Fréchet differentiable at \( t, \forall t \in I^o \), and \( Dg(t) = \Re(Df(t + i\bar{o})) \), \( \forall t \in I^o \). By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), \( \exists t_0 \in I^o \) such that \( g(1) - g(0) = Dg(t_0) = \Re(Df(t_0 + i\bar{o})) \). Then, \( \Re(f(t) - f(0)) = g(1) - g(0) = \Re(Df(t_0 + i\bar{o})) = \Re(Df(t_0)) \). This completes the proof of the lemma.

Theorem 9.23 (Mean Value Theorem) Let \( X \) and \( Y \) be normed linear spaces over \( K, D \subset X, f : D \to Y, x_1, x_2 \in D, \) and the line segment connecting \( x_1 \) and \( x_2 \) be contained in \( D \). Assume that \( f \) is continuous at \( x = tx_1 + (1-t)x_2, \forall t \in I := [0, 1] \subset \mathbb{R} \) and \( Df(x) \in B(X, Y) \) exists at \( x = tx_1 + (1-t)x_2, \forall t \in I^o \). Then, \( \exists t_0 \in I^o \) such that \( \|f(x_1) - f(x_2)\| \leq \|Df(t_0)x_1 + (1-t_0)x_2\| \).

Proof. We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( f(x_1) = f(x_2) \); Case 2: \( f(x_1) \neq f(x_2) \) and \( K = \mathbb{R} \); Case 3: \( f(x_1) \neq f(x_2) \) and \( K = \mathbb{C} \). Case 1: \( f(x_1) = f(x_2) \). Take \( t_0 \) to be any point in \( I^o \). The desired result follows.

Case 2: \( f(x_1) \neq f(x_2) \) and \( K = \mathbb{R} \). By Proposition 7.85, \( \exists y_* \in Y^* \) with \( \|y_*\| = 1 \) such that \( \|\langle y_*, f(x_1) - f(x_2) \rangle\| = \|f(x_1) - f(x_2)\| \). Define \( \varphi : I \to D \) by \( \varphi(t) = tx_1 + (1-t)x_2, \forall t \in I \). By Propositions 9.19, 9.16, and 9.15 and Chain Rule, \( \varphi \) is Fréchet differentiable. Define \( g : I \to \mathbb{R} \) by \( g(t) = \langle y_*, f(\varphi(t)) \rangle, \forall t \in I \). By Proposition 3.12, \( g \) is continuous. By Chain Rule and Propositions 9.17 and 9.19, \( g \) is Fréchet differentiable at \( t, \forall t \in I^o \), and \( Dg(t)(d) = \langle y_*, Df(\varphi(t))(D\varphi(t))(d) \rangle \) for \( \forall t \in I^o, d \in \mathbb{R} \) and \( \forall t \in I^o \). Hence, \( Dg(t) = \langle y_*, Df(\varphi(t))(x_1 - x_2) \rangle, \forall t \in I^o \). By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), there exists \( t_0 \in I^o \) such that \( g(1) - g(0) = Dg(t_0) \). Then, we have \( \|f(x_1) - f(x_2)\| = \|\langle y_*, f(x_1) - f(x_2) \rangle\| = \|g(1) - g(0) = Dg(t_0) \leq \|Dg(t_0)\| = \|\langle y_*, Df(\varphi(t_0))(x_1 - x_2) \rangle\| \leq \|Df(\varphi(t_0))(x_1 - x_2)\| \), where the last inequality follows from Proposition 7.72. The desired result follows.
9.3. CHAIN RULE AND MEAN VALUE THEOREM

Case 3: $f(x_1) \neq f(x_2)$ and $K = \mathbb{C}$. By Proposition 7.85, $\exists \nu \in \mathbb{Y}$ with $\|y\|_2 = 1$ such that $\left\langle \left\langle y, f(x_1) - f(x_2) \right\rangle \right\rangle = \|f(x_1) - f(x_2)\|$. Let $\bar{D} := \{a + i0 \mid a \in I\} \subseteq \mathbb{C}$. Define $\varphi : \bar{D} \to D$ by $\varphi(t) = tx_1 + (1-t)x_2$, $\forall t \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, $\varphi$ is Fréchet differentiable and $D\varphi(t)(d) = d(x_1 - x_2)$, $\forall d \in \mathbb{C}$ and $\forall t \in \bar{D}$. Define $g : \bar{D} \to \mathbb{C}$ by $g(t) = \left\langle \left\langle y, f(\varphi(t)) \right\rangle \right\rangle$, $\forall t \in \bar{D}$. By Proposition 3.12, $g$ is continuous. By Chain Rule and Propositions 9.17 and 9.19, $g$ is Fréchet differentiable at $a + i0$, $\forall a \in I^\circ$, and $Dg(a + i0)(d) = \left\langle \left\langle y, Df(\varphi(a + i0))(d) \right\rangle \right\rangle = \left\langle \left\langle y, Df(\varphi(a))(d(x_1 - x_2)) \right\rangle \right\rangle = \left\langle \left\langle y, Df(\varphi(a))(x_1 - x_2) \right\rangle \right\rangle = \left\langle \left\langle y, f(x_1) - f(x_2) \right\rangle \right\rangle \in \mathbb{R}$. By Lemma 9.22, there exists $t_0 \in I^\circ$ such that $g(1) - g(0) = \text{Re}(Dg(t_0))$. Then, we have $\|f(x_1) - f(x_2)\| = g(1) - g(0) \leq |Dg(t_0)| = \left\langle \left\langle y, Df(\varphi(t_0))(x_1 - x_2) \right\rangle \right\rangle \leq \|Df(\varphi(t_0))(x_1 - x_2)\|$, where the last inequality follows from Proposition 7.72. The desired result follows.

This completes the proof of the theorem.

\[ \square \]

Proposition 9.24 Let $X$, $Y$, and $Z$ be normed linear spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $f : D \to Z$, $(x_0, y_0) \in D$. Assume that the following conditions hold.

(i) $\exists \delta_0 \in (0, \infty) \subseteq \mathbb{R}$, $\forall (x, y) \in D \cap \mathcal{B}_{X \times Y}((x_0, y_0), \delta_0)$, we have $(x, ty + (1-t)y_0) \in D$, $\forall t \in I := [0, 1] \subseteq \mathbb{R}$.

(ii) $f$ is partial differentiable with respect to $x$ at $(x_0, y_0)$ and $\frac{\partial f}{\partial x}(x_0, y_0) \in \mathcal{B}(X, Z)$.

(iii) $\forall (x, y) \in D \cap \mathcal{B}_{X \times Y}((x_0, y_0), \delta_0)$, $f$ is partial differentiable with respect to $y$ at $(x, y)$ and $\frac{\partial f}{\partial y}(x_0, y_0)$ is continuous at $(x_0, y_0)$.

Then, $f$ is Fréchet differentiable at $(x_0, y_0)$ and $Df(x_0, y_0) \in \mathcal{B}(X \times Y, Z)$ is given by $Df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$, $\forall (h, k) \in X \times Y$.

In “matrix” notation, $Df(x_0, y_0) = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix}$.

Proof We will first show that $\text{span}\{(A_D(x_0, y_0))\} = \mathbb{X} \times \mathbb{Y}$. Define $D_{x_0} := \{x \in \mathbb{Y} \mid (x, y) \in D\}$ and $D_{y_0} := \{x \in \mathbb{X} \mid (x, y) \in D\}$. By the partial differentiability of $f$ with respect to $x$ at $(x_0, y_0)$, we have $\text{span}\{(A_{D_{x_0}}(x_0, y_0))\} = \mathbb{X}$. By the partial differentiability of $f$ with respect to $y$ at $(x_0, y_0)$, we have $\text{span}\{(A_{D_{y_0}}(x_0, y_0))\} = \mathbb{Y}$. Hence, $\{\partial x\} \times \mathbb{Y} \subseteq \text{span}\{(A_D(x_0, y_0))\}$. By symmetry, we have $\{\partial y\} \times \mathbb{X} \subseteq \text{span}\{(A_D(x_0, y_0))\}$. By Proposition 7.17, $\text{span}\{(A_D(x_0, y_0))\} = \mathbb{X} \times \mathbb{Y}$.
Define $L : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ by $L(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$, $\forall (h, k) \in \mathcal{X} \times \mathcal{Y}$. Clearly, $L$ is a linear operator. Note that

$$
\| L \| = \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \| (h, k) \| \leq 1} \| L(h, k) \| 
\leq \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \| (h, k) \| \leq 1} \left( \| \frac{\partial f}{\partial x}(x_0, y_0) \| \| h \| + \| \frac{\partial f}{\partial y}(x_0, y_0) \| \| k \| \right) 
\leq \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \| (h, k) \| \leq 1} \sqrt{\| \frac{\partial f}{\partial x}(x_0, y_0) \|^2 + \| \frac{\partial f}{\partial y}(x_0, y_0) \|^2} \cdot \sqrt{\| h \|^2 + \| k \|^2} 
\leq \sqrt{\| \frac{\partial f}{\partial x}(x_0, y_0) \|^2 + \| \frac{\partial f}{\partial y}(x_0, y_0) \|^2} < +\infty
$$

where the first inequality follows from Proposition 7.64 and the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in \mathcal{B}(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the partial differentiability of $f$ with respect to $x$ at $(x_0, y_0)$, $\exists \delta_1 \in (0, \delta_0) \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_1)$, we have $\left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| \leq \epsilon/\sqrt{2} \| x - x_0 \|$. By the continuity of $\frac{\partial f}{\partial y}$ at $(x_0, y_0)$, $\exists \delta_2 \in (0, \delta_1] \subset \mathbb{R}$, such that $\forall (\bar{x}, \bar{y}) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_2)$, we have $\left\| \frac{\partial f}{\partial y}(\bar{x}, \bar{y}) - \frac{\partial f}{\partial y}(x_0, y_0) \right\| < \epsilon/\sqrt{2}$.

**Claim 9.24.1** $\left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(y - y_0) \right\| \leq \epsilon/\sqrt{2} \| y - y_0 \|$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_2)$.

**Proof of claim:** Fix any $(x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_2)$. Let $\bar{D} = \{ a + i0 \mid a \in I \} \subset \mathbb{C}$ if $\mathbb{K} = \mathbb{C}$. Define $\psi : \bar{D} \to \mathcal{Z}$ by $\psi(t) = f(x, ty + (1 - t)y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(t(y - y_0))$, $\forall t \in \bar{D}$. By Chain Rule, each term in the definition of $\psi$ is Fréchet differentiable. By Proposition 9.15, $\psi$ is Fréchet differentiable. By Mean Value Theorem, $\exists t_0 \in I^2$,

$$
\left\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| = \| \psi(1) - \psi(0) \| \leq \| D\psi(t_0) \| 
= \left\| \frac{\partial f}{\partial y}(x, ty + (1 - t_0)y_0)(y - y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| 
\leq \left\| \frac{\partial f}{\partial y}(x, ty + (1 - t_0)y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \right\| \| y - y_0 \| \leq \epsilon/\sqrt{2} \| y - y_0 \|
$$

where the second inequality follows from Proposition 7.64. This completes the proof of the claim.

Therefore, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_2)$, we have

$$
\left\| f(x, y) - f(x_0, y_0) - L(x - x_0, y - y_0) \right\|
$$
\[ \leq \left\| f(x,y) - f(x,y_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \right\| \\
+ \left\| f(x,y_0) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) \right\| \leq \epsilon/\sqrt{2} \| y-y_0 \| + \epsilon/\sqrt{2} \| x-x_0 \| \leq \epsilon \| (x-x_0,y-y_0) \| \]

where the last inequality follows from Cauchy-Schwarz Inequality. Hence, \( Df(x_0,y_0) = L \). This completes the proof of the proposition.

We observe that Conditions (i) of Proposition 9.24 is easily satisfied when \((x_0,y_0) \in D^0\).

### 9.4 Higher Order Derivatives

#### 9.4.1 Basic concept

We introduce the following notation. Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \). Denote \( B(\mathcal{X}, \mathcal{Y}) \) by \( B_1(\mathcal{X}, \mathcal{Y}) \). Recursively, denote \( B(\mathcal{X}, B_k(\mathcal{X}, \mathcal{Y})) \) by \( B_{k+1}(\mathcal{X}, \mathcal{Y}) \), \( \forall k \in \mathbb{N} \). Note that \( B_k(\mathcal{X}, \mathcal{Y}) \) is the set of bounded multilinear \( \mathcal{Y} \)-valued functions on \( \mathcal{X}^k \), \( \forall k \in \mathbb{N} \). Define the subset of symmetric functions by \( B_{S_k}(\mathcal{X}, \mathcal{Y}) := \{ L \in B_k(\mathcal{X}, \mathcal{Y}) \mid L(h_k)\cdots(h_1) = L(v_k)\cdots(v_1), \forall (h_1, \ldots, h_k) \in \mathcal{X}^k, \forall (v_1, \ldots, v_k) = \text{a permutation of} \ (h_1, \ldots, h_k) \} \). Note that \( B_{S_k}(\mathcal{X}, \mathcal{Y}) \) is a closed subspace of \( B_k(\mathcal{X}, \mathcal{Y}) \).

Then, by Proposition 7.13, \( B_{S_k}(\mathcal{X}, \mathcal{Y}) \) is a normed linear space over \( \mathbb{K} \). If \( \mathcal{Y} \) is a Banach space, then, by Proposition 7.66, \( B_k(\mathcal{X}, \mathcal{Y}) \) is a Banach space. Then, by Proposition 4.39, \( B_{S_k}(\mathcal{X}, \mathcal{Y}) \) is a Banach space. For notational consistency, we will denote \( B_{S_0}(\mathcal{X}, \mathcal{Y}) := B_0(\mathcal{X}, \mathcal{Y}) := \mathcal{Y} \).

**Definition 9.25** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \), \( f : D \to \mathcal{Y} \), and \( x_0 \in D \). Let \( f^{(1)} \) be defined with domain of definition \( \text{dom} \ (f^{(1)}) \). We may consider the derivative of \( f^{(1)} \). If \( f^{(1)} \) is differentiable at \( x_0 \in \text{dom} \ (f^{(1)}) \), then \( f \) is said to be twice Fréchet differentiable at \( x_0 \). The second order derivative of \( f \) at \( x_0 \) is \( D(Df)(x_0) := D^2 f(x_0) := f^{(2)}(x_0) \in B(\mathcal{X}, B(\mathcal{X}, \mathcal{Y})) = B_2(\mathcal{X}, \mathcal{Y}) \). \( D^2 f \) or \( f^{(2)} \) will denote the \( B_2(\mathcal{X}, \mathcal{Y}) \)-valued function whose domain of definition is \( \text{dom} \ (f^{(2)}) := \{ x \in \text{dom} \ (f^{(1)}) \mid f^{(2)}(x) \in B_2(\mathcal{X}, \mathcal{Y}) \text{ exists} \} \).

Recursively, if \( f^{(k)} \) is Fréchet differentiable at \( x_0 \in \text{dom} \ (f^{(k)}) \), then \( f \) is said to be \((k+1)\)-times Fréchet differentiable at \( x_0 \) and the \((k+1)\)st order derivative of \( f \) at \( x_0 \) is \( D^{k+1} f(x_0) := f^{(k+1)}(x_0) \in B_{k+1}(\mathcal{X}, \mathcal{Y}) \), where \( k \in \mathbb{N} \). \( D^{k+1} f \) or \( f^{(k+1)} \) will denote the \( B_{k+1}(\mathcal{X}, \mathcal{Y}) \)-valued function whose domain of definition is \( \text{dom} \ (f^{(k+1)}) := \{ x \in \text{dom} \ (f^{(k)}) \mid f^{(k+1)}(x) \in B_{k+1}(\mathcal{X}, \mathcal{Y}) \text{ exists} \} \). For notational consistency, we will let \( f^{(0)} = f \).

Note that \( \text{dom} \ (f^{(k+1)}) \subseteq \text{dom} \ (f^{(k)}) \subseteq D \), \( \forall k \in \mathbb{N} \).
Definition 9.26 Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $f : D \to Y$, and $x_0 \in D$. Assume that $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f$ is $k$-times Fréchet differentiable at $x$, $\forall x \in D \cap B_X(x_0, \delta_0)$, where $k \in \mathbb{N}$, that is $D \cap B_X(x_0, \delta_0) \subseteq \text{dom}(f^{(k)})$, and $f^{(k)}$ is continuous at $x_0$. Then, we say that $f$ is $C_k$ at $x_0$. If $f$ is $C_k$ at $x$, $\forall x \in D$, then, we say $f$ is $C_k$. If $f$ is $C_k$ at $\forall x \in D \cap B_X(x_0, \delta_0)$, $\forall k \in \mathbb{N}$, then, we say that $f$ is $C_\infty$ at $x_0$. If $f$ is $C_\infty$ at $x$, $\forall x \in D$, then, we say that $f$ is $C_\infty$.

Note that $f$ being $(k+1)$-times differentiable at $x_0 \in D$ does not imply that $f$ is $C_{k+1}$ at $x_0$ since dom $(f^{(k)})$ may not contain $D \cap B_X(x_0, \delta)$, $\forall \delta \in (0, \infty) \subset \mathbb{R}$. When dom $(f^{(k)}) \supseteq D \cap B_X(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$, and $f$ is $(k+1)$-times differentiable at $x_0$, then $f$ is $C_{k+1}$ at $x_0$. In particular, if $f$ is $C_{k+1}$ at $x_0$, then $f$ is $C_k$ at $x$, $\forall x \in D \cap B_X(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$. If $f$ is infinitely many times differentiable at $x$, $\forall x \in D$, then $f$ is $C_\infty$.

Proposition 9.27 Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $f : D \to Y$ be $C_{n+m-1}$ at $x_0$ and $(n + m)$-times Fréchet differentiable at $x_0 \in D$, where $n, m \in \mathbb{N}$. Fix $(h_1, \ldots, h_n) \in \mathbb{X}^n$. Define the function $g : \text{dom}(f^{(n)}) \to Y$ by $g(x) = f^{(n)}(x)(h_n) \cdots (h_1)$, $\forall x \in D_n := \text{dom}(f^{(n)})$. Then, the following statements hold.

(i) $g$ is $m$-times Fréchet differentiable at $x_0$ and $g^{(m)}(x_0) \in B_m(X, Y)$ is given by $g^{(m)}(x_0)(h_{n+m}) \cdots (h_{n+1}) = f^{(n+m)}(x_0)(h_{n+m}) \cdots (h_1)$, $\forall (h_{n+1}, \ldots, h_{n+m}) \in \mathbb{X}^m$.

(ii) If $f$ is $C_{n+m}$ at $x_0$, then $g$ is $C_m$ at $x_0$.

Proof We will first prove (i) using mathematical induction on $m$.

1° $m = 1$. Since $f$ is $C_n$ at $x_0$, then $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap B_X(x_0, \delta_0) \subseteq D_n$. By the $(n + 1)$-times differentiability of $f$ at $x_0$, we have span $(A_{D_n}(x_0)) = X$. Define $L : X \to Y$ by, $\forall h \in X$, $L(h) = f^{(n+1)}(x_0)(h)(h_n) \cdots (h_1)$. Clearly, $L$ is a linear operator. Note that

$$
\|L\| = \sup_{h \in X, \|h\| \leq 1} \|L(h)\| = \sup_{h \in X, \|h\| \leq 1} \|f^{(n+1)}(x_0)(h)(h_n) \cdots (h_1)\|
$$

$$
\leq \sup_{h \in X, \|h\| \leq 1} \|f^{(n+1)}(x_0)\| \|h\| \|h_n\| \cdots \|h_1\|
$$

$$
\leq \|f^{(n+1)}(x_0)\| \|h_n\| \cdots \|h_1\| < +\infty
$$

where the first inequality follows from Proposition 7.64. Hence, $L \in B(X, Y)$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the differentiability of $f^{(n)}$ at $x_0$, $\exists \delta \in (0, \delta_0) \subset \mathbb{R}$ such that $\forall x \in D_n \cap B_X(x_0, \delta)$, we have $\|f^{(n)}(x) - f^{(n)}(x_0) - f^{(n+1)}(x_0)(x - x_0)\| \leq \epsilon/(1 + \|h_n\| \cdots \|h_1\|) \|x - x_0\|$. Then, we have

$$
\|g(x) - g(x_0) - L(x - x_0)\|
$$
\[ = \left\| f^{(n)}(x)(h_n) \cdots (h_1) - f^{(n)}(x_0)(h_n) \cdots (h_1) \right\| \\
= \left\| f^{(n)}(x)(x - x_0)(h_n) \cdots (h_1) \right\| \\
\leq \left\| f^{(n)}(x) - f^{(n)}(x_0) \right\| \left\| (x - x_0)(h_n) \cdots (h_1) \right\| \\
\leq \epsilon \left\| x - x_0 \right\|
\]

where the first inequality follows from Proposition 7.64. Hence, \( Dg(x_0) = L \) and \( g \) is Fréchet differentiable at \( x_0 \).

2° Assume that (i) holds for \( m \leq k, \forall k \in \mathbb{N} \).

3° Consider the case \( m = k + 1 \). Since \( f \) is \( C_{n+k+1} \) at \( x_0 \), then, \( \exists \delta_0 \in (0, \infty) \subset \mathbb{R} \) such that \( f \) is \( (n + k) \)-times Fréchet differentiable at \( \bar{x} \) and is \( C_{n+k} \) at \( \bar{x} \). \( \forall x \in D \cap B_{X}(x_0, \delta_0) \). By inductive assumption, \( g \) is \( k \)-times Fréchet differentiable at \( \bar{x} \) and \( g^{(k)}(\bar{x})(h_{n+k}) \cdots (h_{n+1}) = f^{(n+k)}(\bar{x})(h_{n+k}) \cdots (h_{n+1}) \), \( \forall (h_{n+1}, \ldots, h_{n+k}) \in X^k \). Hence, \( \bar{x} \in D_{n+k} := \text{dom} \left( g^{(k)} \right) \). Then, we have \( D \cap B_{X}(x_0, \delta_0) \subseteq D_{n+k} \subseteq D \). Note that \( D_{n+k} \cap B_{X}(x_0, \delta_0) = D \cap B_{X}(x_0, \delta_0) \). Then, span \( (A_{D}(x_0)) = \mathbb{R}, \) since \( f \) is differentiable at \( x_0 \).

Define \( L : X \rightarrow B_{k}(X, y) \) by, \( \forall h \in X \), \( \forall (h_{n+1}, \ldots, h_{n+k}) \in X^k \), \( L(h)(h_{n+k}) \cdots (h_{n+1}) = f^{(n+k+1)}(x_0)(h)(h_{n+k}) \cdots (h_{n+1}) \). Clearly, \( L \) is a linear operator. Note that

\[ \| L \| = \sup_{h \in X, \| h \| \leq 1} \| L(h) \| \\
= \sup_{h \in X, \| h \| \leq 1, h_{n+1} \in X} \| L(h)(h_{n+k}) \cdots (h_{n+1}) \| \\
= \sup_{h \in X, \| h \| \leq 1, h_{n+1} \in X} \| f^{(n+k+1)}(h)(h_{n+k}) \cdots (h_{n+1}) \| \\
\leq \| f^{(n+k+1)}(x_0) \| \| h \| \cdots \| h_1 \| < +\infty
\]

where the first inequality follows from Proposition 7.64. Then, \( L \in B_{k+1}(X, y) \).

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the differentiability of \( f^{(n+k)} \) at \( x_0 \), \( \exists \delta \in (0, \delta_0) \subset \mathbb{R} \) such that \( \forall x \in \text{dom} \left( f^{(n+k)} \right) \cap B_{X}(x_0, \delta) = D \cap B_{X}(x_0, \delta) \), we have

\[ \| f^{(n+k)}(x) - f^{(n+k)}(x_0) - f^{(n+k+1)}(x)(x - x_0) \| \leq \epsilon/(1 + \| h \| \cdots \| h_1 \|) \| x - x_0 \|. \]

Then, \( \forall x \in D_{n+k} \cap B_{X}(x_0, \delta) = D \cap B_{X}(x_0, \delta), \)

\[ \| g^{(k)}(x) - g^{(k)}(x_0) - L(x - x_0) \| \\
= \sup_{h_{n+1} \in X, \| h_{n+1} \| \leq 1, i = 1, \ldots, k} \| g^{(k)}(x) - g^{(k)}(x_0) - L(x - x_0)(h_{n+k}) \cdots (h_{n+1}) \| \\
= \sup_{h_{n+1} \in X, \| h_{n+1} \| \leq 1, i = 1, \ldots, k} \| f^{(n+k)}(x)(h_{n+k}) \cdots (h_{n+1}) - f^{(n+k+1)}(x_0)(x - x_0)(h_{n+k}) \cdots (h_{n+1}) \| \\
\]
where the first inequality follows from Proposition 7.64. Hence, $g^{(k+1)}(x_0) = D_f^{(k)}(x_0) = L$. Therefore, $g$ is $(k+1)$-times differentiable at $x_0$.

This completes the induction process and the proof of (i).

For (ii), let $f$ be $C_{n+m}$ at $x_0$. By Definition 9.26, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap B_X (x_0, \delta_0) \subseteq D_{n+m}$, i.e., $f$ is continuous at $x_0$. Assume that $f$ is $(n+m)$-times differentiable at $x_0$, then, $D \cap B_X (x_0, \delta_0) \subseteq D_{n+m}$, $f$ is $C_{n+m-1}$ at $x_0$. Let $f$ be $C_{m}$ at $x_0$. By the continuity of $f(n+m)$ at $x_0$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\| f(n+m)(x) - f(n+m)(x_0) \| < \epsilon / (1 + \| h_n \| \cdots \| h_1 \|)$, $\forall x \in D_{n+m} \cap B_X (x_0, \delta) = D \cap B_X (x_0, \delta)$. Then, $D \cap B_X (x_0, \delta) \subseteq D_{n+m}$, $f$ is $C_{m}$ at $x_0$.

where the first inequality follows from Proposition 7.64. Hence, $g^{(m)}$ is continuous at $x_0$. Therefore, $g$ is $C_m$ at $x_0$.

This completes the proof of the proposition.  

**Proposition 9.28** Let $X$ and $Y$ be normed linear spaces over $\mathbb{K}$, $D \subseteq X$, $x_0 \in D$, and $f : D \to Y$ be $C_n$ at $x_0$, where $n \in \mathbb{N}$. Assume that $D$ is locally convex at $x_0$. Then, $f^{(n)}(x_0) \in B_{\mathbb{R}^n}(X, Y)$.

**Proof** By the assumption, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap B_X (x_0, \delta_0)$ is convex. Without loss of generality, assume $f$ is $n$-times differentiable at $x_0$. We will prove the proposition by mathematical induction on $n$.

1° $n = 1$. Clearly $f^{(1)}(x_0) \in B(X, Y) = B_1(X, Y) = B_{\mathbb{R}^1}(X, Y)$.

Next, we consider $n = 2$. We will prove this case using an argument of contradiction. Suppose $f^{(2)}(x_0)$ is not symmetric. Then, $\exists h_0, l_0 \in X$ such that $f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0)$. By the differentiability of $f$ at $x_0$, we have span $(A_D(x_0)) = X$. Then, by $f^{(2)}(x_0) \in B_2(X, Y)$ and Proposition 3.66, $\exists h_0, l_0 \in span(A_D(x_0))$ such that $f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0)$. Since $f^{(2)}(x_0)$ is multi-linear, then $\exists h_0, l_0 \in A_D(x_0)$ such that $f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0)$. By continuity of $f^{(2)}(x_0)$,
\[9.4. \text{ HIGHER ORDER DERIVATIVES} \]

Then, by Proposition 9.15, \( \psi_t \in \mathcal{B}_X \left( \tilde{h}_0, \epsilon_1 \right) \), such that \( x_0 + r_h \tilde{h}_0, x_0 + r_l \tilde{l}_0 \in D \cap \mathcal{B}_X (x_0, \delta_0) \). Clearly, we have \( f^{(2)}(x_0)(r_h \tilde{h}_0)(r_l \tilde{l}_0) = r_h r_l f^{(2)}(x_0)(\tilde{h}_0)(\tilde{l}_0) \neq r_h r_l f^{(2)}(x_0)(\tilde{l}_0)(\tilde{h}_0) = f^{(2)}(x_0)(r_l \tilde{l}_0)(r_h \tilde{h}_0) \). Let \( h_0 := r_h / 2 \tilde{h}_0 \) and \( l_0 := r_l / 2 \tilde{l}_0 \). Then, by the convexity of the set \( D \cap \mathcal{B}_X (x_0, \delta_0) \), we have \( x_0, x_0 + h_0, x_0 + l_0, x_0 + h_0 + l_0 \in D \cap \mathcal{B}_X (x_0, \delta_0) \) and \( f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0) \). Clearly, \( h_0 \neq \hat{\vartheta}_X \) and \( l_0 \neq \hat{\vartheta}_X \). Let \( \epsilon_0 := \left\| f^{(2)}(x_0)(h_0)(l_0) \right\| - f^{(2)}(x_0)(l_0)(h_0) \right\| \leq \left\| h_0 \right\| \left\| l_0 \right\| \) \((0, \infty) \subset \mathbb{R} \). Since \( f \) is \( C_2 \) at \( x_0 \), then \( \exists \delta_1 \in (0, \delta_0) \subset \mathbb{R} \) such that \( \forall x \in D \cap \mathcal{B}_X (x_0, \delta_1) \), we have \( \left\| f^{(2)}(x) - f^{(2)}(x_0) \right\| < \epsilon_0 / 2 \). By proper scaling of \( h_0 \) and \( l_0 \), we may assume that \( \forall t_1, t_2 \in I := [0, 1] \subset \mathbb{R}, x_0 + t_1 h_0 + t_2 l_0 \in D \cap \mathcal{B}_X (x_0, \delta_1) \).

In summary, \( \exists h_0, l_0 \in X \setminus \{ \hat{\vartheta}_X \} \) such that \( \epsilon_0 := \left\| f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0) \right\| \leq \left\| h_0 \right\| \left\| l_0 \right\| \) \((0, \infty) \subset \mathbb{R}, \forall t_1, t_2 \in I, f \) is twice differentiable at \( x_0 + t_1 h_0 + t_2 l_0 \in D \cap \mathcal{B}_X (x_0, \delta_0) \) and \( \left\| f^{(2)}(x_0 + t_1 h_0 + t_2 l_0) - f^{(2)}(x_0) \right\| < \epsilon_0 / 2 \).

Let \( D := I \), if \( K = \mathbb{R} \); or \( D := \{ a + i0 \mid a \in I \} \subset \mathbb{C} \), if \( K = \mathbb{C} \). \( \forall t_1 \in I \), define \( \psi_1 : D \to B(X, Y) \) by \( \psi_1(t_2) = f^{(1)}(x_0 + t_1 h_0 + t_2 l_0) - f^{(1)}(x_0 + t_1 h_0) \). \( \forall t_2 \in D \). By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of \( \psi_1 \) is Fréchet differentiable. Then, by Proposition 9.15, \( \psi_1 \) is Fréchet differentiable at \( t_2, \forall t_2 \in D \). By Mean Value Theorem, \( \exists \tilde{l}_2 \in I^r \) such that

\[
\left\| f^{(1)}(x_0 + t_1 h_0 + \tilde{l}_2 l_0) - f^{(1)}(x_0 + t_1 h_0) - f^{(2)}(x_0)(l_0) \right\| = \left\| \psi_1(1) - \psi_1(0) \right\| \leq \left\| D \psi_1(t_2) \right\| = \left\| f^{(2)}(x_0 + t_1 h_0 + \tilde{l}_2 l_0)(l_0) - f^{(2)}(x_0)(l_0) \right\| \leq \left\| f^{(2)}(x_0 + t_1 h_0 + \tilde{l}_2 l_0) - f^{(2)}(x_0) \right\| \left\| l_0 \right\| < \epsilon_0 \left\| l_0 \right\| / 2
\]

where the second inequality follows from Proposition 7.64.

Define \( \gamma : D \to \mathbb{R} \) by \( \gamma(t_1) = f(x_0 + l_0 + t_1 h_0) - f(x_0 + t_1 h_0) - f^{(2)}(x_0)(l_0)(t_1 h_0), \forall t_1 \in D \). By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of \( \gamma \) is Fréchet differentiable. Then, by Proposition 9.15, \( \gamma \) is Fréchet differentiable at \( t_1, \forall t_1 \in D \). By Mean Value Theorem, \( \exists \hat{l}_1 \in I^r \) such that

\[
\left\| f(x_0 + l_0 + \hat{l}_1 h_0) - f(x_0 + l_0) - f(x_0 + l_0 + \hat{l}_1 h_0) - f^{(2)}(x_0)(l_0)(h_0) \right\| = \left\| \gamma(1) - \gamma(0) \right\| \leq \left\| D \gamma(\tilde{l}_1) \right\| = \left\| f^{(1)}(x_0 + l_0 + \hat{l}_1 h_0)(h_0) \right. \\
- f^{(1)}(x_0 + \tilde{l}_1 h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0) \right\| \leq \left\| f^{(1)}(x_0 + l_0 + \hat{l}_1 h_0) - f^{(1)}(x_0 + \tilde{l}_1 h_0) - f^{(2)}(x_0)(l_0) \right\| \left\| h_0 \right\| = \left\| \psi_1(1) - \psi_1(0) \right\| \left\| l_0 \right\| < \epsilon_0 \left\| l_0 \right\| \left\| h_0 \right\| / 2
\]

where the second inequality follows from Proposition 7.64.
By symmetry, we have \( \| f(x_0 + h_0 + l_0) - f(x_0 + l_0) - f(x_0 + h_0) + f(x_0) - f(x_0)(h_0)(l_0) \| < \varepsilon_0 \| h_0 \| \| l_0 \| / 2. \) Then, \( \| f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0) \| < \varepsilon_0 \| h_0 \| \| l_0 \|. \) This leads to the contradiction \( \varepsilon_0 := \| f^{(2)}(x_0)(h_0)(l_0) - f^{(2)}(x_0)(l_0)(h_0) \| / (\| h_0 \| \| l_0 \| ) < \varepsilon_0. \)

Hence, \( f^{(2)}(x_0) \) must be symmetric and \( f^{(2)}(x_0) \in B_{S^2}(X, \mathbb{X}). \)

2° Assume that the result holds \( \forall n \leq k, k \in \{ 2, 3, \ldots \}. \)

3° Consider the case \( n = k + 1. \) \( \forall (h_1, \ldots, h_{k+1}) \in X^{k+1}, \) let \( (v_1, \ldots, v_{k+1}) \) be a permutation of \( (h_1, \ldots, h_{k+1}). \) We need to show that

\[
f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1) = f^{(k+1)}(x_0)(v_{k+1}) \cdots (v_1)
\]

Since any permutation can be arrived at in finite number of steps by interchanging two consecutive elements, then, all we need to show is that, \( \forall i = 1, \ldots, k, \)

\[
f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_{i+1})(h_i) \cdots (h_1) = f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_i)(h_{i+1}) \cdots (h_1)
\]

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( 1 \leq i \leq k; \) Case 2: \( i = k. \) Define \( g : \text{dom}(f^{(k)}) \to \mathbb{Y} \) by \( g(x) = f^{(k)}(x)(h_k) \cdots (h_{i+1})(h_i) \cdots (h_1), \forall x \in \text{dom}(f^{(k)}). \) \( \forall x \in D \cap B_X(x_0, \delta_0) = \text{dom}(f^{(k+1)}) \cap B_X(x_0, \delta_0), f \) is \( (k + 1)-\)times differentiable at \( x, \) and \( \text{dom}(f^{(k)}) \supseteq D \cap B_X(x, \delta_x) \cap B_X(x_0, \delta_0) \) is convex. By the case \( i = 1, \) \( g \) is \( C_{\delta_x} \) at \( x. \)

Clearly, \( D \cap B_X(x, \delta_2) = (D \cap B_X(x_0, \delta_0)) \cap B_X(x, \delta_x) \) is convex. By inductive assumption, \( g(x) = f^{(k)}(x)(h_k) \cdots (h_i)(h_{i+1}) \cdots (h_1). \) By Proposition 9.27, \( g \) is \( C_{\delta_x} \) at \( x_0 \) and \( f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_{i+1})(h_i) \cdots (h_1) = g^{(1)}(x_0)(h_{k+1}) = f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_i)(h_{i+1}) \cdots (h_1). \) This case is proved.

Case 2: \( i = k. \) Define \( g : \text{dom}(f^{(k+1)}) \to \mathbb{Y} \) by, \( \forall x \in \text{dom}(f^{(k+1)}), \)

\[
g(x) = f^{(k+1)}(x)(h_{k+1}) \cdots (h_1).
\]

Then, by Proposition 9.27, \( g \) is \( C_{\delta_x} \) at \( x_0 \) and \( g^{(2)}(x_0)(u)(v) = f^{(k+1)}(x_0)(u)(v)(h_{k+1}) \cdots (h_1), \forall u, v \in \mathbb{X}. \) Note that \( \text{dom}(f^{(k+1)}) \cap B_X(x, \delta_0) = D \cap B_X(x_0, \delta_0) \) is convex. By the case \( n = 2, \) we have \( f^{(k+1)}(x_0)(h_{k+1})(h_k) \cdots (h_1) = g^{(2)}(x_0)(h_{k+1})(h_k) = g^{(2)}(x_0)(h_{k+1}) = f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1). \) This case is also proved.

Hence, \( f^{(k+1)}(x_0) \in B_{S^k+1}(X, \mathbb{X}). \)

This completes the induction process and the proof of the proposition. \( \square \)

**Definition 9.29** Let \( X_1, \ldots, X_p, \) and \( \mathbb{Y} \) be normed linear spaces over \( \mathbb{K}, \)

where \( p \in \{2, 3, \ldots\}, D \subseteq X := \prod_{i=1}^p X_i, f : D \to \mathbb{Y}, \) and \( x_0 := (x_1, \ldots, x_p) \in D. \) Assume that \( \frac{\partial f}{\partial x_{i_1}} : \text{dom}\left(\frac{\partial f}{\partial x_{i_1}}\right) \to \mathbb{B}(X_{i_1}, \mathbb{Y}) \) is partial differentiable with respect to \( x_{i_2} \) at \( x_0 \in \text{dom}(\frac{\partial f}{\partial x_{i_1}}), \) where \( i_1, i_2 \in \{1, \ldots, p\}. \) Then, this partial derivative is denoted by \( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(x_0) \in B(\mathbb{X}_{i_2}, \mathbb{B}(X_{i_1}, \mathbb{Y})), \) which is one of the second order partial derivatives of
9.4. HIGHER ORDER DERIVATIVES

277

f. We will let \( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}} \) denote the \( B(\mathcal{X}_{i_2}, B(\mathcal{X}_{i_1}, \mathcal{Y})) \)-valued function whose domain of definition is \( \text{dom} \left( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}} \right) := \left\{ x \in \text{dom} \left( \frac{\partial f}{\partial x_{i_1}} \right) \mid \frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}(x) \in B(\mathcal{X}_{i_2}, B(\mathcal{X}_{i_1}, \mathcal{Y})) \right\} \). There are a total of \( p^2 \) second order partial derivatives of \( f \). Recursively, we may define \( k \)-th order partial derivatives, \( k \in \{3, 4, \ldots \} \). There are a total of \( p^k \) \( k \)-order partial derivatives. A typical such derivative is denoted by \( \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \), where \( i_1, \ldots, i_k \in \{1, \ldots, p\} \).

Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \) and \( D \subseteq \mathcal{X} \). Define \( \text{Sm} (D) := \bigcup_{n \in \mathbb{K}} nD \). Clearly, \( \text{Sm} (D) \subseteq \text{span} (D) \).

**Proposition 9.30** Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \), and \( x_0 \in \overline{D} \). Then, \( \text{span} (A_D (x_0)) \subseteq \text{span} ((D \cap B(x_0, \delta)) - x_0) \) and \( A_D (x_0) \subseteq \text{Sm} ((D \cap B(x_0, \delta)) - x_0) \). \( \forall \delta \in (0, \infty) \subseteq \mathbb{K} \).

**Proof** \( \forall u \in A_D (x_0) \), we will show that \( u \in \text{Sm} (D - x_0) \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{K} \), by \( u \in A_D (x_0) \), we have \( \exists \epsilon := x_0 + r \bar{u} \in D \) such that \( 0 < r < \epsilon \) and \( \bar{u} \in B(u, \epsilon) \). Then, \( \bar{u} \in \text{Sm} (D - x_0) \cap B(u, \epsilon) \neq \emptyset \). Hence, by Proposition 3.3, \( u \in \text{Sm} (D - x_0) \). By the arbitrariness of \( u \), we have \( A_D (x_0) \subseteq \text{Sm} (D - x_0) \subseteq \text{span} (D - x_0) \). By Proposition 7.17, \( \text{span} (A_D (x_0)) \subseteq \text{span} (D - x_0) \). By Definition 3.2, \( \text{span} (A_D (x_0)) \subseteq \text{span} (D - x_0) \). \( \forall \delta \in (0, \infty) \subseteq \mathbb{K} \), we have \( x_0 \in D \cap B(x_0, \delta) \) and \( A_D (x_0) = A_D \cap \mathbb{B}(x_0, \delta) (x_0) \). Then, \( A_D (x_0) \subseteq \text{Sm} ((D \cap B(x_0, \delta)) - x_0) \). This completes the proof of the proposition. \( \square \)

9.4.2 Interchange order of differentiation

Next, we present two results on the interchangability of order of differentiation. The first result does not assume the existence of the derivative after the interchange, which then requires a stronger assumption on the set \( D = \text{dom} (f) \). The second result assumes the existence of the derivative after the interchange, where the stronger assumption on \( D \) can be removed.

**Proposition 9.31** Let \( \mathcal{X}, \mathcal{Y} \), and \( \mathcal{Z} \) be normed linear spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \times \mathcal{Y} \), \( f : D \to \mathcal{Z} \), and \( (x_0, y_0) \in D \). Assume that the following conditions are satisfied.

(i) \( D \) is locally convex at \( (x_0, y_0) \), that is, \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \) such that \( D \cap B_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_0) =: \hat{D} \) is convex. Furthermore, \( \forall (x, y) \in \hat{D} \), we have \( (x, y_0), (x_0, y) \in \hat{D} \).

(ii) \( \frac{\partial f}{\partial y} (x, y), \frac{\partial f}{\partial x} (x, y), \) and \( \frac{\partial^2 f}{\partial x \partial y} (x, y) \) exist, \( \forall (x, y) \in \hat{D} \), and \( \frac{\partial^2 f}{\partial x \partial y} \) is continuous at \( (x_0, y_0) \).
Then, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap B_{\mathbb{R}^2}((x_0, y_0), \delta) =: \dot{D}_\delta$, we have $\left\| f(x, y) - f(x, y_0) - f(x_0, y_0) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \varepsilon \| x - x_0 \| \| y - y_0 \|$. 

Furthermore, if, in addition, the following condition is satisfied.

(iii) $\exists M \in (0, \infty) \subset \mathbb{R}$, $\forall h \in X$, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in \dot{D}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\exists h_1, h_2 \in \text{Sm}\left( (D_y \cap B_{X}(x_0, \delta)) - x_0 \right)$ such that $\| h - h_1 + h_2 \| \leq \varepsilon \| h \|$, $\| h_1 \| \leq M \| h \|$, and $\| h_2 \| \leq M \| h \|$, where $D_y := \{ x \in X \mid (x, y) \in D \}$, $\forall y \in Y$.

Then, $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$ exists and is given by

$$
\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k); \quad \forall h \in X, \forall k \in Y
$$

Proof

$\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, by the continuity of $\frac{\partial^2 f}{\partial x \partial y}$ at $(x_0, y_0)$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap B_{\mathbb{R}^2}((x_0, y_0), \delta) =: \dot{D}_\delta \subset \mathbb{D}$, we have $\left\| \frac{\partial^2 f}{\partial x \partial y}(x, y) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| < \varepsilon / 3$. Note that, by (i), $\dot{D}_\delta$ is convex and, $\forall (x, y) \in \dot{D}_\delta$, we have $(x_0, y_0), (x, y) \in \dot{D}_\delta$.

Claim 9.31.1 $\forall (x, y) \in \dot{D}_\delta$, $\left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \varepsilon / 3 \| x - x_0 \| \| y - y_0 \|$. 

Proof of claim: Fix $(x, y) \in \dot{D}_\delta$. Let $I := [0, 1] \subset \mathbb{R}$. Define $\bar{D} := I$ if $K = \mathbb{R}$; or $\bar{D} := \{ a + i0 \mid a \in I \} \subset \mathbb{C}$ if $K = \mathbb{C}$.

$\forall t_2 \in I$, define $\psi : \bar{D} \to B(Y, Z)$ by $\psi(t_2) = \frac{\partial f}{\partial y}(t_2x + (1 - t_2)y_0, t_2y + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(t_2(x - x_0))$, $\forall t_2 \in \bar{D}$.

By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of $\psi$ is differentiable. By Proposition 9.15, $\psi$ is differentiable. By Mean Value Theorem, $\exists t_2 \in I$ such that

$$
\left\| \frac{\partial f}{\partial y}(t_2x + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| = \left\| \psi(t_2)(1) - \psi(t_2)(0) \right\| \leq \| D\psi(t_2) \| \\
= \left\| \frac{\partial^2 f}{\partial x \partial y}((t_2x + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| \leq \varepsilon / 3 \| x - x_0 \|
$$

where the last inequality follows from Proposition 7.64.
9.4. HIGHER ORDER DERIVATIVES

Define $\gamma : \tilde{D} \to \mathbb{Z}$ by $\gamma(t_2) = f(x, t_2y + (1 - t_2)y_0) - f(x, t_2y + (1 - t_2)y_0) - \partial^2 f/\partial x \partial y(x_0, y_0)(x - x_0)(y - y_0)$ for all $t_2 \in \tilde{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of $\gamma$ is differentiable. By Proposition 9.15, $\gamma$ is differentiable. By Mean Value Theorem, $\exists \bar{t}_2 \in I^0$ such that

$$\| f(x, y) - f(x, y_0) - f(x_0, y_0) + f(x_0, y_0) - \partial^2 f/\partial x \partial y(x_0, y_0)(x - x_0)(y - y_0) \|
\leq \| \gamma(1) - \gamma(0) \| \leq \| D\gamma(\bar{t}_2) \|
\leq \left| \frac{\partial f}{\partial y}(x, \bar{t}_2y + (1 - \bar{t}_2)y_0) - \frac{\partial f}{\partial y}(x_0, \bar{t}_2y + (1 - \bar{t}_2)y_0)(y - y_0) \right|
- \left| \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right|
\leq \| \gamma(1) - \gamma(0) \| \leq \| \gamma(1) - \gamma(0) \| \leq \frac{\epsilon}{3} \| x - x_0 \| \| y - y_0 \|

where the second inequality follows from Proposition 7.64. This completes the proof of the claim.

Define $D_{x_0} := \{ y \in \mathbb{Y} \mid (x_0, y) \in \tilde{D} \}$; and $\tilde{D} := \{ y \in \mathbb{Y} \mid (x_0, y) \in \text{dom} \left( \frac{\partial f}{\partial x} \right) \}$. By (ii), we have span $\left( A_{D_{x_0}}(y_0) \right) = \mathbb{Y}$ and span $\left( A_{D_y}(x) \right) = \mathbb{X}$, $\forall (x, y) \in \tilde{D}$. Clearly, we have $\tilde{D} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta_0) = D_{x_0} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta_0)$. This implies that span $\left( A_{\tilde{D}}(y_0) \right) = \mathbb{Y}$.

$\forall k \in \mathbb{Y}$, define $L_k : \mathbb{X} \to \mathbb{Z}$ by $L_k(h) = \partial^2 f/\partial x \partial y(x_0, y_0)(h)(k)$, $\forall h \in \mathbb{X}$. Clearly, $L_k$ is a linear operator. Note that $\| L_k \| = \sup_{h \in \mathbb{X}, \| h \| \leq 1} \| L_k(h) \| \leq \| \partial^2 f/\partial x \partial y(x_0, y_0) \| \| k \| < +\infty$, where the first inequality follows from Proposition 7.64. Hence, $L_k \in B(\mathbb{X}, \mathbb{Z})$.

Define $L : \mathbb{Y} \to B(\mathbb{X}, \mathbb{Z})$ by $L(k) = L_k$, $\forall k \in \mathbb{Y}$. Clearly, $L$ is a linear operator. Note that $\| L \| = \sup_{k \in \mathbb{Y}, \| k \| \leq 1} \| L(k) \| \leq \| \partial^2 f/\partial x \partial y(x_0, y_0) \| < +\infty$. Hence, $L \in B(\mathbb{Y}, B(\mathbb{X}, \mathbb{Z}))$.

Now, fix any $y \in \tilde{D} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta/\sqrt{2})$, we will show that $\| \Delta_y \| \leq (2M + 1)\epsilon \| y - y_0 \|$, where $\Delta_y := \partial f/\partial x(x_0, y) - \partial f/\partial x(x_0, y_0) - L(y - y_0) \in B(\mathbb{X}, \mathbb{Z})$. This immediately implies that $\partial^2 f/\partial y \partial x(x_0, y_0) = L$ and completes the proof of the proposition.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $y = y_0$; Case 2: $y \neq y_0$. Case 1: $y = y_0$. The result is immediate. Case 2: $y \neq y_0$. By the existence of $\partial f/\partial x(x_0, y)$, $\exists \delta_{y_1} \in (0, \delta/\sqrt{2}) \subset \mathbb{R}$ such that $\forall x \in D_y \cap \mathcal{B}_{\mathbb{X}}(x_0, \delta_{y_1})$, we have $\| f(x, y) - f(x_0, y_0) - \partial f/\partial x(x_0, y)(x - x_0) \| \leq$
Proposition 9.32

Let $D$ be the set $\mathbb{R}^n$ such that $\forall x \in D_{y_0} \cap \mathcal{B}_X (x_0, \delta y_2)$, we have $\| f(x, y_0) - f(x_0, y_0) - \partial f(x, y_0)(x - x_0) \| \leq \epsilon/3 \| y - y_0 \| \| x - x_0 \|$. By the existence of $\frac{\partial f}{\partial x}(x_0, y_0)$, $\exists \delta y_2 \in (0, \delta/\sqrt{2}] \subset \mathbb{R}$ such that $\forall x \in D_{y_0} \cap \mathcal{B}_X (x_0, \delta y_2)$, we have $\| f(x, y_0) - f(x_0, y_0) - \partial f(x_0, y_0)(x - x_0) \| \leq \epsilon/3 \| y - y_0 \| \| x - x_0 \|$. Let $\delta y = \min \{ \delta y_1, \delta y_2 \} \in (0, \delta/\sqrt{2}] \subset \mathbb{R}$. $\forall x \in D_{y_0} \cap \mathcal{B}_X (x_0, \delta y)$, we have $(x, y), (x, y_0), (x_0, y) \in \tilde{D}_\delta$. Then, $x \in D_{y_0} \cap \mathcal{B}_X (x_0, \delta y)$. By Claim 9.31.1 and the preceding argument, we have

$$\| \Delta_y(x - x_0) \| = \left\| \frac{\partial f}{\partial x}(x_0, y)(x - x_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) - L(y - y_0)(x - x_0) \right\|$$

$$\leq \left\| \frac{\partial f}{\partial x}(x_0, y)(x - x_0) - f(x, y) + f(x_0, y) \right\| + \left\| f(x, y) - f(x_0, y) \right\|$$

$$- \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\| + \left\| f(x, y) - f(x_0, y) + f(x_0, y) \right\|$$

$$\leq \epsilon \| x - x_0 \| \| y - y_0 \|$$

$\forall h \in X$, by (iii) and the continuity of $\Delta_y$, $\exists h_1, h_2 \in \text{Sm} \left( \left( D_{y_0} \cap \mathcal{B}_X (x_0, \delta y) \right) - x_0 \right)$, such that $\| \Delta_y(h - h_1 + h_2) \| \leq \epsilon \| h \| \| y - y_0 \|$, $\| h_1 \| \leq M \| h \|$, and $\| h_2 \| \leq M \| h \|$. Then, $\exists \alpha_1, \alpha_2 \in \mathbb{K}$ and $\exists x_1, x_2 \in D_{y_0} \cap \mathcal{B}_X (x_0, \delta y)$ such that $h_i = \alpha_i (x_i - x_0)$, $i = 1, 2$. Then, we have

$$\| \Delta_y(h) \| \leq \epsilon \| h_1 \| \| y - y_0 \|$$

$$\leq \| \Delta_y(x_1 - x_0) \| + \epsilon \| h_1 \| \| y - y_0 \|$$

Then, we have $\| \Delta_y \| \leq (2M + 1)\epsilon \| y - y_0 \|$. This case is proved.

This completes the proof of the proposition. \( \Box \)

In the preceding proposition, conditions (i) and (iii) are assumptions on the set $D$. It is clear that (i) and (iii) are satisfied if $(x_0, y_0) \in D^0$.

Proposition 9.32 Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $D \subseteq X \times Y$, $f : D \to Z$, and $(x_0, y_0) \in D$. Assume that the following conditions are satisfied.

(i) $D$ is locally convex at $(x_0, y_0)$, that is, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap \mathcal{B}_{X \times Y}((x_0, y_0), \delta_0) =: \tilde{D}$ is convex. Furthermore, $\forall (x, y) \in \tilde{D}$, we have $(x, y_0), (x_0, y) \in \tilde{D}$.

(ii) $\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y), \frac{\partial^2 f}{\partial x^2}(x, y), \frac{\partial^2 f}{\partial x \partial y}(x, y)$, and $\frac{\partial^2 f}{\partial y^2}(x, y)$ exist, $\forall (x, y) \in \tilde{D}$, and $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial y^2}$ are continuous at $(x_0, y_0)$. 

9.4. HIGHER ORDER DERIVATIVES

Then, \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k) \), \( \forall h \in \mathcal{X}, \forall k \in \mathbb{Y} \).

**Proof** Define \( D_{x_0} := \{ y \in \mathbb{Y} \mid (x_0, y) \in D \} \) and \( D_{y_0} := \{ x \in \mathcal{X} \mid (x, y_0) \in D \} \). By (ii), we have \( \text{span}(A_{D_{x_0}}(y_0)) = \mathbb{Y} \) and \( \text{span}(A_{D_{y_0}}(x_0)) = \mathcal{X} \).

For every \( \varepsilon \in (0, \infty) \subset \mathbb{R} \), by Proposition 9.31, \( \exists \delta_1 \in (0, \delta_0] \subset \mathbb{R} \) such that \( \forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathbb{Y}}((x_0, y_0), \delta_1) \), we have \( f(x, y) - f(x_0, y_0) - f(x_0, y_0) - \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(x - x_0)(y - y_0) \leq \varepsilon/2 \| x - x_0 \| \| y - y_0 \| \).

By symmetry, \( \exists \delta_2 \in (0, \delta_0] \subset \mathbb{R} \) such that \( \forall (x, y) \in \mathcal{B}_{\mathcal{X} \times \mathbb{Y}}((x_0, y_0), \delta_2) \cap D =: \tilde{D}_{\delta_2} \subset \tilde{D} \), we have \( f(x, y) - f(x_0, y_0) - f(x_0, y_0) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(y - y_0)(x - x_0) \leq \varepsilon/2 \| x - x_0 \| \| y - y_0 \| \).

Therefore, let \( \delta := \min \{ \delta_1, \delta_2 \} \in (0, \delta_0] \subset \mathbb{R} \), \( \forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathbb{Y}}((x_0, y_0), \delta) =: \tilde{D} = \tilde{D}_{\delta_1} \cap \tilde{D}_{\delta_2} \subset \tilde{D} \), we have \( \| \partial f \|_{\mathcal{X} \times \mathbb{Y}}((x_0, y_0), \delta_0) \subset \mathcal{D}_{\delta_0} \cap \mathcal{D}_{\delta_2} \subset \mathcal{D}. \)

By the arbitrariness of \( \varepsilon \), we have \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(y - y_0)(x - x_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0). \)

Clearly, by (i), \( \tilde{D} \) is convex.

Suppose that the result of the proposition is not true, then \( \exists h_0 \in \mathcal{X} \) and \( \exists k_0 \in \mathbb{Y} \) such that \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k_0)(h_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h_0)(k_0) \). By the fact that \( \text{span}(A_{D_{x_0}}(y_0)) = \mathbb{Y} \) and \( \text{span}(A_{D_{y_0}}(x_0)) = \mathcal{X} \) and the continuity of \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \) and \( \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \), \( \exists h_0 \in \text{span}(A_{D_{x_0}}(x_0)) \) and \( \exists k_0 \in \text{span}(A_{D_{y_0}}(y_0)) \) such that \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k_0)(h_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h_0)(k_0) \).

By the multi-linearity of these two second-order partial derivatives, \( \exists h_0 \in A_{D_{x_0}}(x_0) \) and \( \exists k_0 \in A_{D_{y_0}}(y_0) \) such that \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(h_0)(k_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h_0)(k_0) \). By Proposition 9.30, we have

\[
A_{D_{x_0}}(x_0) \subseteq \text{Sm}(D_{x_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) - x_0)
\]

\[
A_{D_{y_0}}(y_0) \subseteq \text{Sm}(D_{y_0} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta) - y_0)
\]

This implies that, by the continuity of these two second-order partial derivatives, \( \exists h_0 \in \text{Sm}(D_{y_0} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta) - y_0) \) and \( \exists k_0 \in \text{Sm}(D_{x_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta) - x_0) \). Then, \( \exists \bar{x} \in D_{y_0} \cap \mathcal{B}_{\mathbb{Y}}(y_0, \delta), \exists \bar{y} \in D_{x_0} \cap \mathcal{B}_{\mathcal{X}}(x_0, \delta), \) and \( \exists h_0, \alpha_k \in \mathbb{K} \) such that \( h_0 = \alpha_h (\bar{x} - x_0) \) and \( k_0 = \alpha_k (\bar{y} - y_0) \). By the multi-linearity of these two second-order partial derivatives, \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\bar{y} - y_0)(\bar{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\bar{x} - x_0)(\bar{y} - y_0) \). Clearly, \( (\bar{x}, \bar{y}), (x_0, y_0) \in \tilde{D}_{\delta} \). By the convexity of \( \tilde{D}_{\delta} \), we have \( (\bar{x}, \bar{y}) := 0.5 (\bar{x}, \bar{y}) + 0.5 (x_0, y_0) \in \tilde{D}_{\delta} \). By the multi-linearity of these two second-order partial derivatives, we have
\[
\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\dot{y} - y_0)(\dot{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\dot{x} - x_0)(\dot{y} - y_0). \quad \text{This is a contradiction. Hence, the result of the proposition must hold.}
\]

This completes the proof of the proposition. \qed

### 9.4.3 High order derivatives of some common functions

**Proposition 9.33** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D \subseteq X \), \( f : D \to Y \), and \( x_0 \in D \). Assume that \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \) and \( \exists \gamma_0 \in Y \) such that \( f(x) = y_0 \) and \( \text{span}(A_D(x)) = X \), \( \forall x \in D \cap B_X(x_0, \delta_0) \). Then, \( f \) is \( C_\infty \) at \( x_0 \) and \( f^{(i)}(x) = \vartheta_{B_{S_i}(X, y)}, \forall x \in D \cap B_X(x_0, \delta_0), \forall i \in \mathbb{N} \).

**Proof** \( \forall x \in D \cap B_X(x_0, \delta_0), \) let \( \delta_x := \delta_0 - \|x - x_0\| > 0 \). Then, \( \forall \bar{x} \in D \cap B_X(x_0, \delta_x) \subseteq D \cap B_X(x_0, \delta_0) \), we have \( f(\bar{x}) = y_0 \). Note that \( \text{span}(A_{D}(x)) = X \). By Proposition 9.10, \( f^{(i)}(x) = \vartheta_{B_{S_i}(X, y)} \). Let \( D_1 := \text{dom}(f^{(1)}) \). Then, \( D_1 \cap B_X(x_0, \delta_0) = D \cap B_X(x_0, \delta_0) \). Then, \( \text{span}(A_{D_1}(x)) = X \), \( \forall x \in D_1 \cap B_X(x_0, \delta_0) \). By recursively applying the above argument and Proposition 9.28, we have \( f^{(i)}(x) = \vartheta_{B_{S_i}(X, y)}, \forall x \in D \cap B_X(x_0, \delta_0), \forall i \in \mathbb{N} \). This completes the proof of the proposition. \qed

**Proposition 9.34** Let \( X \) and \( Y \) be normed linear spaces over \( K \), \( D_2 \subseteq D_1 \subseteq X \), \( x_0 \in D_2 \), \( f : D_1 \to Y \), \( g := f|_{D_2}, k \in \mathbb{N} \), and \( n \in \mathbb{N} \cup \{\infty\} \). Then, the following statements holds.

(i) If, \( \forall x \in D_2, f^{(k)}(x) \) exists and \( \text{span}(A_{D_2}(x)) = X \), then \( g \) is \( k \)-times Fréchet differentiable and \( g^{(i)}(x) = f^{(i)}(x), \forall x \in D_2, \forall i \in \{1, \ldots, k\} \).

(ii) If, \( \forall x \in D_2, g^{(k)}(x) \) exists and \( \exists \delta_x \in (0, \infty) \subseteq \mathbb{R} \) such that \( D_1 \cap B_X(x, \delta_x) = D_2 \cap B_X(x, \delta_x) \), then \( f^{(k)}(x) \) is exists and \( f^{(i)}(x) = g^{(i)}(x), \forall x \in D_2, \forall i \in \{1, \ldots, k\} \).

(iii) If \( f \) is \( C_n \) at \( x_0 \) and \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \text{span}(A_{D_2}(x)) = X \), \( \forall x \in D_2 \cap B_X(x_0, \delta), \) then, \( g \) is \( C_n \) at \( x_0 \).

(iv) If \( g \) is \( C_n \) at \( x_0 \) and \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( D_1 \cap B_X(x_0, \delta) = D_2 \cap B_X(x_0, \delta) \), then \( f \) is \( C_n \) at \( x_0 \).

**Proof** (i) We will use mathematical induction on \( k \) to prove this statement.

1° \( k = 1 \). \( \forall x \in D_2 \), by Proposition 9.11, we have \( g^{(1)}(x) \) exists and \( g^{(1)}(x) = f^{(1)}(x) \). Hence, the result holds in this case.

2° Assume that the result holds for \( k \leq \bar{k} \in \mathbb{N} \).

3° Consider the case \( k = \bar{k} + 1 \). By inductive assumption, \( g^{(\bar{k})}(x) \) exists and \( g^{(1)}(x) = f^{(1)}(x), \forall x \in D_2, \forall i \in \{1, \ldots, \bar{k}\} \). Then, \( \text{dom}(g^{(\bar{k})}) = \text{dom}(f^{(\bar{k})}) \)
Case 1: \( n \in \mathbb{N} \). Without loss of generality, assume \( f \) is \( n \)-times differentiable at \( x \), \( \forall x \in D_2 \). Let \( D_2 := D_2 \cap B_\mathcal{X}(x_0, \delta) \) and \( g := f|_{D_2} \). Let \( \delta_x := \delta - \|x - x_0\| \in (0, \infty) \subset \mathbb{R} \). Then, \( D_2 \cap B_\mathcal{X}(x, \delta_x) = D_2 \cap B_\mathcal{X}(x, \delta_x) \). Hence, \( \text{span}(A_{D_2}(x)) = \mathcal{X} \). Then, by (i), \( g \) is \( n \)-times differentiable and \( \tilde{g}^{(i)}(x) = f^{(i)}(x) \), \( \forall x \in \tilde{D}_2 \). By (ii), \( g^{(n)}(x) \) exists and \( \tilde{g}^{(i)}(x) = f^{(i)}(x) \), \( \forall x \in \tilde{D}_2 \). By the continuity of \( f^{(n)} \) at \( x_0 \), \( \forall x \in \text{dom}(f^{(n)}) \cap B_\mathcal{X}(x_0, \delta) = D_1 \cap B_\mathcal{X}(x_0, \delta) \), we have \( \|f^{(n)}(x) - f^{(n)}(x_0)\| < \varepsilon \). Therefore, \( g \) is \( C_n \) at \( x_0 \).

Case 2: \( n = \infty \). By the assumption, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( f \) is \( C_\mathcal{C} \) at \( x \), \( \forall x \in D \cap B_\mathcal{X}(x_0, \delta) \). By Case 1, \( g \) is \( C_\mathcal{C} \) at \( x \), \( \forall x \in D \cap B_\mathcal{X}(x_0, \delta) \), \( \forall \delta \in \mathbb{N} \). Hence, \( g \) is \( C_\infty \) at \( x_0 \).

(iv) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: \( n \in \mathbb{N} \). Without loss of generality, assume \( g \) is \( n \)-times differentiable at \( x \), \( \forall x \in D_2 \). Let \( \tilde{g} := g|_{D_2} \). Then, \( \tilde{g}^{(i)}(x) = g^{(i)}(x), \forall x \in \tilde{D}_2 \), \( \forall i \in \{1, \ldots, n\} \). By (ii), \( f^{(n)}(x) \) exists and \( f^{(i)}(x) = \tilde{g}^{(i)}(x) = g^{(i)}(x), \forall x \in \tilde{D}_2 \), \( \forall i \in \{1, \ldots, n\} \). By the continuity of \( g^{(n)} \) at \( x_0 \), \( \forall x \in \text{dom}(g^{(n)}) \cap B_\mathcal{X}(x_0, \delta) = D_1 \cap B_\mathcal{X}(x_0, \delta) \), we have \( \|f^{(n)}(x) - f^{(n)}(x_0)\| < \varepsilon \).
\[\|g^{(n)}(x) - g^{(n)}(x_0)\| < \epsilon, \forall x \in \text{dom} \left( f^{(n)} \right) \cap B_X(x_0, \delta) = D_1 \cap B_X(x_0, \delta),\]
we have \[\|f^{(n)}(x) - f^{(n)}(x_0)\| = \|g^{(n)}(x) - g^{(n)}(x_0)\| < \epsilon.\] Hence, \(f^{(n)}\) is continuous at \(x_0\). Therefore, \(f\) is \(C_\infty\) at \(x_0\).

Case 2: \(n = \infty\). By the assumption, \(g\) is \(C_i\) at \(x\), \(\forall x \in \mathcal{D}_2, \forall i \in \mathbb{N}\). By Case 1, \(f\) is \(C_i\) at \(x\), \(\forall x \in \mathcal{D}_2, \forall i \in \mathbb{N}\). Hence, \(f\) is \(C_\infty\) at \(x_0\).

This completes the proof of the proposition. \(\Box\)

**Proposition 9.35** Let \(X\) and \(Y\) be normed linear spaces over \(\mathbb{K}\), \(D \subseteq X\), \(f : D \to Y\). Assume that there exist open sets \(O_1, O_2 \subseteq X\) such that \(D_1 := D \cap O_1\) and \(D_2 := D \cap O_2\) satisfy \(D_1 \cup D_2 = D\) and \(f|_{D_1}\) and \(f|_{D_2}\) are \(C_k\), where \(k \in \mathbb{N} \cup \{\infty\}\). Then, \(f\) is \(C_k\).

**Proof** \(\forall x \in D\), without loss of generality, assume that \(x \in D_1\). Then, \(\exists \delta \in (0, \infty) \subset \mathbb{R}\) such that \(B_X(x, \delta) \subseteq O_1\), which implies that \(D_1 \cap B_X(x, \delta) = D \cap B_X(x, \delta)\). Since \(f|_{D_1}\) is \(C_k\), then, by Proposition 9.34, \(\forall x \in D_1, f^{(i)}(x) = f^{(i)}(x), \forall i \in \mathbb{N}\) with \(i \leq k\). Then, \(f\) is \(C_k\) at \(x\). By the arbitrariness of \(x\), \(f\) is \(C_k\). This completes the proof of the proposition. \(\Box\)

Note that \(D_1\) and \(D_2\) are open sets in the subset topology of \(D\). This result should be compared with Theorem 3.11, where continuity on \(D\) can be concluded when \(D_1\) and \(D_2\) are relatively open or are relatively closed. For continuous differentiability, \(D_1\) and \(D_2\) must be relatively open for the conclusion to hold.

**Proposition 9.36** Let \(X\) be a normed linear space over \(\mathbb{K}\), and \(f : X \to X\) be given by \(f = \text{id}_X\), that is \(f(x) = x, \forall x \in X\). Then, \(f\) is \(C_\infty\), \(f^{(1)}(x) = \text{id}_X, \text{ and } f^{(i+1)}(x) = \partial_{B_{i+1}(x,x)}\), \(\forall x \in X, \forall i \in \mathbb{N}\).

**Proof** This is straightforward, and is therefore omitted. \(\Box\)

**Proposition 9.37** Let \(X\) and \(Y\) be normed linear spaces over \(\mathbb{K}\), and \(f : X \times Y \to X\) be given by \(f = \pi_X\), that is \(f(x, y) = x, \forall (x, y) \in X \times Y\). Then, \(f\) is \(C_\infty\), \(f^{(1)}(x, y) = [\text{id}_X \partial_{B(y,x)}]\) and \(f^{(i+1)}(x, y) = \partial_{B_{i+1}(x,y,x)}\), \(\forall (x, y) \in X \times Y, \forall i \in \mathbb{N}\).

**Proof** This is straightforward, and is therefore omitted. \(\Box\)

**Proposition 9.38** Let \(X\) be a normed linear space over \(\mathbb{K}\), \(f : X \times X \to X\) be given by \(f(x_1, x_2) = x_1 + x_2, \forall (x_1, x_2) \in X \times X\). Then, \(f\) is \(C_\infty\), \(f^{(1)}(x_1, x_2) = [\text{id}_X \text{id}_X]\), and \(f^{(i+1)}(x_1, x_2) = \partial_{B_{i+1}(x_1,x,x)}\), \(\forall (x_1, x_2) \in X \times X, \forall i \in \mathbb{N}\).

**Proof** This is straightforward, and is therefore omitted. \(\Box\)

**Proposition 9.39** Let \(X\) be a normed linear space over \(\mathbb{K}\), \(f : K \times X \to X\) be given by \(f(\alpha, x) = \alpha x, \forall (\alpha, x) \in K \times X\). Then, \(f\) is \(C_\infty\), \(f^{(1)}(\alpha, x) = [x \alpha \text{id}_X]\), \(f^{(2)}(\alpha, x)(d_2, h_2)(d_1, h_1) = d_1 h_2 + d_2 h_1\), and \(f^{(i+2)}(\alpha, x) = \partial_{B_{i+2}(\mathbb{K} \times X, x)}\), \(\forall (\alpha, x) \in K \times X, \forall i \in \mathbb{N}, \forall (d_1, h_1) \in K \times X, \forall (d_2, h_2) \in K \times X\).
9.4. HIGHER ORDER DERIVATIVES

Proof This is straightforward, and is therefore omitted. □

Proposition 9.40 Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $f_1 : D \to Y$, $f_2 : D \to Y$, $x_0 \in D$, $\alpha_1, \alpha_2 \in K$, $n \in \mathbb{N}$, $k \in \mathbb{N} \cup \{\infty\}$, and $g : D \to Y$ be given by $g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$, $\forall x \in D$. If $f_1$ and $f_2$ are $n$-times differentiable, then, $g$ is $n$-times differentiable and $g^{(i)}(x) = \alpha_1 f_1^{(i)}(x) + \alpha_2 f_2^{(i)}(x)$, $\forall x \in D$, $\forall i \in \{1, \ldots, n\}$. If $f_1$ and $f_2$ are $C_k$ at $x_0$, then $g$ is $C_k$ at $x_0$ and $g^{(i)}(x) = \alpha_1 f_1^{(i)}(x) + \alpha_2 f_2^{(i)}(x)$, $\forall x \in D \cap B(x_0, \delta_0)$, $\forall i \in \mathbb{N}$ with $i \leq k$, for some $\delta_0 \in (0, \infty) \subset \mathbb{R}$.

Proof This is straightforward, and is therefore omitted. □

Proposition 9.41 Let $X$ and $Y$ be normed linear spaces over $K$, $f : B(X, Y) \times X \to Y$ be given by $f(A, x) = Ax$, $\forall (A, x) \in B(X, Y) \times X$. Then, $f$ is $C_\infty$, $f^{(1)}(A, x) = \begin{bmatrix} rot(A) & A \end{bmatrix}$, $f^{(2)}(A, x)(\Delta_2, h_2)(\Delta_1, h_1) = \Delta_1 h_2 + \Delta_2 h_1$, and $f^{(i+2)}(A, x) = \partial_{B_{i+2}}(B(X, Y) \times X, Y)$, $\forall (A, x) \in B(X, Y) \times X$, $\forall i \in \mathbb{N}$, $\forall (\Delta_1, h_1) \in B(X, Y) \times X$, $\forall (\Delta_2, h_2) \in B(X, Y) \times X$.

Proof This is straightforward, and is therefore omitted. □

Proposition 9.42 Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $f : B(Y, Z) \times B(X, Y) \to B(X, Z)$ be given by $f(A_{yz}, A_{xy}) = A_{yz} A_{xy}$, $\forall (A_{yz}, A_{xy}) \in B(Y, Z) \times B(X, Y)$. Then, $f$ is $C_\infty$, $f^{(1)}(A_{yz}, A_{xy}) = \begin{bmatrix} rot(A_{xy}) & A_{yz} \end{bmatrix}$, $f^{(2)}(A_{yz}, A_{xy})(\Delta_{yz2}, \Delta_{xy2})(\Delta_{yz1}, \Delta_{xy1}) = \Delta_{yz1} \Delta_{xy2} + \Delta_{yz2} \Delta_{xy1}$, and $f^{(i+2)}(A_{yz}, A_{xy}) = \partial_{B_{i+2}}(B(Y, Z) \times B(X, Y), B(X, Z))$, $\forall (A_{yz}, A_{xy}) \in B(Y, Z) \times B(X, Y)$, $\forall i \in \mathbb{N}$, $\forall (\Delta_{yz1}, \Delta_{xy1}) \in B(Y, Z) \times B(X, Y)$, $\forall (\Delta_{yz2}, \Delta_{xy2}) \in B(Y, Z) \times B(X, Y)$.

Proof This is straightforward, and is therefore omitted. □

Proposition 9.43 Let $X_1, \ldots, X_p$ and $Y_1, \ldots, Y_m$ be normed linear spaces over $K$, where $p, m \in \mathbb{N}$, $Z_{ji} := B(X_i, Y_j)$, $i = 1, \ldots, p$, $j = 1, \ldots, m$, $Z := \prod_{i=1}^p \prod_{j=1}^m Z_{ji}$, and $f : Z \to B\left(\prod_{i=1}^p X_i, \prod_{j=1}^m Y_j\right)$ be given by $f(A_{11}, \ldots, A_{mp}) = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mp} \end{bmatrix}$, $\forall (A_{11}, \ldots, A_{mp}) \in Z$.

Then, $f$ is $C_\infty$ and $f^{(1)}(A_{11}, \ldots, A_{mp})(\Delta_{11}, \ldots, \Delta_{mp}) = \begin{bmatrix} \Delta_{11} & \cdots & \Delta_{1p} \\ \vdots & \ddots & \vdots \\ \Delta_{m1} & \cdots & \Delta_{mp} \end{bmatrix}$, $f^{(i+1)}(A_{11}, \ldots, A_{mp}) = \partial_{B_{i+1}}(Z, B(\prod_{i=1}^p X_i, \prod_{j=1}^m Y_j))$, $\forall (A_{11}, \ldots, A_{mp}) \in Z$, $\forall i \in \mathbb{N}$, $\forall (\Delta_{11}, \ldots, \Delta_{mp}) \in Z$.

Proof This is straightforward, and is therefore omitted. □
9.4.4 Properties of high order derivatives

**Proposition 9.44** Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $D \subseteq X$, $f_1 : D \to Y$, $f_2 : D \to Z$, $x_0 \in D$, $k \in \mathbb{N}$, and $g : D \to Y \times Z$ be given by $g(x) = (f_1(x), f_2(x))$, $\forall x \in D$. Then, the following statements hold.

(i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f_1^{(k)}(x)$ and $f_2^{(k)}(x)$ exist, $\forall x \in D \cap B_{X}(x_0, \delta_0)$ if, and only if, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(k)}(x)$ exists, $\forall x \in D \cap B_{X}(x_0, \delta_0)$. In this case $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_{X}(x_0, \delta_0)$, $\forall i \in \{1, \ldots, k\}$.

(ii) Let $n \in \mathbb{N} \cup \{\infty\}$. Then, $f_1$ and $f_2$ are $C_n$ at $x_0$ if, and only if, $g$ is $C_n$ at $x_0$.

**Proof** (i) We will use mathematical induction on $k$ to prove this statement.

1° $k = 1$. The statement holds by Proposition 9.19.

2° Assume that the result holds for $k = \bar{k} \in \mathbb{N}$.

3° Consider the case $k = \bar{k} + 1$. “Necessity” Let $f_1^{(\bar{k} + 1)}(x)$ and $f_2^{(\bar{k} + 1)}(x)$ exist, $\forall x \in D \cap B_{X}(x_0, \delta_0)$. By inductive assumption, $g^{(\bar{k})}(x)$ exists and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_{X}(x_0, \delta_0)$, $\forall i \in \{1, \ldots, \bar{k}\}$. Let $\bar{D}_1 := \text{dom} \left( f_1^{(\bar{k})} \right)$, $\bar{D}_2 := \text{dom} \left( f_2^{(\bar{k})} \right)$ and $\bar{D} := \text{dom} \left( g^{(\bar{k})} \right)$. Then, $\bar{D}_1 \subseteq D$, $\bar{D}_2 \subseteq D$, $\bar{D} \supseteq D \cap B_{X}(x_0, \delta_0)$, and $D \cap B_{X}(x_0, \delta_0) \subseteq \bar{D}_1 \cap \bar{D}_2$. This implies that $D \cap B_{X}(x_0, \delta_0) = D \cap B_{X}(x_0, \delta_0) =: \bar{D}$ and $\text{span} (A_{\bar{D}}(x)) = \text{span} (A_{\bar{D}}(x)) = X$, $\forall x \in \bar{D}$. By the assumption, $f_1^{(\bar{k})}$ and $f_2^{(\bar{k})}$ are differentiable at $x$, $\forall x \in \bar{D}$. By Propositions 9.19 and 9.11, $\forall x \in \bar{D}$, $g^{(\bar{k})}$ is differentiable at $x$ and $g^{(\bar{k} + 1)}(x) = D \frac{d}{dx} g^{(\bar{k})}(x) = \begin{bmatrix} D f_1^{(\bar{k})} \bigg|_{\bar{D}} (x) \\ D f_2^{(\bar{k})} \bigg|_{\bar{D}} (x) \end{bmatrix}$.

“Sufficiency” Let $g^{(\bar{k} + 1)}(x)$ exist, $\forall x \in D \cap B_{X}(x_0, \delta_0)$. By inductive assumption, $f_1^{(\bar{k})}(x)$ and $f_2^{(\bar{k})}(x)$ exist and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_{X}(x_0, \delta_0)$, $\forall i \in \{1, \ldots, \bar{k}\}$. Let $\bar{D}_1 := \text{dom} \left( f_1^{(\bar{k})} \right)$, $\bar{D}_2 := \text{dom} \left( f_2^{(\bar{k})} \right)$ and $\bar{D} := \text{dom} \left( g^{(\bar{k})} \right)$. Then, $\bar{D}_1 \subseteq D$, $\bar{D}_2 \subseteq D$, $\bar{D} \supseteq D \cap B_{X}(x_0, \delta_0)$, and $D \cap B_{X}(x_0, \delta_0) \subseteq \bar{D}_1 \cap \bar{D}_2$. This implies that $D \cap B_{X}(x_0, \delta_0) = D \cap B_{X}(x_0, \delta_0) =: \bar{D}$ and $\text{span} (A_{\bar{D}}(x)) = \text{span} (A_{\bar{D}}(x)) = X$, $\forall x \in \bar{D}$. By Propositions 9.11 and 9.19, we have, $\forall x \in \bar{D}$, $f_1^{(\bar{k})} \bigg|_{\bar{D}}$ and $f_2^{(\bar{k})} \bigg|_{\bar{D}}$ are different-
9.4. **HIGHER ORDER DERIVATIVES**

Differentiable at $x$ and $g^{(k+1)}(x) = \frac{D g^{(k)}}{D} (x) = \begin{bmatrix} D f_1^{(k)}(x) \\ D f_2^{(k)}(x) \end{bmatrix}$. By Proposition 9.11, we have $D f_1^{(k)}(x) = D f_1^{(k+1)}(x)$ and $D f_2^{(k)}(x) = D f_2^{(k+1)}(x)$, $\forall x \in \tilde{D}$. Thus, $g^{(k+1)}(x) = \begin{bmatrix} f_1^{(k+1)}(x) \\ f_2^{(k+1)}(x) \end{bmatrix}$, $\forall x \in \tilde{D}$.

This completes the induction process.

(ii) We will distinguish two exhaustive and mutually exclusive cases:

**Case 1: $n \in \mathbb{N}$**

Case 1: $n \in \mathbb{N}$. “Sufficiency” Let $g$ be $C_n$ at $x_0$. Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $g$ is $n$-times differentiable at $x$, $\forall x \in D \cap B(x_0, \delta_0) =: \tilde{D}$. By (i), $f_1$ and $f_2$ are $n$-times differentiable at $x$, $\forall x \in \tilde{D}$ and $g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix}$, $\forall x \in \tilde{D}$. By the continuity of $g^{(n)}$ at $x_0$ and Proposition 3.32, $f_1^{(n)}$ and $f_2^{(n)}$ are continuous at $x_0$. Then, $f_1$ and $f_2$ are $C_n$ at $x_0$.

“Necessity” Let $f_1$ and $f_2$ be $C_n$ at $x_0$. Then, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $f_1$ and $f_2$ are $n$-times differentiable at $x$, $\forall x \in D \cap B(x_0, \delta_0) =: \tilde{D}$. By (i), $g$ is $n$-times differentiable at $x$, $\forall x \in \tilde{D}$ and $g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix}$, $\forall x \in \tilde{D}$. By the continuity of $f_1^{(n)}$ and $f_2^{(n)}$ at $x_0$ and Proposition 3.32, $g^{(n)}$ is continuous at $x_0$. Then, $g$ is $C_n$ at $x_0$.

**Case 2: $n = \infty$**

“Sufficiency” Let $g$ be $C_\infty$ at $x_0$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $g$ is $C_\infty$ at $x$, $\forall x \in D \cap B(x_0, \delta)$, $\forall i \in \mathbb{N}$. By Case 1, $f_1$ and $f_2$ are $C_i$ at $x$, $\forall x \in D \cap B(x_0, \delta)$, $\forall i \in \mathbb{N}$. Then, $f_1$ and $f_2$ are $C_\infty$ at $x_0$.

“Necessity” Let $f_1$ and $f_2$ be $C_\infty$ at $x_0$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f_1$ and $f_2$ are $C_i$ at $x$, $\forall x \in D \cap B(x_0, \delta)$, $\forall i \in \mathbb{N}$. By Case 1, $g$ is $C_i$ at $x$, $\forall x \in D \cap B(x_0, \delta)$, $\forall i \in \mathbb{N}$. Then, $g$ is $C_\infty$ at $x_0$.

This completes the proof of the proposition.

**Proposition 9.45** Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $D_1 \subseteq X$, $D_2 \subseteq Y$, $f : D_1 \to D_2$, $g : D_2 \to Z$, $x_0 \in D_1$, and $y_0 := f(x_0) \in D_2$. Then, the following statements hold.

(i) Assume that $f$ is $C_k$ at $x_0$ and $g$ is $C_k$ at $y_0$, for some $k \in \mathbb{N} \cup \{\infty\}$.

Then, $h := g \circ f$ is $C_k$ at $x_0$.

(ii) Let $k \in \mathbb{N}$. Assume that $f$ is $k$-times differentiable and $g$ is $k$-times differentiable. Then, $h$ is $k$-times differentiable.

**Proof** (i) We will first use mathematical induction on $k$ to show that the result holds if $k \in \mathbb{N}$.

1° $k = 1$. By $g$ being $C_1$ at $y_0$, then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(1)}(y)$ exists, $\forall y \in D_2 \cap B(y_0, \delta_1)$, and $g^{(1)}$ is continuous at $y_0$. By
Proposition 9.46 Let $X$, $Y$, and $Z$ be normed linear spaces over $k$, $D \subseteq X \times Y$, $f : D \to Z$ be partial differentiable with respect to $x$, and partial differentiable with respect to $y$. Then, the following statements hold.

(i) If $f$ is $(n+1)$-times differentiable, where $n \in \mathbb{N}$, then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are $n$-times differentiable.

(ii) If $f$ is $C_1$ at $(x_0, y_0) \in D$, then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at $(x_0, y_0)$.

(iii) If $f$ is $C_n$ at $(x_0, y_0) \in D$, where $n \in \{2, 3, \ldots\} \cup \{\infty\}$, then $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are $C_{n-1}$ at $(x_0, y_0)$. 


Proof By Proposition 9.9, \( f^{(1)}(x, y) = \left[ \frac{\partial f}{\partial x}(x, y) \quad \frac{\partial f}{\partial y}(x, y) \right], \)
\( \forall (x, y) \in D. \) Define \( g : \mathcal{X} \times \mathcal{Y} \to B(\mathcal{X}, \mathcal{X} \times \mathcal{Y}) \) by \( g(x, y) = \left[ \text{id}_\mathcal{X} \quad \theta_{B(\mathcal{X}, \mathcal{Y})} \right]. \)

By Proposition 9.33, \( g \) is \( C_\infty. \) It is clear that \( \frac{\partial f}{\partial x}(x, y) = f^{(1)}(x, y)g(x, y), \)
\( \forall (x, y) \in D. \)

(i) Since \( f \) is \( (n + 1) \)-times differentiable, then \( f^{(1)} \) is \( n \)-times differentiable. By Propositions 9.42, 9.44, and 9.45, we have \( \frac{\partial f}{\partial x} \) is \( n \)-times differentiable. By symmetry, \( \frac{\partial f}{\partial y} \) is \( n \)-times differentiable. This completes the proof of the proposition.

(ii) Since \( f \) is \( C_1 \) at \((x_0, y_0)\), then \( f^{(1)} \) is continuous at \((x_0, y_0)\). By Propositions 9.42, 9.44, and 9.45, we have \( \frac{\partial f}{\partial x} \) is continuous at \((x_0, y_0)\). By symmetry, \( \frac{\partial f}{\partial y} \) is continuous at \((x_0, y_0)\).

(iii) Since \( f \) is \( C_n \) at \((x_0, y_0)\), then \( f^{(1)} \) is \( C_{n-1} \) at \((x_0, y_0)\). By Propositions 9.42, 9.44, and 9.45, we have \( \frac{\partial f}{\partial x} \) is \( C_{n-1} \) at \((x_0, y_0)\). By symmetry, \( \frac{\partial f}{\partial y} \) is \( C_{n-1} \) at \((x_0, y_0)\).

This completes the proof of the proposition.

Proposition 9.47 Let \( \mathcal{X}_1, \ldots, \mathcal{X}_p, \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K}, \)
where \( p \in \{2, 3, \ldots\}, D \subseteq \mathcal{X} := \prod_{i=1}^p \mathcal{X}_i, f : D \to \mathcal{Y}, x_0 \in D^o, \) and \( k \in \mathbb{N}. \)
Assume that \( \exists \delta_0 \in (0, \infty) \subset \mathbb{R} \) such that all partial derivatives of \( f \) up to \( k \)th order exist and are continuous at \( x, \forall x \in \hat{D} := B(\mathcal{X}, x_0, \delta_0) \subseteq D. \) Then, \( f \) is \( C_k \) at \( x, \forall x \in \hat{D}. \)

Proof We will prove this using mathematical induction on \( k. \)

1° \( k = 1. \) \( \forall x \in \hat{D}, \) by repeated application of Proposition 9.24, we have \( f^{(1)}(x) \) exists and \( f^{(1)}(x) = \left[ \frac{\partial f}{\partial x_1}(x) \quad \cdots \quad \frac{\partial f}{\partial x_p}(x) \right]. \)

2° Assume that the result holds for \( k = \tilde{k} \in \mathbb{N}. \)

3° Consider the case \( k = \tilde{k} + 1. \) \( \forall x \in \hat{D}, \) by the case \( k = 1, \) \( f^{(1)}(x) \)
exists and \( f^{(1)}(x) = \left[ \frac{\partial f}{\partial x_1}(x) \quad \cdots \quad \frac{\partial f}{\partial x_p}(x) \right]. \) \( \forall i \in \{1, \ldots, p\}, \) by the assumption, all partial derivatives of \( \frac{\partial f}{\partial x_i} \) up to \( k \)th order exist and are continuous at \( x. \) Then, by inductive assumption, \( \frac{\partial f}{\partial x_i} \) is \( C_{\tilde{k}} \) at \( x. \) Define the function \( g : \hat{D} \to \prod_{i=1}^p B(\mathcal{X}_i, \mathcal{Y}) \) by \( g(x) = \left( \frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_p}(x) \right), \) \( \forall x \in \hat{D} := \bigcap_{i=1}^p \text{dom} \left( \frac{\partial f}{\partial x_i} \right). \) Clearly, \( \hat{D} \subseteq \hat{D} \subseteq D. \) Then, by Proposition 9.44,

\( g \) is \( \tilde{k} \)-times differentiable at \( x \) and \( g^{(j)}(x) = \left[ \frac{D^j}{\partial x_1}(x) \quad \cdots \quad \frac{D^j}{\partial x_p}(x) \right], \forall j = 1, \ldots, \tilde{k}. \)

By Proposition 3.32, \( g \) is \( C_{\tilde{k}} \) at \( x. \) By Propositions 9.45 and 9.43, \( f^{(1)} \) is \( C_{\tilde{k}} \) at \( x. \) Therefore, \( f \) is \( C_{\tilde{k}+1} \) at \( x. \) This completes the induction process and the proof of the proposition. \( \square \)
Theorem 9.48 (Taylor’s Theorem) Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $f : D \to Y$, $x_0, x_1 \in D$, and $n \in \mathbb{N}$. Let $\bar{D} := I := [0,1] \subset \mathbb{R}$ and $\tilde{D} := I^\circ$ if $K = \mathbb{R}$ or $\bar{D} := \{a + i0 \mid a \in I\} \subset \mathbb{C}$ and $\tilde{D} := \{a + i0 \mid a \in I^\circ\} \subset \mathbb{C}$. Let $\varphi : \tilde{D} \to D$ be given by $\varphi(t) = tx_1 + (1-t)x_0$, $\forall t \in \tilde{D}$. Assume that $\text{dom}(f^{(n)}) \supseteq \varphi(\tilde{D})$, $\text{dom}(f^{(n+1)}) \supseteq \varphi(\tilde{D})$, and $f^{(n)}$ is continuous at $x = \varphi(t)$, $\forall t \in \tilde{D}$. Let $R_n \in \mathcal{Y}$ be given by

\[ R_n := f(x_1) - \left( f(x_0) + \frac{1}{1!}f^{(1)}(x_0)(x_1 - x_0) + \cdots + \frac{1}{n!}f^{(n)}(x_0) (x_1 - x_0) \cdot \cdots (x_1 - x_0) \right) \]

Then, the following statements hold.

(i) If $\mathcal{Y} = \mathbb{R}$ and $K = \mathbb{R}$, then $\exists \tilde{t}_0 \in I^\circ$ such that

\[ R_n = \frac{1}{(n+1)!}f^{(n+1)}(\varphi(\tilde{t}_0)) (x_1 - x_0) \cdots (x_1 - x_0) \]

(ii) $\exists \tilde{t}_0 \in I^\circ$ such that

\[ \|R_n\| \leq \frac{1}{(n+1)!} \|f^{(n+1)}(\varphi(\tilde{t}_0))\| \|x_1 - x_0\|^{n+1} \]

Proof

(i) Let $\mathcal{Y} = \mathbb{R}$ and $K = \mathbb{R}$. Define $F : I \to \mathbb{R}$ by

\[ F(t) = f(\varphi(1)) - \left( f(\varphi(1-t)) + \frac{t}{1!}f^{(1)}(\varphi(1-t))(x_1 - x_0) + \cdots + \frac{t}{n!}f^{(n)}(\varphi(1-t)) (x_1 - x_0) \cdot \cdots (x_1 - x_0) + R_n t^{n+1} \right); \quad \forall t \in I \]

By Propositions 3.12, 3.32, 9.7, 7.23, and 7.65, $F$ is continuous. Clearly, $\varphi$ is differentiable. By Chain Rule and Propositions 9.10, 9.15–9.17 and 9.19, $F$ is differentiable at $t$, $\forall t \in I$. Clearly, $F(0) = F(1) = 0$. By Mean Value Theorem 9.20, $\exists \tilde{t}_0 \in I^\circ$ such that $0 = F(1) - F(0) = DF(\tilde{t}_0)$. Then, we have

\[ 0 = - \left( - \frac{t}{1!}f^{(1)}(\varphi(1-t_0))(x_1 - x_0) + \frac{1}{n!}f^{(n)}(\varphi(1-t_0))(x_1 - x_0) \right) \]

\[ 0 = \left( - \frac{t}{1!}f^{(2)}(\varphi(1-t_0))(x_1 - x_0)(x_1 - x_0) + \cdots + \frac{t}{(n-1)!}f^{(n)}(\varphi(1-t_0))(x_1 - x_0) \cdot \cdots (x_1 - x_0) \right) \]

\[ + \frac{t}{n!}f^{(n+1)}(\varphi(1-t_0)) (x_1 - x_0) \cdots (x_1 - x_0) \]
9.4. Higher order derivatives

\[-\frac{t_0^n}{n!} f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0) + (n + 1) R_n t_0^n\]

\[= \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0) - (n + 1) R_n t_0^n\]

Hence,

\[R_n = \frac{1}{(n + 1)!} f^{(n+1)}(\varphi(\hat{t}_0))(x_1 - x_0) \cdots (x_1 - x_0)\]

\[(n+1)\text{-times}\]

where \(\hat{t}_0 = 1 - t_0 \in I^o\).

(ii) By Proposition 7.85, \(\exists y_* \in \mathcal{Y}^*\) with \(\|y_*\| \leq 1\) such that \(\|R_n\| = \langle \langle y_*, R_n \rangle \rangle\). Define \(G: \bar{D} \to K\) by

\[G(t) = \langle \langle y_*, f(x_1) - f(\varphi(1 - t)) \rangle \rangle - \sum_{i=1}^{n} \int_0^t f^{(i)}(\varphi(1 - t))(x_1 - x_0) \cdots (x_1 - x_0)\rangle \rangle - R_n \| t^{n+1}\]

\(\forall t \in \bar{D}\). By Propositions 3.12, 3.32, 9.7, 7.23, and 7.65, \(G\) is continuous. Clearly, \(\varphi\) is differentiable. By Chain Rule and Propositions 9.10, 9.15–9.17 and 9.19, \(G\) is differentiable at \(t, \forall t \in \bar{D}\). Clearly, \(G(0) = G(1) = 0\).

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(K = R\); Case 2: \(K = C\).

Case 1: \(K = R\). By Mean Value Theorem 9.20, \(\exists t_0 \in I^o\) such that

\(0 = G(1) - G(0) = D G(t_0)\). Then, we have

\[0 = \langle \langle y_*, f^{(1)}(\varphi(1 - t_0))(x_1 - x_0) \rangle \rangle - \sum_{i=1}^{n} \int_0^t f^{(i-1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0)\rangle \rangle - \frac{n}{n!} f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0)\rangle \rangle - (n + 1) R_n \| t_0^n\]

\[= \langle \langle y_*, t_0^n f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0)\rangle \rangle - (n + 1) R_n \| t_0^n\]

Then, we have

\[\|R_n\| = \frac{1}{(n + 1)!} \langle \langle y_*, f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0)\rangle \rangle \]

\[(n+1)-\text{times}\]
where the last two inequalities follows from Proposition 7.64 and \( \bar{t}_0 = 1 - t_0 \in I^0 \).

Case 2: \( K = \mathbb{C} \). By Lemma 9.22, \( \exists t_0 \in I^0 \) such that \( \Re(G(1) - G(0)) = \Re(DG(t_0)) \). Then, we have

\[
0 = \Re \left( \langle y_*, f^{(1)}(\varphi(1 - t_0))(x_1 - x_0) \rangle \right) \\
- \Re \left( \langle y_*, \sum_{i=1}^{n} \frac{n!}{i!(n-i)!} f^{(i)}(\varphi(1 - t_0))(x_1 - x_0) \cdot (x_1 - x_0) \rangle \right) \\
+ \Re \left( \langle y_*, \sum_{i=1}^{n} \frac{n!}{i!(n-i)!} f^{(i+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdot (x_1 - x_0) \rangle \right) \\
- (n+1) \| R_n \| t_0^n \\
= \Re \left( \langle y_*, \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdot (x_1 - x_0) \rangle \right) \\
- (n+1) \| R_n \| t_0^n \\
\]

Hence,

\[
\| R_n \| = \frac{1}{(n+1)!} \Re \left( \langle y_*, f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \cdot (x_1 - x_0) \rangle \right) \\
\leq \frac{1}{(n+1)!} \| y_* \| \| f^{(n+1)}(\varphi(1 - t_0))(x_1 - x_0) \| (n+1) \text{-times} \\
\leq \frac{1}{(n+1)!} \| f^{(n+1)}(\varphi(\bar{t}_0)) \| (n+1) \text{-times} \\
\| x_1 - x_0 \|^{n+1} \\
\]

where the last two inequalities follows from Proposition 7.64 and \( \bar{t}_0 = 1 - t_0 \in I^0 \).

This completes the proof of the theorem. \( \square \)

### 9.5 Mapping Theorems

**Definition 9.49** Let \( \mathcal{X} := (X, \rho) \) be a metric space, \( S \subseteq X \), and \( T : S \to X \). \( T \) is said to be a contraction mapping on \( S \) if \( T(S) \subseteq S \) and \( \exists \alpha \in [0, 1) \subseteq \mathbb{R} \) such that \( \forall x_1, x_2 \in S \), we have \( \rho(T(x_1), T(x_2)) \leq \alpha \rho(x_1, x_2) \). Then, \( \alpha \) is called an contraction index for \( T \).
Theorem 9.50 (Contraction Mapping Theorem) Let $S \neq \emptyset$ be a closed subset of a complete metric space $X := (X, \rho)$ and $T$ be a contraction mapping on $S$ with contraction index $\alpha \in [0, 1) \subset \mathbb{R}$. Then, the following statements hold.

(i) $\exists x_0 \in S$ such that $x_0 = T(x_0)$.

(ii) $\forall x_1 \in S$, recursively define $x_{n+1} = T(x_n)$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \in \mathbb{N}} x_n = x_0$.

(iii) $\rho(x_n, x_0) \leq \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)$, $\rho(x_n, x_0) \leq \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)$, and $\rho(x_n, x_0) \leq \alpha^{n-1} \rho(x_2, x_1)$, $\forall n \in \{2, 3, \ldots\}$.

Proof Fix any $x_1 \in S \neq \emptyset$. Recursively define $x_{n+1} = T(x_n)$, $\forall n \in \mathbb{N}$. Then, $\rho(x_n, x_{n-1}) = \rho(T(x_{n-1}), T(x_{n-2})) \leq \alpha \rho(x_{n-1}, x_{n-2}) \leq \cdots \leq \alpha^{n-2} \rho(x_2, x_1)$, $\forall n \in \{3, 4, \ldots\}$. Therefore, $(x_n)_{n=1}^{\infty} \subseteq S$ is a Cauchy sequence and converges to $x_0 \in S$ by $X$ being complete, $S$ being closed, and Proposition 4.39. Clearly, $T$ is continuous. Then, by Proposition 3.66, $x_0 = T(x_0)$, $\forall x \in S$ such that $x = T(x)$. Then, $\rho(x_0, x) \leq \alpha \rho(x_0, x)$ and $\rho(x_0, x) = 0$. Hence, $x = x_0$. Hence the statements (i) and (ii) are true.

Note that, by Propositions 3.66 and 4.30, $\forall n \in \{2, 3, \ldots\}$,

$$
\rho(x_n, x_0) = \lim_{m \in \mathbb{N}} \rho(x_n, x_m) \leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=n+1}^{\infty} \alpha^{i-2} \rho(x_2, x_1)
$$

$$
= \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)
$$

$$
\rho(x_n, x_0) \leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=1}^{\infty} \alpha^{i} \rho(x_n, x_{n-1}) = \frac{\alpha}{1-\alpha} \rho(x_n, x_{n-1})
$$

$$
\rho(x_n, x_0) = \rho(T(x_{n-1}), T(x_0)) \leq \alpha \rho(x_{n-1}, x_0)
$$

This completes the proof of the theorem.

Lemma 9.51 Let $X$ and $Y$ be normed linear spaces over $\mathbb{K}$, $D \subseteq X$, $x_0 \in D$, $f : D \to Y$ be $C_1$ at $x_0$. Assume that $D$ is locally convex at $x_0$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that, $\forall x_1, x_2 \in D \cap B_X (x_0, \delta)$, we have $\|f(x_1) - f(x_2) - f^{(1)}(x_0)(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|$.

Proof By the assumption, $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D \cap B_X (x_0, \delta_0) =: \tilde{D}$ is convex. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by $f$ being $C_1$ at $x_0$, $\exists \delta \in (0, \delta_0) \subset \mathbb{R}$ such that $\forall x \in D \cap B_X (x_0, \delta) =: \tilde{D}$, we have $\|f^{(1)}(x) - f^{(1)}(x_0)(x - x_0)\| < \epsilon$. Define $\gamma : \tilde{D} \to Y$ by $\gamma(x) = f(x) - f^{(1)}(x_0)(x - x_0)$, $\forall x \in D$. By Propositions 9.45, 9.38, 9.41, and 9.44, $\gamma$ is $C_1$ at $x_0$ and $\gamma^{(1)}(x) = f^{(1)}(x) - f^{(1)}(x_0)$, $\forall x \in \tilde{D}$. Clearly, $\tilde{D}$ is convex. Then, $\forall x_1, x_2 \in \tilde{D}$, by Mean Value Theorem 9.23 and Proposition 7.64, we have $\|f(x_1) - f(x_2) - f^{(1)}(x_0)(x_1 - x_2)\| = \|\gamma(x_1) - \gamma(x_2)\| \leq \|\gamma^{(1)}(t_0 x_1 + (1-t_0) x_2)(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|$, for some $t_0 \in (0, 1) \subset \mathbb{R}$. This completes the proof of the lemma.

$\square$
Theorem 9.52 (Injective Mapping Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \to \mathcal{Y}$ be $C_1$ at $x_0 \in D^o$. Assume that $F^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y})$ is injective and $M := R(F^{(1)}(x_0)) \subseteq \mathcal{Y}$ is closed. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ with $U := B_X(x_0, \delta) \subseteq D$ such that $F_{|U} : U \to F(U)$ is bijective and admits a continuous inverse $F_1 : F(U) \to U$.

Proof By Proposition 7.13 and Proposition 4.39, $M \subseteq \mathcal{Y}$ is a Banach space. Then, $F^{(1)}(x_0) : \mathcal{X} \to M$ is bijective. By Open Mapping Theorem 7.103, the inverse $A$ of $F^{(1)}(x_0) : \mathcal{X} \to M$ belongs to $B(M, \mathcal{X})$. By $h \in \mathcal{X}$, we have $\|h\| = \|AF^{(1)}(x_0)h\| \leq \|A\| \|F^{(1)}(x_0)h\|$. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $r \|A\| \leq 1$ and $r \|h\| \leq \|F^{(1)}(x_0)h\|$, $\forall h \in \mathcal{X}$.

By Lemma 9.51, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ with $U := B_X(x_0, \delta) \subseteq D$ such that $\forall x_1, x_2 \in U$, we have $\|F(x_1) - F(x_2)\| \leq \|F^{(1)}(x_0)(x_1 - x_2)\| \leq r \|x_1 - x_2\|/2$. $\forall x_1, x_2 \in U$, $r \|x_1 - x_2\| \leq \|F^{(1)}(x_0)(x_1 - x_2)\| \leq \|F(x_1) - F(x_2)\| + \|F(x_1) - F(x_2) - F^{(1)}(x_0)(x_1 - x_2)\| \leq \|F(x_1) - F(x_2)\| + r \|x_1 - x_2\|/2$. This implies that $\|F(x_1) - F(x_2)\| \geq r \|x_1 - x_2\|/2$. Hence, $F_{|U} : U \to F(U)$ is injective and surjective. Then, $F_{|U} : U \to F(U)$ is bijective and admits a inverse $F_1 : F(U) \to U$. $\forall y_1, y_2 \in F(U)$, we have $\|F_1(y_1) - F_1(y_2)\| \leq 2/r \|F(F_1(y_1)) - F(F_1(y_2))\| = 2/r \|y_1 - y_2\|$. Hence, $F_1$ is uniformly continuous. This completes the proof of the theorem.

Theorem 9.53 (Surjective Mapping Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $F : D \to \mathcal{Y}$, $x_0 \in D^o$, and $y_0 := F(x_0) \in \mathcal{Y}$. Assume that $F$ is $C_1$ at $x_0$, and $F^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y})$ is surjective. Then, $\exists r \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1 \delta \leq r$ such that $\forall y \in B_y(y_0, \delta/2)$, $\forall \bar{x} \in B_X(x_0, r/2)$ with $\bar{y} = F(\bar{x})$, $\forall \bar{y} \in B_{\bar{y}}(\bar{y}, \delta/2)$, $\exists \bar{x} \in B_X(x_0, r) \subseteq D$ with $\|x - \bar{x}\| \leq c_1 \|y - \bar{y}\|$, we have $y = F(x)$.

Proof Let $\mathcal{M} := N(F^{(1)}(x_0))$, which is a closed subspace by Proposition 7.68. By Proposition 7.45, the quotient space $\mathcal{X}/\mathcal{M}$ is a Banach space. By Proposition 7.70, $F^{(1)}(x_0) = A \circ \phi$, where $\phi : \mathcal{X} \to \mathcal{X}/\mathcal{M}$ is the natural homomorphism, and $A \in B(\mathcal{X}/\mathcal{M}, \mathcal{Y})$ is injective. Since $F^{(1)}(x_0)$ is surjective, then $A$ is injective and, by Open Mapping Theorem 7.103, $A^{-1} \in B(\mathcal{Y}, \mathcal{X}/\mathcal{M})$. Let $c_1 := 4 \|A^{-1}\| \in (0, \infty) \subset \mathbb{R}$.

Define $\gamma : D \to \mathcal{X}/\mathcal{M}$ by $\gamma(x) = A^{-1}(F(x) - F^{(1)}(x_0)(x - x_0))$. $\forall x \in D$. By Propositions 9.45, 9.38, 9.41, 9.34, and 9.44, $\gamma$ is $C_1$ at $x_0$ and $\gamma^{(1)}(x) = A^{-1}(F^{(1)}(x) - F^{(1)}(x_0))$, $\forall x \in B_X(x_0, r_0) \subseteq D$, for some $r_0 \in (0, \infty) \subset \mathbb{R}$. Clearly, $\gamma^{(1)}(x_0) = \partial_{B_X(x_0, r_0)}$. Then, by Lemma 9.51, $\exists r \in (0, r_0) \subset \mathbb{R}$ such that $\forall x_1, x_2 \in B_X(x_0, r) \subseteq D$, we have $\|\gamma(x_1) - \gamma(x_2)\| \leq \|x_1 - x_2\|/4$.

Let $\delta \in (0, \infty) \subset \mathbb{R}$ be such that $c_1 \delta \leq r$. Fix any $\bar{y} \in B_{\bar{y}}(\bar{y}_0, \delta/2)$ and any $\bar{x} \in B_X(x_0, r/2)$ with $\bar{y} = F(\bar{x})$. Fix any $y \in B_y(y, \delta/2)$. Recursively define $x_1 := \bar{x}$, $[x_{k+1}] = [x_k] + A^{-1}(y - F(x_k))$ and select $x_{k+1} \in [x_{k+1}]$ such that $\|x_{k+1} - x_k\| \leq 2 \|x_{k+1} - x_k\|$, $\forall k \in \mathbb{N}$, where $[x_k] = \phi(x_k) = x_k + M$ is the coset containing $x_k$. Clearly, $x_1 \in B_X(x_0, r)$. Note that
\[ \| x_2 - x_1 \| \leq 2 \left\| x_2 - x_1 \right\| = 2 \left\| A^{-1}(y - \tilde{y}) \right\| \leq 2 \left\| A^{-1} \right\| \| y - \tilde{y} \| = c_1 \| y - \tilde{y} \| / 2 < r / 4, \] 
where the second inequality follows from Proposition 7.64. 

Then, \( x_2 \in B_X(x_0, r) \). Assume that \( x_1, \ldots, x_k \in B_X(x_0, r) \) for some \( k \in \{2, 3, \ldots\} \). Note that for \( \forall i \in \{2, \ldots, k\}, \| x_{i+1} - x_i \| \leq 2 \left\| A^{-1}(y - F(x_i) + A\phi(x_i)) - A^{-1}(y - F(x_{i-1}) + A\phi(x_{i-1})) \right\| = 2 \| \gamma(x_i) - \gamma(x_{i-1}) \| \leq \| x_i - x_{i-1} \| / 2. \] 

Then, \( \| x_{i+1} - x_i \| \leq \| x_2 - x_1 \| / 2^k \). Hence, \( \| x_{k+1} - x_i \| \leq \sum_{i=1}^{k} \| x_{i+1} - x_i \| \leq \sum_{i=1}^{k} \| x_2 - x_1 \| / 2^k = (2 - 1/2^k) \| x_2 - x_1 \| / 2 = \| x_2 - x_1 \| / 2 - \| x_2 - x_1 \| / 2^k, \) \forall k \in \mathbb{N}.

Hence, \( (x_k)_{k=0}^\infty \subseteq B_X(x_0, r) \) and \( \| x_{k+1} - x_k \| \leq \| x_2 - x_1 \| / 2^k, \forall k \in \mathbb{N}. \) 

Hence, \( (x_k)_{k=0}^\infty \) is a Cauchy sequence. By the completeness of \( X \), we have \( \lim_{k \to \infty} x_k = x \in X \). Note that, by Propositions 3.66 and 4.30, \( \| x - \bar{x} \| = \lim_{k \to \infty} \| x_k - x_1 \| \leq \lim_{k \to \infty} (2 - 1/2^{k-2}) \| x_2 - x_1 \| = 2 \| x_2 - x_1 \| / 2 < r / 2. \)

Hence, \( x \in B_X(x_0, r) \). By the differentiability of \( F \) and Propositions 9.7, 3.66, and 7.69, we have \( \phi(x) = \lim_{k \to \infty} \phi(x_{k+1}) = \lim_{k \to \infty} (A^{-1}(y - F(x_k)) + \phi(x_k)) = \phi(x) + A^{-1}(y - F(x)). \) 

This implies that \( y = F(x). \) Note that \( \| x - \bar{x} \| \leq 2 \| x_2 - x_1 \| \leq c_1 \| y - \tilde{y} \|. \) This completes the proof of the theorem.

**Theorem 9.54 (Open Mapping Theorem)** Let \( X \) and \( Y \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq X \) be open, and \( F : D \to Y \) be \( C_1 \). Assume that \( F^{(1)}(x) \in B(X, Y) \) is surjective, \( \forall x \in D. \) Then, \( F \) is an open mapping.

**Proof** 
Fix any open subset \( U \subseteq D \), where \( U \) is open in the subset topology of \( D \). Since \( D \) is open, then \( U \) is open in \( X \). We will show that \( F(U) \) is an open subset of \( Y \). Fix any \( y_0 \in F(U) \), there exists \( x_0 \in U \) such that \( y_0 = \tilde{y}_0 \). Then, \( \exists r \in (0, \infty) \subseteq \mathbb{R} \) such that \( B_Y(x_0, r) \subseteq U \). It is easy to check that all assumptions of Surjective Mapping Theorem are satisfied at \( x_0 \). Then, there exist an open set \( V \subseteq Y \) with \( y_0 \in V \) and \( c_1 \in [0, \infty) \subseteq \mathbb{R} \) such that \( \forall y \in V, \exists \bar{x} \in D \) with \( \| \bar{x} - x_0 \| \leq c_1 \| \tilde{y} - y_0 \| \), we have \( \bar{y} = F(\bar{x}). \) 

Take \( \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( c_1 \delta \leq r \) and \( B_Y(y_0, \delta) \subseteq V \). Then, \( \forall \tilde{y} \in B_Y(y_0, \delta), \| \tilde{y} - x_0 \| \leq c_1 \| \tilde{y} - y_0 \| \leq r \). Then, \( \tilde{x} \in B_X(x_0, r) \subseteq U \) and \( \tilde{y} \in F(U) \). Hence, \( B_Y(y_0, \delta) \subseteq F(U) \). Therefore, \( y_0 \in (F(U))^o \). By the arbitrariness of \( y_0 \), we have \( F(U) \) is open in \( Y \). By the arbitrariness of \( U, F \) is an open mapping. This completes the proof of the theorem.

**Proposition 9.55** Let \( X \) and \( Y \) be Banach spaces over \( \mathbb{K} \) and \( A \in B(X, Y) \) be bijective. Then, \( \forall T \in B(X, Y) \) with \( \| T - A \| \left\| A^{-1} \right\| < 1, \) we have \( T \) is bijective and \( \left\| T^{-1} - A^{-1} \right\| \leq \frac{\left\| A^{-1}\right\|^2 \| T - A \|}{1 - \| A^{-1} \| \| T - A \|}. \)

**Proof** 
By Open Mapping Theorem 7.103, \( A^{-1} \in B(Y, X) \). We will first prove the result for the special case \( Y = X \) and \( A = \text{id}_X \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( X \) is a singleton set; Case 2: \( \exists \bar{x} \in X \) such that \( \bar{x} \neq \emptyset \).

Case 1: \( X \) is a singleton set. Then, \( \| A^{-1} \| = \| T \| = 0 \). \( \forall T \in B(X, X), \) we have \( T = \text{id}_X \). Then, \( T \) is bijective and \( \| T^{-1} - \text{id}_X \| = 0. \) The result holds for this case.
CHAPTER 9. DIFFERENTIATION IN BANACH SPACES

296

Case 2: \( \exists \bar{x} \in X \) such that \( \bar{x} \neq \vartheta_X \). Then, \( \|A\| = \|A^{-1}\| = 1 \). \( \forall T \in B(X,X) \) with \( \|T - \text{id}_X\| < 1 \), let \( \Delta := T - \text{id}_X \). We will show that \( T \) is bijective. \( \forall x_1, x_2 \in X \) with \( T(x_1) = T(x_2) \), we have \( x_1 + \Delta(x_1) = x_2 + \Delta(x_2) \), which implies that \( \Delta(x_1 - x_2) = x_2 - x_1 \). By Proposition 7.64, we have \( \|\Delta\| \|x_1 - x_2\| \geq \|x_1 - x_2\| \). Since \( \|\Delta\| < 1 \) then \( \|x_1 - x_2\| = 0 \) and \( x_1 = x_2 \). Therefore, \( T \) is injective. \( \forall x_0 \in X \), define \( \phi : X \rightarrow X \) by \( \phi(x) = x_0 - \Delta(x) \), \( \forall x \in X \). Clearly, \( \phi \) is a contraction mapping on \( X \) with contraction index \( \|\Delta\| \). By Contraction Mapping Theorem, there exists a unique \( \bar{x} \in X \) such that \( \bar{x} = \phi(\bar{x}) \). Then, \( x_0 = \bar{x} + \Delta(\bar{x}) = T(\bar{x}) \). Hence, \( T \) is surjective. Then, \( T \) is bijective. By Open Mapping Theorem 7.103, \( T^{-1} \in B(X,X) \).

\[ \forall y \in Y, \text{ let } x = T^{-1}y. \text{ Then, } y = Tx = x + \Delta x \text{ and } x = y - \Delta x. \]

By Proposition 7.64, we have \( \|x\| \leq \|y\| + \|\Delta\| \|x\| \) and \( \|x\| \leq \frac{\|v\|}{1 - \|\Delta\|} \). By the arbitrariness of \( y \), we have \( \|T^{-1}\| \leq 1/(1 - \|\Delta\|) \). This further implies that \( \|T^{-1} - \text{id}_X\| = \|T^{-1}(\text{id}_X - T)\| \leq \|T^{-1}\| \|\Delta\| \leq \|\Delta\|/(1 - \|\Delta\|) \).

The result holds in this case.

Hence, the result holds for the special case \( Y = X \) and \( A = \text{id}_X \). Now consider the general case. \( \forall T \in B(X,Y) \) with \( \|T - A\| \|A^{-1}\| < 1 \), we have \( \bar{T} := A^{-1}T \in B(X,X) \). Note that \( \|\bar{T} - \text{id}_X\| = \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\| < 1 \). Then, \( \|\bar{T} - \text{id}_X\| \|\text{id}_X\| < 1 \). By the special case, we have \( \bar{T} \) is injective and \( \|\bar{T}^{-1} - \text{id}_X\| \leq \frac{\|\text{id}_X\|^2}{1 - \|\text{id}_X\| \|\bar{T} - \text{id}_X\|} \). Then, \( T \) is bijective and, by Proposition 7.64,

\[
\begin{align*}
\|\bar{T} - \text{id}_X\| & = \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\| \\
\|A^{-1}\| \|\text{id}_X\| & = \|A^{-1}\| \\
\|T^{-1} - A^{-1}\| & = \|(\bar{T}^{-1} - \text{id}_X)A^{-1}\| \leq \|A^{-1}\| \frac{\|\text{id}_X\|^2 \|\bar{T} - \text{id}_X\|}{1 - \|\text{id}_X\| \|\bar{T} - \text{id}_X\|} \\
& \leq \frac{\|A^{-1}\|^2 \|T - A\|}{1 - \|A^{-1}\| \|T - A\|}
\end{align*}
\]

This completes the proof of the proposition. \( \Box \)

**Proposition 9.56** Let \( X \) and \( Y \) be Banach spaces over \( K \), \( D := \{ L \in B(X,Y) \mid L \text{ is bijective} \} \), and \( f : D \rightarrow B(Y,X) \) be given by \( f(A) = A^{-1}, \forall A \in D \). Then, \( D \) is open in \( B(X,Y) \), \( f \) is \( C_\infty \), and \( f^{(1)}(A)(\Delta) = -A^{-1} \Delta A^{-1}, \forall A \in D, \forall \Delta \in B(X,Y) \).

**Proof** By Proposition 9.55 and Open Mapping Theorem 7.103, \( D \) is open and \( f \) is continuous. \( \forall A \in D, \text{ span}(A_D(A)) = B(X,Y) \) since \( D \) is open and \( A \in D^\circ \). Define \( L : B(X,Y) \rightarrow B(Y,X) \) by \( L(\Delta) = -A^{-1} \Delta A^{-1}, \forall \Delta \in B(X,Y) \). Clearly, \( L \) is a linear operator. Note that

\[
\|L\| = \sup_{\Delta \in B(X,Y), \|\Delta\| \leq 1} \|L(\Delta)\| \leq \|A^{-1}\|^2
\]
where the inequality follows from Proposition 7.64. Hence, \( L \) is a bounded linear operator.

\[ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \text{by the continuity of } f, \exists \delta \in (0, \infty) \subset \mathbb{R} \text{ such that} \]
\[ \| f(\bar{A}) - f(A) \| < \epsilon, \forall \bar{A} \in \mathcal{B}_{\mathcal{Y}(\mathcal{X}, \mathcal{Y})}(A, \delta). \]

Then, \( \forall \bar{A} \in \mathcal{B}_{\mathcal{Y}(\mathcal{X}, \mathcal{Y})}(A, \delta) \), we have
\[ \| f(\bar{A}) - f(A) - L(\bar{A} - A) \| = \| \bar{A}^{-1} - A^{-1} + A^{-1}(\bar{A} - A)A^{-1} \| \]
\[ = \| A^{-1}(\bar{A} - A)\bar{A}^{-1} + A^{-1}(\bar{A} - A)A^{-1} \| \]
\[ \leq \| A^{-1} \| \| A - \bar{A} \| \| f(\bar{A}) - f(A) \| \leq \epsilon \| A^{-1} \| \| \bar{A} - A \| \]

where the first inequality follows from Proposition 7.64. Hence, we have \( f^{(1)}(A) = L \). Then, \( f \) is differentiable. Note that \( f^{(1)}(A) = -f(A)ro(f(A)), \forall A \in D \). By Propositions 9.42, 9.47, 9.42, 9.44, \( f^{(1)} \) is continuous. Hence, \( f \) is \( C_1 \).

Assume that \( f \) is \( C_k \), for some \( k \in \mathbb{N} \). We will show that \( f \) is \( C_{k+1} \).

Theorem 9.57 (Inverse Function Theorem) Let \( \mathcal{X} \) and \( \mathcal{Y} \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \), \( F : D \to \mathcal{Y} \) be \( C_1 \) at \( x_0 \in D^0 \). Assume that \( F^{(1)}(x_0) \in \mathcal{B}(\mathcal{X}, \mathcal{Y}) \) is bijective. Then, \( \exists \) open set \( U \subseteq D \) with \( x_0 \in U \) and \( \exists \) open set \( V \subseteq \mathcal{Y} \) with \( y_0 := F(x_0) \in V \) such that

(i) \( F|_U : U \to V \) is bijective;

(ii) the inverse mapping \( F_i : V \to U \) of \( F|_U \) is differentiable, \( F_i^{(1)} : V \to \mathcal{B}(\mathcal{X}, \mathcal{Y}) \) is given by \( F_i^{(1)}(y) = (F^{(1)}(F_i(y)))^{-1}, \forall y \in V \), and \( F_i^{(1)} \) is continuous at \( y_0 \);

(iii) if \( F \) is \( k \)-times differentiable for some \( k \in \mathbb{N} \), then \( F_i \) is \( k \)-times differentiable;

(iv) if \( F \) is \( C_k \) at \( x_0 \) for some \( k \in \mathbb{N} \cup \{ \infty \} \), then \( F_i \) is \( C_k \) at \( y_0 \).

Proof By Open Mapping Theorem 7.103, \( (F^{(1)}(x_0))^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X}). \)

By \( F \) being \( C_1 \) at \( x_0 \in D^0 \), \( \exists \bar{\epsilon} \in (0, \infty) \subset \mathbb{R} \) such that \( F^{(1)}(x) \) exists, \( \forall x \in B_{\mathcal{Y}}(x_0, \bar{\epsilon}) \subseteq D \). Define \( T : B_{\mathcal{X}}(\bar{x}_X, \bar{\epsilon}) \to \mathcal{X} \) by \( T(x) = (F^{(1)}(x_0))^{-1}(F(x + x_0) - y_0), \forall x \in B_{\mathcal{X}}(\bar{x}_X, \bar{\epsilon}) \subseteq D - x_0 \). Clearly, \( T(\bar{x}_X) = \bar{x}_X \). By Propositions 9.45, 9.38, 9.34, and 9.44, \( T \) is \( C_1 \) at \( \bar{x}_X \), \( T \) is differentiable, and \( T^{(1)}(x) = (F^{(1)}(x_0))^{-1}F^{(1)}(x + x_0), \forall x \in B_{\mathcal{X}}(\bar{x}_X, \bar{\epsilon}) \).

Clearly, \( T^{(1)}(\bar{x}_X) = \text{id}_\mathcal{X} \).

Define \( \psi : B_{\mathcal{X}}(\bar{x}_X, \bar{\epsilon}) \to \mathcal{X} \) by \( \psi(x) = T(x) - x \). Then, by Propositions 9.38, 9.45, and 9.44, \( \psi \) is differentiable, \( \psi \) is \( C_1 \) at \( \bar{x}_X \), and \( \psi^{(1)}(x) = T^{(1)}(x) - \text{id}_\mathcal{X}, \forall x \in B_{\mathcal{X}}(\bar{x}_X, \bar{\epsilon}) \). Clearly, \( \psi(\bar{x}_X) = \bar{x}_X \) and \( \psi^{(1)}(\bar{x}_X) = \bar{x}_X \). Fix any \( \alpha \in (0, 1) \subset \mathbb{R} \). Then, \( \exists \bar{r}_1 \in (0, \bar{\epsilon}) \subset \mathbb{R} \).
such that \( \overline{B}_X(\vartheta_X, r_1) \subseteq D - x_0 \) and \( \|\psi^{(1)}(x)\| \leq \alpha, \forall x \in \overline{B}_X(\vartheta_X, r_1). \)

\( \forall x_1, x_2 \in \overline{B}_X(\vartheta_X, r_1), \) by Mean Value Theorem, \( \|\psi(x_1) - \psi(x_2)\| \leq \sup_{t_0 \in (0, 1)} \|\psi^{(1)}(t_0 x_1 + (1 - t_0)x_2)(x_1 - x_2)\| \leq \alpha \|x_1 - x_2\|, \) where the last inequality follows from Proposition 7.64.

\( \forall \bar{x} \in \overline{B}_X(\vartheta_X, (1 - \alpha)r_1), \) define \( \phi : \overline{B}_X(\vartheta_X, r_1) \to X \) by \( \phi(x) = \bar{x} - \psi(x), \forall x \in \overline{B}_X(\vartheta_X, r_1). \) \( \forall x \in \overline{B}_X(\vartheta_X, r_1), \) \( \|\phi(x)\| \leq \|\bar{x}\| + \|\psi(x)\| < (1 - \alpha)r_1 + \|\psi(x) - \psi(\vartheta_X)\| \leq (1 - \alpha)r_1 + \alpha \|x\| \leq r_1. \) Hence, \( \phi : \overline{B}_X(\vartheta_X, r_1) \to B_X(\vartheta_X, r_1) \subseteq \overline{B}_X(\vartheta_X, r_1). \) It is easy to see that \( \phi \) is a contraction mapping with contraction index \( \alpha. \) By Contraction Mapping Theorem, \( \exists \hat{x} \in \overline{B}_X(\vartheta_X, r_1) \) such that \( \hat{x} = \phi(\hat{x}) \in \overline{B}_X(\vartheta_X, r_1), \) which is equivalent to \( \hat{x} = T(\hat{x}). \) Let \( \hat{U} := T_{\text{inv}}(B_X(\vartheta_X, (1 - \alpha)r_1)) \cap B_X(\vartheta_X, r_1) \) and \( \hat{V} := B_X(\vartheta_X, (1 - \alpha)r_1). \) Note that \( \hat{U} \) and \( \hat{V} \) are open sets in \( X \) since \( T \) is continuous by Proposition 9.7. Then, \( T|_{\hat{U}} : \hat{U} \to \hat{V} \) is bijective. Since \( T(\vartheta_X) = \vartheta_X, \) then \( \vartheta_X \in \hat{U}. \)

Hence, there exists an inverse mapping \( T_i : \hat{V} \to \hat{U} \) of \( T|_{\hat{U}}. \) Let \( U := \hat{U} + x_0 \) and \( V := F^{(1)}(x_0)(\hat{V}) + y_0. \) Clearly, \( U \) and \( V \) are open sets in \( X \) and \( Y, \) respectively. Note that \( F|_U(x) = F^{(1)}(x_0)T_i(\psi(T_i(x)) + y_0, \forall x \in U. \)

Then, \( F|_U : U \to V \) is bijective, whose inverse function is \( F_i : V \to U. \) The inverse function \( F_i \) is given by \( F_i(y) = T_i((F^{(1)}(x_0))^{-1}(y - y_0)) + x_0, \forall y \in V. \) Hence, the statement (i) holds.

Next, we will show that \( F_i : V \to U \) is differentiable. Note that \( \forall x \in \hat{V}, \) \( x = T(T_i(x)) = T_i(x) + \psi(T_i(x)) \) and \( T_i(x) = x - \psi(T_i(x). \) \( \forall x_1, x_2 \in \hat{V}, \) we have \( T_i(x_1), T_i(x_2) \in U \subseteq B_X(\vartheta_X, r_1) \) and

\[
\|T_i(x_1) - T_i(x_2)\| = \|x_1 - x_2 - \psi(T_i(x_1)) + \psi(T_i(x_2))\|
\leq \|x_1 - x_2\| + \|\psi(T_i(x_1)) - \psi(T_i(x_2))\|
\leq \|x_1 - x_2\| + \alpha \|T_i(x_1) - T_i(x_2)\|
\]

which further implies that \( \|T_i(x_1) - T_i(x_2)\| \leq \|x_1 - x_2\|/(1 - \alpha). \) Therefore, \( T_i \) is continuous. By Propositions 3.12, 7.23, and 3.32, \( F_i \) is continuous. We need the following intermediate result.

**Claim 9.57.1** \( \forall x \in U, \) let \( y = F(x). \) Then, \( F_i \) is differentiable at \( F(x) \) and \( F_i^{(1)}(y) = (F^{(1)}(x))^{-1}. \)

**Proof of claim:** Fix any \( x \in U \) and let \( y = F(x). \) Then, \( x - x_0 \in \hat{U} \subseteq B_X(\vartheta_X, r_1) \) and \( \|\psi^{(1)}(x - x_0)\| \leq \alpha. \) Note that \( T = \text{id}_X + psd \) and \( T^{(1)}(x - x_0) = \text{id}_X + psd(x - x_0). \) By Proposition 9.55, \( T^{(1)}(x - x_0) \) is bijective and \( \|T^{(1)}(x - x_0)\| < \infty. \) Note that \( F^{(1)}(x) = F^{(1)}(x_0)T^{(1)}(x - x_0). \) Then, \( F^{(1)}(x) \) is bijective and \( c_1 := \|F^{(1)}(x))^{-1}\| < \infty.

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \) with \( \epsilon c_1 < 1, \) by the differentiability of \( F \) at \( x, \) \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that \( \forall h \in X \) with \( \|h\| < \delta_1, \) we have \( \|F(x + h) - F(x) - F^{(1)}(x)(h)\| < \epsilon \|h\|. \) By the continuity of \( F_i, \) \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \forall u \in Y \) with \( \|u\| < \delta, \) we have \( y + u \in V \) and \( \|F_i(y + u) - F_i(y)\| < \delta_1. \)
\[ \forall u \in \mathcal{Y} \text{ with } \|u\| < \delta, \text{ let } \beta := \|F_i(y + u) - F_i(y) - (F_i^{(1)}(x))^{-1}u\| \geq 0. \text{ Let } h := F_i(y + u) - x \in \mathcal{X}. \text{ Then, } \|h\| = \|F_i(y + u) - F_i(y)\| < \delta_1. \text{ Note that } \]

\[
\beta = \|x + h - x - (F_i^{(1)}(x))^{-1}u\| = \|(F_i^{(1)}(x))^{-1}(u - (F_i^{(1)}(x))h)\|
\]

\[
\leq \|(F_i^{(1)}(x))^{-1}\| \|u - (F_i^{(1)}(x))h\|
\]

\[
= c_1 \|F(x + h) - F(x) - (F_i^{(1)}(x))h\|
\]

\[
\leq c_1 \|h\| = c_1 \|F_i(y + u) - F_i(y)\|
\]

\[
\leq \epsilon c_1 \left( \|F_i(y + u) - F_i(y) - (F_i^{(1)}(x))^{-1}u\| + \|(F_i^{(1)}(x))^{-1}u\| \right)
\]

\[
\leq \epsilon c_1 (\beta + c_1 \|u\|)
\]

Then, \( \beta \leq \frac{\epsilon^2 c_1^2}{1-c_1} \|u\| \). Hence, \( F_i \) is differentiable at \( y \) and \( F_i^{(1)}(y) = (F_i^{(1)}(x))^{-1} \). This completes the proof of the claim. \( \Box \)

Then, \( \forall y \in V, F_i^{(1)}(y) = (F_i^{(1)}(F_i(y)))^{-1} \). By Propositions 9.56 and 3.12, the continuity of \( F_i \), and the continuity of \( F_i^{(1)} \) at \( x_0 \), we have \( F_i^{(1)} \) is continuous at \( y_0 \). Then, the statement (ii) holds.

For (iii), we will use mathematical induction on \( k \).

1° \( k = 1 \). The result holds by (ii).

2° Assume that the result holds for \( k = \bar{k} \in \mathbb{N} \).

3° Consider the case \( k = \bar{k} + 1 \). By inductive assumption, \( F_i \) is \( \bar{k} \)-times differentiable. Clearly, \( F_i^{(1)} \) is \( \bar{k} \)-times differentiable. By (ii) and Propositions 9.45 and 9.56, \( F_i^{(1)} \) is \( \bar{k} \)-times differentiable. Then, \( F_i \) is \( (\bar{k}+1) \)-times differentiable. This completes the induction process and the proof of the statement (iii).

For (iv), we will use mathematical induction on \( k \) to show that the result holds when \( k \in \mathbb{N} \).

1° \( k = 1 \). The result holds by (ii).

2° Assume that the result holds for \( k = \bar{k} \in \mathbb{N} \).

3° Consider the case \( k = \bar{k} + 1 \). Clearly, \( F_i^{(1)} \) is \( C_{\bar{k}} \) at \( x_0 \). By the inductive assumption, \( F_i \) is \( C_{\bar{k}} \) at \( y_0 \). By (ii) and Propositions 9.45 and 9.56, \( F_i^{(1)} \) is \( C_{\bar{k}} \) at \( y_0 \). Hence, \( F_i \) is \( C_{\bar{k}+1} \) at \( y_0 \). This completes the induction process.

When \( k = \infty \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( F \) is \( C_{\bar{k}} \) at \( x \), \( \forall x \in B_{\mathcal{X}}(x_0, \delta) \subseteq D, \forall i \in \mathbb{N} \). By taking \( r_i < \delta \) in the proof of (i), we have \( U \subseteq B_{\mathcal{X}}(x_0, \delta) \).

Then, by the induction conclusion, \( F_i \) is \( C_{\bar{k}} \) at \( y \), \( \forall y \in F(U) = V, \forall i \in \mathbb{N} \). Hence, \( F_i \) is \( C_{\infty} \) at \( y_0 \).

This completes the proof of the theorem. \( \Box \)

**Theorem 9.58 (Implicit Function Theorem)** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( \mathcal{Y} \) and \( \mathcal{Z} \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \times \mathcal{Y} \), \( F : D \to \mathcal{Z} \) be continuous. Assume that \( F \) is partial differentiable with respect to \( y \) and \( \frac{\partial F}{\partial y} \) is continuous at \((x_0, y_0) \in D^0, F(x_0, y_0) = \partial_{\mathcal{Z}}, \frac{\partial F}{\partial y}(x_0, y_0) \in B(\mathcal{Y}, \mathcal{Z}) \) is bijective. Then, the following statements hold.

(i) There exist an open set \( U_0 \in \mathcal{O} \) with \( x_0 \in U_0 \) and \( r_1 \in (0, \infty) \subset \mathbb{R} \) such that \( U_0 \times B_{\mathcal{Y}}(y_0, r_1) \subseteq D \) and \( \forall x \in U_0, \exists ! y \in B_{\mathcal{Y}}(y_0, r_1) \)
satisfying \( F(x, y) = \partial z \). This defines a function \( \phi : U_0 \rightarrow B(y, r_1) \) by \( \phi(x) = y \), \( \forall x \in U_0 \). Then, \( \phi \) is continuous.

(ii) \( \forall (x, y) \in U_0 \times B(y, r_1), \frac{\partial F}{\partial y}(x, y) \) is bijective and \( \left\| \left( \frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty \).

**Proof** By Open Mapping Theorem 7.103, \( \left( \frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \in B(\mathbb{Z}, y) \).

Define a mapping \( \psi : D \rightarrow \mathbb{Y} \) by

\[
\psi(x, y) = y - \left( \frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} F(x, y); \quad \forall (x, y) \in D
\]

Note that \( \psi(x_0, y_0) = y_0 \). Then, by Propositions 7.23, 3.12, 3.27, and 3.32, \( \psi \) is continuous. By the partial differentiability of \( F \) with respect to \( y \), Chain Rule, and Propositions 9.41, 9.15, and 9.19, \( \psi \) is partial differentiable with respect to \( y \) and

\[
\frac{\partial \psi}{\partial y}(x, y) = \text{id}_y - \left( \frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \frac{\partial F}{\partial y}(x, y); \quad \forall (x, y) \in D
\]

By the continuity of \( \frac{\partial F}{\partial y} \) at \( (x_0, y_0) \), then \( \frac{\partial \psi}{\partial y} \) is continuous at \( (x_0, y_0) \in D^\circ \) and \( \frac{\partial \psi}{\partial y}(x_0, y_0) = \partial \psi_{\mathbb{B}(y,y)} \). Fix any \( \alpha \in (0, 1) \subset \mathbb{R} \). Then, \( \exists U_1 \in \mathcal{O} \) with \( x_0 \in U_1 \) and \( \exists r_1 \in (0, \infty) \subset \mathbb{R} \) such that \( U_1 \times \overline{B_y}(y_0, r_1) \subseteq D \) and

\[
\left\| \frac{\partial \psi}{\partial y}(x, y) \right\| < \alpha, \forall (x, y) \in U_1 \times \overline{B_y}(y_0, r_1).
\]

By the continuity of \( \psi \), \( \exists U_0 \in \mathcal{O} \) with \( x_0 \in U_0 \subseteq U_1 \) such that \( \| \psi(x, y) - y_0 \| = \| \psi(x, y) - \psi(x_0, y_0) \| < (1 - \alpha)r_1, \forall x \in U_0 \).

Fix any \( x \in U_0 \), define mapping \( \gamma_x : \overline{B_y}(y_0, r_1) \rightarrow \mathbb{Y} \) by \( \gamma_x(y) = \psi(x, y) \), \( \forall y \in \overline{B_y}(y_0, r_1) \). We will show that \( \gamma_x \) is a contraction mapping with contraction index \( \alpha \). \( \forall y \in \overline{B_y}(y_0, r_1) \), we have

\[
\| \gamma_x(y) - y_0 \| = \| \psi(x, y) - \psi(x_0, y_0) \|
\leq \| \psi(x, y) - \psi(x, y_0) \| + \| \psi(x, y) - \psi(x_0, y_0) \|
\leq \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0)(y - y_0) \right\| + (1 - \alpha)r_1
\leq \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0) \right\| ||y - y_0|| + (1 - \alpha)r_1
\leq \alpha r_1 + (1 - \alpha)r_1 = r_1
\]

where the second inequality follows from Mean Value Theorem and the third inequality follows from Proposition 7.64. Then, \( \gamma_x : \overline{B_y}(y_0, r_1) \rightarrow \overline{B_y}(y_0, r_1) \subseteq \overline{B_y}(y_0, r_1) \). \( \forall y_1, y_2 \in \overline{B_y}(y_0, r_1) \), we have

\[
\| \gamma_x(y_1) - \gamma_x(y_2) \| = \| \psi(x, y_1) - \psi(x, y_2) \|
\]
where the first inequality follows from Mean Value Theorem and the second inequality follows from Proposition 7.64. Hence, $\gamma_x$ is a contraction mapping with contraction index $\alpha$.

By Contraction Mapping Theorem, $\exists y \in \overline{B}_y(y_0, r_1)$ such that $y = \gamma_x(y) \in B_y(y_0, r_1)$, $y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, where $\gamma_{x,n}(y_0)$ is recursively defined by $\gamma_{x,1}(y_0) = y_0$ and $\gamma_{x,k+1}(y_0) = \gamma_x(\gamma_{x,k}(y_0))$, $\forall k \in \mathbb{N}$, and $\|\gamma_{x,n}(y_0) - y\| \leq \frac{r_1}{2-\alpha}\|\alpha^n - 1\| < r_1\alpha^{n-1}$, $\forall n \in \{2, 3, \ldots\}$. By $\gamma_x(y) = y$, we conclude that $F(x, y) = \vartheta_Z$. Hence, $\forall x \in U_0$, $\exists y \in B_y(y_0, r_1)$ such that $F(x, y) = \vartheta_Z$, since $F(x, y) = \vartheta_Z \iff y = \gamma_x(y)$. Then, we may define $\phi : U_0 \to B_y(y_0, r_1)$ by $\phi(x) = y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, $\forall x \in U_0$. Hence, $F(x, \phi(x)) = \vartheta_Z, \forall x \in U_0$.

Next, we show that $\phi$ is continuous. Fix any $\bar{x} \in U_0$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ with $n_0 > 1$ such that $\alpha^{n_0-1}r_1 < \epsilon/3$. By the continuity of $\psi$, we have that $\gamma_{x,n_0}(y_0)$ is continuous with respect to $x$, that is, $\exists \bar{U} \in \mathcal{O}$ with $\bar{x} \in \bar{U} \subseteq U_0$ such that $\forall x_1 \in \bar{U}$, we have $\|\gamma_{x_1,n_0}(y_0) - \gamma_{x,n_0}(y_0)\| < \epsilon/3$. Then,

$$
\|\phi(x_1) - \phi(\bar{x})\| \leq \|\gamma_{x_1,n_0}(y_0) - \gamma_{x,n_0}(y_0)\| + \|\gamma_{x_1,n_0}(y_0) - \psi(\bar{x})\| + \|\gamma_{x,n_0}(y_0) - \phi(\bar{x})\| \leq r_1\alpha^{n_0-1} + \epsilon/3 + r_1\alpha^{n_0-1} < \epsilon
$$

Hence, $\phi$ is continuous. Thus, the statement (i) is proved.

(ii) Note that, $\forall (x, y) \in U_1 \times \overline{B}_y(y_0, r_1)$, we have $\|\frac{\partial F}{\partial y}(x, y)\| \leq \alpha$. By Proposition 9.55, $\text{id}_y - \frac{\partial \psi}{\partial y}(x, y)$ is bijective with continuous inverse. Note that $\psi(y) - \frac{\partial \psi}{\partial y}(x, y) = \left(\frac{\partial F}{\partial y}(x_0, y_0)\right)^{-1}\frac{\partial F}{\partial y}(x, y)$. Therefore, $\frac{\partial F}{\partial y}(x, y)$ is bijective with continuous inverse.

This completes the proof of the theorem. □

**Theorem 9.59 (Implicit Function Theorem)** Let $X$ be a normed linear space over $K$, $Y$ and $Z$ be Banach spaces over $K$, $D \subseteq X \times Y$, $F : D \to Z$ be continuous. Assume that $F$ is partial differentiable with respect to $y$ and $\frac{\partial F}{\partial y}$ is continuous at $(x_0, y_0) \in D^+$, $F(x_0, y_0) = \vartheta_Z$, and $\frac{\partial F}{\partial y}(x_0, y_0) \in B(Y, Z)$ is bijective. Then, the following statements hold.

(i) There exist $r_0, r_1 \in (0, \infty) \subset \mathbb{R}$ such that $B_X(x_0, r_0) \times B_y(y_0, r_1) \subseteq D$ and $\forall x \in B_X(x_0, r_0)$, $\exists ! y \in B_y(y_0, r_1)$ satisfying $F(x, y) = \vartheta_Z$. This defines a function $\phi : B_X(x_0, r_0) \to B_y(y_0, r_1)$ by $\phi(x) = y$, $\forall x \in B_X(x_0, r_0)$. Then, $\phi$ is continuous.
(ii) If $F$ is Fréchet differentiable at $(x, \phi(x)) \in B_X(x_0, r_0) \times B_Y(y_0, r_1)$ for some $x \in B_X(x_0, r_0)$, then $\phi$ is Fréchet differentiable at $x$ and

$$
\phi^{(1)}(x) = -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x))
$$

(iii) Let $n \in \mathbb{N}$. If $F$ is $n$-times Fréchet differentiable, then $\phi$ is $n$-times Fréchet differentiable.

(iv) Let $n \in \mathbb{N} \cup \{\infty\}$ and $\bar{x} \in B_X(x_0, r_0)$. If $F$ is $C_n$ at $(\bar{x}, \phi(\bar{x}))$, then $\phi$ is $C_n$ at $\bar{x}$.

**Proof**

By Implicit Function Theorem 9.58, the statement (i) holds. Furthermore, $\forall (x, y) \in B_X(x_0, r_0) \times B_Y(y_0, r_1)$, $\frac{\partial F}{\partial y}(x, y)$ is bijective and $\left\| \left( \frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty$.

(ii) Fix some $x \in B_X(x_0, r_0)$ such that $F$ is differentiable at $(x, \phi(x))$. Let $y := \phi(x) \in B_Y(y_0, r_1)$. By Proposition 9.9, $\left[ \frac{\partial F}{\partial x}(x, y), \frac{\partial F}{\partial y}(x, y) \right] = F^{(1)}(x, y)$. Let $L := -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x)) \in B(X, Y)$ and $c_1 := \left\| \left( \frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$ with $\epsilon c_1 < 1$, by the differentiability of $F$ at $(x, y, k)$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $\forall (h, k) \in B_X \times B_Y \left( \left( \partial_x, \partial_y \right), \delta_1 \right)$, we have $(x + h, y + k) \in D$ and $\left\| F(x + h, y + k) - F(x, y) - \frac{\partial F}{\partial y}(x, y)h - \frac{\partial F}{\partial x}(x, y)k \right\| \leq \epsilon \left\| (h, k) \right\|$. By the continuity of $\phi$ at $x$, $\exists \delta \in (0, \min \left\{ r_0 - \left\| x - x_0 \right\|, \delta_1 / \sqrt{2} \right\}) \subset \mathbb{R}$ such that $\forall h \in B_X \left( \partial_x, \delta \right)$, we have $\left\| \phi(x + h) - \phi(x) \right\| = \left\| \phi(x + h) - y \right\| < \delta_1 / \sqrt{2}$.

$\forall h \in B_X \left( \partial_x, \delta \right)$, let $\beta := \left\| \phi(x + h) - \phi(x) - Lh \right\| \geq 0$. Then, $\left\| (h, \phi(x + h) - y) \right\| < \delta_1$. Note that, by Proposition 7.64,

$$
\beta = \left\| \left( \frac{\partial F}{\partial y}(x, y) \right)^{-1} \left( \frac{\partial F}{\partial y}(x, y)(\phi(x + h) - \phi(x)) + \frac{\partial F}{\partial x}(x, y)h \right) \right\|
\leq c_1 \left\| F(x + h, \phi(x + h)) - F(x, y) - \frac{\partial F}{\partial x}(x, y)h \right\|
- \frac{\partial F}{\partial y}(x, y)(\phi(x + h) - y) \right\|
\leq c_1 \epsilon \left\| (h, \phi(x + h) - y) \right\| \leq c_1 \epsilon \left( \left\| h \right\| + \left\| \phi(x + h) - y \right\| \right)
\leq c_1 \epsilon \left( \left\| h \right\| + \left\| \phi(x + h) - \phi(x) - Lh \right\| + \left\| Lh \right\| \right)
\leq c_1 \epsilon \left( \beta + (1 + \left\| L \right\|) \left\| h \right\| \right)
$$

Then, we have $\beta \leq \frac{\epsilon c_1 c_2 \left( 1 + \left\| L \right\| \right)}{1 - c_1 \epsilon} \left\| h \right\|$. Hence, $\phi^{(1)}(x) = L$. Then, the statement (ii) holds.

For (iii), we will use mathematical induction on $n$ to prove this result.
9.5. MAPPING THEOREMS

Proposition 9.60

Let \( n = 1 \). The result follows from (ii).

2° Assume that the result holds for \( n = \bar{n} \in \mathbb{N} \).

3° Consider the case \( n = \bar{n} + 1 \). By (ii), we have \( \phi \) is Fréchet differentiable and \( \phi^{(1)}(x) = -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial y}(x, \phi(x)), \forall x \in B_X(x_0, r_0) \). By inductive assumption, \( \phi \) is \( \bar{n} \)-times differentiable. By Proposition 9.46, \( \frac{\partial F}{\partial y} \) and \( \frac{\partial F}{\partial y} \) are \( \bar{n} \)-times differentiable. By Propositions 9.45, 9.56, 9.44, and 9.42, \( \phi^{(1)} \) is \( \bar{n} \)-times differentiable. Then, \( \phi \) is \( (\bar{n} + 1) \)-times differentiable. This completes the induction process.

For (iv), let \( \bar{y} = \phi(\bar{x}) \). We will first use mathematical induction on \( n \) to prove the result for \( n \in \mathbb{N} \).

1° \( n = 1 \). By \( F \) being \( C_1 \) at \((\bar{x}, \bar{y})\), then \( \exists \bar{r} \in (0, \infty) \subset \mathbb{R} \) such that \( F \) is differentiable at \((x, y)\), \( \forall (x, y) \in \bar{D} := B_{X \times Y}((\bar{x}, \bar{y}), \bar{r}) \subseteq D \). By Propositions 9.34 and 9.9, \( F^{(1)}(x, y) = \left[ \frac{\partial F}{\partial x}(x, y) \quad \frac{\partial F}{\partial y}(x, y) \right] \), \( \forall (x, y) \in \bar{D} \). By Proposition 9.46, \( \frac{\partial F}{\partial x} \) and \( \frac{\partial F}{\partial y} \) are continuous at \((\bar{x}, \bar{y})\). Since \( \bar{D} \) is open, then \( \frac{\partial F}{\partial x} \) and \( \frac{\partial F}{\partial y} \) are continuous at \((\bar{x}, \bar{y})\). By the continuity of \( \phi \), \( \exists \bar{\delta} \in (0, \min \{r_0 - \|x - x_0\|, \bar{r}\}] \subset \mathbb{R} \) such that \( F \) is differentiable at \((x, \phi(x))\), \( \forall x \in B_X(\bar{x}, \bar{\delta}) \). By (ii), \( \phi \) is differentiable at \( x \) and \( \phi^{(1)}(x) = -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial y}(x, \phi(x)), \forall x \in B_X(\bar{x}, \bar{\delta}) \). By Propositions 3.12, 3.32, 9.56, 9.42, and 9.7, \( \phi^{(1)} \) is continuous at \( \bar{x} \). Hence, \( \phi \) is \( C_1 \) at \( \bar{x} \).

2° Assume that the result holds for \( n = \bar{n} \in \mathbb{N} \).

3° Consider the case \( n = \bar{n} + 1 \). By \( F \) being \( C_{\bar{n} + 1} \) at \((\bar{x}, \bar{y})\), then \( \exists \bar{r} \in (0, \infty) \subset \mathbb{R} \) such that \( F \) is differentiable at \((x, y)\), \( \forall (x, y) \in \bar{D} := B_{X \times Y}((\bar{x}, \bar{y}), \bar{r}) \subseteq D \). By Propositions 9.34 and 9.9, \( F^{(1)}(x, y) = \left[ \frac{\partial F}{\partial x}(x, y) \quad \frac{\partial F}{\partial y}(x, y) \right] \), \( \forall (x, y) \in \bar{D} \). By Proposition 9.46, \( \frac{\partial F}{\partial x} \) and \( \frac{\partial F}{\partial y} \) are \( C_{\bar{n}} \) at \((\bar{x}, \bar{y})\). Since \( \bar{D} \) is open, then, by Proposition 9.34, \( \frac{\partial F}{\partial x} \) and \( \frac{\partial F}{\partial y} \) are \( C_{\bar{n}} \) at \((\bar{x}, \bar{y})\). By the inductive assumption, \( \phi \) is \( C_{\bar{n}} \) at \( \bar{x} \). By the continuity of \( \phi \), \( \exists \bar{\delta} \in (0, \min \{r_0 - \|x - x_0\|, \bar{r}\}] \subset \mathbb{R} \) such that \( F \) is differentiable at \((x, \phi(x))\), \( \forall x \in B_X(\bar{x}, \bar{\delta}) \). By (ii), \( \phi \) is differentiable at \( x \) and \( \phi^{(1)}(x) = -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial y}(x, \phi(x)), \forall x \in B_X(\bar{x}, \bar{\delta}) \). By Propositions 9.45, 9.44, 9.56, and 9.42, \( \phi^{(1)} \) is \( C_{\bar{n}} \) at \( \bar{x} \). Hence, \( \phi \) is \( C_{\bar{n} + 1} \) at \( \bar{x} \).

This completes the induction process. Hence, (iv) holds when \( n \in \mathbb{N} \).

When \( n = \infty \), there exists \( \bar{\delta} \in (0, \infty) \subset \mathbb{R} \), such that \( F \) is \( C_1 \) at \((\bar{x}, \bar{y})\), \( \forall (x, y) \in B_{X \times Y}((\bar{x}, \bar{y}), \bar{r}) \subseteq B_X(x_0, r_0) \times B_Y(y_0, r_0), \forall i \in \mathbb{N} \). By the induction result, we have \( \phi \) is \( C_1 \) at \( \bar{x} \), \( \forall x \in B_X(\bar{x}, \delta/\sqrt{2}) \cap \phi_{\text{inv}}(B_Y(\bar{y}, \delta/\sqrt{2})) \). Hence, \( \phi \) is \( C_{\infty} \) at \( \bar{x} \). Hence, (iv) holds.

This completes the proof of the proposition. \( \square \)

Proposition 9.60

Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( Y \) and \( Z \) be
normed linear spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $f : D \to \mathbb{Z}$ be partial differentiable with respect to $y$ and $\frac{\partial f}{\partial y}(x,y) = \delta_{D(y,z)}$, $\forall (x,y) \in D$. Let $D := \pi_X(D)$, where $\pi_X$ is the projection function of $X \times Y$ to $X$. Assume that $\forall x \in D$, the set $D_x := \{ y \in Y \mid (x,y) \in D \} \subseteq Y$ is convex. Then, there exists a function $\phi : D \to \mathbb{Z}$ such that $f(x,y) = \phi(x)$, $\forall (x,y) \in D$. Furthermore, the following statements hold.

(i) If $f$ is continuous at $(x_0, y_0) \in D^*$, then $\phi$ is continuous at $x_0$.

(ii) If $X$ is a normed linear space $\mathcal{X}$ over $\mathbb{K}$ and $f$ is $C_k$ at $(x_0, y_0) \in D^*$, where $k \in \mathbb{N} \cup \{\infty\}$, then $\phi$ is $C_k$ at $x_0$.

Proof \hspace{0.5cm} \forall x \in D = \pi_X(D)$, $D_x \neq \emptyset$. By Axiom of Choice, $\exists g : D \to Y$ such that $g(x) \in D_x$, $\forall x \in D$. Define $\phi : D \to \mathcal{X}$ by $\phi(x) = f(x, g(x))$, $\forall x \in D$. $\forall (x,y) \in D$, we have $x \in D$ and $y, g(x) \in D_x$. By the convexity of $D_x$, the line segment connecting $y$ and $g(x)$ is contained in $D_x$. By Mean Value Theorem 9.23, $\exists \theta \in (0, 1) \subset \mathbb{R}$ such that $\|f(x,y) - \phi(x)\| = \|f(x,y) - f(x, g(x))\| \leq \left\| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0) g(x)) (y - g(x)) \right\| = 0$. Hence, $f(x, y) = \phi(x)$.

(i) Let $f$ be continuous at $(x_0, y_0) \in D^*$. Then, $\exists U \in \mathcal{O}$ with $x_0 \in U$ and $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $U \times B_Y(y_0, \delta) \subseteq D$. $\forall x \in U$, we have $\phi(x) = f(x, y_0)$. Hence, $\phi$ is continuous at $x_0$.

(ii) Let $X$ be a normed linear space $\mathcal{X}$ over $\mathbb{K}$ and $f$ be $C_k$ at $(x_0, y_0) \in D^*$, where $k \in \mathbb{N} \cup \{\infty\}$. Then, $\exists \delta_x, \delta_y \in (0, \infty) \subset \mathbb{R}$ such that $B_X(x_0, \delta_x) \times B_Y(y_0, \delta_y) \subseteq D$. $\forall x \in B_X(x_0, \delta_x)$, we have $\phi(x) = f(x, y_0)$. By Proposition 9.45, $\phi|_{B_X(x_0, \delta_x)}$ is $C_k$ at $x_0$. By Proposition 9.34, $\phi$ is $C_k$ at $x_0$.

This completes the proof of the proposition. □

9.6 Global Inverse Function Theorem

Definition 9.61 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be topological spaces, $F : \mathcal{X} \to \mathcal{Y}$ and $\sigma : \mathcal{Z} \to \mathcal{Y}$. We will say $\theta : \mathcal{Z} \to \mathcal{X}$ inverts $F$ along $\sigma$ if $\sigma = F \circ \theta$.

Lemma 9.62 Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $F : \mathcal{X} \to \mathcal{Y}$ be continuous and countably proper, $x_0 \in \mathcal{X}$, and $y_0 := F(x_0) \in \mathcal{Y}$. Assume that $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}_\mathcal{X}$ with $x \in U$ and $\exists V \in \mathcal{O}_\mathcal{Y}$ with $F(x) \in V$ such that $F|_U : U \to V$ is a homeomorphism. Then, given any continuous mapping $\sigma : [a, b] \to \mathcal{Y}$ with $\sigma(t_0) = y_0$, where $a, t_0, b \in \mathbb{R}$ and $a \leq t_0 \leq b$, there exists a unique continuous mapping $\theta : [a, b] \to \mathcal{X}$ with $\theta(t_0) = x_0$ that inverts $F$ along $\sigma$.

Proof We will distinguish three exhaustive and mutually exclusive cases: Case 1: $a = b = t_0$; Case 2: $a = t_0 < b$; Case 3: $a < t_0 \leq b$.

Case 1: $a = b = t_0$. Clearly, $\theta$ exists and is unique. This case is proved.
Case 2: $a = t_0 < b$. “Uniqueness” Let $\theta_1 : [a, b] \to \mathcal{X}$ and $\theta_2 : [a, b] \to \mathcal{X}$ be continuous mappings that inverts $F$ along $\sigma$ with $\theta_1(a) = \theta_2(a) = x_0$. Let $S := \{s \in [a, b] \subset \mathbb{R} \mid \theta_1(s) = \theta_2(s) \forall t \in [a, s] \subset \mathbb{R}\}$ and $\xi := \sup S$. Clearly, $a \in S$ and $a \leq \xi \leq b$. It is easy to show that $\theta_1(t) = \theta_2(t)$, $\forall t \in \mathbb{R}$ with $a \leq t < \xi$. Therefore, we must have $\theta_1(t) = \theta_2(t)$, where the limit operator makes sense since $\mathcal{X}$ is Hausdorff. Then, $\xi \in S$.

We will next show that $\xi = b$ by an argument of contradiction. Suppose that $\xi < b$. Let $x := \theta_1(\xi) = \theta_2(\xi)$ and $y := F(x)$. Then, $\exists U \in \mathcal{O}_\mathcal{X}$ with $x \in U$ and $\exists \mathcal{V} \in \mathcal{O}_\mathcal{Y}$ with $y \in \mathcal{V}$ such that $F\vert_U : U \to \mathcal{V}$ is a homeomorphism. By the continuity of $\theta_1$ and $\theta_2$, $\exists [\xi, \xi] \subset \mathbb{R}$ such that $\theta_1(\xi), \theta_2(\xi) \in U$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Then, $\sigma(t) = F(\theta_1(t)) = F(\theta_2(t)) \in V$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Since $F\vert_U : U \to \mathcal{V}$ is a homeomorphism, then $\theta_1(t) = \theta_2(t)$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Then, $\xi \in S$ and $\xi < \xi < \sup S = \xi$. This is a contradiction. Therefore, we must have $\xi = b$.

Therefore, $\theta_1(t) = \theta_2(t)$, $\forall t \in [a, \xi] = [a, b] \subset \mathbb{R}$, since $\xi \in S$. This shows that $\theta_1 = \theta_2$. Hence, if $\theta$ exists then it must be unique.

“Existence” Let $S := \{s \in [a, b] \subset \mathbb{R} \mid \text{there exists a continuous } \theta : [a, s] \to \mathcal{X} \text{ that inverts } F \text{ along } \sigma|_{[a, s]} \text{ with } \theta(a) = x_0 \} \subset \mathbb{R}$ and $\xi := \sup S$. Clearly, $a \in S$ and $a \leq \xi \leq b$.

We will show that $\xi \in S$ by an argument of contradiction. Suppose $\xi \not\in S$. Then, $a < \xi < b$ and $\exists (t_n)_{n=1}^\infty \subseteq S$, which is nondecreasing, such that $\lim_{n \in \mathbb{N}} t_n = \xi$. Then, $\exists \mathcal{V} \in \mathcal{O}_\mathcal{Y}$ with $y \in \mathcal{V}$ such that $F\vert_U : U \to \mathcal{V}$ is a homeomorphism. By the continuity of $\sigma$ and Proposition 3.66, we have $\lim_{n \in \mathbb{N}} \sigma(t_n) = \sigma(\xi) \in \mathcal{Y}$, where the limit operator makes sense since $\mathcal{Y}$ is Hausdorff. Then, $(\theta(t_n))_{n=1}^\infty \subseteq F_{\text{inv}}(M)$, where $M := \{\sigma(t_n) \in \mathcal{Y} \mid n \in \mathbb{N}\} \cup \{\sigma(\xi) \} \subseteq \mathcal{Y}$. Clearly, $M$ is compact in $\mathcal{Y}$. Since $F$ is countably proper, then $F_{\text{inv}}(M)$ is countably compact. By Proposition 5.26, $F_{\text{inv}}(M)$ have the Bolzano-Weierstrass property and $(\theta(t_n))_{n=1}^\infty$ admits a cluster point $x \in F_{\text{inv}}(M)$. By the continuity of $F$ and Proposition 3.66, $(\sigma(F(\theta(t_n))))_{n=1}^\infty \subseteq \{\sigma(t_n)\}_{n=1}^\infty$ admits a cluster point $F(x)$. Since $\mathcal{Y}$ is Hausdorff, then $F(x) = \sigma(\xi) =: y$.

By the assumption of the lemma, $\exists U \in \mathcal{O}_\mathcal{X}$ with $x \in U$ and $\exists \mathcal{V} \in \mathcal{O}_\mathcal{Y}$ with $y \in \mathcal{V}$ such that $F\vert_U : U \to \mathcal{V}$ is a homeomorphism. Since $\sigma$ is continuous, then $\sigma_{\text{inv}}(V)$ is open in $[a, b] \subset \mathbb{R}$. Since $\xi \in \sigma_{\text{inv}}(V)$, then $\exists [\xi, \xi] \subset \mathbb{R}$ such that $[\xi, \xi] \subseteq \sigma_{\text{inv}}(V)$. By $\lim_{n \in \mathbb{N}} t_n = \xi$, $\exists N \in \mathbb{N}$ such that $\forall n \geq N, t_n \in [\xi, \xi] \subset \mathbb{R}$. Since $(\theta(t_n))_{n=1}^\infty$ admits a cluster point $x \in U$, then $\exists n_0 \in \mathbb{N}$ with $n_0 \geq N$ such that $\theta(t_{n_0}) \in \mathcal{V}$. Clearly, $\sigma([t_{n_0}, \xi]) \subseteq \mathcal{V}$. Define $\hat{\theta} := (F\vert_U)^{\text{inv}} \circ \sigma|_{[t_{n_0}, \xi]} : [t_{n_0}, \xi] \to U$. Clearly, $\hat{\theta}$ is continuous and inverts $F$ along $\sigma|_{[t_{n_0}, \xi]}$ with $\hat{\theta}(t_{n_0}) = \theta(t_{n_0})$. Define
Chapter 9. Differentiation in Banach Spaces

Let \( \vartheta : [a, \xi] \to \mathcal{X} \) by \( \vartheta(t) = \theta(t), \forall t \in [a, t_0] \subset \mathbb{R} \), \( \vartheta(t) = \theta_1(t), \forall t \in [t_0, \xi] \). By Theorem 3.11, \( \vartheta \) is continuous. Clearly, \( \vartheta \) inverts \( F \) along \( \sigma \) such that \( F(\sigma) = 0 \). This completes the proof of the lemma.

Next, we will show that \( \xi = b \) by an argument of contradiction. Suppose \( \xi < b \). Since \( \xi \in S \), then \( \exists \) a continuous function \( \theta : [a, \xi] \to \mathcal{X} \) that is continuous and inverts \( F \) along \( \sigma \) such that \( \theta(a) = x_0 \). Let \( x := \theta(\xi) \in \mathcal{X} \) and \( y := F(x) = \sigma(\xi) \in \mathcal{Y} \). By Theorem 3.11, \( \sigma \) is continuous and \( \sigma \) inverts \( F \) along \( \sigma \) with \( \sigma(a) = x_0 \). Then, \( \xi \in S \). This implies that \( \xi < \tilde{\xi} \leq \sup S = \xi \), which is a contradiction. Hence, \( \xi = b \in S \).

This case is proved.

Case 3: \( a < t_0 \leq b \). By Case 1 and 2, \( \exists! \theta_1 : [t_0, b] \to \mathcal{X} \) that is continuous and inverts \( F \) along \( \sigma \) with \( \theta_1(t_0) = x_0 \). Define \( \sigma : [t_0, 2t_0 - a] \to \mathcal{Y} \) by \( \sigma(t) = \sigma(2t_0 - t), \forall t \in [t_0, 2t_0 - a] \subset \mathbb{R} \). By Proposition 3.12, \( \sigma \) is continuous and \( \sigma(t_0) = y_0 \). By Case 2, \( \exists! \theta_2 : [t_0, 2t_0 - a] \to \mathcal{X} \) that is continuous and inverts \( F \) along \( \sigma \) with \( \theta_2(t_0) = x_0 \). Define \( \vartheta : [a, b] \to \mathcal{X} \) by \( \vartheta(t) = \theta_1(t), \forall t \in [t_0, b] \subset \mathbb{R} \), \( \vartheta(t) = \theta_2(2t_0 - t), \forall t \in [a, t_0] \subset \mathbb{R} \). By Theorem 3.11, \( \vartheta \) is continuous and inverts \( F \) along \( \sigma \) with \( \vartheta(t_0) = x_0 \). The uniqueness of \( \vartheta \) follows from the uniqueness of \( \theta_1 \) and \( \theta_2 \). This case is proved.

This completes the proof of the lemma.

\[ \square \]

**Lemma 9.63** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be Hausdorff topological spaces, \( F : \mathcal{X} \to \mathcal{Y} \) be continuous and countably proper, \( x_0 \in \mathcal{X} \), and \( y_0 = F(x_0) \in \mathcal{Y} \). Assume that \( \forall x \in \mathcal{X}, \exists U \in \mathcal{O}_x \) with \( x \in U \) and \( \exists V \in \mathcal{O}_{y_0} \) with \( F(x) \in V \) such that \( F_U : U \to V \) is a homeomorphism. Then, given any continuous mapping \( \sigma : [a, b] \times [c, d] \to \mathcal{Y} \) with \( \sigma(a, c) = y_0 \), where \( a, b, c, d \in \mathbb{R} \), \( a \leq b \), and \( c \leq d \), there exists a unique continuous mapping \( \theta : [a, b] \times [c, d] \to \mathcal{X} \) with \( \theta(a, c) = x_0 \) that inverts \( F \) along \( \sigma \).

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( a = b \); Case 2: \( a < b \). Case 1: \( a = b \). The result holds by Lemma 9.62.

Case 2: \( a < b \). “Uniqueness” Let \( \theta_1 : [a, b] \times [c, d] \to \mathcal{X} \) and \( \theta_2 : [a, b] \times [c, d] \to \mathcal{X} \) be continuous mappings that invert \( F \) along \( \sigma \) with \( \theta_1(a, c) = \theta_2(a, c) = x_0 \). Fix any \( (t, r) \in [a, b] \times [c, d] \subset \mathbb{R}^2 \). Let \( \sigma : [0, 1] \to \mathcal{Y}, \vartheta_1 : [0, 1] \to \mathcal{X}, \vartheta_2 : [0, 1] \to \mathcal{X} \) be defined by, \( \forall \lambda \in [0, 1] \subset \mathbb{R} \),

\[ \begin{align*}
\vartheta(\lambda) &= \sigma(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c) \\
\vartheta_1(\lambda) &= \theta_1(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c) \\
\vartheta_2(\lambda) &= \theta_2(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c)
\end{align*} \]
Since \( \sigma = F \circ \theta_1 \) and \( \sigma = F \circ \theta_2 \), then \( \sigma = F \circ \theta_1 \) and \( \sigma = F \circ \theta_2 \). By Proposition 3.12, \( \sigma \), \( \sigma \), \( \theta_1 \), and \( \theta_2 \) are continuous. This implies that \( \theta_1 \) inverts \( F \) along \( \sigma \) with \( \theta_1(0) = \theta_1(a,c) = x_0 \); and \( \theta_2 \) inverts \( F \) along \( \sigma \) with \( \theta_1(0) = x_0 \). By Lemma 9.62, we have \( \theta_1 = \theta_2 \). Then, \( \theta_1(t,r) = \theta_1(1) = \theta_2(1) = \theta_2(t,r) \). Hence, \( \theta_1 = \theta_2 \). This shows that \( \theta : [a, b] \times [c, d] \to \mathcal{X} \) is unique when it exists.

"Existence" Define \( \sigma : [a, b] \to \mathcal{Y} \) by \( \sigma_c(t) = \sigma(t,c), \forall t \in [a, b] \subset \mathbb{R} \). By Proposition 3.12, \( \sigma \) is continuous with \( \sigma_c(a) = y_0 \). By Lemma 9.62, there exists a unique continuous mapping \( \theta : [a, b] \to \mathcal{X} \) that inverts \( F \) along \( \sigma \) with \( \theta_{\xi}(a) = x_0 \). Fix any \( t \in [a, b] \subset \mathbb{R} \). Define \( \sigma_t : [c, d] \to \mathcal{Y} \) by \( \sigma_t(r) = \sigma(t,r), \forall r \in [c, d] \subset \mathbb{R} \). By Proposition 3.12, \( \sigma_t \) is continuous with \( \sigma_t(c) = \sigma(t,c) = \sigma_c(t) = F(\theta_c(t)) \). By Lemma 9.62, there exists a unique continuous mapping \( \theta_t : [c, d] \to \mathcal{X} \) that inverts \( F \) along \( \sigma_t \) with \( \theta_t(c) = \theta_{\xi}(t) \).

Define \( \theta : [a, b] \times [c, d] \to \mathcal{X} \) by \( \theta(t,r) = \theta_t(r), \forall (t, r) \in [a, b] \times [c, d] \). Clearly, \( \forall (t, r) \in [a, b] \times [c, d] \), \( \sigma(t,r) = \sigma_c(r) = F(\theta_c(r)) = F(\theta(t,r)) \). Hence, \( \theta \) inverts \( F \) along \( \sigma \). \( \theta(a,c) = \theta_a(c) = \theta_{\xi}(a) = x_0 \). All we need to show is that \( \theta \) is continuous to complete the proof of the lemma. Define \( S := \{ r \in [c, d] \subset \mathbb{R} : \theta_{[a,b] \times [c,r]} \text{ is continuous} \} \subset \mathbb{R} \) and \( \xi = \sup S \). Clearly, \( c \in S \) a \( c \leq \xi \leq d \).

We will show that \( \xi \in S \) by an argument of contradiction. Suppose \( \xi \notin S \). Then, \( c < \xi \leq d \) and \( \exists (r_n)_{n=1}^\infty \subset S \), which is nondecreasing, such that \( \lim_{n \in \mathbb{N}} r_n = \xi \), \( \forall (t, r) \in [a, b] \times [c, \xi] \subset \mathbb{R}^2 \), there exists \( n_0 \in \mathbb{N} \) such that \( \forall n \geq n_0 \), we have \( r_n > r \). Then, \( \theta_{[a,b] \times [c,r_n]} \) is continuous implies that \( \theta \) is continuous at \( (t, r) \). Hence, \( \theta_{[a,b] \times [c,\xi]} \) is continuous.

Fix any \( t \in [a, b] \), let \( x = \theta(t, \xi) \in \mathcal{X} \) and \( y = F(x) = \sigma(t, \xi) \in \mathcal{Y} \). Then, \( \exists U \in \mathcal{O}_x \) with \( x \in U \) and \( \exists V \in \mathcal{O}_y \) with \( y \in V \) such that \( F \mid_U : U \to V \) is a homeomorphism. Since \( \sigma \) is continuous, then \( \exists a_t, b_t, c_t, d_t \in \mathbb{R} \) with \( a_t < t < b_t \) and \( c_t < \xi < d_t \) such that \( \sigma(D_t) \subseteq V \), where \( D_t := ((a_t, b_t) \times [c_t, d_t]) \cap ([a, b] \times [c, d]) \subset \mathbb{R}^2 \). Define \( \tilde{\theta} : D_t \to U \) by \( \tilde{\theta}(t, r) = F(\sigma(t, r)), \forall (t, r) \in D_t \). By Proposition 3.12, \( \tilde{\theta} \) is continuous.

Claim 9.63.1 \( \tilde{\theta}(\bar{t}, \bar{r}) = \bar{\theta}(\tilde{t}, \tilde{r}), \forall (\tilde{t}, \tilde{r}) \in D_t \).

Proof of claim: Fix any \( (\tilde{t}, \tilde{r}) \in D_t \). Note that \( \tilde{\theta} \) inverts \( F \) along \( \sigma_{|\tilde{D}_1} \).

Define \( \tilde{\theta} : D_{t,2} \to \mathcal{X} \) by \( \tilde{\theta}(\tilde{t}, \tilde{r}) = \tilde{\theta}(t, r), \forall (t, r) \in D_{t,2} := [c_t, d_t] \cap [c, d] \subset \mathbb{R} \).

Then, \( \tilde{\theta} \) is continuous and inverts \( F \) along \( \sigma_{|\tilde{D}_{t,2}} \) with \( \tilde{\theta}(\xi) = x \).

Note that \( \theta_{|\tilde{D}_{t,2}} \) is also continuous and inverts \( F \) along \( \sigma_{|\tilde{D}_{t,2}} \) with \( \theta_t(\xi) = x \).

By Lemma 9.62, we have \( \tilde{\theta} = \tilde{\theta}_{|\tilde{D}_{t,2}} \) and, in particular, \( \tilde{\theta}(t, c_t) = \tilde{\theta}(\tilde{t}, \tilde{c}_t) = \theta(t, c_t) \).

Define \( \tilde{\theta}_{c_t} : D_{t,1} \to \mathcal{X} \) by \( \tilde{\theta}_{c_t}(\tilde{t}) = \tilde{\theta}(\tilde{t}, c_t) \), \( \forall \tilde{t} \in D_{t,1} := (a_t, b_t) \cap [a, b] \subset \mathbb{R} \). Define \( \sigma_{c_t} : [a, b] \to \mathcal{Y} \) by \( \sigma_{c_t}(t) = \sigma(t, c_t) \), \( \forall t \in [a, b] \).

Define \( \tilde{\theta}_{c_t} : [a, b] \to \mathcal{X} \) by \( \tilde{\theta}_{c_t}(\tilde{t}) = \theta(\tilde{t}, c_t) \), \( \forall \tilde{t} \in [a, b] \).

Then, \( \tilde{\theta}_{c_t} \) is continuous and inverts \( F \) along \( \sigma_{c_t} \) with \( \tilde{\theta}_{c_t}(t) = \theta(t, c_t) \).

Since \( c_t \in [c, \xi] \subset \mathbb{R} \), then \( \theta_{c_t} \mid_{D_{t,1}} \) is continuous and inverts \( F \) along \( \sigma_{c_t} \mid_{D_{t,1}} \).
with \( \theta_{c(t)}(t) = \theta(t, c_t) \). By Lemma 9.62, we have \( \theta_{c(t)}|_{D_t,1} = \bar{\theta}_{c(t)} \) and, in particular, \( \theta_{c(t)} = \theta_{c(t)}(t) = \bar{\theta}_{c(t)}(t) = \bar{\theta}(t, c_t) \). Define \( \bar{\theta}_t : D_{t,2} \to \mathcal{X} \) by \( \bar{\theta}_t(\bar{r}) = \bar{\theta}(t, \bar{r}) \), \( \forall \bar{r} \in D_{t,2} \). Then, \( \bar{\theta}_t \) is continuous and inverts \( F \) along \( \sigma_t|_{D_{t,2}} \) with \( \bar{\theta}_t(c_t) = \theta(t, c_t) \). Note that \( \bar{\theta}_t|_{D_{t,2}} \) is also continuous and inverts \( F \) along \( \sigma_t|_{D_{t,2}} \) with \( \bar{\theta}_t(c_t) = \theta(t, c_t) \). By Lemma 9.62, we have \( \bar{\theta}_t = \theta_t|_{D_{t,2}} \) and, in particular, \( \theta_t(\bar{r}) = \bar{\theta}_t(\bar{r}) = \theta_t(\bar{r}) = \theta_t(\bar{r}) \). This completes the proof of the claim. \( \square \)

Then, \( \theta_t|_{D_t} \) is continuous. Then, \( \theta \) is continuous at \( (t, \xi) \). Then, \( \theta|_{[a,b] \times [e,\xi]} \) is continuous and \( \xi \in S \). This contradicts with the hypothesis \( \xi \notin S \). Therefore, \( \xi \in S \).

Now, we show \( \xi = d \) by an argument of contradiction. Suppose \( \xi < d \). Fix any \( t \in [a, b] \), let \( x = \theta(t, \xi) \in \mathcal{X} \) and \( y = F(x) = \sigma(t, \xi) \in \mathcal{Y} \). Then, \( \exists U \in \mathcal{O}_X \) with \( x \in U \) and \( \exists V \in \mathcal{O}_Y \) with \( y \in V \) such that \( F|_U : U \to V \) is a homeomorphism. Since \( \sigma \) is continuous, then \( \exists \alpha_t, b_t, c_t, d_t \in \mathbb{R} \) with \( a_t < t < b_t \) and \( c_t < \xi < d_t \leq d \) such that \( \sigma(D_t) \subseteq V \), where \( D_t := ((a_t, b_t) \times (c_t, d_t)) \cap ([a, b] \times [c, d]) \subseteq \mathbb{R}^2 \). Define \( \theta : D_t \to U \) by \( \theta(t, \bar{r}) = (F|_U)_{\text{inv}}(\bar{r}) \), \( \forall (t, \bar{r}) \in D_t \). By Proposition 3.12, \( \theta \) is continuous.

Claim 9.63.2 \( \theta(t, \bar{r}) = \theta(t, \bar{r}), \forall (t, \bar{r}) \in D_t \).

**Proof of claim:** Fix any \( (t, \bar{r}) \in D_t \). Note that \( \theta \) inverts \( F \) along \( \sigma|_{D_t} \).

Define \( \theta_x : D_{t,1} \to \mathcal{X} \) by \( \theta_x(t) = \theta(t, \xi) \), \( \forall t \in D_{t,1} := (a_t, b_t) \cap [a, b] \subseteq \mathbb{R} \). Define \( \theta_x : [a, b] \to \mathcal{X} \) by \( \theta_x(t) = \sigma(t, \xi), \forall t \in [a, b] \). Define \( \theta_x : [a, b] \to \mathcal{X} \) by \( \theta_x(t) = \sigma(t, \xi), \forall t \in [a, b] \). Then, \( \theta_x \) is continuous and inverts \( F \) along \( \sigma_x|_{D_{t,1}} \) with \( \theta_x(t) = x \). Since \( \xi \in S \), then \( \theta_x|_{D_{t,1}} \) is continuous and inverts \( F \) along \( \sigma_x|_{D_{t,1}} \) with \( \theta_x(t) = \theta(t, \xi) = x \). By Lemma 9.62, we have \( \theta_x|_{D_{t,1}} = \bar{\theta}_x \) and, in particular, \( \theta_x(t) = \bar{\theta}_x(t) = \theta(t, \xi) \).

Define \( \theta_t : D_{t,2} \to \mathcal{X} \) by \( \theta_t(\bar{r}) = \theta(t, \bar{r}) \), \( \forall \bar{r} \in D_{t,2} := (c_t, d_t) \cap [c, d] \subseteq \mathbb{R} \). Then, \( \theta_t \) is continuous and inverts \( F \) along \( \sigma_t|_{D_{t,2}} \) with \( \theta_t(\bar{r}) = \theta(t, \xi) = \theta(t, \bar{r}) \). Note that \( \theta_t|_{D_{t,2}} \) is also continuous and inverts \( F \) along \( \sigma_t|_{D_{t,2}} \) with \( \theta_t(\xi) = \theta(t, \xi) \). By Lemma 9.62, we have \( \theta_t = \theta_t|_{D_{t,2}} \) and, in particular, \( \theta(t, \bar{r}) = \theta_t(\bar{r}) = \theta_t(\bar{r}) = \theta(t, \bar{r}) \). This completes the proof of the claim. \( \square \)

Then, \( \theta|_{D_t} \) is continuous. Note that \( [a, b] \subseteq \bigcup_{t \in [a,b]} \Theta(a_t, b_t) \).

By the compactness of \( [a, b] \subseteq \mathbb{R} \), there exists a finite set \( T_N \subseteq [a, b] \subseteq \mathbb{R} \) such that \( [a, b] \subseteq \bigcup_{t \in T_N} (a_t, b_t) \). Note that \( \theta|_{D_{t}} \), \( \forall t \in T_N \) and \( \theta|_{[a,b] \times [e,\xi]} \) are continuous. Then, by Theorem 3.11, \( \theta|_{D} \) is continuous, where \( D := ([a, b] \times [c, \xi]) \cup \bigcup_{t \in T_N} D_t \subseteq [a, b] \times [c, d] \subseteq \mathbb{R}^2 \) and all of the sets are relatively open. Set \( d = (\min_{t \in T_N} d_t + \xi)/2 \in (\xi, d] \). It is clear that \( [a, b] \times [c, d] \subseteq D \subseteq \mathbb{R}^2 \). Then, \( \theta|_{[a,b] \times [c,d]} \) is continuous and \( d \in S \). This leads to the contradiction \( \xi < d \leq \sup S = \xi \). Therefore, \( \xi = d \). Then, \( \theta \) is continuous. This completes the proof of Case 2.

This completes the proof of the lemma. \( \square \)
Theorem 9.64 (Global Inverse Function Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $\mathcal{X} \neq \emptyset$, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably proper. Assume that $\forall x \in \mathcal{X}$, $\exists U \in \mathcal{O}_\mathcal{X}$ with $x \in U$ and $\exists V \in \mathcal{O}_\mathcal{Y}$ with $F(x) \in V$ such that $F|_U : U \rightarrow V$ is a homeomorphism, $\mathcal{X}$ is arcwise connected, and $\mathcal{Y}$ is simply connected. Then, $F : \mathcal{X} \rightarrow \mathcal{Y}$ is a homeomorphism.

Proof  Fix $x_0 \in \mathcal{X} \neq \emptyset$, let $y_0 = F(x_0) \in \mathcal{Y}$, $\forall y \in \mathcal{Y}$, since $\mathcal{Y}$ is simply connected, then $\mathcal{Y}$ is arcwise connected. Then, there exists a curve $\sigma : I \rightarrow \mathcal{Y}$, where $I := [0,1] \subset \mathbb{R}$, such that $\sigma(0) = y_0$ and $\sigma(1) = y$. By Lemma 9.62, there exists a continuous mapping $\theta : I \rightarrow \mathcal{X}$ that inverts $F$ along $\sigma$ with $\theta(0) = x_0$. Then, $y = \sigma(1) = F(\theta(1))$. Hence, $F$ is surjective.

Fix $x_1, x_2 \in \mathcal{X}$ such that $F(x_1) = F(x_2) = y$. Since $\mathcal{X}$ is arcwise connected, then there exists a curve $\delta : I \rightarrow \mathcal{X}$ such that $\delta(0) = x_1$ and $\delta(1) = x_2$. Consider the curve $\eta := F \circ \delta$, which is continuous by Proposition 3.12. $\eta$ is a closed curve since $\eta(0) = F(x_1) = y = F(x_2) = \eta(1)$. Since $\mathcal{Y}$ is simply connected, then $\eta$ is homotopic to a single point $\bar{y} \in \mathcal{Y}$. Then, there exists a continuous mapping $\gamma : I \times I \rightarrow \mathcal{Y}$ such that $\gamma(t,0) = \eta(t)$, $\gamma(t,1) = \bar{y}$, and $\gamma(0,t) = \gamma(1,t)$, $\forall t \in I$. By Lemma 9.63, there exists a continuous function $\zeta : I \times I \rightarrow \mathcal{X}$ that inverts $F$ along $\gamma$ with $\zeta(0,0) = x_1$.

Define $\gamma_0 : I \rightarrow \mathcal{Y}$ by $\gamma_0(t) = \gamma(0,t)$, $\gamma_1 : I \rightarrow \mathcal{Y}$ by $\gamma_1(t) = \gamma(1,t)$, $\gamma_0 : I \rightarrow \mathcal{Y}$ by $\gamma_0(t) = \gamma(t,0)$, and $\gamma_1 : I \rightarrow \mathcal{Y}$ by $\gamma_1(t) = \gamma(t,1)$, $\forall t \in I$. Then, $\gamma_0 = \gamma_1$. Define $\zeta_0 : I \rightarrow \mathcal{X}$ by $\zeta_0(t) = \zeta(0,t)$, $\zeta_1 : I \rightarrow \mathcal{X}$ by $\zeta_1(t) = \zeta(1,t)$, $\zeta_0 : I \rightarrow \mathcal{X}$ by $\zeta_0(t) = \zeta(t,0)$, and $\zeta_1 : I \rightarrow \mathcal{X}$ by $\zeta_1(t) = \zeta(t,1)$, $\forall t \in I$. Then, $\zeta_0$ is continuous and inverts $F$ along $\gamma_0$. Set $\bar{x} = \zeta_0(1) = \zeta(0,1) \in \mathcal{X}$, then, $\bar{y} = F(\bar{x})$. $\zeta_1$ is continuous and inverts $F$ along $\gamma_1$ with $\zeta_1(0) = \zeta(0,1) = \bar{x}$. Since $\gamma_1$ is a constant function with value $\bar{y}$, then, the constant mapping $\lambda : I \rightarrow \mathcal{X}$ given by $\lambda(t) = \bar{x}$, $\forall t \in I$ is continuous and inverts $F$ along $\gamma_1$ with $\lambda(0) = \bar{x}$. By Lemma 9.62, we have $\lambda = \zeta_1$ and, in particular, $\bar{x} = \lambda(1) = \zeta_1(1) = \zeta(1,1)$. Note that $\zeta_1$ is continuous and inverts $F$ along $\gamma_1$ with $\zeta_1(1) = \zeta(1,1) = \bar{x}$. Since $\gamma_1 = \gamma_0$. $\zeta_0$ is continuous and inverts $F$ along $\gamma_1$ with $\zeta_0(1) = \bar{x}$. By Lemma 9.62, we have $\zeta_0 = \zeta_1$ and, in particular $\zeta_1(0) = \zeta_0(0) = \zeta(0,0) = x_1$. Note that $\zeta_0$ is continuous and inverts $F$ along $\gamma_0 = \eta$ with $\zeta_0(0) = \zeta(0,0) = x_1$. By construction, $\delta$ is continuous and inverts $F$ along $\eta$ with $\delta(0) = x_1$. By Lemma 9.62, $\delta = \zeta_0$ and, in particular, $x_2 = \delta(1) = \zeta_0(1) = \zeta(1,0) = x_1$. Hence, $x_1 = x_2$. Therefore, $F$ is injective.

Hence, $F$ is bijective and admits inverse $F_i : \mathcal{Y} \rightarrow \mathcal{X}$. $\forall y \in \mathcal{Y}$, let $x = F_i(y) \in \mathcal{X}$. Then, $\exists U \in \mathcal{O}_\mathcal{X}$ with $x \in U$ and $\exists V \in \mathcal{O}_\mathcal{Y}$ with $y \in V$ such that $F_i|_U : U \rightarrow V$ is a homeomorphism. Then, $F_i|_V$ is the inverse of $F|_U$ and is continuous. Then, $F_i$ is continuous at $y$ since $V$ is open. By the arbitrariness of $y$, $F_i$ is continuous. Hence, $F$ is a homeomorphism.

This completes the proof of the theorem. \hfill $\square$

Theorem 9.65 (Global Inverse Function Theorem) Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces, $F : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous and countably proper.
Thus, let 

\[ F := \{ x \in X \mid \exists U \in \mathcal{O} \text{ with } x \in U \ni F|_U : U \to F(U) \in \mathcal{Y} \text{ is a homeomorphism} \} \subseteq \mathcal{X}, \quad \Sigma := \mathcal{X} \setminus H, \quad \Sigma_0 := \mathcal{F}(\Sigma), \]

\[ X_0 := \mathcal{X} \setminus \Sigma_0, \text{ and } Y_0 := \mathcal{Y} \setminus F(\Sigma). \]

Assume that \( X_0 \neq \emptyset \) is arcwise connected, and \( Y_0 \) is simply connected. Then, \( G := F|_{X_0} : X_0 \to Y_0 \) is a homeomorphism.

**Proof** Let \( \mathcal{O}_{X_0} \) and \( \mathcal{O}_{Y_0} \) be the subset topology on \( X_0 \) and \( Y_0 \), respectively.

**Claim 9.65.1** \( \Sigma \subseteq \mathcal{X} \) is closed.

**Proof of claim:** \( \forall x \in H, \exists U_x \in \mathcal{O}_X \) with \( x \in U_x \) such that \( F|_{U_x} : U_x \to F(U_x) \in \mathcal{Y} \) is a homeomorphism. \( \forall \bar{x} \in U_x, \bar{x} \in H. \) Then, \( U_x \subseteq H \) and \( H = \bigcup_{x \in H} U_x. \) Hence, \( H \in \mathcal{O}_X. \) Then, \( \Sigma := \mathcal{X} \setminus H \) is closed. \( \square \)

**Claim 9.65.2** \( F(\Sigma_0) = X_0. \) \( G : X_0 \to Y_0 \) is continuous and countably proper.

**Proof of claim:** By Proposition 2.5, \( F(\Sigma_0) = F(\Sigma) \setminus F(\Sigma) = \mathcal{X} \setminus \Sigma_0 = X_0 \) and \( F(X_0) = F(F(\Sigma_0)) \subseteq Y_0. \) Hence, \( G \) is a function of \( X_0 \) to \( Y_0. \)

Fix any \( K \subseteq Y_0 \) such that \( K \) is compact in \( \mathcal{O}_{Y_0}. \) Then \( K \) is a compact set in \( \mathcal{O}_Y. \) Then, \( G(K) = F(K) \subseteq X_0. \) By the countable properness of \( F, \) we have \( F(K) \subseteq \mathcal{X} \) is countably compact in \( \mathcal{O}_X. \) Then, it is easy to show that \( G(K) \) is countably compact in \( \mathcal{O}_{X_0}. \) Hence, \( G \) is countably proper.

\( \forall V_0 \in \mathcal{O}_{Y_0}, V_0 = Y_0 \cap V, \) where \( V \in \mathcal{O}_Y. \) \( G(V_0) = F(V_0) = F(\Sigma_0) \cap F(V) = X_0 \cap F(V). \) Since \( F \) is continuous, \( F(V) \in \mathcal{O}_X. \) Thus, \( G(V_0) \in \mathcal{O}_{X_0}. \) Hence, \( G \) is continuous. \( \square \)

By Proposition 2.5, \( \Sigma_0 = F(F(\Sigma)) \supset \Sigma \cap \text{dom } F = \Sigma. \) Then, \( X_0 = \mathcal{X} \setminus \Sigma_0 \subseteq \mathcal{X} \setminus \Sigma = H. \) \( \forall x \in X_0 \subseteq H, \exists U \in \mathcal{O}_X \) with \( x \in U \) such that \( F|_U : U \to F(U) \in \mathcal{O}_Y \) is a homeomorphism. Let \( U_0 := X_0 \cap U \in \mathcal{O}_{X_0}, \) \( V := F(U) \in \mathcal{O}_Y, \) and \( V_0 := F(U_0) \cap Y_0 \in \mathcal{O}_{Y_0}. \) Clearly, \( x \in U_0. \) By Proposition 2.5, \( G(U_0) = F(V_0) \subseteq F(X_0) \cap F(U) \subseteq Y_0 \cap V = V_0. \) Then, \( G(U_0) : U_0 \to V_0. \) Note that \( G|_{U_0} = (F|_U)|_{U_0}. \) Since \( F|_U \) is injective, then \( G|_{U_0} \) is injective. \( \forall \bar{y} \in V_0, \bar{y} \in V \) then \( \exists x \in U \) such that \( \bar{y} = F(\bar{x}). \) Note that \( \bar{y} \in V_0 \) and \( F(\Sigma_0) = X_0, \) then \( \bar{x} \in X_0. \) Hence, \( \bar{x} \in U_0 \) and \( G(\bar{x}) = F(\bar{x}) = \bar{y}. \) Then, \( G|_{U_0} : U_0 \to V_0 \) is surjective. Hence, \( G|_{U_0} \) is bijective with inverse \( G_i = F_i|_{V_0}, \) where \( F_i \) is the inverse of \( F|_U : U \to V. \) Since \( F|_U \) is homeomorphism, then \( F|_U \) and \( F_i \) are continuous. Then, \( G|_{U_0} \) and \( G_i \) are continuous. This shows that \( G|_{U_0} : U_0 \to V_0 \) is a homeomorphism.

By Global Inverse Function Theorem 9.64, \( G : X_0 \to Y_0 \) is a homeomorphism. This completes the proof of the theorem. \( \square \)

**Proposition 9.66** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space and \( A_{\alpha} \subseteq X \) be arcwise connected (in subset topology), \( \forall \alpha \in \Lambda, \) where \( \Lambda \) is an index set. Assume that \( A_{\alpha_1} \cap A_{\alpha_2} \neq \emptyset, \forall \alpha_1, \alpha_2 \in \Lambda. \) Then, \( A := \bigcup_{\alpha \in \Lambda} A_{\alpha} \) is arcwise connected (in subset topology).
Proposition 9.67 Let $X$ be a normed linear space and $O \subseteq X$ be open and connected. Then, $O$ is arcwise connected.

Proof The result is trivial if $O = \emptyset$. Let $x_0 \in O \neq \emptyset$. Let $\mathcal{A} := \{ A \subseteq O \mid x_0 \in A \text{ and } A \text{ is arcwise connected} \}$ and $A_0 := \bigcup A \in \mathcal{A}$. Then, $x_0 \in A_0 \subset O$ and, by Proposition 9.66, $A_0$ is arcwise connected. \( \forall x \in A_0 \subset O \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that $B_X(x, \delta) \subseteq O$. Clearly, $B_X(x, \delta)$ is arcwise connected. Then, by Proposition 9.66, $A_0 \cup B_X(x, \delta)$ is arcwise connected. This implies that $A_0 \cup B_X(x, \delta) \subset \mathcal{A}$. Then, $B_X(x, \delta) \subseteq A_0$. By the arbitrariness of $x$, we have $A_0$ is open in $X$. We will show that $A_0 = O$ by an argument of contradiction. Suppose $A_0 \subset O$. Let $E := O \setminus A_0 \neq \emptyset$. Let $\partial E$ be the boundary of $E$ in $X$.

Claim 9.67.1 $\partial E \cap E \neq \emptyset$.

Proof of claim: Suppose $\partial E \cap E = \emptyset$. By Proposition 3.3, $E = \overline{E} \cap E = (\partial E \cup E^o) \cap E = E^o \cap E = E^o$. Hence, $E$ is open in $X$. Then, $A_0$ and $E$ form a separation of $O$. This contradicts with the assumption that $O$ is connected. Hence, the claim holds.

Let $x_1 \in \partial E \cap E \subseteq O$. Then, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $B_X(x_1, \delta_1) \subseteq O$. \( x_1 \in \partial E \) implies that $B_X(x_1, \delta_1) \cap \overline{E} \neq \emptyset$. This implies that $\exists x_2 \in B_X(x_1, \delta_1) \cap (O \setminus E) = B_X(x_1, \delta_1) \cap A_0$. By Proposition 9.66, $A_0 \cup B_X(x_1, \delta_1)$ is arcwise connected. Then, by the definition of $A_0$, we have $B_X(x_1, \delta_1) \subseteq A_0$ and $x_1 \in A_0$. This contradicts with the fact that $x_1 \in E = O \setminus A_0$. Therefore, $A_0 = O$ and $O$ is arcwise connected. This completes the proof of the proposition.

9.7 Interchange Differentiation and Limit

Proposition 9.68 Let $X$ and $Y$ be normed linear spaces over $K$, $D \subseteq X$, $x_0 \in D$, and $F_n : D \rightarrow Y$, $\forall n \in \mathbb{N}$. Assume that

(i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $\tilde{D} := D \cap B_X(x_0, \delta_0) - x_0$ is a conic segment;
(ii) $F_n$ is differentiable, $\forall n \in \mathbb{N}$;

(iii) $\left( F^{(1)}_n \right)_{n=1}^{\infty}$ converges uniformly to $G : D \to B(X, Y)$;

(iv) $\forall x \in D$, $\lim_{n \in \mathbb{N}} F_n(x) = F(x)$, where $F : D \to Y$.

Then, $F$ is differentiable at $x_0$ and $F^{(1)}(x_0) = G(x_0) = \lim_{n \in \mathbb{N}} F^{(1)}_n(x_0)$.

**Proof** By the differentiability of $F_1$, we have $\text{span}(A_D(x)) = X$, $\forall x \in D$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by (iii), $\exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq n_0$, $\| F^{(1)}_n(x) - G(x) \| < \epsilon / 4$, $\forall x \in D$. $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, by Proposition 9.15, $g : D \to \mathbb{Y}$, defined by $g(x) = F_n(x) - F_{n_0}(x)$, $\forall x \in D$, is differentiable and $g^{(1)}(x) = F^{(1)}_n(x) - F^{(1)}_{m}(x)$, $\forall x \in D$. By Mean Value Theorem 9.23, $\forall x \in D + x_0$ and $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, $\| F_n(x) - F_m(x) - F_n(x_0) + F_m(x_0) \| = \| g(x) - g(x_0) \| \leq \| g^{(1)}(\bar{x})(x - x_0) \| \leq \| F^{(1)}_n(\bar{x}) - F^{(1)}_{m}(\bar{x}) \| \| x - x_0 \| \leq \epsilon \| x - x_0 \| / 2$, where $\bar{x} = t_0 x + (1 - t_0) x_0 \in D + x_0$ and $t_0 \in (0, 1) \subset \mathbb{R}$. By $F^{(1)}_n(x_0) \in B(X, Y)$, $\exists \delta \in (0, \delta_0] \subset \mathbb{R}$ such that, $\forall x \in D := D \cap B_X(x_0, \delta) \subseteq D + x_0$, we have $\| F_{n_0}(x) - F_{n_0}(x_0) - F^{(1)}_{n_0}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|/4$.

Fix any $x \in \bar{D}$. $\forall m \in \mathbb{N}$ with $m \geq n_0$, we have $\| F_{m}(x) - F_{m}(x_0) - G(x_0)(x - x_0) \| \leq \| F_{m}(x) - F_{m}(x_0) - F_{n_0}(x) + F_{n_0}(x_0) \| + \| F_{n_0}(x) - F_{n_0}(x_0) - F^{(1)}_{n_0}(x_0)(x - x_0) \| + \| F^{(1)}_{n_0}(x_0)(x - x_0) - F^{(1)}_{n_0}(x_0)(x - x_0) \| + \| F^{(1)}_{n_0}(x_0)(x - x_0) - F^{(1)}_{n_0}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|/4 + \epsilon \| x - x_0 \|/4 + \epsilon \| x - x_0 \|/4 = \epsilon \| x - x_0 \|$. Take limit as $m \to \infty$, we have $\| F(x) - F(x_0) - G(x_0)(x - x_0) \| \leq \lim_{m \in \mathbb{N}} \| F_m(x) - F_{n_0}(x) - G(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|$. By the arbitrariness of $x$, $F$ is differentiable at $x_0$ and $F^{(1)}(x_0) = G(x_0) = \lim_{n \in \mathbb{N}} DF_n(x_0)$. This completes the proof of the proposition.

**Example 9.69** Let $X$ and $Y$ be normed linear spaces over $\mathbb{K}$, $\Omega \subseteq X$ be a compact set, which satisfies $\text{span}(A_D(x)) = X$, $\forall x \in \Omega$, $k \in \mathbb{N}$, and $W := C(\Omega, Y)$ be the normed linear space defined in Example 7.31. Define $Z := \{ f \in C(\Omega, Y) \mid f \text{ is } C^k \}$. By Proposition 9.40, $Z$ is a subspace of $C(\Omega, Y)$. Then, $Z := (Z, \oplus_W, \otimes_W, \partial_W)$ is vector space over the field $\mathbb{K}$. $\forall f \in Z$, $\forall i \in \{ 0, \ldots, k \}$, $f^{(i)} \in C(\Omega, B_i(X, Y))$ with norm $\| f^{(i)} \|_C \in \mathbb{R}$.

Now, define a norm on $Z$ by $\| f \|_{C_k} := \left( \sum_{i=0}^{k} \| f^{(i)} \|_C^2 \right)^{1/2} \in [0, \infty) \subset \mathbb{R}$, $\forall f \in Z$. If $f = \partial_W$, by Proposition 9.33, $f$ is $C_\infty$ and $f^{(i)}(x) = \partial_{B_i(X, Y)}$, $\forall x \in \Omega$, $\forall i \in \mathbb{N}$, then, $\| f \|_{C_k} = 0$. On the other hand, if $\| f \|_{C_k} = 0$, then $f = \partial_W$. $\forall f_1, f_2 \in Z$, by Proposition 9.40, $\| f_1 + f_2 \|_{C_k} = \left( \sum_{i=0}^{k} \| f_1^{(i)} + f_2^{(i)} \|_C^2 \right)^{1/2} \leq \left( \sum_{i=0}^{k} \| f_1^{(i)} \|_C^2 \right)^{1/2} + \left( \sum_{i=0}^{k} \| f_2^{(i)} \|_C^2 \right)^{1/2} = \| f_1 \|_{C_k} + \| f_2 \|_{C_k}$, where the inequality follows from Minkowski’s Inequality, Theorem 7.9. $\forall \alpha \in \mathbb{K}$, $\forall f \in Z$, by Proposition 9.40, $\| \alpha f \|_{C_k} = \| f \|_{C_k}$.
linear space, which will be denoted by \( C \). This shows that \( (Z, K, \|\cdot\|_{C}) \) is a normed linear space, which will be denoted \( C_{k}(\Omega, y) \).

**Example 9.70** Let \( X \) be a normed linear space over \( K \) and \( y \) be a Banach space over \( K \), \( \Omega \subseteq X \) be a compact set, and \( k \in \mathbb{N} \). Assume that, \( \forall x \in \Omega \), \( \exists x_{0} \in (0, \infty) \subset \mathbb{R} \) such that \( \Omega \cap B_{X}(x, \delta_{x}) = x \) is a conic segment and \( \text{span}(A_{\Omega}(x)) = X \). Let \( C_{k}(\Omega, y) \) be the normed linear space defined in Example 9.69. We will show that \( C_{k}(\Omega, y) \) is also a Banach space over \( K \).

Fix any Cauchy sequence \( (f_{n})_{n=1}^{\infty} \subseteq C_{k}(\Omega, y) \). By the definition of the norm \( \|\cdot\|_{C_{k}} \), \( (f_{n})_{n=1}^{\infty} \subseteq C(\Omega, B_{1}(X, y)) \) is a Cauchy sequence, \( \forall i \in \{0, \ldots, k\} \), \( \forall i \in \{0, \ldots, k\} \), by Example 7.32 and Proposition 7.66, \( C(\Omega, B_{1}(X, y)) \) is a Banach space. Then, \( \exists g_{i} \in C(\Omega, B_{1}(X, y)) \) such that \( \lim_{n \in \mathbb{N}} f_{n}^{(i)} = g_{i} \) in \( C(\Omega, B_{1}(X, y)) \). Then, \( (f_{n})_{n=1}^{\infty} \) converges uniformly to \( g_{i} \) and \( g_{i} \) is continuous. By Proposition 9.68, we have \( g_{i}^{(1)} = g_{i+1}, \forall i \in \{0, \ldots, k-1\} \). Then, we have \( g_{0}^{(i)} = g_{i}, \forall i \in \{0, \ldots, k\} \), and \( g_{0} \in C_{k}(\Omega, y) \). Then, \( \|f_{n} - g_{0}\|_{C_{k}} = \left( \sum_{i=0}^{k} \|f_{n}^{(i)} - g_{0}^{(i)}\|_{C}^{2} \right)^{1/2} \rightarrow 0 \) as \( n \rightarrow \infty \), where the first equality follows from Proposition 9.40. Then, \( \lim_{n \in \mathbb{N}} f_{n} = g_{0} \) in \( C_{k}(\Omega, y) \). Hence, \( C_{k}(\Omega, y) \) is complete and therefore a Banach space.

Sometimes, we need to consider a normed linear space of continuous functions on a topological space which is not necessarily compact. This leads us to the following examples.

**Example 9.71** Let \( \mathcal{X} := (X, \mathcal{C}) \) be a topological space, \( y \) be a normed linear space over the field \( K \), and \( C(\mathcal{X}, y) \) be the vector space of all continuous functions of \( \mathcal{X} \) to \( y \) as defined in Example 7.50 with null vector \( \vartheta \). Define a function \( \|\cdot\| : C(\mathcal{X}, y) \rightarrow \mathbb{R}_{e} \) by \( \|f\| = \max \{ \sup_{x \in \mathcal{X}} \|f(x)\|_{y}, 0 \} \), \( \forall f \in C(\mathcal{X}, y) \). Consider the set \( \mathcal{M} := \{ f \in C(\mathcal{X}, y) \mid \|f\| < +\infty \} \). Clearly, \( \vartheta \in \mathcal{M} \). \( \forall f_{1}, f_{2} \in \mathcal{M}, \forall \alpha, \beta \in K, \|\alpha f_{1} + \beta f_{2}\| = \max \{ \sup_{x \in \mathcal{X}} \|\alpha f_{1}(x) + \beta f_{2}(x)\|_{y}, 0 \} \leq \max \{ \sup_{x \in \mathcal{X}} \|\alpha f_{1}(x)\|_{y} + \|\beta f_{2}(x)\|_{y}, 0 \} \leq +\infty \). Then, \( \alpha f_{1} + \beta f_{2} \in \mathcal{M} \). Hence, \( \mathcal{M} \) is a subspace of \( C(\mathcal{X}, y) \).

Clearly, \( \forall f \in \mathcal{M}, \|f\| \in [0, \infty) \subset \mathbb{R} \) and \( \|f\| = 0 \Leftrightarrow f = \vartheta \). \( \forall f_{1}, f_{2} \in \mathcal{M}, \forall \alpha \in K, \|f_{1} + f_{2}\| = \max \{ \sup_{x \in \mathcal{X}} \|f_{1}(x) + f_{2}(x)\|_{y}, 0 \} \leq \max \{ \sup_{x \in \mathcal{X}} \|f_{1}(x)\|_{y} + \sup_{x \in \mathcal{X}} \|f_{2}(x)\|_{y}, 0 \} = \|f_{1}\| + \|f_{2}\| \) for all \( \alpha \neq 0 \), where the first inequality follows from Proposition 3.81. \( \|\alpha f_{1}\| = \max \{ \sup_{x \in \mathcal{X}} \|\alpha f_{1}(x)\|_{y}, 0 \} = \max \{ \sup_{x \in \mathcal{X}} \|\alpha f_{1}(x)\|_{y}, 0 \} = \{ \begin{align*} \max \{ \|\alpha\| \sup_{x \in \mathcal{X}} \|f_{1}(x)\|_{y}, 0 \} & \quad \alpha \neq 0, \\ 0 & \quad \alpha = 0 \end{align*} \} = |\alpha| \|f_{1}\| \), where the third equality follows from Proposition 3.81. Hence, \( (\mathcal{M}, K, \|\cdot\|) \) is a normed linear space, which will be denoted by \( C_{0}(\mathcal{X}, y) \).
Example 9.72  Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space and $\mathcal{Y}$ be a Banach space over the field $\mathbb{K}$ (with norm $\| \cdot \|_Y$). Consider the normed linear space $C_b(\mathcal{X}, \mathcal{Y})$ (with norm $\| \cdot \|$) defined in Example 9.71. We will show that this space is a Banach space. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \neq \emptyset$. Case 1: $\mathcal{X} = \emptyset$. Then, $C_b(\mathcal{X}, \mathcal{Y})$ is a singleton set. Hence, any Cauchy sequence in $C_b(\mathcal{X}, \mathcal{Y})$ must converge. Thus, $C_b(\mathcal{X}, \mathcal{Y})$ is a Banach space. Case 2: $\mathcal{X} \neq \emptyset$. Take a Cauchy sequence $(f_n)_{n=1}^{\infty} \subseteq C_b(\mathcal{X}, \mathcal{Y})$. Let $\epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $\forall n, m \geq N$, $0 \leq \| f_n(x) - f_m(x) \|_Y \leq \| f_n - f_m \| < \epsilon$, $\forall x \in \mathcal{X}$. This shows that, $\forall x \in \mathcal{X}$, $(f_n(x))_{n=1}^{\infty} \subseteq \mathcal{Y}$ is a Cauchy sequence, which converges to $f(x) \in \mathcal{Y}$ since $\mathcal{Y}$ is complete. It is easy to show that $(f_n)_{n=1}^{\infty}$, viewed as a sequence of functions of $\mathcal{X}$ to $\mathcal{Y}$, converges uniformly to $f$. By Proposition 4.26, $f$ is continuous. $\forall x \in \mathcal{X}$, $\| f(x) \|_Y \leq \| f_n(x) - f(x) \|_Y + \| f_n(x) \|_Y \leq \epsilon + \| f_n \|$. Hence, $\| f \| \leq \| f_n \| + \epsilon$. Then, $f \in C_b(\mathcal{X}, \mathcal{Y})$. It is easy to show that $\lim_{n \in \mathbb{N}} \| f_n - f \| = 0$. Hence, $\lim_{n \in \mathbb{N}} f_n = f$ in $C_b(\mathcal{X}, \mathcal{Y})$. Hence, $C_b(\mathcal{X}, \mathcal{Y})$ is a Banach space. In both cases, we have shown that $C_b(\mathcal{X}, \mathcal{Y})$ is a Banach space when $\mathcal{Y}$ is a Banach space.

Example 9.73  Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be endowed with the subset topology, which satisfies $\text{span}(A_{\Omega}(x)) = \mathcal{X}$, $\forall x \in \Omega$, $k \in \mathbb{N}$, and $W := C_b(\Omega, \mathcal{Y})$ be the normed linear space defined in Example 9.71 with null vector $\theta_W$. Define $Z := \{ f \in C_b(\Omega, \mathcal{Y}) \mid f$ is $C_k$ and $f^{(i)} \in C_b(\Omega, B_i(\mathcal{X}, \mathcal{Y}))$, $i = 1, \ldots, k \}$. By Proposition 9.40, $Z$ is a subspace of $W$. Then, $Z := (Z, \oplus, W, \ominus, \theta_W)$ is vector space over the field $\mathbb{K}$. $\forall f \in Z$, $\forall i \in \{0, \ldots, k\}$, $f^{(i)} \in C_b(\Omega, B_i(\mathcal{X}, \mathcal{Y}))$ with norm $\| f^{(i)} \|_{C_k} \in \mathbb{R}$. Now, define a norm on $Z$ by $\| f \|_{C_k} := \left( \sum_{i=0}^{k} \| f^{(i)} \|_{C_k}^2 \right)^{1/2} \in [0, \infty) \subseteq \mathbb{R}$, $\forall f \in Z$. If $f = \theta_W$, by Proposition 9.33, $f$ is $C_\infty$ and $f^{(i)}(x) = \theta_{B_i(\mathcal{X}, \mathcal{Y})}$, $\forall x \in \Omega$, $\forall i \in \mathbb{N}$, then $\| f \|_{C_k} = 0$. On the other hand, if $\| f \|_{C_k} = 0$, then $f = \theta_W$. $\forall f_1, f_2 \in Z$, by Proposition 9.40, $\| f_1 + f_2 \|_{C_k} = \left( \sum_{i=0}^{k} \| f^{(i)}_1 + f^{(i)}_2 \|^2_{C_k} \right)^{1/2} \leq \left( \sum_{i=0}^{k} \| f^{(i)}_1 \|^2_{C_k} \right)^{1/2} + \left( \sum_{i=0}^{k} \| f^{(i)}_2 \|^2_{C_k} \right)^{1/2} = \| f_1 \|_{C_k} + \| f_2 \|_{C_k}$, where the inequality follows from Minkowski's Inequality, Theorem 7.9. $\forall \alpha \in \mathbb{K}$, $\forall f \in Z$, by Proposition 9.40, $\| \alpha f \|_{C_k} = \left( \sum_{i=0}^{k} \| \alpha f^{(i)} \|^2_{C_k} \right)^{1/2} = \left( \sum_{i=0}^{k} | \alpha |^2 \| f^{(i)} \|^2_{C_k} \right)^{1/2} = | \alpha | \| f \|_{C_k}$, which shows that $(Z, \mathbb{K}, \| \cdot \|_{C_k})$ is a normed linear space, which will be denoted $C_{k}(\Omega, \mathcal{Y})$.

Example 9.74  Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$ and $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be endowed with the subset topology, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subseteq \mathbb{R}$ such that $\Omega \cap B_{\delta_x}(x, \delta_x) - x$ is a conic segment and $\text{span}(A_{\Omega}(x)) = \mathcal{X}$. Let $C_{b,k}(\Omega, \mathcal{Y})$ be the normed
9.8. TENSOR ALGEBRA

Definition 9.75 Let \( m \in \mathbb{Z}_+ \), \( X_i \) be a normed linear space over \( \mathbb{K} \), \( i = 1, \ldots, m \), \( \mathcal{Z} \) be a normed linear space over \( \mathbb{K} \). A bounded linear operator \( A \in \mathcal{B}(X_m, \mathcal{B}(X_{m-1}, \ldots, \mathcal{B}(X_1, \mathcal{Z})\cdots)) \) is said to be an \( m \)th order \( \mathcal{Z} \)-valued tensor. Let \( B \in \mathcal{B}(Y_n, \mathcal{B}(Y_{n-1}, \ldots, \mathcal{B}(Y_1, X_m)\cdots)) \) be another \( n \)th order \( X_m \)-valued tensor. We define \( AB := A \cdot B \in \mathcal{B}(Y_n, \mathcal{B}(Y_{n-1}, \ldots, \mathcal{B}(Y_1, X_{m-1}, \ldots, \mathcal{B}(X_1, \mathcal{Z})\cdots))\cdots) \) to be an \((n+m-1)\)st order \( \mathcal{Z} \)-valued tensor such that \((AB)(y_n)\cdots(y_1)(x_{n-1})\cdots(x_1) = A(B(y_n)\cdots(y_1))x_{n-1}\cdots(x_1) \in \mathcal{Z}, \forall y_i \in Y_i, i = 1, \ldots, n, \forall x_j \in X_j, j = 1, \ldots, m - 1\). Let \((n_1, \ldots, n_m)\) be any permutation of \((1, \ldots, m)\). Then, we may define the transpose of tensor \( A \) with permutation \((n_1, \ldots, n_m)\) to be \( A^{T_{n_1,\ldots,n_m}} \in \mathcal{B}(X_{n_m}, \mathcal{B}(X_{n_{m-1}}, \mathcal{B}(X_{n_{m-2}}, \mathcal{Z})\cdots)) \) such that

\[
A^{T_{n_1,\ldots,n_m}}(x_{n_m})\cdots(x_{n_1}) = A(x_m)\cdots(x_1) \quad \forall x_i \in X_i, i = 1, \ldots, m
\]

Proposition 9.76 Let \( m, n \in \mathbb{Z}_+ \), \( X_i, i = 1, \ldots, m \), \( Y_j, j = 1, \ldots, n \), \( \mathcal{Z} \) be normed linear spaces over \( \mathbb{K} \), \( A, A_k \in \mathcal{W}_1 := \mathcal{B}(X_m, \mathcal{B}(X_{m-1}, \mathcal{Z})\cdots), k = 1, 2 \), be \( m \)th order \( \mathcal{Z} \)-valued tensors, \( B, B_l \in \mathcal{W}_2 := \mathcal{B}(Y_n, \mathcal{B}(Y_{n-1}, X_m)\cdots) \) be \( n \)th order \( X_m \)-valued tensors, \( l = 1, 2 \), and \( \mathcal{W}_3 := \mathcal{B}(Y_n, \mathcal{B}(Y_{n-1}, \mathcal{B}(X_{m-1}, \mathcal{B}(X_1, \mathcal{Z})\cdots))\cdots) \). Then, the following statements hold.

(i) \( \| AB \| \leq \| A \| \| B \| \).

(ii) \( \forall \alpha_k, \beta_l \in \mathbb{K}, k = 1, 2, l = 1, 2, \) we have \( (\alpha_1 A_1 + \alpha_2 A_2)(\beta_1 B_1 + \beta_2 B_2) = \alpha_1 \beta_1 A_1 B_1 + \alpha_1 \beta_2 A_1 B_2 + \alpha_2 \beta_1 A_2 B_1 + \alpha_2 \beta_2 A_2 B_2 \).
(iii) Let \( A \in B(\mathcal{X}_B, B(\mathcal{X}_I, \mathbb{K})) \) be a second order \( \mathbb{K} \)-valued tensor. Then, 
\[
A^T_{2,1} = A'\phi_{X_I}, \quad \text{where} \quad \phi_{X_I} : \mathcal{X}_I \to \mathcal{X}_I^{**} \quad \text{is the natural mapping as defined in Remark 7.88.}
\]

(iv) Let \( f : W_1 \times W_2 \to W_3 \) be defined by \( f(A, B) = AB \in W_3, \forall A \in W_1, \forall B \in W_2 \). Then, \( f \) is continuous, \( f^{(1)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = A_0\Delta_{2,1} + \Delta_{1,1}B_0, \) \( f^{(2)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1})(\Delta_{1,2}, \Delta_{2,2}) = \Delta_{1,2}\Delta_{2,1} + \Delta_{1,1}\Delta_{2,2}, \) and \( f^{(i+2)}(A_0, B_0) = \vartheta_B \iota_{\mathbb{R}_{i+2}[W_1 \times W_2, W_3]}, \forall (A_0, B_0) \in W_1 \times W_2, \forall (\Delta_{1,1}, \Delta_{2,1}) \in W_1 \times W_2, \forall (\Delta_{1,2}, \Delta_{2,2}) \in W_1 \times W_2 \), \( \forall i \in \mathbb{N} \).

(v) Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( x_0 \in D \subseteq \mathcal{X} \), \( A : D \to W_1 \) and \( B : D \to W_2 \) be tensor-valued functions that are Fréchet differentiable at \( x_0 \), and \( C : D \to W_3 \) be defined by \( C(x) = A(x)B(x), \forall x \in D \). Then,
\[
C^{(1)}(x_0) = \left( (A^{(1)}(x_0))^{T_1,\ldots,m-1,m+1;\ldots,n+1} B(x_0) \right)^{T_1,\ldots,m-1,m+1;\ldots,n+1} + A(x_0)B^{(1)}(x_0)
\]

(vi) Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( x_0 \in D \subseteq \mathcal{X} \), \( A : D \to W_1 \) be a tensor-valued function that is Fréchet differentiable at \( x_0 \), \( (n_1, \ldots, n_m) \) be a permutation of \( (1, \ldots, m) \), and \( C : D \to W_p := B(\mathcal{X}_{n_m}, \ldots, B(\mathcal{X}_{n_1}, \mathbb{Z}) \ldots) \) be defined by \( C(x) = (A(x))^{T_{n_1,\ldots,n_m}}, \forall x \in D \). Then, we have
\[
C^{(1)}(x_0) = (A^{(1)}(x_0))^{T_{n_1,\ldots,n_m}, m+1}
\]

**Proof** (i) and (ii) These are straightforward, and are therefore omitted.
(iii) Note that \( A \in B(\mathcal{X}_2, \mathcal{X}_2^*) \). Then, \( A' \in B(\mathcal{X}_I^*, \mathcal{X}_2^*) \) and \( A^T_{2,1} \in B(\mathcal{X}_B, B(\mathcal{X}_I, \mathbb{K})) = B(\mathcal{X}_I, \mathcal{X}_2^*). \) Then, \( A^T_{2,1}(x_1)(x_2) = A(x_2)(x_1) = (A(x_2)(x_1))^{(\phi_{X_I}(x_1), A(x_2))} = (A'(\phi_{X_I}(x_1), x_2)) = A'(\phi_{X_I}(x_1))(x_2), \forall x_i \in \mathcal{X}_I, i = 1, 2. \) Then, we have \( A^T_{2,1} = A'\phi_{X_I}. \)
(iv) is straightforward, and is therefore omitted.
(v) follows directly from (iv), the Chain Rule and Proposition 9.19.
(vi) is straightforward, and is therefore omitted.

**Definition 9.77** Let \( \mathcal{X}_i, i = 1, \ldots, m, \mathcal{Z}, y_j, j = 1, \ldots, n, \) be normed linear spaces over \( \mathbb{K} \), \( A \in B(\mathcal{X}_m, B(\mathcal{X}_I, \mathcal{Z}) \ldots) \) be an \( m \)-th order \( \mathcal{Z} \)-valued tensor, and \( B \in B(\mathcal{Y}_n, \ldots, B(\mathcal{Y}_1, \mathcal{Z}) \ldots) \) be an \( n \)-th order \( \mathbb{K} \)-valued tensor. Define the outer product of \( A \) and \( B \) to be an \( (n+m) \)-th order \( \mathcal{Z} \)-valued tensor \( C := A \otimes B \in B(\mathcal{Y}_n, \ldots, B(\mathcal{Y}_1, \mathcal{Z}) \ldots) \) such that \( C(y_n) \ldots (y_1)(x_m) \ldots (x_1) = B(y_n) \ldots (y_1)A(x_m) \ldots (x_1) \in \mathcal{Z}, \forall x_i \in \mathcal{X}_i, i = 1, \ldots, m, \forall y_j \in \mathcal{Y}_j, j = 1, \ldots, n. \) Similarly, we may define the outer product of \( B \) and \( A \) to be an \( (n+m) \)-th order \( \mathcal{Z} \)-valued tensor \( \bar{C} := B \otimes A \in B(\mathcal{X}_m, \ldots, B(\mathcal{X}_I, B(\mathcal{Y}_n, \ldots, B(\mathcal{Y}_1, \mathcal{Z}) \ldots)) \ldots) \) such that
\textbf{Theorem 9.80 (Taylor Series)}

Let \( A(x) \) be analytic at \( x_0 \), then \( A(x) \) can be written as a Taylor series at \( x_0 \):

\[ A(x) = A(x_0) + \sum_{n=1}^{\infty} \frac{1}{n!} (x - x_0)^n A^{(n)}(x_0) \]

\textbf{Proof} These are straightforward, and are therefore omitted. \hfill \Box

### 9.9 Analytic Functions

\textbf{Definition 9.79} Let \( X \) and \( Y \) be normed linear spaces over \( \mathbb{K} \), \( f : X \to Y \), and \( x_0 \in D^o \). \( f \) is said to be analytic at \( x_0 \) if \( f \) is \( C_\infty \) at \( x_0 \), and \( \exists \delta < 0, \in \mathbb{R} \), such that \( \| f^{(n)}(x) \| \leq cn!M^n, \forall n \in \mathbb{Z}_+ \), \( x \in B_X(x_0, \delta) \). In this case, \( \forall x \in D_1 \), \( f \) is analytic on \( D \). When \( f \) is analytic on \( D \), then we say that \( f \) is analytic or an analytic function.

\textbf{Theorem 9.80 (Taylor Series)} Let \( X \) and \( Y \) be normed linear spaces over \( \mathbb{K} \), \( f : X \to Y \), and \( x_0 \in D^o \). Assume that \( f \) is analytic...
at \( x_0 \) with an analytic radius \( \delta \in (0, \infty) \subset \mathbb{R} \). Then, \( \forall x \in B_X (x_0, \delta) \),

\[
f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0) (x-x_0)^n
\]

**Proof** Since \( f \) is analytic at \( x_0 \in D^c \) with an analytic radius \( \delta \in (0, \infty) \subset \mathbb{R} \), then \( \exists \varepsilon \in (0, \infty) \subset \mathbb{R} \) and \( \exists M \in (0, \infty) \subset \mathbb{R} \) such that \( \| f^{(n)}(x) \| \leq c \varepsilon M^n, \forall n \in \mathbb{Z}, \forall x \in B_X (x_0, \delta) \subseteq D \), and \( M \delta < 1 \). Then, \( \forall x \in B_X (x_0, \delta), \forall n \in \mathbb{Z} \), let

\[
R_n(x, x_0; f) := f(x) - \sum_{i=0}^{n} \frac{1}{i!} f^{(i)}(x_0) (x-x_0)^i
\]

By Taylor Theorem 9.48, \( \exists \varepsilon \in (0, 1) \subset \mathbb{R} \) such that \( \| R_n(x, x_0; f) \| \leq \frac{1}{(n+1)!} \| f^{(n+1)}(t_0 x + (1-t_0)x_0) \| \| x-x_0 \|^{n+1} < c(M\delta)^{n+1} \). Clearly, \( \lim_{n \in \mathbb{N}} \| R_n(x, x_0; f) \| = 0 \). Hence, the desired equality holds. This completes the proof of the theorem. \( \square \)

Next, we establish a result that is needed in the further analysis of analytic functions.

**Proposition 9.81** Let \( X, Y, Z, \) and \( W \) be normed linear spaces over \( \mathbb{K} \), \( D_1 \subseteq X, f : D_1 \rightarrow B (Y, Z), g : D_1 \rightarrow B (Z, W) \), \( n \in \mathbb{Z}^+ \), and \( x_0 \in D_1^c \). Assume that \( f \) and \( g \) are \( C_n \) at \( x_0 \). Define \( h : D_1 \rightarrow B (Y, W) \) by \( h(x) = g(x)f(x), \forall x \in D_1 \). Then, \( h \) is \( C_n \) at \( x_0 \), and \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that

\[
h^{(k)}(x)(\xi_{x_1}) \cdots (\xi_{x_l})(\xi_y) = \sum_{i=0}^{k} \sum_{\Lambda \in \{1, \ldots, k\} : |\Lambda| = l} g^{(i)}(x)(\xi_{x_{\Lambda_2}}) \cdots (\xi_{x_{\Lambda_k}}) (f^{(k-i)}(x)(\xi_{x_{\Lambda_1}}) \cdots (\xi_{x_{\Lambda_{l-1}}}) (\xi_y))
\]

\( \forall x \in B_X (x_0, \delta), \forall \xi_{x_1} \in X, l = 1, \ldots, k, \forall \xi_y \in Y, \forall k \in \mathbb{Z}^+ \) with \( l \leq n \).

**Proof** By Propositions 9.45, 9.44, 3.12, and 9.42, \( h \) is \( C_n \) at \( x_0 \). Since \( f \) and \( g \) are \( C_n \) at \( x_0 \), then \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( f^{(n)}(x) \in B_{B_{\mathbb{K}}}(X, B(Y, Z)) \) and \( g^{(n)}(x) \in B_{B_{\mathbb{K}}}(X, B(Z, W)) \), \( \forall x \in B_X (x_0, \delta) \). We will use mathematical induction on \( k \) to prove the result.

1. \( k = 0 \). We have \( \forall x \in B_X (x_0, \delta), \forall \xi_y \in Y \),

\[
LHS = h^{(0)}(x)(\xi_y) = h(x)(\xi_y) = g(x)f(x)\xi_y = g(x)(f(x)(\xi_y)) = RHS
\]

This case is proved. If \( n = 0 \), the proposition holds. If \( n \in \mathbb{N} \), continue to the following steps.

2. Assume that the result holds for \( k = l \in \mathbb{Z}^+ \) with \( l < n \).
3. Consider the case when \( k = l + 1 \). By the inductive assumption, 
\( \forall x \in \mathcal{B}_X(x_0, \delta), \forall \xi_{x_s} \in X, s = 1, \ldots, l, \forall \xi_y \in \mathcal{Y}, \)
\[
h^{(l)}(x)(\xi_{x_1}) \cdots (\xi_{x_l})(\xi_y) = \\
\sum_{i=0}^l \sum_{\Lambda_1 = \{j_1, \ldots, j_i\} \subset \Lambda} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}})
\]
Then, by Proposition 9.27, we have 
\( \forall x \in \mathcal{B}_X(x_0, \delta), \forall \xi_{x_s} \in X, s = 1, \ldots, l+1, \forall \xi_y \in \mathcal{Y}, \)
\[
h^{(l+1)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_l}}) = \bar{h}^{(1)}(x)(\xi_{x_{j+1}})
\]
where \( \bar{h} : \mathcal{B}_X(x_0, \delta) \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{W}) \) is defined by \( \bar{h}(x) = h^{(l)}(x)(\xi_{x_1}) \cdots (\xi_{x_l}) \), 
\( \forall x \in \mathcal{B}_X(x_0, \delta) \). This implies that, by the Chain Rule (Theorem 9.18) and Proposition 9.42,
\[
h^{(l+1)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_l}})(\xi_y) = \\
\sum_{i=0}^{l+1} \sum_{\Lambda_1 = \{j_1, \ldots, j_i\} \subset \Lambda} g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}})
\]
where the last equality follows since all \( i \)-element subsets of \( \{1, \ldots, l+1\} \) 
is divided into two disjoint classes: one with \( l + 1 \) as an element; and the other without \( l + 1 \) as an element. This completes the induction process.
This completes the proof of the proposition. □

Now, we establish the important result that the composition of analytic functions is an analytic function.

**Theorem 9.82** Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $D_1 \subseteq X$, $D_2 \subseteq Y$, $f : D_1 \to D_2$, $g : D_2 \to Z$, $x_0 \in D_1^0$, and $y_0 := f(x_0) \in D_2^0$. Assume that $f$ and $g$ are analytic at $x_0$ and $y_0$, respectively. Then, $h := g \circ f : D_1 \to Z$ is analytic at $x_0$.

**Proof** Since $g$ is analytic at $y_0$, then $\exists \delta_y \in (0, \infty) \subseteq \mathbb{R}$, $\exists c_g \in (0, \infty) \subseteq \mathbb{R}$, and $\exists M_g \in (0, \infty) \subseteq \mathbb{R}$ such that $\|g^{(n)}(y)\| \leq c_g n! M_g^n$, $\forall n \in \mathbb{Z}^+$, $\forall y \in B_Y(y_0, \delta_y) \subseteq D_2$. Since $f$ is analytic at $x_0$, then $\exists \delta_f \in (0, \infty) \subseteq \mathbb{R}$, $\exists c_f \in (0, \infty) \subseteq \mathbb{R}$, and $\exists M_f \in (0, \infty) \subseteq \mathbb{R}$ such that $\|f^{(n)}(x)\| \leq c_f n! M_f^n$ and $f(x) \in B_Y(y_0, \delta_y)$, $\forall n \in \mathbb{Z}^+$, $\forall x \in B_X(x_0, \delta_f)$. By Proposition 9.45, $h$ is $C_X$ at $\forall x \in B_X(x_0, \delta_f)$.

Define nonnegative constants $K_{n, i} := \sup_{x \in B_X(x_0, \delta_f)} \|D^n(g^{(i)} \circ f)(x)\|$, $\forall i, n \in \mathbb{Z}^+$. Clearly, $\forall i \in \mathbb{Z}^+$, by Chain Rule (Theorem 9.18),

$$K_{n+1, i} = \sup_{x \in B_X(x_0, \delta_f)} \|D^{n+1}(g^{(i)} \circ f)(x)\| = \sup_{x \in B_X(x_0, \delta_f)} \|D^n((g^{(i+1)} \circ f)(f^{(1)}))(x)\|$$

Define $l_{i+1} : B_X(x_0, \delta_f) \to B_{S^{i+1}}(Y, Z)$ by $l_{i+1}(x) := g^{(i+1)}(f(x))$, $\forall x \in B_X(x_0, \delta_f)$, $\forall i \in \mathbb{Z}^+$. Then, $\forall i, n \in \mathbb{Z}^+$,

$$K_{n+1, i} = \sup_{x \in B_X(x_0, \delta_f)} \|D^n((g^{(i+1)} \circ f)(f^{(1)}))(x)\| = \sup_{x \in B_X(x_0, \delta_f)} \|D^n(l_{i+1}f^{(1)})(x)\| = \sup_{x \in B_X(x_0, \delta_f)} \sup_{\xi_x \in X, x = 1, \ldots, n, \|\xi_x\| \leq 1} \|D^n(l_{i+1}f^{(1)})(x)(\xi_x)\| = \sup_{x \in B_X(x_0, \delta_f)} \sup_{\xi_x \in X, x = 1, \ldots, n, \|\xi_x\| \leq 1} \left\| \sum_{k=0}^n \sum_{\Lambda_k = \{j_1, \ldots, j_k\} \text{ is any } k \text{-element subset of } \Lambda = \{1, \ldots, n\}, \Lambda_2 = \{j_{k+1}, \ldots, j_n\} = \Lambda \setminus \Lambda_1} l_{i+1}^{(k)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_k}}) \right\|$$

$$(f^{(n-k+1)}(x)(\xi_{x_{j_{n-k+1}}} \cdots (\xi_{x_{j_n}}))(\xi_x))$$
9.9. ANALYTIC FUNCTIONS

where the fourth equality follows from Proposition 9.81. Hence, \( 0 \leq K_{n+1,i} \leq \sum_{k=0}^{n-1} K_{k,i} c_f (n - k + 1)! M_f^{n-k+1} \), \( \forall i, n \in \mathbb{Z}_+ \).

We will use mathematical induction on \( n \in \mathbb{N} \) to show that \( \forall n \in \mathbb{N}, \forall i \in \mathbb{Z}_+ \),

\[
0 \leq K_{n,i} \leq (n + i)! c_g M_g^{i+1} M_f^{n+1} \sum_{j=0}^{n-1} \binom{n}{j} (n-1) \left( n-i \right) \left( \frac{(n-i)-(j-1)}{i} \right) c_f^{j+1} M_f^j
\]

1° \( n = 1 \). By the recursive formula and the inequality for \( K_{0,i}, \forall i \in \mathbb{Z}_+ \),

\[
K_{1,i} \leq K_{0,i+1} c_f M_f \leq (i+1)! c_f M_f c_g M_g^{i+1} = \text{RHS}
\]

This case is proved.

2° Assume that the result holds for \( n = 1, \ldots, l \in \mathbb{N} \).

3° Consider the case \( n = l + 1 \). \( \forall i \in \mathbb{Z}_+ \), by the recursive formula,

\[
K_{l+1,i} \leq \sum_{k=0}^{l} \binom{l}{k} K_{k,i+1} c_f (l - k + 1)! M_f^{l-k+1}
\]

\[
\leq c_f (l+1)! M_f^{l+1} c_g (i+1)! M_g^{i+1} + \sum_{k=1}^{l} \binom{l}{k} c_f (l - k + 1)! M_f^{l-k+1}
\]

\[
\cdot (k + i + 1)! c_g M_g^{i+2} M_f^k \sum_{j=0}^{k-1} \binom{k}{j} \left( k-1 \right) \left( k+i+1 \right) \left( j-1 \right) c_f^{j+1} M_f^j
\]

\[
= c_f c_g M_g^{i+1} M_f^{i+1} (l + i + 1)! (l+1) \left( \frac{l+i+1}{l} \right) c_f^{i+1} M_f^{i+1}
\]

\[
+ \sum_{k=1}^{l} \sum_{j=0}^{k-1} \frac{k!(l-k+1)!}{k!(j+1)!} \frac{(i+j+2)!}{(k-j-1)!} c_f^{j+2} c_g M_f^{i+1} M_g^{i+2}
\]

\[
= c_f c_g M_g^{i+1} M_f^{i+1} (l + i + 1)! (l+1) \left( \frac{l+i+1}{l} \right) c_f^{i+1} M_f^{i+1}
\]

where the fourth equality follows from Proposition 9.81. Hence, \( 0 \leq K_{n+1,i} \leq \sum_{k=0}^{n-1} K_{k,i} c_f (n - k + 1)! M_f^{n-k+1} \), \( \forall i, n \in \mathbb{Z}_+ \).
This completes the induction process.

Hence, \( \forall n \in \mathbb{N} \),

\[
\sup_{x \in B_{\varepsilon} \left( x_0, \delta \right)} \| h^{(n)}(x) \| = \sup_{x \in B_{\varepsilon} \left( x_0, \delta \right)} \| D^n (g \circ f)(x) \| = K_{n,0}
\]

\[
\leq n! c_g M_g M_f^n \sum_{j=0}^{n-1} \binom{n}{j+1} \binom{n-1}{j} \binom{n}{n-j-1} c_f^{j+1} M_g^j
\]

\[
= n! c_g M_g M_f^n \sum_{j=0}^{n-1} \binom{n-1}{j} c_f^{j+1} M_g^j = n! c_f c_g M_g M_f^n \left( 1 + c_f M_g \right)^{n-1}
\]
\[ \frac{c_\varepsilon c_g M_g}{1 + c_\varepsilon M_g} (M_f(1 + c_f M_g))^n \]

Hence, \( h \) is analytic at \( x_0 \). This completes the proof of the theorem. \( \square \)

**Proposition 9.83** The constant mapping as defined in Proposition 9.33, the identity mapping as delineated in Proposition 9.36, the projection mapping as delineated in Proposition 9.37, the vector addition, the scalar multiplication, the operation of a linear operator on a vector as delineated in Proposition 9.41, the operation of composition of two linear operators as delineated in Proposition 9.42, and the building up of linear operator on the product space using linear operators on individual spaces as delineated in Proposition 9.43 are analytic on the interior of the domain of the mapping.

**Proof** This is immediate from Definition 9.79. \( \square \)

**Proposition 9.84** Let \( X \) and \( Y \) be Banach spaces over \( K \), \( D := \{ L \in B(X,Y) \mid L \text{ is bijective} \} \), and \( f : D \to B(Y,X) \) be defined by \( f(A) = A^{-1}, \forall A \in D \subseteq B(X,Y) \). Then, \( f \) is analytic.

**Proof** By Proposition 9.56, \( D \) is open in \( B(X,Y) \), \( f \) is \( C_\varepsilon \), and \( f^{(1)}(A) (\Delta) = -f(A) \Delta f(A), \forall A \in D, \forall \Delta \in B(X,Y) \). Fix any \( A \in D \).

Since \( D \) is open, then \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( U := B_{B(X,Y)} (A, \delta) \subseteq D \).

Without loss of generality, assume that \( \delta \| f(A) \| < 1 \). By Proposition 9.55, \( \forall T \in U \), we have \( \| f(T) - f(A) \| \leq \frac{\| f(A) \| \| T - A \|}{1 - \| f(A) \| \| \Delta \|} \leq \| f(A) \|^2 \delta (1 - \| f(A) \| \| \Delta \|)^{-1} \). This implies that \( \| f(T) \| \leq \| f(A) \| + \| f(A) \|^2 \delta (1 - \| f(A) \| \| \Delta \|)^{-1} = \| f(A) \| (1 - \| f(A) \| \| \Delta \|)^{-1} =: \alpha \in [0, \infty) \subseteq \mathbb{R} \). This leads to

\[ \| f^{(1)}(T) \| = \sup_{\Delta \in B(X,Y), \| \Delta \| \leq 1} \| f^{(1)}(T)(\Delta) \| \leq \| f(T) \| \leq \alpha^2 \]

We will use mathematical induction on \( n \in \mathbb{Z}_+ \) to show that \( \| f^{(n)}(T) \| \leq n! \alpha^{n+1}, \forall n \in \mathbb{Z}_+, \forall T \in U \).

1° Clearly, \( n = 0 \) and \( n = 1 \) cases are proved.

2° Assume that the result holds for \( n = 0, \ldots, k \in \mathbb{N} \).

3° Consider the case when \( n = k + 1 \). \( \forall T \in U, \forall \Delta_1 \in B(X,Y), \)

\[ g^{(k)}_{\Delta_1}(T)(\Delta_{k+1}) \cdots (\Delta_2) = f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_2)(\Delta_1) \]

This implies that

\[ \| f^{(k+1)}(T) \| = \sup_{\Delta_i \in B(X,Y), \| \Delta_i \| \leq 1, i = 1, \ldots, k+1} \| f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_1) \| \]
By Proposition 9.81, we have the following equality,

\[
g^{(k)}_{\Delta}(T)(\Delta_{k+1}) \cdots (\Delta_2) = \sum_{i=0}^{k} \binom{k}{i} \sum_{\Lambda:=\{\Lambda_1\mid \Lambda_1:\{j_1,\ldots,j_k\}\text{ is any }\nu\text{-element subset of }\Lambda:=\{2,\ldots,k+1\}\}:=A_\Lambda \backslash A_{\Lambda_1}} \| f^{(i)}(T)(\Delta_{j_1}) \cdots (\Delta_{j_k}) \|
\]

Thus,

\[
\| f^{(k+1)}(T) \| \leq \sum_{i=0}^{k} \binom{k}{i} \| f^{(i)}(T) \| \| f^{(k-i)}(T) \|
\]

\[
= \sum_{i=0}^{k} \binom{k}{i} i! \alpha^{i+1} (k-i)! \alpha^{k+1-i} = \sum_{i=0}^{k} k! \alpha^{k+2} = (k+1)! \alpha^{k+2}
\]

where the last inequality follows from the inductive assumption. This completes the induction process.

Hence, \( f \) is analytic at \( T \). By the arbitrariness of \( T \), \( f \) is an analytic function. This completes the proof of the proposition. \( \square \)

Finally, we establish the analytic versions of the Inverse Function Theorem and the Implicit Function Theorem.

**Theorem 9.85 (Inverse Function Theorem)** Let \( X \) and \( Y \) be Banach spaces over \( K \), \( D \subseteq X \), \( F: D \rightarrow Y \) be analytic at \( x_0 \in D^n \), and \( y_0 = F(x_0) \in Y \). Assume that \( F^{(1)}(x_0) \in B(X,Y) \) is bijective. Then, \( \exists \delta \in (0,\infty) \subseteq K \), \( \exists U \subseteq D \) with \( x_0 \in U \subseteq \mathcal{O}_X \) such that

(i) \( F|_U: U \rightarrow B_Y(y_0,\delta) =: V \subseteq Y \) is bijective;

(ii) the inverse mapping \( F_i: V \rightarrow U \) is \( C_\infty \) and \( F_i^{(1)}: V \rightarrow B(Y,X) \) is given by \( F_i^{(1)}(y) = (F^{(1)}(F_i(y)))^{-1}, \forall y \in V \);

(iii) the inverse mapping \( F_i \) is analytic at \( y_0 \in V \).

**Proof** The result (i) and (ii) are implied by Inverse Function Theorem 9.57. All we need to show is (iii). Since \( F \) is analytic at \( x_0 \), then \( \exists \delta_1 \in (0,\infty) \subseteq K \), \( \exists \in [0,\infty) \subseteq K \), and \( \exists M \in (0,\infty) \subseteq K \) such that \( \bar{U} := B_X(x_0,\delta_1) \subseteq U \), \( \| F^{(n)}(x) \| \leq cn!M^n, \forall n \in \mathbb{Z}_+, \forall x \in \bar{U} \). Then, \( \exists \delta_2 \in (0,\delta] \subseteq K \) such that \( F_i(B_Y(y_0,\delta_2)) \subseteq \bar{U} \).
Define \( h : \bar{U} \rightarrow B(y, X) \) by \( h(x) = (F^{(1)}(x))^{-1}, \forall x \in \bar{U} \). Since 
\( F|_{\bar{U}} \) is an analytic function, then \( F^{(1)}|_{\bar{U}} \) is an analytic function. By

Proposition 9.84 and Theorem 9.82, \( h \) is an analytic function. Then, 
\( \exists \delta_1 \in (0, \delta] \subset \mathbb{R}, \exists c_h \in [0, \infty) \subset \mathbb{R} \) and \( \exists M_h \in (0, \infty) \subset \mathbb{R} \) such that 
\( \| h^{(n)}(x) \| \leq c_h M_h^n, \forall n \in \mathbb{Z}_+, \forall x \in B(x_0, \delta_1) \). By choosing \( \delta_1 \)

sufficiently small, without loss of generality, we may take \( \delta_1 = \delta \).

Note that \( F^{(1)}_i \circ F(x) = h(x), \forall x \in \bar{U} \). Then, \( h^{(1)}(x) = D(F^{(1)}_i \circ F)(x) = 
(F^{(2)}_i \circ F(x))F^{(1)}(x) \) and \( F^{(2)}_i \circ F(x) = (h^{(1)}(x))h(x), \forall x \in \bar{U} \). Define 
\( h_k : \bar{U} \rightarrow B_{S_k}(y, X) \) by \( h_k(x) = F^{(k)}_i \circ F(x), \forall x \in \bar{U}, \forall 2 \leq k \in \mathbb{N} \). This

leads to the recursive relationship

\[
(Dh_k(x))h(x) = F^{(k+1)}_i \circ F(x) = h_{k+1}(x), \; \forall x \in \bar{U}, \forall 2 \leq k \in \mathbb{N}
\]

with the initial condition \( h_2(x) = h^{(1)}(x)h(x), \forall x \in \bar{U} \). Define the nonnegative

constants \( K_{k,j} := \sup_{x \in \bar{U}} \| D^j h_k(x) \|, \forall j \in \mathbb{Z}_+, \forall 2 \leq k \in \mathbb{N} \). We will

use mathematical induction on \( 2 \leq k \in \mathbb{N} \) to show that

\[
0 \leq K_{k,j} \leq \frac{(j + 2k - 2)!}{(k - 1)!k^{2j}c_h^jM_h^{j+k-1}}, \; \forall j \in \mathbb{Z}_+, \forall 2 \leq k \in \mathbb{N}
\]

1° \( k = 2, \forall j \in \mathbb{Z}_+, \) by the initial condition of the recursion,

\[
K_{2,j} = \sup_{x \in \bar{U}} \| D^j h_2(x) \| = \sup_{x \in \bar{U}} \| D^j (h^{(1)} h)(x) \|
\]

\[
= \sup_{x \in \bar{U}} \sup_{\xi_x \in X, \|\xi_x\| \leq 1} \| D^j (h^{(1)} h)(x)(\xi_{x_1}) \cdots (\xi_{x_j})(\xi_y) \|
\]

\[
= \sup_{x \in \bar{U}} \sup_{\xi_x \in X, \|\xi_x\| \leq 1} \left\| \sum_{\substack{L = (\xi_{x_1}, \ldots, \xi_{x_j}) \in \mathbb{N}^j \setminus \{0\} \setminus \{(0, \ldots, 0)\} \text{ is any } \ell \text{-element subset of } \\ \Lambda_1 = \{1, \ldots, j\}, \Lambda_2 = \{1, \ldots, j\} \cup \Lambda_1 \}} \\
\right. \\
\left. (h^{(j-l)}(x)(\xi_{x_{j+1}}) \cdots (\xi_{x_j})(\xi_y)) \right\|
\]

\[
\leq \sup_{x \in \bar{U}} \sum_{l=0}^{j} \sum_{\substack{L = (\xi_{x_1}, \ldots, \xi_{x_j}) \in \mathbb{N}^j \setminus \{0\} \setminus \{(0, \ldots, 0)\} \text{ is any } \ell \text{-element subset of } \\ \Lambda_1 = \{1, \ldots, j\}, \Lambda_2 = \{1, \ldots, j\} \cup \Lambda_1 \}} \\
\left\| h^{(l+1)}(x) \right\| \left\| h^{(j-l)}(x) \right\|
\]

\[
= \sup_{x \in \bar{U}} \sum_{l=0}^{j} \left( \frac{j!}{l!(j-l)!} \right) \left\| h^{(l+1)}(x) \right\| \left\| h^{(j-l)}(x) \right\|
\]

\[
\leq \sum_{l=0}^{j} \frac{j!}{l!(j-l)!} c_h (l + 1)!M_h^{l+1} c_h (j - l)!M_h^{j-l}
\]
where the fourth equality follows from Proposition 9.81. This case is proved.

2° Assume that the result holds for $k = n \geq 2$ with $n \in \mathbb{N}$.

3° Consider the case when $k = n + 1$. \forall j \in \mathbb{Z}_+$, by the recursive relationship,

\[
K_{n+1,j} = \sup_{x \in U} \| D^j h_{n+1}(x) \| = \sup_{x \in U} \| D^j (h^{(1)}_n) h(x) \| = \sup_{x \in U} \sup_{\{\xi \} \subseteq \mathbb{R}^d, \|\xi\| \leq 1} \left\| D^j (h^{(1)}_n) h(x) (\xi_x) \cdots (\xi_x) (\xi_y) \right\|
\]

\[
= \sup_{x \in U} \sup_{\{\xi \} \subseteq \mathbb{R}^d, \|\xi\| \leq 1} \left\| \sum_{l=0}^{j} \sum_{\Lambda = \{x_1, \ldots, x_l\} \text{ is any } l\text{-element subset of } \Lambda = \{1, \ldots, j\}; A_2 = (x_{j+1}, \ldots, x_{j+1}) = \Lambda \setminus \Lambda_1} \right\| \left\| h^{(l+1)}_n (x) (\xi_{x_1}) \cdots (\xi_{x_1}) \right\|
\]

\[
\leq \sup_{x \in U} \sum_{l=0}^{j} \sum_{\Lambda_1 = \{x_1, \ldots, x_l\} \text{ is any } l\text{-element subset of } \Lambda = \{1, \ldots, j\}; A_2 = (x_{j+1}, \ldots, x_{j+1}) = \Lambda \setminus \Lambda_1} \right\| h^{(l+1)}_n (x) \right\| \left\| h^{(j-l)}(x) \right\|
\]

\[
= \sup_{x \in U} \sum_{l=0}^{j} \binom{j}{l} \left\| h^{(l+1)}_n (x) \right\| \left\| h^{(j-l)}(x) \right\|
\]

\[
\leq \sup_{x \in U} \sum_{l=0}^{j} \binom{j}{l} K_{n,l+1} \left\| h^{(j-l)}(x) \right\|
\]

\[
= \sum_{l=0}^{j} \binom{j}{l} \frac{l! j!}{(l+2n-1)!} \frac{(j+2n-1)!}{(n-1)!2^{n-1}} \epsilon^2 c_n^2 c_h (j-l)! M_n^{j-l}
\]

\[
= \epsilon_n^2 M_n^{j+n} \frac{j!}{(n-1)!2^{n-1}} \sum_{l=0}^{j} \prod_{s=1}^{2n} (l+s)
\]

\[
= \epsilon_n^2 M_n^{j+n} \frac{j!}{(n-1)!2^{n-1}} \frac{1}{2n} \frac{(j+2n)!}{j!} = \frac{(j+2n)!}{n!2^n} \epsilon_n^2 M_n^{j+n} = \text{RHS}
\]

where the fourth equality follows from Proposition 9.81; the second inequality follows the inductive assumption. This completes the induction process.
9.9. ANALYTIC FUNCTIONS

Based on this inequality, we have the following bounds for the high order derivatives of \( F_1 \) around \( y_0 \), \( \forall y \in B_y(y_0, \delta_2), \forall 2 \leq n \in \mathbb{N}, \)

\[
\left\| F_1^{(n)}(y) \right\| \leq \sup_{x \in U} \left\| F_1^{(n)} \circ F(x) \right\| = \sup_{x \in U} \left\| h_n(x) \right\| \leq K_{n,0} \\
\leq \frac{(2n-2)!}{(n-1)!2^{n-1}c_h^nM^n} \leq \frac{(n-1)!2^{n-1}c_h^nM^n}{2M^n} = \frac{1}{2M^n}n!(2c_hM)^n
\]

Hence, \( F_1 \) is analytic at \( y_0 \). This completes the proof of the theorem. \( \square \)

Theorem 9.86 (Implicit Function Theorem) Let \( X, Y, \) and \( Z \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq X \times Y \), \( F : D \to Z \), and \( (x_0, y_0) \in D^\circ \). Assume that \( F \) is analytic at \( (x_0, y_0) \), \( F(x_0, y_0) = \vartheta_Z \), and \( \frac{\partial F}{\partial y}(x_0, y_0) \in B(Y, Z) \) is bijective. Then, the following statements hold.

(i) There exists \( r_0, r_1 \in (0, \infty) \subseteq \mathbb{R} \) such that \( U \times V := B_X(x_0, r_0) \times B_Y(y_0, r_1) \subseteq D \) and \( \forall x \in U \), \( \exists y \in V \) satisfying \( F(x, y) = \vartheta_Z \). This defines a function \( \phi : U \to V \) by \( \phi(x) = y \), \( \forall x \in U \). Then, \( \phi \) is \( C_\infty \).

(ii) \( \phi^{(1)}(x) = -\left( \frac{\partial F}{\partial y}(x, \phi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \phi(x)) \), \( \forall x \in U \).

(iii) \( \phi \) is analytic at \( x_0 \).

Proof The results (i) and (ii) follows immediately from Implicit Function Theorem 9.59. All we need to show is (iii). Since \( F \) is analytic at \( (x_0, y_0) \), we may choose \( r_0 \) and \( r_1 \) to be sufficiently small such that \( \exists \epsilon \in (0, \infty) \subseteq \mathbb{R} \) and \( \exists M \in (0, \infty) \subseteq \mathbb{R} \), \( \left\| F^{(n)}(x, y) \right\| \leq cn!M^n \), \( \forall n \in \mathbb{Z}_+ \), \( \forall (x, y) \in U \times V \).

Define mapping \( \tilde{F} : D \to X \times Z \) by \( \tilde{F}(x, y) = (x, F(x, y)) \), \( \forall (x, y) \in U \times V \). By Proposition 9.83 and Theorem 9.82, \( \tilde{F} \) is analytic at \( (x_0, y_0) \in D^\circ \). By Propositions 9.44, 9.13, and 9.24,

\[
\tilde{F}^{(1)}(x_0, y_0) = \begin{bmatrix} \text{id}_X & \vartheta_B(y, \cdot) \\
\frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \end{bmatrix}
\]

Clearly, \( \tilde{F}^{(1)}(x_0, y_0) \) is bijective. By Proposition 4.31, \( X \times Y \) and \( X \times Z \) are Banach spaces over \( \mathbb{K} \). By the analytic version of the Inverse Function Theorem 9.85, \( \exists U \subseteq D \) with \( (x_0, y_0) \in U \in \mathcal{O}_{X \times Y} \), \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \tilde{F}|_U : U \to B_{X \times Z}((x_0, \vartheta_Z), \delta) \) is bijective, whose inverse function \( \tilde{F}_1 : B_{X \times Z}((x_0, \vartheta_Z), \delta) \to \tilde{U} \) is analytic at \( (x_0, \vartheta_Z) \). By taking \( r_0 \) and \( r_1 \) sufficiently small, we may assume that \( U \times V \subseteq \tilde{U} \). Then, \( \exists \delta_1 \in (0, \min\{\delta, r_0\}) \subseteq \mathbb{R} \) such that \( \tilde{F}|_W : \tilde{B}_{X \times Z}((x_0, \vartheta_Z), \delta_1) \) is bijective. By the uniqueness of \( y \in V \) that solves \( F(x, y) = \vartheta_Z \), we have \( \phi(x) = \pi_y \circ \tilde{F}_1(x, \vartheta_Z), \forall x \in \tilde{B}_X(x_0, \delta_1) \). By Theorem 9.82 and Proposition 9.83, \( \phi \) is analytic at \( x_0 \). This completes the proof of the theorem. \( \square \)
**Definition 9.87** Let \( X \) be a Banach space over \( K \), and \( A \in B(X,X) \). We will define \( \exp(A) \in B(X,X) \) by \( \exp(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!} \), where \( A^0 := \text{id}_X \) by notation.

This function is well-defined since \( B(X,X) \) is a Banach space over \( K \) by Proposition 7.66, and the series is absolutely summable since \( \|A\|_{B(X,X)} < \infty \). Thus, the series converges in \( B(X,X) \) by Proposition 7.27. The following proposition lists a number of properties for the exponential function.

**Proposition 9.88** Let \( X \) be a Banach space over \( K \), and \( A \in B(X,X) \). \( \forall t, \tau \in K \), we have

(i) \( h : K \to B(X,X) \) defined by \( h(t) = \exp(At) \), \( \forall t \in K \), is an analytic function. \( \forall t \in K \), \( h \) admits arbitrarily large analytic radius at \( t \).

(ii) \( \exp(A(t + \tau)) = \exp(At) \exp(A\tau) \) and \( \exp(\partial B(X,X)) = \text{id}_X \);

(iii) \( \frac{d}{dt} \exp(At) = A \exp(At) = \text{id}_X \);

(iv) if \( B \in B(X,X) \) and \( AB = BA \), then \( \exp(At)B = B \exp(At) \), and \( \exp(A) \exp(B) = \exp(B) \exp(A) = \exp(A + B) \).

(v) \( (\exp(A))^{-1} = \exp(-A) \) and \( (\exp(A))' = \exp(A) \).

**Proof** (iii) \( \exp(At) = \sum_{n=0}^{\infty} \frac{1}{n!} A^n t^n \). By Propositions 9.16 and 9.68, it can be differentiated term by term and \( \frac{d}{dt} \exp(At) = \frac{d}{dt} \sum_{n=0}^{\infty} \frac{1}{n!} A^n t^n = \sum_{n=1}^{\infty} \frac{n}{n!} A^n t^{n-1} = A \exp(At) = \exp(At)A \).

(iii) By repeated application of (iii), we have \( h^{(n)}(t) = \frac{d^n}{dt^n} \exp(At) = A^n \exp(At) \), \( \forall n \in \mathbb{N} \). Then, \( h \) is an analytic function by Definition 9.79. Furthermore, \( \forall t \in K \), \( h \) admits arbitrarily large analytic radius at \( t \).

(iv) \( \exp(At)B = \left( \sum_{n=0}^{\infty} \frac{1}{n!} A^n t^n \right)B = \sum_{n=0}^{\infty} \frac{1}{n!} A^n B t^n = \sum_{n=0}^{\infty} \frac{1}{n!} B A^n t^n \). \( t^n = B \exp(At) \), where the third equality follows from \( AB = BA \). Furthermore, \( \exp(A) \exp(B) = \exp(A) \left( \sum_{n=0}^{\infty} \frac{1}{n!} B^n \right) = \sum_{n=0}^{\infty} \frac{1}{n!} \exp(A) B^n = \sum_{n=1}^{\infty} \frac{n}{n!} B^n \exp(A) = \exp(B) \exp(A) \), where the third equality follows from \( \exp(A)B = B \exp(A) \). \( \exp(A + B) = \sum_{n=0}^{\infty} \frac{1}{n!} (A + B)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \sum_{i=0}^{n} \frac{1}{(n-i)!} A^i B^{n-i} \right) = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \sum_{i=0}^{n} \frac{1}{(n-i)!} A^i B^{n-i} \right) = \exp(A) \cdot \exp(B) \), where the second equality follows from \( AB = BA \); and the third through sixth equality follows from the absolute summability of the series involved.

(ii) This directly follows from (iv) and Definition 9.87.

(v) The inverse relationship follows from (ii). \( (\exp(A))' = \left( \sum_{n=0}^{\infty} \frac{1}{n!} A^n \right)' = \sum_{n=0}^{\infty} \frac{1}{n!} (A^n)' = \sum_{n=0}^{\infty} \frac{1}{n!} (A')^n = \exp(A') \), where the second and third equalities follow from Proposition 7.110. This completes the proof of the proposition. \( \square \)
Chapter 10

Local Theory of Optimization

In this chapter, we will develop a number of tools for optimization of sufficiently many times differentiable functions. As in Chapter 8, we will be mainly concerned with real spaces, rather than complex ones.

10.1 Basic Notion

Definition 10.1 Let $\mathcal{X} := (\mathcal{X}, \mathcal{O})$ be a topological space, $f : \mathcal{X} \to \mathbb{R}$, and $x_0 \in \mathcal{X}$. $x_0$ is said to be a point of minimum for $f$ if $f(x_0) \leq f(x)$, $\forall x \in \mathcal{X}$. It is said to be the point of strict minimum for $f$ if $f(x_0) < f(x)$, $\forall x \in \mathcal{X} \setminus \{x_0\}$. It is said to be a point of relative minimum for $f$ if $\exists O \in \mathcal{O}$ with $x_0 \in O$ such that $f(x_0) \leq f(x)$, $\forall x \in O$. It is said to be a point of relative strict minimum for $f$ if $\exists O \in \mathcal{O}$ with $x_0 \in O$ such that $f(x_0) < f(x)$, $\forall x \in O \setminus \{x_0\}$. Similar definitions for points of maxima. Moreover, $x_0$ is said to be a point of relative extremum if it is a point of relative minimum or relative maximum. It is said to be a point of relative strict extremum if it is a point of relative strict minimum or relative strict maximum.

Proposition 10.2 Let $\mathcal{X}$ be a real normed linear space, $D \subseteq \mathcal{X}$, $x_0 \in D$, $f : D \to \mathbb{R}$, $u \in A_D (x_0)$. Assume that the directional derivative of $f$ at $x_0$ along $u$ exists and $x_0$ is a point of relative minimum for $f$. Then, $Df(x_0; u) \geq 0$.

Proof This is immediate from Definition 9.4. \(\square\)

Definition 10.3 Let $\mathcal{X}$ be a real normed linear space and $A \in \mathcal{B}_{\mathcal{S}_2} (\mathcal{X}, \mathbb{R})$. $A$ is said to be positive definite if $\exists m \in (0, \infty) \subseteq \mathbb{R}$ such that $A(x)(x) = \langle \langle Ax, x \rangle \rangle \geq m \|x\|^2$, $\forall x \in \mathcal{X}$. Let the set of all such positive definite
operators be denoted by \( S_+ X \). \( A \) is said to be positive semi-definite if \( A(x)(x) \geq 0, \forall x \in X \). Let the set of all such positive semi-definite operators be denoted by \( S_{\text{psd}} X \). \( A \) is said to be negative definite if \( \exists m \in (0, \infty) \subseteq \mathbb{R} \) such that \( A(x)(x) \leq -m \| x \|^2, \forall x \in X \). Let the set of all such negative definite operators be denoted by \( S_- X \). \( A \) is said to be negative semi-definite if \( A(x)(x) \leq 0, \forall x \in X \). Let the set of all such negative semi-definite operators be denoted by \( S_{\text{npsd}} X \). We will denote \( B_{S_2}(X, \mathbb{R}) \) by \( S_X \).

**Proposition 10.4** Let \( X \) be a real normed linear space. Then,
(i) \( S_+ X = -S_- X \) and \( S_{\text{npsd}} X = -S_{\text{psd}} X \);
(ii) \( S_+ X \) and \( S_- X \) are open sets in \( B_{S_2}(X, \mathbb{R}) = S_X \);
(iii) \( S_{\text{psd}} X \) and \( S_{\text{npsd}} X \) are closed convex cones in \( S_X \);
(iv) \( S_+ X \subseteq S_{\text{psd}}^o X \) and \( S_- X \subseteq S_{\text{npsd}}^o X \).

**Proof**  (i) This is clear. (ii) Fix \( A \in S_+ X \). Then, \( \exists m \in (0, \infty) \subseteq \mathbb{R} \) such that \( A(x)(x) \geq m \| x \|^2, \forall B \in B_{S_2}(X, \mathbb{R}) \) with \( \| B - A \| < m/2, \forall x \in X \), we have \( B(x)(x) = A(x)(x) + (B - A)(x)(x) \geq m \| x \|^2 - \| B - A \| \| x \|^2 \geq m \| x \|^2 / 2 \), where the first inequality follows from Proposition 7.64. This implies that \( B \in S_+ X \). Then, \( A \in S_+ X \). Hence, \( S_+ X \) is open in \( S_X \).

(iii) Clearly, \( \partial B_{S_2}(X, \mathbb{R}) \subseteq S_{\text{psd}} X \). \( \forall A \in S_{\text{psd}} X, \forall \alpha \in (0, \infty) \subseteq \mathbb{R}, \alpha A \in S_{\text{psd}} X \). Hence, \( S_{\text{psd}} X \) is a cone. \( \forall A, B \in S_{\text{psd}} X, \forall x \in X, (A + B)(x)(x) = A(x)(x) + B(x)(x) \geq 0 \). Hence, \( A + B \in S_{\text{psd}} X \). Then, \( S_{\text{psd}} X \) is a convex cone.

Let \( M := S_X \setminus S_{\text{psd}} X \). \( \forall A \in M, \exists x_0 \in X \) such that \( A(x_0)(x_0) < 0 \). Then, there exists \( m \in (0, \infty) \subseteq \mathbb{R} \) such that \( A(x_0)(x_0) < -m \| x_0 \|^2 < 0 \). \( \forall B \in B_{S_2}(X, \mathbb{R}) \) with \( \| B - A \| < m/2, \) we have \( B(x_0)(x_0) = A(x_0)(x_0) + (B - A)(x)(x) \geq -m \| x_0 \|^2 + \| B - A \| \| x_0 \|^2 \leq -m \| x_0 \|^2 / 2 < 0 \), where the first inequality follows from Proposition 7.64. This implies that \( B \in M \) and \( A \in M^c \). Then, \( M = M^c \) is open. \( S_{\text{psd}} X \) is closed.

\( S_{\text{npsd}} X = -S_{\text{psd}} X \) is clearly also a closed convex cone.

(iv) Clearly, \( S_+ X \subseteq S_{\text{psd}} X \). Then, \( S_+ X \subseteq S_{\text{psd}}^o X \). This further implies that \( S_- X \subseteq S_{\text{npsd}}^o X \). This completes the proof of the proposition.

**Definition 10.5** Let \( (\mathcal{X}, K) \) be a vector space and \( K \subseteq \mathcal{X} \) be convex. \( M \subseteq K \) is said to be an extreme subset of \( K \) if \( M \neq \emptyset \) and \( \forall x_1, x_2 \in K, \forall \alpha \in (0, 1) \subseteq \mathbb{R}, \alpha x_1 + (1 - \alpha) x_2 \in M \) implies that \( x_1, x_2 \in M \). If a singleton set \( \{ x_0 \} \subseteq K \) is an extreme subset, then \( x_0 \) is called an extreme point of \( K \).

**Proposition 10.6** Let \( X \) be a real normed linear space, \( K \subseteq X \) be a nonempty convex set, and \( H := \{ x \in X \mid \langle x, x_0 \rangle = c \} \) be a supporting hyperplane of \( K \), where \( x_0 \in X^* \) with \( x_0 \neq 0 \) and \( c \in \mathbb{R} \). Then, any extreme subset of \( K_1 := K \cap H \) is also an extreme subset of \( K \).
10.1. BASIC NOTION

Proof: Without loss of generality, assume that \( \inf_{k \in K} \langle \langle x_0, x \rangle \rangle = c \). Let \( M \subseteq K_1 \) be an extreme subset of \( K_1 \). Then, \( M \neq \emptyset \). \( \forall x_1, x_2 \in K, \forall \alpha \in (0, 1) \subseteq \mathbb{R} \), let \( \bar{x} := \alpha x_1 + (1 - \alpha) x_2 \in M \subseteq H \). Then, \( \langle \langle x_0, x_1 \rangle \rangle \geq c \) and \( \langle \langle x_0, x_2 \rangle \rangle \geq c \) and \( \langle \langle x_0, \bar{x} \rangle \rangle = c \). This implies that \( \alpha (\langle \langle x_0, x_1 \rangle \rangle - c) + (1 - \alpha) (\langle \langle x_0, x_2 \rangle \rangle - c) = 0 \) and \( \langle \langle x_0, x_1 \rangle \rangle = c = \langle \langle x_0, x_2 \rangle \rangle \). Hence, \( x_1, x_2 \in K_1 \). Since, \( M \) is an extreme subset of \( K_1 \), then \( x_1, x_2 \in M \). Therefore, \( M \) is an extreme subset of \( K \). This completes the proof of the proposition.

Proposition 10.7 (Krein-Milman) Let \( X \) be a real reflexive Banach space, \( K \subseteq X \) be a nonempty bounded closed convex set, and \( M \subseteq K \) be a weakly compact extreme subset of \( K \). Then, \( M \) contains at least one extreme point of \( K \).

Proof: By Proposition 8.11, \( K \) is compact in \( X_{\text{weak}} \). Let \( M = \{ E \subseteq M \mid E \text{ is a weakly compact extreme subset of } K \} \). Clearly, \( M \subseteq M \neq \emptyset \). Clearly, \( \supseteq \) defines an antisymmetric partial ordering on \( M \), where smaller sets are further down the stream. Next, we will use Zorn’s Lemma to show that \( M \) admits a maximal element.

Let \( E \subseteq M \) be a nonempty totally ordered (by \( \supseteq \) ) subcollection. Let \( E_0 := \bigcap_{E \in E} E \). \( \forall E \in E, E \subseteq M \) is weakly compact extreme subset of \( K \). Then, by Propositions 7.116, 5.5, and 3.61, \( E \) is weakly closed. By Proposition 5.5, \( E_0 \) is weakly compact. \( \forall x_1, x_2 \in K, \forall \alpha \in (0, 1) \subseteq \mathbb{R} \), let \( \alpha x_1 + (1 - \alpha) x_2 \in E_0 \). \( \forall E \in E, \alpha x_1 + (1 - \alpha) x_2 \in E \). Since \( E \) is an extreme subset of \( K \), then \( x_1, x_2 \in E \). By the arbitrariness of \( E \), we have \( x_1, x_2 \in E_0 \). Hence, \( E_0 \) is an extreme subset of \( K \) if \( E_0 \neq \emptyset \). \( \forall E \in E, E \) is nonempty. Since \( E \) is totally ordered by \( \supseteq \), then the intersection of finite number of sets in \( E \) is again in \( E \), and hence nonempty. By Proposition 5.12, \( E_0 \neq \emptyset \). Then, \( E_0 \) is a weakly compact extreme subset of \( K \) and \( E_0 \in M \). Clearly, \( E_0 \) is an upper bound of \( E \) (in terms of \( \supseteq \) ). Then, by Zorn’s Lemma, \( M \) admits a maximal element \( E_M \). Then, \( E_M \subseteq M \) is a weakly compact extreme subset of \( K \) and \( E_M \neq \emptyset \).

We will show that \( E_M \) is a singleton set, which then proves that \( M \) contains an extreme point of \( K \). Suppose that \( \exists x_1, x_2 \in E_M \) with \( x_1 \neq x_2 \). Let \( N := \text{span} \{ x_2 - x_1 \} \) and define a functional \( f : N \rightarrow \mathbb{R} \) by \( f(\alpha (x_2 - x_1)) = \alpha / \sup \alpha \in \mathbb{R} \). Clearly, \( f \) is a linear functional on \( N \), and \( \| f \|_N = 1 / \| x_2 - x_1 \| < \infty \). By Hahn-Banach Theorem 7.83, there exists \( x_0 \in X^* \) with \( \| x_0 \| = 1 / \| x_2 - x_1 \| \) such that \( \langle \langle x_0, \alpha (x_2 - x_1) \rangle \rangle \geq \alpha, \forall \alpha \in \mathbb{R} \). Clearly, \( \langle \langle x_0, x_1 \rangle \rangle \neq \langle \langle x_0, x_2 \rangle \rangle \). Note that \( E_M \) is nonempty and compact in \( X_{\text{weak}} \) and \( x_0 \) is weakly continuous. By Proposition 5.29, \( c := \langle \langle x_0, x_0 \rangle \rangle = \inf_{x \in E_M} \langle \langle x_0, x \rangle \rangle \in \mathbb{R} \) for some \( x \in E_M \). Define \( H := \{ x \in X \mid \langle \langle x_0, x \rangle \rangle = c \} \). Let \( E_m := E_M \cap H \). Then, at least one of \( x_1 \) and \( x_2 \) is not in \( E_m \). Hence, \( E_M \supseteq E_m \). Clearly, \( E_m \) is nonempty. Note that \( H \) is weakly closed. Then, by Proposition 5.5, \( E_m \subseteq M \) is weakly compact. \( \forall \bar{x}_1, \bar{x}_2 \in K, \forall \bar{\alpha} \in (0, 1) \subseteq \mathbb{R} \), let \( \bar{\alpha} \bar{x}_1 + (1 - \bar{\alpha}) \bar{x}_2 \in E_m \subseteq E_M \). Since \( E_M \) is an extreme subset of \( E_M \), then we have \( \bar{x}_1, \bar{x}_2 \in E_M \).
This further implies that $\langle \langle x_0, \bar{x}_1 \rangle \rangle \geq c$ and $\langle \langle x_0, \bar{x}_2 \rangle \rangle \geq c$. Note that $\langle \langle x_0, \alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2 \rangle \rangle = c$. Then, we must have $\langle \langle x_0, \bar{x}_1 \rangle \rangle = c = \langle \langle x_0, \bar{x}_2 \rangle \rangle$ and $\bar{x}_1, \bar{x}_2 \in E_m$. This shows that $E_m$ is a extreme subset of $K$. Case 2: $E_m \notin \mathcal{M}$. This contradicts with the fact that $E_M$ is maximal with respect to $\supseteq$. Therefore, $E_M$ is a singleton set.

This completes the proof of the proposition.

**Proposition 10.8** Let $\mathcal{X}$ be a real reflexive Banach space, $K \subseteq \mathcal{X}$ be a nonempty bounded closed convex set, and $E$ be the set of extreme points of $K$. Then, $K = \overline{co}(E)$.

**Proof** Let $C := \overline{co}(E)$. By Proposition 7.15, $C$ is closed and convex and $C \subseteq K$. We will prove the result by an argument of contradiction. Suppose $K \supset C$. Then, $\exists x_0 \in K \setminus C$. By Proposition 8.10, there exists $x_0 \in \mathcal{X}^*$ such that $\langle \langle x_0, x \rangle \rangle < \inf_{x \in C} \langle \langle x_0, x \rangle \rangle$. By Proposition 8.11, $K$ is compact in the weak topology and $x_0$ is continuous in the weak topology. Then, by Proposition 5.29, $c_0 := \langle \langle x_0, x_1 \rangle \rangle = \inf_{x \in K} \langle \langle x_0, x \rangle \rangle \in \mathbb{R}$ for some $x_1 \in K$. Clearly $c_0 \leq \langle \langle x_0, x_0 \rangle \rangle < \inf_{x \in C} \langle \langle x_0, x \rangle \rangle$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0 = \partial x^*$; Case 2: $x_0 \neq \partial x^*$. Case 1: $x_0 = \partial x^*$. Then, $c_0 = 0$ and $C = \emptyset$. By Proposition 10.7, $E \neq \emptyset$. Then, $C \neq \emptyset$. This is a contradiction. Case 2: $x_0 \neq \partial x^*$. Let $H := \{x \in \mathcal{X} \mid \langle \langle x_0, x \rangle \rangle = c_0\}$. $H$ is a supporting hyperplane of $K$ and $H \cap C = \emptyset$. Let $C_m := K \cap H \neq \emptyset$. Clearly, $C_m$ is bounded closed and convex. By Proposition 10.7, there is an extreme point $x_m \in C_m$ of $C_m$. Then, $\{x_m\} \subseteq C_m$ is an extreme subset of $C_m$. By Proposition 10.6, $\{x_m\}$ is an extreme subset of $K$. Then, $x_m$ is an extreme point of $K$ and $x_m \in E$. This leads to the contradiction $x_m \in C_m \subseteq H$, $x_m \in E \subseteq C$ and $C \cap H = \emptyset$. Thus, in both cases, we have arrived at a contradiction. Then, the hypothesis must be false. Hence, $K = \overline{co}(E)$.

This completes the proof of the proposition.

**Proposition 10.9** Let $\mathcal{X}$ be a real vector space, $\Omega \subseteq \mathcal{X}$ be a convex set, and $f_1 : \Omega \to \mathbb{R}$ and $f_2 : \Omega \to \mathbb{R}$ be convex functionals. Then, the following statements hold.

(i) $\forall n \in \mathbb{N}, \forall x_1, \ldots, x_n \in \Omega, \forall \alpha_1, \ldots, \alpha_n \in [0, 1] \subseteq \mathbb{R}$ with $\sum_{i=1}^{n} \alpha_i = 1$, we have $f_1(\sum_{i=1}^{n} \alpha_i x_i) \leq \sum_{i=1}^{n} \alpha_i f_1(x_i)$. If, in addition, $f_1$ is strictly convex, $x_1, \ldots, x_n$ are distinct, and $\alpha_1, \ldots, \alpha_n \in (0, 1) \subseteq \mathbb{R}$, then $f_1(\sum_{i=1}^{n} \alpha_i x_i) < \sum_{i=1}^{n} \alpha_i f_1(x_i)$.

(ii) $\forall \alpha_1, \alpha_2 \in [0, \infty) \subseteq \mathbb{R}$, $\alpha_1 f_1 + \alpha_2 f_2$ is convex. If, in addition, $f_1$ is strictly convex and $\alpha_1 \in (0, \infty) \subseteq \mathbb{R}$, then $\alpha_1 f_1 + \alpha_2 f_2$ is strictly convex.

(iii) $\forall c \in \mathbb{R}, \{x \in \Omega \mid f_1(x) \leq c\}$ is convex.

(iv) Let $\mathcal{Y}$ be a real vector space, $A : \mathcal{Y} \to \mathcal{X}$ be an affine operator, $D \subseteq \mathcal{Y}$ be convex, and $A(D) \subseteq \Omega$. Then, $f_1 \circ A|_D : D \to \mathbb{R}$ is convex.
Proof  This is straightforward, and is therefore omitted. □

Proposition 10.10 Let $f : [a, b] \to \mathbb{R}$, where $a, b \in \mathbb{R}$ with $a < b$. Then, the following statements hold.

(i) $f$ is convex if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$f(x_2) - f(x_1) \leq \frac{x_2 - x_1}{x_3 - x_2} f(x_3) - f(x_2)$$

if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$f(x_2) - f(x_1) \leq \frac{x_2 - x_1}{x_3 - x_1} f(x_3) - f(x_1)$$

(ii) $f$ is strictly convex if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$f(x_2) - f(x_1) < \frac{x_2 - x_1}{x_3 - x_2} f(x_3) - f(x_2)$$

if, and only if, $\forall x_1, x_2, x_3 \in [a, b] \subset \mathbb{R}$ with $x_1 < x_2 < x_3$, we have

$$f(x_2) - f(x_1) < \frac{x_2 - x_1}{x_3 - x_1} f(x_3) - f(x_1)$$

(iii) If $f$ is convex and $c \in (a, b) \subset \mathbb{R}$, then, the one-sided derivatives

$$\lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}; \quad \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c}$$

exist.

Proof  This is straightforward, and is therefore omitted. □

Proposition 10.11 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$ be convex, and $f : \Omega \to \mathbb{R}$ be differentiable. Then, the following statements hold.

(i) $f$ is convex if, and only if, $\forall x, y \in \Omega$, we have $f(y) \geq f(x) + f^{(1)}(x)(y - x)$.

(ii) $f$ is strictly convex if, and only if, $\forall x, y \in \Omega$ with $x \neq y$, we have $f(y) > f(x) + f^{(1)}(x)(y - x)$.

Proof  (i) “Necessity” Let $f$ be convex. $\forall x, y \in \Omega$, by the convexity of $\Omega$, we have $y - x \in A_\Omega (x)$. $\forall \alpha \in (0, 1] \subset \mathbb{R}$, we have $f(\alpha y + (1-\alpha)x) \leq \alpha f(y) + (1-\alpha)f(x)$, which further implies that $f(y) - f(x) \geq (f(x + \alpha (y - x)) - f(x))/\alpha$. Then, we have

$$f(y) - f(x) = \lim_{\alpha \to 0^+} (f(x + \alpha (y - x)) - f(x))/\alpha = Df(x; y - x)$$

$$= f^{(1)}(x)(y - x)$$
where the first equality follows from Definition 9.4 and the last equality follows from Proposition 9.5.

“Sufficiency” Let \( f(y) \geq f(x) + f^{(1)}(x)(y - x) \), \( \forall x, y \in \Omega \). \( \forall x_1, x_2 \in \Omega \), \( \forall \alpha \in [0, 1] \in \mathbb{R} \), let \( x := \alpha x_1 + (1 - \alpha)x_2 \in \Omega \). Then, \( f(x_2) = f(x + \alpha (x_2 - x_1)) \geq f(x) + \alpha f^{(1)}(x)(x_2 - x_1) \). Note also, \( f(x_1) = f(x + (1 - \alpha)(x_1 - x_2)) \geq f(x) + (1 - \alpha)f^{(1)}(x)(x_1 - x_2) \). Therefore, \( \alpha f(x_1) + (1 - \alpha)f(x_2) \geq f(x) \). Hence, \( f \) is convex.

(ii) “Necessity” Let \( f \) be strictly convex. \( \forall x, y \in \Omega \) with \( x \neq y \), by the convexity of \( \Omega \), we have \( y - x \in A_{\Omega}(x) \). \( \forall \alpha \in (0, 1) \in \mathbb{R} \), we have \( f(\alpha y + (1 - \alpha)x) < \alpha f(y) + (1 - \alpha)f(x) \), which further implies that \( f(y) - f(x) > (f(x + \alpha(y - x)) - f(x))/\alpha \). Define \( A : \mathbb{R} \to \mathcal{X} \) by \( A(\beta) = x + \beta(y - x) \), \( \forall \beta \in \mathbb{R} \). Clearly, \( A \) is an affine operator and \( A(I) \subseteq \Omega \), where \( I := [0, 1] \subseteq \mathbb{R} \). By Proposition 10.9, \( g := f \circ A \) is convex. Then, we have

\[
\begin{align*}
f(y) - f(x) &> 2(f(x + 0.5(y - x)) - f(x)) = \frac{g(0.5) - g(0)}{0.5} \\
&\geq \lim_{\alpha \to 0^+} \frac{g(\alpha) - g(0)}{\alpha} = \lim_{\alpha \to 0^+} \frac{f(x + \alpha(y - x)) - f(x)}{\alpha} \\
&= Df(x; y - x) = f^{(1)}(x)(y - x)
\end{align*}
\]

where the second inequality follows from Proposition 10.10 and the third equality follows from Definition 9.4 and the last equality follows from Proposition 9.5.

“Sufficiency” Let \( f(y) > f(x) + f^{(1)}(x)(y - x) \), \( \forall x, y \in \Omega \) with \( x \neq y \). \( \forall x_1, x_2 \in \Omega \) with \( x_1 \neq x_2 \), \( \forall \alpha \in (0, 1) \in \mathbb{R} \), let \( x := \alpha x_1 + (1 - \alpha)x_2 \in \Omega \). Then, \( f(x_2) = f(x + \alpha(x_2 - x_1)) > f(x) + \alpha f^{(1)}(x)(x_2 - x_1) \). Note also, \( f(x_1) = f(x + (1 - \alpha)(x_1 - x_2)) > f(x) + (1 - \alpha)f^{(1)}(x)(x_1 - x_2) \). Then, \( \alpha f(x_1) + (1 - \alpha)f(x_2) > f(x) \). Hence, \( f \) is strictly convex.

This completes the proof of the proposition. \( \square \)

**Proposition 10.12** Let \( \mathcal{X} \) be a real normed linear space, \( \Omega \subseteq \mathcal{X} \) be convex, and \( f : \Omega \to \mathbb{R} \) be twice differentiable. Then, the following statements hold.

(i) If \( f^{(2)}(x) \) is positive semi-definite, \( \forall x \in \Omega \), then \( f \) is convex.

(ii) If \( f \) is convex and \( C_2 \), then \( f^{(2)}(x) \) is positive semi-definite, \( \forall x \in \overline{\Omega^0} \cap \Omega \).

(iii) If \( f \) is convex and \( f^{(2)}(x) \) is positive definite for all \( x \in \Omega \setminus E \), where \( E \subseteq \Omega \) does not contain any line segment, then \( f \) is strictly convex.

**Proof**

(i) \( \forall x, y \in \Omega \), by Taylor’s Theorem, \( \exists t_0 \in (0, 1) \in \mathbb{R} \) we have

\[
f(y) = f(x) + f^{(1)}(x)(y - x) + \frac{1}{2}f^{(2)}(t_0y + (1 - t_0)x)(y - x)(y - x)
\]

By the assumption, \( f^{(2)}(t_0y + (1 - t_0)x)(y - x)(y - x) \geq 0 \). Therefore, \( f(y) \geq f(x) + f^{(1)}(x)(y - x) \). By the arbitrariness of \( x, y \) and Proposition 10.11, \( f \) is convex.
(ii) We will prove this statement by an argument of contradiction. Suppose \( \exists x_0 \in \mathbb{R}^n \cap \Omega \) such that \( f^{(2)}(x_0) \) is not positive semi-definite. Then, \( \forall h \in X \) such that \( f^{(2)}(x_0)(h)(h) < 0 \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_0 \in \Omega \); Case 2: \( x_0 \notin \Omega \). Case 1: \( x_0 \in \Omega \). By the continuity of \( f^{(2)} \) and \( x_0 \in \Omega \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( x_1 := x_0 + \delta h \in \Omega \) and \( f^{(2)}(x_0 + \alpha \delta h)(h)(h) < 0, \forall \alpha \in (0, 1) \subset \mathbb{R} \). By Taylor’s Theorem, \( \exists t_0 \in (0, 1) \subset \mathbb{R} \) we have

\[
f(x_1) = f(x_0) + f^{(1)}(x_0)(x_1 - x_0) + \frac{\delta^2}{2} f^{(2)}(x_0 + t_0 \delta h)(h)(h)
\]

Then, \( f(x_1) < f(x_0) + f^{(1)}(x_0)(x_1 - x_0) \). This contradicts with the fact that \( f \) is convex and Proposition 10.11.

Case 2: \( x_0 \notin \Omega \). Then, \( x_0 \in (\Omega^c \cap \Omega) \setminus \Omega \). Then, by Proposition 4.13, there exists \( (x_n)_{n=1}^{\infty} \subset \Omega^c \) such that \( \lim_{n \to \infty} x_n = x_0 \). By the continuity of \( f^{(2)} \), we have \( \exists n_0 \in \mathbb{N} \) such that \( f^{(2)}(x_{n_0})(h)(h) < 0 \). Hence, \( f^{(2)}(x_{n_0}) \) is not positive semi-definite and \( x_{n_0} \in \Omega \). By Case 1, there is a contradiction. Hence, the statement must be true.

(iii) We will prove this statement by an argument of contradiction. Suppose \( f \) is not strictly convex. By Proposition 10.11, \( \exists x, y \in \Omega \) with \( x \neq y \) such that \( f(y) = f(x) + f^{(1)}(x)(y - x) \). By the convexity of \( f \) and Proposition 10.11, \( \forall \alpha \in [0, 1] \subset \mathbb{R} \), we have \( f(x + \alpha(y - x)) \geq f(x) + \alpha f^{(1)}(x)(y - x) = f(x) + \alpha(f(y) - f(x)) \). By the convexity of \( f \), we have \( f(x + \alpha(y - x)) = f(x) + \alpha(f(y) - f(x)) \). Define \( g : [0, 1] \to \mathbb{R} \) by \( g(\alpha) = f(\alpha y + (1 - \alpha)x) - \alpha f(y) - (1 - \alpha)f(x), \forall \alpha \in I := [0, 1] \subset \mathbb{R} \). Then, \( g(\alpha) = 0, \forall \alpha \in I \). This implies that \( g^{(2)}(\alpha) = 0, \forall \alpha \in I, \) and \( f^{(2)}(\alpha y + (1 - \alpha)x)(y - x)(y - x) = 0, \forall \alpha \in I \). Hence, \( f^{(2)}(\alpha y + (1 - \alpha)x) \) is not positive definite, \( \forall \alpha \in I \). Hence, \( E \) contains the line segment connecting \( x \) and \( y \). This contradicts with the assumption. Therefore, the statement must be true.

This completes the proof of the proposition. \( \square \)

## 10.2 Unconstrained Optimization

The basic problem to be considered in this section is

\[
\mu_0 := \inf_{x \in \Omega} f(x)
\]

where \( X \) is a real normed linear space, \( \Omega \subseteq X \) is a set, and \( f : \Omega \to \mathbb{R} \) is a functional.

**Proposition 10.13** Let \( X \) be a real normed linear space, \( \Omega \subseteq X \) be convex, \( f : \Omega \to \mathbb{R} \) be a convex functional, and \( \mu_0 := \inf_{x \in \Omega} f(x) \in \mathbb{R} \). Then, the following statements hold.

(i) The set of all points of minimum for \( f \), which is given by \( \{ x \in \Omega \mid f(x) = \mu_0 \} \), is convex.
(ii) Any point of relative minimum for $f$ is a point of minimum for $f$.

(iii) Any point of relative strict minimum for $f$ is the point of strict minimum for $f$.

**Proof**

(i) Note that \( \{ x \in \Omega \mid f(x) = \mu_0 \} = \{ x \in \Omega \mid f(x) \leq \mu_0 \} \), which is convex by Proposition 10.9.

(ii) Fix any $x_0 \in \Omega$ that is a point of relative minimum for $f$. Then, \( \exists \epsilon \in (0, \infty) \subset \mathbb{R} \) such that $f(x) \geq f(x_0)$, \( \forall x \in \Omega \cap B_{\mathcal{X}}(x_0, \epsilon) \). \( \forall y \in \Omega \), \( \exists \alpha \in (0, 1) \subset \mathbb{R} \) such that $x_0 + \alpha (y - x_0) \in B_{\mathcal{X}}(x_0, \epsilon)$. Then, $f(x_0) \leq f(x_0 + \alpha (y - x_0)) = f(\alpha y + (1 - \alpha)x_0) \leq \alpha f(y) + (1 - \alpha)f(x_0)$. This implies that $f(x_0) \leq f(y)$. Hence, $x_0$ is a point of minimum for $f$.

(iii) Fix any $x_0 \in \Omega$ that is a point of relative strict minimum for $f$. Then, \( \exists \epsilon \in (0, \infty) \subset \mathbb{R} \) such that $f(x) > f(x_0)$, \( \forall x \in (\Omega \cap B_{\mathcal{X}}(x_0, \epsilon)) \setminus \{ x_0 \} \). \( \forall y \in \Omega \) with \( y \neq x_0 \), \( \exists \epsilon \in (0, 1) \subset \mathbb{R} \) such that $x_0 + \alpha (y - x_0) \in B_{\mathcal{X}}(x_0, \epsilon)$. Then, $f(x_0) < f(x_0 + \alpha (y - x_0)) = (\alpha y + (1 - \alpha)x_0) \leq \alpha f(y) + (1 - \alpha)f(x_0)$. This implies that $f(x_0) < f(y)$. Hence, $x_0$ is a point of strict minimum for $f$. This completes the proof of the proposition.

**Proposition 10.14** Let $\mathcal{X}$ be a real reflexive Banach space, $\Omega \subseteq \mathcal{X}$ be a nonempty bounded closed convex set, and $f : \Omega \to \mathbb{R}$ be a weakly upper semicontinuous convex functional. Then, there exist an extreme point $x_0 \in \Omega$ such that $x_0$ is a point of maximum for $f$.

**Proof** By Proposition 8.11, $\Omega$ is weakly compact. Let $\mathcal{O}_{\text{weak}}(\mathcal{X})$ be the weak topology on $\mathcal{X}$ and $\mathcal{O}_{\text{weak}, \Omega}$ be the subset topology on $\Omega$ with respect to $\mathcal{O}_{\text{weak}}(\mathcal{X})$. By Proposition 5.30, there exists $x_1 \in \Omega$ that is a point of maximum for $f$. Let $M := \{ x \in \Omega \mid f(x) \geq f(x_1) \}$. Then $x_1 \in M \neq \emptyset$. Note that, by Proposition 2.5, $M = f_{\text{inv}}(\mathbb{R} \setminus I) = \Omega \setminus f_{\text{inv}}(I)$, where $I = (-\infty, f(x_1)) \subset \mathbb{R}$. Since $I$ is open in $\mathbb{R}$ and $f$ is weakly upper semicontinuous, then, $f_{\text{inv}}(I) \in \mathcal{O}_{\text{weak}, \Omega}$. Then, $M$ is closed in $\mathcal{O}_{\text{weak}, \Omega}$. By Proposition 5.5, $M$ is weakly compact. Since $f(x_1) = \max_{x \in \Omega} f(x)$, then $M$ is an extreme subset of $\Omega$. By Proposition 10.7, there exists $x_0 \in M$ that is an extreme point of $\Omega$. This completes the proof of the proposition.

**Proposition 10.15** Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \to \mathbb{R}$ be differentiable at $x_0$. Assume that $x_0$ is a point of relative extremum of $f$, then $f^{(1)}(x_0) = \vartheta_{\mathcal{X}}^\circ$.

**Proof** Without loss of generality, assume that $x_0$ is a point of relative minimum for $f$. The case when $x_0$ is a point of relative maximum for $f$ can be proved similarly. Since $x_0 \in \Omega^\circ$, then $A_{\Omega}(x_0) = \mathcal{X}$. \( \forall u \in \mathcal{X} \), by Propositions 10.2 and 9.5, $Df(x_0; u) = f^{(1)}(x_0)u \geq 0$ and $Df(x_0; -u) = -f^{(1)}(x_0)u \geq 0$. Then, $f^{(1)}(x_0)u = 0$. By the arbitrariness of $u$, we have $f^{(1)}(x_0) = \vartheta_{\mathcal{X}}^\circ$. This completes the proof of the proposition.
Proposition 10.16 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \to \mathbb{R}$ be $C^2$ at $x_0$. Assume that $x_0$ is a point of relative minimum for $f$, then $f^{(1)}(x_0) = \partial_{\mathcal{X}} f$ and $f^{(2)}(x_0)$ is positive semi-definite.

Proof By Proposition 10.15, $f^{(1)}(x_0) = \partial_{\mathcal{X}} f$. We will show that $f^{(2)}(x_0)$ is positive semi-definite by an argument of contradiction. Suppose $f^{(2)}(x_0)$ is not positive semi-definite. Then, $\exists h \in \mathcal{X}$ such that $f^{(2)}(x_0)(h)(h) < 0$. By the continuity of $f^{(2)}$ at $x_0$, there exists $\delta \in (0, \infty) \subseteq \mathbb{R}$ with $D := \mathcal{B}_\mathcal{X}(x_0, \delta) \subseteq \Omega$ such that $f^{(2)}(x)(h)(h) < 0$, $\forall x \in D$. Let $x_1 := x_0 + \alpha h \in D$, where $\alpha \in (0, \infty) \subseteq \mathbb{R}$ is an arbitrary constant. By Taylor’s Theorem 9.48, there exists $t_0 \in (0, 1) \subseteq \mathbb{R}$ such that

$$f(x_1) = f(x_0) + \frac{1}{2} f^{(2)}(x_0 + t_0 \alpha h)(x_1 - x_0)(x_1 - x_0) = f(x_0) + \frac{\alpha^2}{2} f^{(2)}(x_0 + t_0 \alpha h)(h)(h) < f(x_0)$$

This contradicts with the fact that $x_0$ is a point of relative minimum for $f$. Hence, the result must be true. This completes the proof of the proposition.

\[ \square \]

Proposition 10.17 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \to \mathbb{R}$ be twice differentiable at $x$ for $\forall x \in D := \mathcal{B}_\mathcal{X}(x_0, \delta_0) \subseteq \Omega$, where $\delta_0 \in (0, \infty) \subseteq \mathbb{R}$ is some constant. Assume that $f^{(1)}(x_0) = \partial_{\mathcal{X}} f$ and $f^{(2)}(x)$ is positive semi-definite, $\forall x \in D$. Then, $x_0$ is a point of relative minimum for $f$.

Proof $\forall x \in D$, by Taylor’s Theorem 9.48, there exists $t_0 \in (0, 1) \subseteq \mathbb{R}$ such that $f(x) = f(x_0) + \frac{1}{2} f^{(2)}(x_0 + t_0 (x - x_0))(x - x_0)(x - x_0) \geq f(x_0)$

Hence, $x_0$ is a point of relative minimum for $f$. This completes the proof of the proposition.

\[ \square \]

Proposition 10.18 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^\circ$, and $f : \Omega \to \mathbb{R}$ be $C^2$ at $x_0$. Assume that $f^{(1)}(x_0) = \partial_{\mathcal{X}} f$ and $f^{(2)}(x_0)$ is positive definite. Then, $x_0$ is a point of relative strict minimum for $f$.

Proof By Proposition 10.4 and the continuity of $f^{(2)}$ at $x_0$, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $f^{(2)}(x)$ is positive definite, $\forall x \in D := \mathcal{B}_\mathcal{X}(x_0, \delta) \subseteq \Omega$. Let $x \in D \setminus \{x_0\}$, by Taylor’s Theorem, there exists $t_0 \in (0, 1) \subseteq \mathbb{R}$ such that $f(x) = f(x_0) + \frac{1}{2} f^{(2)}(x_0 + t_0 (x - x_0))(x - x_0)(x - x_0) > f(x_0)$

Hence, $x_0$ is a point of relative strict minimum for $f$. This completes the proof of the proposition.

\[ \square \]
10.3 Optimization with Equality Constraints

The basic problem to be considered in this section is

\[ \mu_0 := \inf_{x \in \Omega} f(x) \quad \text{subject to} \quad H(x) = \vartheta \]  

(10.2)

where \( \mathcal{X} \) and \( \mathcal{Y} \) are real normed linear spaces, \( \Omega \subseteq \mathcal{X} \) is a set, \( f : \Omega \to \mathbb{R} \) is a functional, and \( H : \Omega \to \mathcal{Y} \) is a function.

**Definition 10.19** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real normed linear spaces, \( \Omega \subset \mathcal{X} \), \( H : \Omega \to \mathcal{Y} \) be \( C_1 \) at \( x_0 \in \Omega^0 \). \( x_0 \) is said to be a regular point of \( H \) if \( H'(x_0) \in B(\mathcal{X}, \mathcal{Y}) \) is surjective.

**Lemma 10.20** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real Banach spaces, \( \Omega \subset \mathcal{X} \), \( f : \Omega \to \mathbb{R} \) be \( C_1 \) at \( x_0 \in \Omega^0 \), and \( H : \Omega \to \mathcal{Y} \) be \( C_1 \) at \( x_0 \). Consider the optimization problem (10.2). Assume that \( x_0 \) is a point of relative minimum for \( f \) on the set \( \Omega_c := \{ x \in \Omega \mid H(x) = \vartheta \} \); and \( x_0 \) is a regular point of \( H \). Then, \( \forall u \in \mathcal{X} \) with \( H'(x_0)u = \vartheta \), we have \( f'(x_0)u = 0 \).

**Proof** Define \( T : \Omega \to \mathcal{X} \times \mathcal{Y} \) by \( T(x) = (f(x), H(x)) \), \( \forall x \in \Omega \). By Proposition 9.44, \( T \) is \( C_1 \) at \( x_0 \) and \( T'(x_0) = \begin{bmatrix} f'(x_0) & H'(x_0) \end{bmatrix} \).

We will prove the result using an argument of contradiction. Suppose the result is not true. Then, \( \exists u_0 \in \mathcal{X} \) with \( H'(x_0)u_0 = \vartheta \), we have \( f'(x_0)u_0 \neq 0 \). We will show that \( T'(x_0) \) is surjective. \( \forall (r, \vartheta) \in \mathcal{X} \times \mathcal{Y} \), \( x_0 \) being a regular point of \( H \), \( \exists v_1 \in \mathcal{X} \) such that \( H'(x_0)v_1 = \vartheta \). Let \( u = \frac{-f'(x_0)v_1}{\|f'(x_0)v_1\|} + u_1 \). Then, \( T'(x_0)u = (r, \vartheta) \). Hence, \( T'(x_0) \) is surjective.

Note that \( T(x_0) = (f(x_0), \vartheta) \) and \( \mathcal{X} \times \mathcal{Y} \) is a real Banach space, by Propositions 7.22 and 4.31. By Surjective Mapping Theorem 9.53, \( \exists \delta_r \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta_x \in (0, \infty) \subset \mathbb{R} \), and \( \delta_1 \in (0, \infty) \subset \mathbb{R} \) with \( c_1 \delta \leq \delta_x \), such that \( \forall (r, \vartheta) \in \mathcal{X} \times \mathcal{Y} (T(x_0), \delta/2) \), \( \exists x \in \mathcal{X} (x_0, \delta_x) \subseteq \Omega \) with \( \|x - x_0\| \leq c_1 \|r, \vartheta - T(x_0)\| \), we have \( T(x) = (r, \vartheta) \). Then, \( \forall \delta_r \in (0, \delta/2) \subset \mathbb{R} \), let \( r_1 = f(x_0) - \delta_x \in \mathbb{R} \). \( (r_1, \vartheta) \in \mathcal{X} (x_0, \delta_r) \). Then, \( \exists x_1 \in \Omega \) with \( \|x_1 - x_0\| \leq c_1 \|r_1, \vartheta - T(x_0)\| \), \( c_1 \|r_1 - f(x_0)\| = c_1 \delta_x < (1+c_1) \delta_r \) such that \( T(x_1) = (r_1, \vartheta) \). Then, \( H(x_1) = \vartheta \) and \( x_1 \in \Omega_c \cap \mathcal{B}_\mathcal{X} (x_0, (1+c_1) \delta_r) \). Furthermore, \( f(x_1) = r_1 < f(x_0) \). This contradicts with the assumption that \( x_0 \) is a point of relative minimum for \( f \) on \( \Omega_c \). Therefore, the result must be true.

This completes the proof of the lemma. \( \square \)

**Proposition 10.21** (Lagrange Multiplier) Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real Banach spaces, \( \Omega \subset \mathcal{X} \), \( f : \Omega \to \mathbb{R} \) be \( C_1 \) at \( x_0 \in \Omega^0 \), and \( H : \Omega \to \mathcal{Y} \) be \( C_1 \) at \( x_0 \). Consider the optimization problem (10.2). Assume that \( x_0 \) is a
point of relative minimum for $f$ on the set $\Omega_c := \{ x \in \Omega \mid H(x) = \vartheta_y \}$; and $x_0$ is a regular point of $H$. Then, there exists a Lagrange multiplier $y_{\ast} \in \mathbb{Y}^\ast$ such that the Lagrangian $L : \Omega \times \mathbb{Y}^\ast \to \mathbb{R}$ defined by $L(x, y_{\ast}) = f(x) + \langle \langle y_{\ast}, H(x) \rangle \rangle$, $\forall (x, y_{\ast}) \in \Omega \times \mathbb{Y}^\ast$, is stationary at $(x_0, y_{\ast})$, that is $L^{(1)}(x_0, y_{\ast}) = \vartheta_{B(\mathbb{X} \times \mathbb{Y}^\ast, \mathbb{R})}$.

**Proof** By Lemma 10.20, $\forall u \in \mathbb{X}$ with $H^{(1)}(x_0)u = \vartheta_y$, we have $f^{(1)}(x_0)u = 0$. Then, $f^{(1)}(x_0) \in (\mathcal{N}(H^{(1)}(x_0)))^\perp$. Since $x_0$ is a regular point of $H$, then $\mathcal{R}(H^{(1)}(x_0)) = \mathbb{Y}$ is closed. By Proposition 10.114, we have $(\mathcal{N}(H^{(1)}(x_0)))^\perp = \mathcal{R}\left(\mathcal{H}^{(1)}(x_0)\right)^{\perp}$. Then, there exists $y_{\ast} \in \mathbb{Y}^\ast$ such that $f^{(1)}(x_0) = (H^{(1)}(x_0))'y_{\ast}$. By Propositions 9.34, 9.41, 9.38, 9.37, 9.44, and 9.45, the Lagrangian $L$ is $C_1$ at $(x_0, y_{\ast})$ and, $\forall (u, v_{\ast}) \in \mathbb{X} \times \mathbb{Y}^\ast$,

$L^{(1)}(x_0, y_{\ast})(u, v_{\ast}) = f^{(1)}(x_0)u + \langle \langle v_{\ast}, H(x_0) \rangle \rangle + \langle \langle y_{\ast}, H^{(1)}(x_0)u \rangle \rangle$

$= f^{(1)}(x_0)u + \langle \langle \left(H^{(1)}(x_0)\right)'y_{\ast}, u \rangle \rangle = 0$

where the second equality follows from that fact that $x_0 \in \Omega_c$. Hence, $L^{(1)}(x_0, y_{\ast}) = \vartheta_{B(\mathbb{X} \times \mathbb{Y}^\ast, \mathbb{R})}$. This completes the proof of the proposition. □

**Proposition 10.22 (Generalized Lagrange Multiplier) Let $\mathbb{X}$ and $\mathbb{Y}$ be real Banach spaces, $\Omega \subseteq \mathbb{X}$, $f : \Omega \to \mathbb{R}$ be $C_1$ at $x_0 \in \Omega^\ast$, and $H : \Omega \to \mathbb{Y}$ be $C_1$ at $x_0$. Consider the optimization problem (10.2). Assume that $x_0$ is a point of relative minimum for $f$ on the set $\Omega_c := \{ x \in \Omega \mid H(x) = \vartheta_y \}$; and $\mathcal{R}(H^{(1)}(x_0)) \subseteq \mathbb{Y}$ is closed. Then, there exists a Lagrange multiplier $(r_0, y_{\ast}) \in \mathbb{R} \times \mathbb{Y}^\ast$ with $(r_0, y_{\ast}) \neq (0, \vartheta_y)$ such that the Lagrangian $L : \Omega \times \mathbb{Y}^\ast \to \mathbb{R}$ defined by $L(x, y_{\ast}) = r_0f(x) + \langle \langle y_{\ast}, H(x) \rangle \rangle$, $\forall (x, y_{\ast}) \in \Omega \times \mathbb{Y}^\ast$, is stationary at $(x_0, y_{\ast})$, that is $L^{(1)}(x_0, y_{\ast}) = \vartheta_{B(\mathbb{X} \times \mathbb{Y}^\ast, \mathbb{R})}$.

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{R}(H^{(1)}(x_0)) = \mathbb{Y}$. Take $r_0 = 1$. Clearly, $x_0$ is a regular point of $H$. The result follows from Proposition 10.21.

Case 2: $\mathcal{R}(H^{(1)}(x_0)) \subset \mathbb{Y}$. Take $r_0 = 0$. Clearly, $x_0$ is not a regular point of $H$. Then, $\exists y_0 \in \mathbb{Y} \setminus \mathcal{R}(H^{(1)}(x_0))$. Let $M = \mathcal{R}(H^{(1)}(x_0))$, which is a closed subspace of $\mathbb{Y}$ by the assumption. Then, by Proposition 4.10, $\delta := \inf_{m \in M} \| y_0 - m \| > 0$. By Proposition 7.97, we have $\delta = \max_{y_{\ast} \in M^\perp, \| y_{\ast} \| \leq 1} \langle \langle y_{\ast}, y_0 \rangle \rangle$, where the maximum is achieved at $y_0 \in M^\perp$. Then, $y_{\ast} \neq \vartheta_y$ and, by Proposition 7.112, $y_{\ast} \in \mathcal{N} \left( \left(H^{(1)}(x_0)\right)' \right)$. The Lagrangian $L$ is given by $L(x, y_{\ast}) = \langle \langle y_{\ast}, H(x) \rangle \rangle$, $\forall (x, y_{\ast}) \in \Omega \times \mathbb{Y}^\ast$. By Propositions 9.34, 9.41, 9.37, 9.44, and 9.45, $L$ is $C_1$ at $(x_0, y_{\ast})$ and, $\forall (u, v_{\ast}) \in \mathbb{X} \times \mathbb{Y}^\ast$,

$L^{(1)}(x_0, y_{\ast})(u, v_{\ast}) = \langle \langle v_{\ast}, H(x_0) \rangle \rangle + \langle \langle y_{\ast}, H^{(1)}(x_0)u \rangle \rangle$

$= \langle \langle \left(H^{(1)}(x_0)\right)'y_{\ast}, u \rangle \rangle = \langle \vartheta_{\mathbb{X}^\ast}, u \rangle = 0$
where the second equality follows from that fact that \( x_0 \in \Omega_c \). Hence, 
\[ L^{(1)}(x_0, y_{s0}) = \vartheta_{B(X \times Y^*, R)} \]. This case is proved.

This completes the proof of the proposition. \( \square \)

**Proposition 10.23** Let \( X \) and \( Y \) be real Banach spaces, \( \Omega \subseteq X \), \( f : \Omega \to \mathbb{R} \)
be \( C_2 \) at \( x_0 \in \Omega^0 \), \( H : \Omega \to Y \) be \( C_2 \) at \( x_0 \). Consider the optimization problem
(10.2). Assume that

(i) \( H(x_0) = \vartheta_Y \) and \( \mathcal{R}(H^{(1)}(x_0)) \subseteq Y \) is closed;

(ii) the Lagrangian \( L : \Omega \times Y^* \to \mathbb{R} \) defined by \( L(x, y_*) = f(x) + \langle (y_*, H(x)) \rangle \), \( \forall (x, y_*) \in \Omega \times Y^* \), is stationary at \((x_0, y_{s0})\), where \( y_{s0} \in Y^* \) is a Lagrange multiplier;

(iii) \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{s0}) \) is positive definite on the subspace \( M := \mathcal{N}(H^{(1)}(x_0)) \),
that is, \( \exists m \in (0, \infty) \subset \mathbb{R} \) such that \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{s0})(h) \geq m \| h \|^2 \),
\( \forall h \in \mathcal{N}(H^{(1)}(x_0)) \).

Then, \( x_0 \) is a point of relative strict minimum for \( f \) on the set \( \Omega_c := \{ x \in \Omega \mid H(x) = \vartheta_Y \} \).

**Proof** By Propositions 9.34, 9.37, 9.38, 9.41, 9.44, and 9.45, \( L \) is \( C_2 \) at \((x_0, y_{s0})\). By Proposition 9.9 and (ii),
\[ L^{(1)}(x_0, y_{s0}) = \left[ \frac{\partial L}{\partial x}(x_0, y_{s0}) \right] \right] = \vartheta_{B(X \times Y^*, R)} \]
Then, \( \exists \delta_0 \in (0, \infty) \subset \mathbb{R} \) such that \( f^{(2)}(h) \), \( H^{(2)}(h) \), and \( \frac{\partial^2 L}{\partial x^2}(x, y_*) \) exists, \( \forall x \in B_X(x_0, \delta_0) \subseteq \Omega \) and \( \forall y_* \in B_{Y^*}(y_{s0}, \delta_0) \), and, by Proposition 9.46, \( \frac{\partial^2 L}{\partial x^2} \) is continuous at \((x_0, y_{s0})\). Then, \( \exists \delta_1 \in (0, \delta_0) \subset \mathbb{R} \) and \( \exists \varepsilon \in [0, \infty) \subset \mathbb{R} \) such that \( \left\| \frac{\partial^2 L}{\partial x^2}(x, y_*) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{s0}) \right\| < \varepsilon \) and \( \| H^{(2)}(h) \| \leq c_1 \), 
\( \forall x \in D_1 := B_X(x_0, \delta_1) \). By Propositions 7.114 and 7.98 and (i), \( M^\perp = \mathcal{R}\left((H^{(1)}(x_0))'\right) \) is closed. Then, by Proposition 7.113, \( \exists \varepsilon_2 \in [0, \infty) \subset \mathbb{R} \)
such that, \( \forall x_* \in \mathcal{R}\left((H^{(1)}(x_0))'\right) \), there exists \( y_* \in Y^* \) such that \( x_* = (H^{(1)}(x_0))' y_* \) and \( \| y_* \| \leq c_2 \| x_* \| \).

**Claim 10.23.1** \( \forall x \in \Omega_c \cap D_1 \), \( \exists h_0 \in M \) such that \( \| x - x_0 - h_0 \| \leq c_1 \| x - x_0 \|^2 \), \( \| x - x_0 \| - c_1 \| x - x_0 \|^2 \leq \| h_0 \| \leq \| x - x_0 \| + c_2 \| x - x_0 \|^2 \).

**Proof of claim:** Fix any \( x \in \Omega_c \cap D_1 \). Then, \( H(x) = H(x_0) = \vartheta_Y \). By Taylor’s Theorem 9.48, \( \exists \theta \in (0, 1) \subset \mathbb{R} \) such that \( \| H^{(1)}(x_0)(x - x_0) \| = \| H(x) - H(x_0) - H^{(1)}(x_0)(x - x_0) \| \leq \frac{\varepsilon}{2} \| H^{(2)}(x_0)(x - x_0) \| \| x - x_0 \|^2 \leq \varepsilon \| x - x_0 \|^2 / 2 \). By Proposition 7.97, we have \( \inf_{h \in M} \| x - x_0 - h \| = \max_{x_* \in \mathcal{R}(H^{(1)}(x_0)), \| x_* \| \leq 1} \langle (x_*, x - x_0) \rangle \), where the maximum is achieved at \( x_{s1} \in \mathcal{M}^\perp = \mathcal{R}(H^{(1)}(x_0))' \) with \( \| x_{s1} \| \leq 1 \). Then, \( \exists y_{s1} \in Y^* \) such that
10.3. Optimization with Equality Constraints

\[ x_{s+1} = (H^{(1)}(x_0))' y_{s+1} \text{ and } \| y_{s+1} \| \leq c_2. \] By Proposition 7.68, \( M \) is closed. Then, \( \exists h_0 \in M \) such that \( \| x - x_0 - h_0 \| \leq 2 \inf_{h \in M} \| x - x_0 - h \| = 2 \langle (x_{s+1}, x - x_0) \rangle = 2 \langle (y_{s+1}, H^{(1)}(x_0)(x - x_0)) \rangle \leq 2 \| y_{s+1} \| \| H^{(1)}(x_0)(x - x_0) \| \leq c_1 c_2 \| x - x_0 \|^2. \]

Then, \( \| h_0 \| \geq \| x - x_0 \| - \| x - x_0 - h_0 \| \geq \| x - x_0 \| - c_1 c_2 \| x - x_0 \|^2; \) and \( \| h_0 \| \leq \| x - x_0 \| + \| x - x_0 - h_0 \| \leq \| x - x_0 \| + c_1 c_2 \| x - x_0 \|^2. \) This completes the proof of the claim.

Let \( c_3 = \| \bar{\partial} L(x_0, y_0) \| \). Let \( \delta \in (0, \delta_1] \subseteq \mathbb{R} \) such that \( c_1 c_2 \delta \leq 1/4, \)
\( c_1^2 c_2^2 (c_3 + m/5) \delta^2 \leq m/5, \) and \( 5c_1 c_2 (c_3 + m/5) \delta/2 \leq m/5. \) \( \forall x \in \Omega \cap B_{\mathbb{R}}(x_0, \delta), \) by Claim 10.23.1, \( \exists h_0 \in M \) such that \( \| x - x_0 - h_0 \| \leq c_1 c_2 \| x - x_0 \|^2 \leq \| x - x_0 \| / 4, 3 \| x - x_0 \| / 4 \leq \| h_0 \| \leq 5 \| x - x_0 \| / 4. \) By Taylor’s Theorem 9.48, \( \exists t_1 \in (0, 1) \subseteq \mathbb{R} \) such that

\[
f(x) - f(x_0) = L(x, y_0) - L(x_0, y_0) - \partial L / \partial x(x_0, y_0)(x - x_0)
= \frac{1}{2} \partial^2 L / \partial x^2(x_1 x + (1 - t_1) x_0, y_0)(x - x_0)(x - x_0)
= \frac{1}{2} \left( \partial^2 L / \partial x^2(\bar{x}, y_0)(h_0) + \partial^2 L / \partial x^2(\bar{x}, y_0)(x - x_0 - h_0)(h_0) + \partial^2 L / \partial x^2(\bar{x}, y_0)(x - x_0 - h_0) \right)
\geq \frac{1}{2} \left( \partial^2 L / \partial x^2(x_0, y_0)(h_0) + \partial^2 L / \partial x^2(\bar{x}, y_0) - \partial^2 L / \partial x^2(x_0, y_0)(h_0)(h_0) \right)
\geq \frac{1}{2} \left( m \| h_0 \|^2 \| x - x_0 - h_0 \| + \| h_0 \| \right) \| x - x_0 - h_0 \| \| x - x_0 - h_0 \| / \| x - x_0 - h_0 \| \right) \| x - x_0 - h_0 \| \| x - x_0 - h_0 \| \| x - x_0 - h_0 \| \right)
\geq \frac{1}{2} \left( \frac{4m}{5} \| h_0 \|^2 - (c_3 + m/5) \| x - x_0 \| \| x - x_0 \| \right)
\geq \frac{1}{2} \left( \frac{4m}{5} \| h_0 \|^2 - (c_3 + m/5) \right) \geq \frac{m}{5} \| x - x_0 \| \}
\]

Hence, \( x_0 \) is a point of relative strict minimum for \( f \) on the set \( \Omega_c. \) This completes the proof of the proposition.

**Proposition 10.24** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real Banach spaces, \( \Omega \subseteq \mathcal{X}, f : \Omega \to \mathbb{R} \) be \( C^2 \) at \( x_0 \in \Omega, H : \Omega \to \mathcal{Y} \) be \( C^2 \) at \( x_0. \) Consider the optimization problem (10.2). Assume that \( x_0 \) is a regular point of \( H \) and is a point of relative minimum for \( f \) on the set \( \Omega_c := \{ x \in \Omega \mid H(x) = \partial \mathcal{Y} \}. \) Then, there exists a Lagrange multiplier \( y_{s+1} \in \mathcal{Y}^* \) such that
(i) the Lagrangian \( L : \Omega \times Y^* \to \mathbb{R} \) defined by \( L(x, y_*) = f(x) + \langle (y_*, H(x)) \rangle \), \( \forall (x, y_*) \in \Omega \times Y^* \), is stationary at \((x_0, y_{*0})\); 

(ii) \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \) is positive semi-definite on the subspace \( N(H(1)(x_0)) = M \), that is, \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h) \geq 0 \), \( \forall h \in N(H(1)(x_0)) \).

**Proof**

Under the assumption of the proposition, by Proposition 10.21, there exists a Lagrange multiplier \( y_{*0} \in Y^* \) such that (i) holds. We will show that (ii) also holds by an argument of contradiction. Suppose (ii) does not hold. Then, \( \exists h_0 \in M \) with \( \| h_0 \| = 1 \) such that \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0})(h_0) < -m < 0 \) for some \( m \in (0, \infty) \subset \mathbb{R} \). By Surjective Mapping Theorem 9.53, \( \exists r_1 \in (0, \infty) \subset \mathbb{R}, \exists \delta_1 \in (0, \infty) \subset \mathbb{R}, \) and \( \exists r_2 \in (0, \infty) \subset \mathbb{R} \) with \( c_1 \delta_1 \leq r_1 \), such that \( \forall \gamma \in B_{x}(\theta_y, \delta_1/2), \forall \delta \in B_{X}(x_0, r_1/2) \) with \( \gamma = H(\delta, \delta), \forall y \in B_{y}(\gamma, \delta_1/2) \), \( \exists x \in B_{X}(x_0, r_1) \subseteq \Omega \) with \( \| x_0 - \delta \delta \| \leq c_1 \| y - \gamma \| \), we have \( y = H(x) \). By Propositions 9.34, 9.37, 9.38, 9.41, 9.44, and 9.45, \( L \) is \( C_2 \) at \((x_0, y_{*0})\). By Proposition 9.9 and (i),

\[
L^{(1)}(x_0, y_{*0}) = \left[ \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right] = \vartheta_{B(x \times y^*, \mathbb{R})}
\]

By Proposition 9.46, \( \exists r_2 \in (0, r_1) \subset \mathbb{R} \) such that \( \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \) exists and

\[
\left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| < m/5, \forall x \in B_{X}(x_0, r_2). \] Since \( H \) is \( C_2 \) at \( x_0 \), then \( \exists \delta_2 \in (0, \infty) \subset \mathbb{R} \) and \( \exists \delta_3 \in (0, r_2) \subset \mathbb{R} \) such that \( \| H^{(2)}(x_0) \| \leq c_2 \), \( \forall \delta \in B_{X}(x_0, r_3) \). Let \( c_3 := \left\| \frac{\partial^2 L}{\partial x^2}(x_0, y_{*0}) \right\| \).

\( \forall \delta \in (0, r_3/2) \subset \mathbb{R} \) such that \( c_2 \delta^2 < \delta_1, c_1 c_2 \delta^2 / 2 \leq 1/4, c_1^2 \delta^2 (c_3 + m/5) \delta^2 / 4 \leq m/5, \) and \( c_1 c_2 (c_3 + m/5) \delta \leq m/5 \). By Taylor's Theorem 9.48, \( \exists h_0 \in (0, 1) \subset \mathbb{R} \) such that

\[
\| H(x_0 + \delta h_0) \| = \| H(x_0 + \delta h_0) - H(x_0) - \delta H^{(1)}(x_0) h_0 \| 
\leq \frac{1}{2} \| H^{(2)}(x_0 + \delta h_0) \| \delta^2 \leq c_2 \delta^2 / 2 < \delta_1 / 2
\]

Let \( \bar{y} := H(x_0 + \delta h_0) \in B_{y}(\theta_y, \delta_1/2) \) and \( \bar{x} := x_0 + \delta h_0 \in B_{X}(x_0, r_1/2) \). Note that \( \bar{y} = H(\bar{x}) \) and \( \bar{x} \in B_{X}(x_0, \delta_1/2) \). Then, \( \exists x_\delta \in B_{X}(x_0, r_1) \) with \( \| x_\delta - \bar{x} \| \leq c_1 \| \bar{y} \| \leq c_1 c_2 \delta^2 / 2 \) such that \( H(x_\delta) = \bar{y} \). Then, we have \( \| x_\delta - x_0 \| \geq \| \delta h_0 \| - \| x_\delta - \bar{x} \| \geq \delta - c_1 c_2 \delta^2 / 2 \geq \delta \delta / 4 \) and \( \| x_\delta - x_0 \| \leq \| \delta h_0 \| + \| x_\delta - \bar{x} \| \leq \delta + c_1 c_2 \delta^2 / 2 \leq 5 \delta^2 / 4 < r_3 \). Hence, \( x_\delta \in \Omega_c \cap B_{X}(x_0, r_3) \). By Taylor's Theorem 9.48, \( \exists \delta_1 \in (0, 1) \subset \mathbb{R} \) such that

\[
f(x_\delta) - f(x_0) = L(x_\delta, y_{*0}) - L(x_0, y_{*0}) - \frac{\partial L}{\partial x}(x_0, y_{*0})(x_\delta - x_0)
= \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(t_1 x_\delta + (1 - t_1) x_0, y_{*0})(x_\delta - x_0)(x_\delta - x_0)
\]
The basic problem to be considered in this section is

\[ \mu_0 := \inf_{x \in \Omega} f(x) \quad \text{subject to} \quad x \in \Omega, \quad G(x) = \vartheta_Z \quad (10.3) \]

where \( X \) is a real normed linear spaces, \( \Omega \subseteq X \) is a set, \( f : \Omega \to \mathbb{R} \) is a functional, \( Z \) is a real normed linear space with a positive cone \( P \subseteq Z \), and \( G : \Omega \to Z \) is a function.

**Theorem 10.25 (Generalized Kuhn-Tucker Theorem)** Let \( X \) be a real Banach space, \( \Omega \subseteq X \), \( x_0 \in \Omega^\circ \), \( Z \) be a real Banach space with a positive cone \( P \subseteq Z \), and \( f : \Omega \to \mathbb{R} \) and \( G : \Omega \to Z \) be \( C_1 \) at \( x_0 \). Consider the optimization problem (10.3). Assume that \( x_0 \) is a regular point of \( G \) and is a point of relative minimum for \( f \) on the set \( \Omega_c := \{ x \in \Omega \mid G(x) = \vartheta_Z \} \).

Then, there exists a Lagrange multiplier \( z_{*0} \in Z^* \) with \( z_{*0} \vartrianglerighteq \vartheta_Z^* \) such that
the Lagrangian $L : \Omega \rightarrow \mathbb{R}$ defined by $L(x) = f(x) + \langle z_0, G(x) \rangle$, $\forall x \in \Omega$, is stationary at $x_0$ and $\langle z_0, G(x_0) \rangle = 0$.

**Proof** Consider the small variation equivalent problem of (10.3) at $x_0$:

$$\bar{\mu}_0 := \inf_{u \in X, G(x_0) + G^{(1)}(x_0)u \leq \vartheta_Z} f^{(1)}(x_0)u \quad (10.4)$$

This problem is convex, and belongs to the class of problems studied in Section 8.8. The primal function for (10.4) is $\omega : \bar{\Gamma} \rightarrow \mathbb{R}$, where

$$\bar{\Gamma} := \left\{ \bar{\varepsilon} \in \mathbb{Z} \mid \exists u \in X : G(x_0) + G^{(1)}(x_0)u \leq \bar{\varepsilon} \right\} \quad (10.5a)$$

$\omega(\bar{\varepsilon}) = \inf_{u \in X, G(x_0) + G^{(1)}(x_0)u \leq \bar{\varepsilon}} f^{(1)}(x_0)u; \quad \forall \bar{\varepsilon} \in \bar{\Gamma} \quad (10.5b)$

Since $G^{(1)}(x_0) \in \mathcal{B}(X, \mathbb{Z})$ is surjective, then $\bar{\Gamma} = \mathbb{Z}$. Note that $\bar{\mu}_0 = \omega(\vartheta_Z)$. Since $x_0 \in \Omega_0$, then $z_0 := G(x_0) \leq \vartheta_Z$. Hence, $\bar{\mu}_0 \leq 0$.

**Claim 10.25.1** $\bar{\mu}_0 = 0$.

**Proof of claim:** We will prove this using an argument of contradiction. Suppose $\bar{\mu}_0 < 0$. Then, $\exists u_0 \in X$ with $\|u_0\| < 0$ and $G(x_0) + G^{(1)}(x_0)u_0 \leq \vartheta_Z$. By Surjective Mapping Theorem 9.53, $\exists r_1 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, and $\exists c_1 \in [0, \infty) \subset \mathbb{R}$ with $c_1 \delta_1 \leq r_1$ such that $\forall \bar{\varepsilon} \in \mathcal{B}_Z(0, \delta_1/2)$, $\forall \bar{\varepsilon} \in \mathcal{B}_X(x_0, r_1/2)$ with $\bar{\varepsilon} = G(\bar{x})$, $\exists \bar{z} \in \mathcal{B}_Z(\bar{z}, \delta_1/2)$, $\exists \bar{x} \in \mathcal{B}_X(x_0, r_1) \subseteq \Omega$ with $\|x - \bar{x}\| \leq c_1 \|z - \bar{z}\|$, we have $z = G(x)$. Let $c_2 := \|G^{(1)}(x_0)\|$ and $c_3 := \|f^{(1)}(x_0)\|$. By $f$ and $G$ being $C_1$ at $x_0$, $\forall \bar{\varepsilon} \in (0, r_1/2) \subset \mathbb{R}$ such that $\forall \bar{\varepsilon} \in \mathcal{B}_X(x_0, r_2)$, $\|f^{(1)}(x) - f^{(1)}(x_0)\| < m/3$ and $\|G^{(1)}(x) - G^{(1)}(x_0)\| < m/3/(1+c_3)$. Let $\delta \varepsilon := (0, 1) \subset \mathbb{R}$ such that $\delta \|u_0\| < r_2/2$, $(c_2 + m/3)\delta \|u_0\| < \delta_1/2$, and $\delta m \|u_0\|/3 < r_2/2$. By Mean Value Theorem 9.23, $\exists \theta_0 \in (0, 1) \subset \mathbb{R}$ such that $\|G(x_0 + \delta u_0) - G(x_0) - \delta G^{(1)}(x_0)u_0\| \leq \|(G^{(1)}(x_0 + \theta_0 \delta u_0) - G^{(1)}(x_0))\| \leq (c_1 \|G^{(1)}(x_0)\| \leq (c_2 \|u_0\| + m)/(1+c_3) \delta \|u_0\| < \delta_1/2$. Let $\bar{x}_\delta := x_0 + \delta u_0 \in \mathcal{B}_X(x_0, r_2/2)$, $\bar{z}_\delta := G(\bar{x}_\delta)$, and $z_\delta := G(x_0) + \delta G^{(1)}(x_0)u_0 \in \mathcal{B}_Z(\bar{z}_\delta, \delta_1/2)$. Note that $\|z_\delta - z_0\| \leq \|G(x_0 + \delta u_0) - G(x_0) - \delta G^{(1)}(x_0)u_0\| + \|\delta G^{(1)}(x_0)\|u_0\| < c_2 \delta \|u_0\| + \|(c_1 + c_3 + m/3)\delta \|u_0\| < \delta_1/2$. Then, $\exists x_\delta \in \mathcal{B}_X(x_0, r_1)$ with $\|x_\delta - \bar{x}_\delta\| \leq c_1 \|z_\delta - \bar{z}_\delta\| \leq (c_1 + c_3 + m/3)\delta \|u_0\| < r_2/2$ such that $z_\delta = G(x_\delta)$. Note that $G(x_\delta) = z_\delta = (1 - \delta)G(x_0) + \delta G(x_0) + G^{(1)}(x_0)u_0 \leq \vartheta_Z$. Then, $x_\delta \in \Omega_0$. Note also that $x_\delta \in \mathcal{B}_X(x_0, r_2)$. By Mean Value Theorem 9.20, $\exists t_1, t_2 \in (0, 1) \subset \mathbb{R}$ such that

$$f(x_\delta) - f(x_0) = f(x_\delta) - f(x_\delta) + f(x_\delta) - f(x_0)$$
Then, we have
\[
L = f(1)(t_1 x_\delta + (1 - t_1) \bar{x}_\delta) (x_\delta - \bar{x}_\delta) + \delta f(1)(x_0 + t_2 \delta u_0) u_0
\]
\[
\leq \left( \| f(1)(x_0) \| + \| f(1)(t_1 x_\delta + (1 - t_1) \bar{x}_\delta) - f(1)(x_0) \| \right) \| x_\delta - \bar{x}_\delta \|
\]
\[
+ \delta f(1)(x_0) u_0 + \delta \| f(1)(x_0 + t_2 \delta u_0) - f(1)(x_0) \| \| u_0 \|
\]
\[
\leq -m\delta \| u_0 \| + (c_3 + m/3) \frac{c_1 m}{3(1 + c_1 + c_3 + m/3)} \| u_0 \|
\]
\[
+ m\delta \| u_0 \| /3 < -m\delta \| u_0 \| /3 < 0
\]
Note that we have shown that \( f(x_\delta) < f(x_0) \), \( x_\delta \in \Omega_c \), and \( \lim_{\delta \to 0^+} x_\delta = x_0 \). This contradicts the assumption that \( x_0 \) is a point of relative minimum for \( f \) on \( \Omega_c \). Hence, \( \bar{u} = 0 \).

This completes the proof of the claim.

Then, \( \omega(\bar{\theta}_z) = 0 \). By Fact 8.52, \( \omega : \mathbb{Z} \to \mathbb{R} \) is real-valued and convex.

Next, we will show that \( \omega \) is continuous at \( \bar{\theta}_z \) by Proposition 8.22.

By Proposition 7.113, \( \exists c_4 \in [0, \infty) \subseteq \mathbb{R} \), \( \forall z \in \mathbb{Z} \), \( \exists \bar{u} \in X \) such that \( \bar{z} = G(1)(x_0) \bar{u} \) and \( \| \bar{u} \| \leq c_4 \| \bar{z} \| \). Then, \( G(x_0) + G(1)(x_0) \bar{u} = G(x_0) + \bar{z} \leq \bar{z} \) and \( f(1)(x_0) \bar{u} \leq \| f(1)(x_0) \| \| \bar{u} \| \leq c_4 \| f(1)(x_0) \| \| \bar{z} \| \). This implies that \( \forall \bar{z} \in B_{\mathbb{R}}(\bar{\theta}_z, 1/2), \| \omega(\bar{z}) \| \leq c_4 \| f(1)(x_0) \| \| \bar{z} \| \). Hence, \( (r_0, \bar{\theta}_z, 1/2) \subseteq [\omega, \mathbb{Z}] \). By Proposition 8.22, \( \omega \) is continuous at \( \bar{\theta}_z \). By Proposition 8.23, \( \omega \) is continuous.

By Proposition 8.34, we have
\[
0 = \bar{\mu}_0 = \omega(\bar{\theta}_z) = \max_{z_* \in \bar{\Gamma}_{\text{conj}}} (-\omega_{\text{conj}}(z_*))
\]
where \( \bar{\Gamma}_{\text{conj}} \) and \( \omega_{\text{conj}} : \bar{\Gamma}_{\text{conj}} \to \mathbb{R} \) are the conjugate set and conjugate functional of \( \omega \). By Fact 8.53, \( \bar{\Gamma}_{\text{conj}} \subseteq P^\Theta \). Let the maximum be achieved at \(-z_0 \in \bar{\Gamma}_{\text{conj}}, \exists z_0 \geq \bar{\theta}_{z_*} \). Define \( \bar{\omega} : P^\Theta \to \mathbb{R} \) by \( \bar{\omega}(z_*) = \sup_{z \in \bar{\Gamma}_{\text{conj}}} \langle \langle z_*, z \rangle \rangle - \omega(z), \forall z_* \in P^\Theta \). By Fact 8.54, we have \( \max_{z_* \in P^\Theta} (-\omega_{\text{conj}}(z_*)) = \max_{z_* \in P^\Theta} (-\bar{\omega}(z_*)) \), where both maximums are achieved at \(-z_0 \), and \( \forall z_* \in P^\Theta \),
\[
-\bar{\omega}(z_*) = \inf_{u \in X} (f(1)(x_0)u + \langle \langle z_*, G(x_0) + G(1)(x_0)u \rangle \rangle)
\]
Then, we have
\[
0 = \bar{\mu}_0 = \max_{z, \bar{z} \geq \bar{\theta}_{z_*}} \inf_{u \in X} (f(1)(x_0)u + \langle \langle z_*, G(x_0) + G(1)(x_0)u \rangle \rangle)
\]
where the maximum is achieved at \( z_0 \). This implies that \( \langle \langle z_0, G(x_0) \rangle \rangle + \inf_{u \in X} \left\langle \langle f(1)(x_0) + (G(1)(x_0))', z_0, u \rangle \rangle \right\rangle = 0 \) by Proposition 8.37. For the infimum to be finite, we must have \( f(1)(x_0) + (G(1)(x_0))' \) \( z_0 = \bar{\theta}_{z_*} \). Hence, the Lagrangian \( L \) is stationary at \( x_0 \). Then, \( 0 = \langle \langle z_0, G(x_0) \rangle \rangle \).

This completes the proof of the proposition. \( \square \)
Proposition 10.26 Let $\mathcal{X}$ be a real Banach space, $\mathcal{Y}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Y}$, and $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. Then,

$$(A^{\text{inv}}(P))^{\oplus} = A'(P^{\oplus})$$

Proof \hspace{1em} $\forall x_{\ast} \in A'(P^{\oplus}), \exists y_{\ast} \in P^{\oplus}$ such that $x_{\ast} = A' y_{\ast}$. $\forall x \in A^{\text{inv}}(P)$, we have $Ax \in P$. Then, $\langle \langle x_{\ast}, x \rangle \rangle = \langle \langle A' y_{\ast}, x \rangle \rangle = \langle \langle y_{\ast}, Ax \rangle \rangle \geq 0$. By the arbitrariness of $x$, $x_{\ast} \in (A^{\text{inv}}(P))^{\oplus}$. By the arbitrariness of $x_{\ast}$, we have $A'(P^{\oplus}) \subseteq (A^{\text{inv}}(P))^{\oplus}$.

On the other hand, fix any $x_{\ast} \in (A^{\text{inv}}(P))^{\oplus}$, $\forall x \in \mathcal{N}(A)$, $x \in A^{\text{inv}}(P)$, since $\vartheta y \in P$. Then, $\langle \langle x_{\ast}, x \rangle \rangle \geq 0$. Since $-x \in \mathcal{N}(A)$ as well, then $\langle \langle x_{\ast}, x \rangle \rangle = 0$. Hence, $x_{\ast} \in (\mathcal{N}(A))_{\perp}$. By Proposition 7.114, $(\mathcal{N}(A))_{\perp} = \mathcal{R}(A')$. Then, $\exists y_{\ast} \in \mathcal{Y}^{\ast}$ such that $x_{\ast} = A' y_{\ast}$. $\forall y \in P$, since $A$ is surjective, then $\exists \overline{x} \in A^{\text{inv}}(P)$ such that $y = Ax$. Then, we have $\langle \langle y_{\ast}, y \rangle \rangle = \langle \langle y_{\ast}, Ax \rangle \rangle = \langle \langle A' y_{\ast}, x \rangle \rangle = \langle \langle x_{\ast}, x \rangle \rangle \geq 0$. By the arbitrariness of $y$, we have $y_{\ast} \in P^{\oplus}$. Then, $x_{\ast} \in A'(P^{\oplus})$. By the arbitrariness of $x_{\ast}$, we have $(A^{\text{inv}}(P))^{\oplus} \subseteq A'(P^{\oplus})$.

Therefore, we have $A'(P^{\oplus}) = (A^{\text{inv}}(P))^{\oplus}$. This completes the proof of the proposition. \hfill \Box

Next, we present the second-order sufficient condition for a relative strict minimum point in the optimization problem (10.3) with inequality constraints.

Proposition 10.27 Let $\mathcal{X}$ be a real Banach space, $\Omega \subseteq \mathcal{X}$, $x_{0} \in \Omega^{o}$, $\mathcal{Z}$ be a real Banach space with a positive cone $P \subseteq \mathcal{Z}$, and $f : \Omega \rightarrow \mathcal{R}$ and $G : \Omega \rightarrow \mathcal{Z}$ be $C_{2}$ at $x_{0}$. Consider the optimization problem (10.3). Assume that

(i) $G(x_{0}) \leq \vartheta_{\mathcal{Z}}$ and $x_{0}$ is regular point of $G$;

(ii) $z_{0} \in P^{\oplus}$ is the Lagrange multiplier, the Lagrangian $L : \Omega \rightarrow \mathcal{R}$ defined by $L(x) = f(x) + \langle \langle z_{0}, G(x) \rangle \rangle$, $\forall x \in \Omega$, is stationary at $x_{0}$;

(iii) $\langle \langle z_{0}, G(x_{0}) \rangle \rangle = 0$;

(iv) $\exists m \in (0, \infty) \subset \mathcal{R}$ such that $L^{(2)}(x_{0})(u)(u) \geq m \| u \|^{2}, \forall u \in M_{c} := \left\{ u \in \mathcal{X} \mid G(x_{0}) + G^{(1)}(x_{0})u \leq \vartheta_{\mathcal{Z}} \right\}$, that is, $L^{(2)}(x_{0})$ is positive definite on the set $M_{c}$.

Then, $x_{0}$ is a point of relative strict minimum of $f$ on the set $\Omega_{c} := \left\{ x \in \mathcal{X} \mid G(x) \leq \vartheta_{\mathcal{Z}} \right\}$.

Proof \hspace{1em} By Propositions 9.34, 9.37, 9.40, 9.41, 9.44, and 9.45, $L$ is $C_{2}$ at $x_{0}$. By (ii), $L^{(1)}(x_{0}) = \vartheta_{\mathcal{B}(\mathcal{X}, \mathcal{R})} = \vartheta_{\mathcal{X}}$. Then, $\exists \delta_{0} \in (0, \infty) \subset \mathcal{R}$ such that $f^{(2)}(x), G^{(2)}(x)$, and $L^{(2)}(x)$ exist, $\forall x \in \mathcal{B}_{\mathcal{X}}(x_{0}, \delta_{0}) \subseteq \Omega$ and $L^{(2)}$ is continuous at $x_{0}$. Then, $\exists \delta_{1} \in (0, \delta_{0}] \subset \mathcal{R}$ and $\exists c_{1} \in [0, \infty) \subset \mathcal{R}$
such that \( \| L^{(2)}(x) - L^{(2)}(x_0) \| < m/5 \) and \( \| G^{(2)}(x) \| \leq c_1, \forall x \in D_1 := B_{\delta}(x_0, \delta_1) \). By Propositions 7.114 and 7.98 and (i), \( \langle \mathcal{N}(G^{(1)}(x_0)) \rangle^* = \mathcal{R} \left( \left( G^{(1)}(x_0) \right)' \right) \) is closed. Then, by Proposition 7.113, \( \exists c_2 \in [0, \infty) \subset \mathbb{R} \) such that, \( \forall x_\ast \in \mathcal{R} \left( \left( G^{(1)}(x_0) \right)' \right) \), there exists \( z_\ast \in \mathbb{Z}^* \) such that \( x_\ast = \left( G^{(1)}(x_0) \right)' z_\ast \) and \( \| z_\ast \| \leq c_2 \| x_\ast \| . \)

**Claim 10.27.1** \( \forall x \in \Omega_c \cap D_1 \), \( \exists h_0 \in M_c \) such that \( \| x - x_0 - h_0 \| \leq c_1 c_2 \| x - x_0 \| ^2 \) and \( \| x - x_0 \| - c_1 c_2 \| x - x_0 \| ^2 \leq \| h_0 \| \leq \| x - x_0 \| + c_1 c_2 \| x - x_0 \| ^2 . \)

**Proof of claim:** Fix any \( x \in \Omega_c \cap D_1 \). Then, \( G(x) \leq \vartheta_\mathbb{Z} \). It is easy to see that \( M_c \ni \vartheta_\mathbb{X} \) is a nonempty closed convex set. Then, by Proposition 8.15, we have \( 0 \leq \delta := \inf_{x \in M_c} \| x - x_0 - h \| = \max_{x \in M_c, \| x \| \leq 1} \left( \| x - x_0 \| - g(x) \right) \geq 0 \), where \( g : M_{c_{\text{supp}}} \to \mathbb{R} \) is the support functional of \( M_c \), \( M_{c_{\text{supp}}} := \{ x \in \mathbb{X} : \sup_{u \in M_c} \langle x, u \rangle < + \infty \} \), and \( g(x) = \sup_{u \in M_c} \langle x, u \rangle \), \( \forall x \in M_{c_{\text{supp}}} \).

We will show that \( M_{c_{\text{supp}}} \subseteq \left( (G^{(1)}(x_0))_{\text{inv}}(P) \right)^{\oplus} \). Fix any \( x_\ast \in M_{c_{\text{supp}}} \), \( \forall x \in (G^{(1)}(x_0))_{\text{inv}}(P) \), we have \( G^{(1)}(x_0) u \leq \vartheta_\mathbb{Z} \). \( \forall \alpha \in [0, \infty) \subset \mathbb{R} \), it is easy to show that \( -\alpha u \in M_c \). Suppose \( \langle x_\ast, u \rangle < 0 \). Then, \( \sup_{u \in M_c} \langle x_\ast, u \rangle \geq \sup_{\alpha \in [0, \infty) \subset \mathbb{R}} \langle x_\ast, -\alpha u \rangle = + \infty \). This implies that \( x_\ast \notin M_{c_{\text{supp}}} \), which is a contradiction. Hence, we must have \( \langle x_\ast, u \rangle \geq 0 \). By the arbitrariness of \( u \), we have \( x_\ast \in (G^{(1)}(x_0))_{\text{inv}}(P) \). By the arbitrariness of \( x_\ast \), we have \( M_{c_{\text{supp}}} \subseteq \left( (G^{(1)}(x_0))_{\text{inv}}(P) \right)^{\oplus} \).

Then, \( \delta = \langle x_\ast, x - x_0 \rangle - g(x_0) \) for some \( x_\ast \in M_{c_{\text{supp}}} \subseteq \left( (G^{(1)}(x_0))_{\text{inv}}(P) \right)^{\oplus} \) with \( \| x_\ast \| \leq 1 \). By (i) and Proposition 10.26, we have \( x_\ast \in (G^{(1)}(x_0))(P^{\oplus}) \subseteq \mathcal{R} \left( \left( G^{(1)}(x_0) \right)' \right) \). Then, \( \exists z_\ast \in \mathbb{Z}^* \) such that \( x_\ast = \left( G^{(1)}(x_0) \right)' z_\ast \) and \( \| z_\ast \| \leq c_2 \| x_\ast \| \leq c_2 \). This further implies that

\[
\delta = \langle z_\ast, 1, G^{(1)}(x_0)(x - x_0) \rangle - \sup_{u \in M_c} \langle z_\ast, 1, G^{(1)}(x_0)u \rangle
\]

\[
= \langle \langle z_\ast, 1, -G(x) + G(x_0) + G^{(1)}(x_0)(x - x_0) \rangle \rangle
\]

\[
- \sup_{u \in M_c} \langle \langle z_\ast, 1, -G(x) + G(x_0) + G^{(1)}(x_0)u \rangle \rangle
\]

where the second equality follows from Proposition 8.37. By (i), we have \( \mathcal{R} \left( G^{(1)}(x_0) \right) = \mathbb{Z} \) and \( \exists u_1 \in \mathbb{X} \) such that \( G^{(1)}(x_0)u_1 = G(x) - G(x_0) \). It is easy to see that \( u_1 \in M_c \) since \( x \in \Omega_c \) and \( G(x) \leq \vartheta_\mathbb{Z} \). Then, \( \delta \leq \langle \langle z_\ast, 1, -G(x) + G(x_0) + G^{(1)}(x_0)(x - x_0) \rangle \rangle \).

By Taylor’s Theorem 9.48, \( \exists \delta_1 \in (0, 1) \subset \mathbb{R} \) such that \( \| G(x) - G(x_0) - G^{(1)}(x)(x - x_0) \| = \delta_1 \| G^{(2)}(t_0 x + (1 - t_0)x_0) \| \| x - x_0 \| ^2 \leq c_1 \| x - x_0 \| ^2 / 2 \). Then, we have \( \delta \leq \| z_\ast \| \| G(x) - G(x_0) - G^{(1)}(x)(x - x_0) \| \leq c_1 c_2 \| x - x_0 \| ^2 / 2 \). Since \( M_c \) is closed, then \( \exists h_0 \in M_c \) such that \( \| x - x_0 - h_0 \| \leq 2 \delta \leq c_1 c_2 \| x - x_0 \| ^2 \).
Then, \( \| h_0 \| \geq \| x - x_0 \| - \| x - x_0 - h_0 \| \geq \| x - x_0 \| - c_1 c_2 \| x - x_0 \|^2 \); and \( \| h_0 \| \leq \| x - x_0 \| + \| x - x_0 - h_0 \| \leq \| x - x_0 \| + c_1 c_2 \| x - x_0 \|^2 \). This completes the proof of the claim.

Let \( c_3 = \| L(2)(x_0) \| \). Let \( \delta \in (0, \delta_1] \subset \mathbb{R} \) such that \( c_1 c_2 \delta \leq 1/4 \), \( c_1^2 c_2^2 (c_3 + m/5) \delta^2 \leq m/5 \), and \( 5c_1 c_2 (c_3 + m/5) \delta^2/2 \leq m/5 \). \( \forall x \in \Omega \cap B(x, \delta) \), by Claim 10.27.1, \( \exists h_0 \in M \) such that \( \| x - x_0 - h_0 \| \leq c_1 c_2 \| x - x_0 \|^2 \leq \| x - x_0 \| /4 \) and \( 3 \| x - x_0 \| /4 \leq \| h_0 \| \leq 5 \| x - x_0 \| /4 \). By Taylor’s Theorem 9.48, \( \exists \delta_1 \in (0, 1) \subset \mathbb{R} \) such that

\[
f(x) - f(x_0) \geq L(x) - L(x_0) - L(1)(x_0)(x - x_0)
= \frac{1}{2} L(2)(\bar{t}x + (1 - \bar{t})x_0)(x - x_0)(x - x_0)
\geq \frac{1}{2} (L(2)(h_0)(h_0) + L(2)(\bar{x})(x - x_0 - h_0)(h_0)
+ L(2)(\bar{x})(h_0)(x - x_0 - h_0) + L(2)(\bar{x})(x - x_0 - h_0)(x - x_0 - h_0) - \| L(2)(\bar{x}) \| (2 \| x - x_0 - h_0 \| \| h_0 \| + \| x - x_0 - h_0 \|^2))
\geq \frac{1}{2} \left( m \| h_0 \|^2 - m \| h_0 \|^2 /5 - (\| L(2)(\bar{x}) - L(2)(x_0) \| + \| L(2)(x_0) \|)
\cdot (2 \| x - x_0 - h_0 \| \| h_0 \| + \| x - x_0 - h_0 \|^2) \right)
\geq \frac{1}{2} \left( \frac{4m}{5} \| h_0 \|^2 - (c_3 + m/5) \left( \frac{5}{2} c_1 c_2 \| x - x_0 \|^3 + c_1^2 c_2^2 \| x - x_0 \|^4 \right) \right)
\geq \frac{1}{2} \left( \frac{4m}{5} \frac{9}{16} - m \frac{m}{5} \right) \| x - x_0 \|^2 = \frac{m}{40} \| x - x_0 \|^2
\]

Hence, \( x_0 \) is a point of relative strict minimum for \( f \) on the set \( \Omega_c \). This completes the proof of the proposition. \( \square \)

**Proposition 10.28** Let \( X \) be a real Banach space, \( \Omega \subseteq X \), \( x_0 \in \Omega^0 \), \( Z \) be a real Banach space with a positive cone \( P \subseteq Z \), and \( f : \Omega \to \mathbb{R} \) and \( G : \Omega \to Z \) be \( C_2 \) at \( x_0 \). Consider the optimization problem (10.3). Assume that \( x_0 \) is a regular point of \( G \) and a point of relative minimum for \( f \) on the set \( \Omega_c := \{ x \in \Omega \mid G(x) \perp \partial Z \} \). Then, there exists a Lagrange multiplier \( z_{*0} \in P^\circ \) such that

(i) the Lagrangian \( L : \Omega \to \mathbb{R} \) defined by \( L(x) = f(x) + \langle (z_{*0}, G(x)) \rangle \), \( \forall x \in \Omega \), is stationary at \( x_0 \) and \( \langle (z_{*0}, G(x_0)) \rangle = 0 \);

(ii) \( L(2)(x_0) \) is positive semi-definite on the subspace \( N(G^{(1)}(x_0)) = M \), that is, \( G^{(2)}(x_0)(h)(h) \geq 0 \), \( \forall h \in N(G^{(1)}(x_0)) \).

**Proof** By the Generalized Kuhn-Tucker Theorem 10.25, there exists a Lagrange multiplier \( z_{*0} \in Z^\ast \) with \( z_{*0} \geq 0 \) such that (i) holds. We
will show that (ii) also holds by an argument of contradiction. Suppose (ii) does not hold. Then, \( \exists h_0 \in M \) with \( \| h_0 \| = 1 \), such that \( L^{(2)}(x_0)(h_0)(h_0) < -m < 0 \) for some \( m \in (0, \infty) \subset \mathbb{R} \). Let \( z_0 := G(x_0) \in \mathbb{Z} \). By Surjective Mapping Theorem 9.53, \( \exists \tau_1 \in (0, \infty) \subset \mathbb{R} \), \( \delta_1 \in (0, \infty) \subset \mathbb{R} \), and \( \exists_2 \in [0, \infty) \subset \mathbb{R} \) with \( c_1 \delta_1 \leq r_1 \) such that \( \forall z \in B_2 \left( z_0, \delta_1 / 2 \right), \forall \bar{x} \in B_\chi \left( x_0, r_1 / 2 \right) \) with \( \bar{z} = G(\bar{x}), \forall z \in B_2 \left( \bar{z}, \delta_1 / 2 \right), \exists x \in B_\chi(x_0,r_1) \subset \Omega \) with \( \| x - \bar{x} \| \leq c_1 \| z - \bar{z} \| \), we have \( z = G(x) \). By Propositions 9.34, 9.37, 9.40, 9.41, 9.44, and 9.45, \( L \) is \( C_2 \) at \( x_0 \). By (i), \( L^{(1)}(x_0) = \delta_0 B_\chi \in \mathbb{S}_2 \). Then, \( \exists r_2 \in (0, r_1] \subset \mathbb{R} \) such that \( L^{(2)}(x) \) exists and \( \| L^{(2)}(x) - L^{(2)}(x_0) \| < m/5, \forall x \in B_\chi \left( x_0, r_2 \right) \). Since \( G \) is \( C_2 \) at \( x_0 \), then \( \exists c_2 \in [0, \infty) \subset \mathbb{R} \) and \( \exists r_3 \in (0, r_2] \subset \mathbb{R} \) such that \( \| G^{(2)}(x) \| \leq c_2, \forall x \in B_\chi \left( x_0, r_3 \right) \). Let \( c_3 := \| L^{(2)}(x_0) \| \).

\[ \forall \delta \in (0, \delta_3 / 2] \subset \mathbb{R} \] such that \( c_2 \delta^2 < \delta_1, c_1 c_2 \delta^2 / 2 \leq 1 / 4, c_1^2 c_2^2 (c_3 + m/5) \delta^2 / 4 \leq m/5, \text{and} c_1 c_2 (c_3 + m/5) \delta \leq m/5 \). By Taylor’s Theorem 9.48, \( \exists h_0 \in (0, 1) \subset \mathbb{R} \) such that

\[
\| G(x_0 + \delta h_0) - z_0 \| = \| G(x_0 + \delta h_0) - G(x_0) - \delta G^{(1)}(x_0)h_0 \| 
\leq \frac{1}{2} \| G^{(2)}(x_0 + t_0 \delta h_0) \| \delta^2 \leq c_2 \delta^2 / 2 < \delta_1 / 2
\]

Let \( z_\delta := G(x_0 + \delta h_0) \in B_\chi \left( z_0, \delta_1 / 2 \right) \) and \( x_\delta := x_0 + \delta h_0 \in B_\chi \left( x_0, r_1 / 2 \right) \). Note that \( z_\delta = G(x_\delta) \) and \( z_0 = G(x_0) \). Then, \( \exists x_\delta \in B_\chi(x_0,r_1) \) with \( \| x_\delta - x_\delta \| \leq c_1 \| z_0 - z_\delta \| \leq c_1 c_2 \delta^2 / 2 \) such that \( G(x_\delta) = z_\delta \). Then, we have \( \| x_\delta - x_\delta \| \geq \| \delta h_0 \| - \| x_\delta - x_\delta \| \geq \delta - c_1 c_2 \delta^2 / 2 \geq 3 \delta / 4 \); and \( \| x_\delta - x_\delta \| \leq \| \delta h_0 \| + \| x_\delta - x_\delta \| \leq \delta + c_1 c_2 \delta^2 / 2 \leq 5 \delta / 4 < r_3 \). Hence, \( x_\delta \in \Omega \cap B_\chi(x_0, r_3) \). By Taylor’s Theorem 9.48, \( \exists \tilde{f}_1 \in (0, 1) \subset \mathbb{R} \) such that

\[
f(x_\delta) - f(x_0) = L(x_\delta) - L(x_0) = L^{(1)}(x_\delta)(x_\delta - x_0) - \langle z_0, G(x_\delta) \rangle
+ \langle z_0, G(x_\delta) \rangle - L^{(1)}(x_\delta)(x_\delta - x_0)
= \frac{1}{2} L^{(2)}(t_1 x_\delta + (1 - t_1) x_0)(x_\delta - x_0)(x_\delta - x_0)
= \frac{1}{2} \left( L^{(2)}(x)(x_\delta - x_0)(x_\delta - x_0) + L^{(2)}(x)(x_\delta - x_\delta)(x_\delta - x_\delta)ight)
+ \frac{1}{2} \left( \delta^2 L^{(2)}(x_\delta)(h_0)(h_0) + \delta^2 (L^{(2)}(x_\delta) - L^{(2)}(x_\delta))(h_0) \right)
\leq \frac{1}{2} \left( \delta^2 L^{(2)}(x_\delta)(h_0)(h_0) + \delta^2 (L^{(2)}(x_\delta) - L^{(2)}(x_\delta))(h_0) \right)
+ \frac{1}{2} \left( -m\delta^2 + m\delta^2 / 5 + \left( \| L^{(2)}(x_\delta) - L^{(2)}(x_\delta) \| + \| L^{(2)}(x_\delta) \| \right)
\cdot (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4 / 4) \right)
\leq \frac{1}{2} \left( -\frac{4m}{5} \delta^2 + (c_3 + m/5)(c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4 / 4) \right)
\]
Thus, \( \lim_{\delta \to 0^+} x_\delta = x_0 \), \( x_\delta \in \Omega_c \), and \( f(x_\delta) < f(x_0) \), for sufficiently small \( \delta \). Then, \( x_0 \) is not a point of relative minimum for \( f \) on the set \( \Omega_c \). This contradicts with the assumption. Hence, (ii) must hold. This completes the proof of the proposition. \( \square \)

**Proposition 10.29** Let \( X \) be a real reflexive Banach space and \( L \in S_X \).
Then, the following statements holds.

(i) \( L' = L \).

(ii) If \( L \in S_{++} \), that is, \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( L(x)(x) \geq \delta \| x \|_X^2 \),
\[ \forall x \in X, \text{ then, } L \in B(X, X^*) \text{ is bijective, } L^{-1} \in S_{++}, \text{ that is, } \]
\[ L^{-1}(x_*) (x) \geq \frac{\delta}{c + \| L_L^{\#}(x_*) \|_2} \| x_* \|_X^2, \forall x_* \in X^*, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}. \]

**Proof** (i) \( \forall x, y \in X \), we have
\[ L(x)(y) = L(y)(x) = \langle L(x), y \rangle = \langle \phi_X(y), L(x) \rangle = \langle L'(\phi_X(y)), x \rangle \]
where the first equality follows from the fact that \( L \in S_X = B_{S_2}(X, \mathbb{R}) \); and \( \phi_X : X \to X^* \) is the natural mapping introduced in Remark 7.88. Since \( X \) is reflexive, then \( X^{**} \) and \( X \) is isometrically isomorphic and \( \phi_X = \text{id}_X \). This implies that
\[ L(x)(y) = \langle L'(y), x \rangle = L(y)(x) = \langle L(y), x \rangle \]
By the arbitrariness of \( x \) and \( y \), we have \( L = L' \).

(ii) By the assumption, \( L \in B_{S_2}(X, \mathbb{R}) \subseteq B(X, X^*) \). \( \forall x \in X \) with \( x \neq \partial_X \), \( L(x) \neq \partial_X \), since \( L(x)(x) > 0 \). Hence, \( \mathcal{N}(L) = \{ \partial_X \} \) and \( L \) is injective. By (i), \( L = L' \). Then, \( \mathcal{N}'(L') = \{ \partial_X \} \). By Proposition 7.112, \( (\mathcal{R}(L))^\perp = \mathcal{N}'(L') \). This shows that \( (\mathcal{R}(L))^\perp = X^* \). By Proposition 7.98, \( \mathcal{R}(L) = X^* \). Hence, \( \mathcal{R}(L) \) is dense in \( X^* \).

Clearly, \( \partial_X \in \mathcal{R}(L) \). \( \forall x_* \in X^* \) with \( x_* \neq \partial_X \), there exists \( (x_i)_i \subseteq \mathcal{R}(L) \) such that \( x_* = \lim_{i \to \infty} x_i \). Without loss of generality, we may assume that \( x_i \neq \partial_X, \forall i \in \mathbb{N} \). Let \( x_i = L(x_i), \forall i \in \mathbb{N} \), where \( x_i \in X \). Clearly, \( x_i \neq \partial_X \) since \( x_i \neq \partial_X, \forall i \in \mathbb{N} \). Then, we have
\[ \| x_i \|_{X^*} \leq \| x_i \|_X \geq \frac{\delta}{c} \| x_i \|_X \]
topology. Since $X$ is reflexive, the weak* topology on $X^{**} = X$ is identical to the weak topology on $X$, and the weak* topology on $X$ is identical to the weak topology on $X^{**}$. By Proposition 7.123, $L \in B(X, X^{**})$ is a continuous as a mapping $L : X_{\text{weak}} \to X_{\text{weak}}^{**}$. By Proposition 3.65, $(x_{si})_{i=1}^{\infty} = (L(x_i))_{i=1}^{\infty}$ has a cluster point $L(x) \in X_{\text{weak}}^{**}$ in weak* topology. By Proposition 7.120, $X_{\text{weak}}^{**}$ is completely regular, then it is Hausdorff by Proposition 3.61. By Proposition 3.65, $L(x) = x_s$, since $x_s = \lim_{i \in \mathbb{N}} x_{si}$ and $x_s = \lim_{i \in \mathbb{N}} x_{si}$ weak*. This shows that $x_s \in \mathcal{R}(L)$. By the arbitrariness of $x_s$, we have $\mathcal{R}(L) = X^*$ and $L$ is surjective.

Hence, $L$ is bijective. By Open Mapping Theorem 7.103, $L^{-1} \in B(X^*, X) = B(X^*, X^{**}) = B(X^*, B(X^*, \mathbb{R}))$.

\forall x_s, y_s \in X^*, let $x = L^{-1}(x_s)$ and $y = L^{-1}(y_s)$. We have

\[ L^{-1}(x_s)(y_s) = \langle \phi_X(L^{-1}(x_s)), y_s \rangle = \langle y_s, L^{-1}(x_s) \rangle = \langle L(y), x \rangle \]

\[ = L(y)(x) = L(x)(y) = \langle \langle L(x), y \rangle, \langle x_s, L^{-1}(y_s) \rangle \rangle = L^{-1}(y_s)(x_s) \]

where $\phi_X : X \to X^{**}$ is the natural mapping introduced in Remark 7.88; and the fifth equality follows from the fact that $L \in \mathcal{S}_X$. By the arbitrariness of $x_s$ and $y_s$, we have $L^{-1} \in \mathcal{S}_X^{**}$.

\forall x_s \in X^*, let $x = L^{-1}(x_s)$. Then, $x_s = L(x)$ and $\|x_s\|_{X^*} \leq \|L\|_{B(X^*, X^*)} \|x\|_X$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|L\|_{B(X^*, X^*)} = 0$; Case 2: $\|L\|_{B(X^*, X^*)} > 0$.

Case 1: $\|L\|_{B(X^*, X^*)} = 0$, we have $\|x_s\|_{X^*} = 0$, $\forall x_s \in X^*$. Hence, $X^* = \{0\}$. Then, $L^{-1}(x_s)(x_s) = 0 \geq \frac{\delta}{\epsilon + \|L\|_{B(X^*, X^*)}} \|x_s\|_{X^*}^2 = 0, \forall \epsilon \in (0, \infty) \subset \mathbb{R}$.

Case 2: $\|L\|_{B(X^*, X^*)} > 0$. Then, we have

\[ L^{-1}(x_s)(x_s) = L(x)(x) \geq \delta \|x\|_X^2 \geq \frac{\delta}{\|L\|_{B(X^*, X^*)}^2} \|x_s\|_{X^*}^2, \quad \forall \epsilon \in (0, \infty) \subset \mathbb{R} \]

Hence, we have shown, in both cases, that $L^{-1}(x_s)(x_s) \geq \frac{\delta}{\epsilon + \|L\|_{B(X^*, X^*)}^2}, \|x_s\|_{X^*}^2, \forall x_s \in X^*$ and $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$. This proves that $L^{-1} \in \mathcal{S}_X^{**}$.

This completes the proof of the proposition. □
Chapter 11

General Measure and Integration

11.1 Measure Spaces

Definition 11.1 A measurable space \((X, \mathcal{B})\) is a set \(X\) and a \(\sigma\)-algebra \(\mathcal{B}\) (of subsets of \(X\)) on \(X\). A subset \(A \subseteq X\) is said to be \(\mathcal{B}\)-measurable if \(A \in \mathcal{B}\), or simply measurable when no confusion may arise from the context.

Definition 11.2 Let \((X, \mathcal{B})\) be a measurable space. A measure \(\mu\) on \((X, \mathcal{B})\) is a nonnegative extended real-valued function on \(\mathcal{B}\), i.e. a mapping \(\mu : \mathcal{B} \to [0, \infty] \subseteq \mathbb{R}_e\), such that

(i) \(\mu(\emptyset) = 0\);
(ii) \(\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)\), where the sequence \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B}\) are disjoint.

Then, the triple \((X, \mathcal{B}, \mu)\) is called a measure space.

Fact 11.3 Let \((X, \mathcal{B}, \mu)\) be a measure space. Then, \(\mu\) is finitely additive, i.e., \(\mu\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \mu(E_i)\), where \(n \in \mathbb{Z}_+\) and \((E_i)_{i=1}^{n} \subseteq \mathcal{B}\) is pairwise disjoint.

Proof Consider the sequence of measurable sets \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B}\), with \(E_i = \emptyset, i = n+1, n+2, \ldots\) Clearly, the sets in \((E_i)_{i=1}^{\infty}\) are pairwise disjoint. Therefore, \(\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu(E_i)\). This completes the proof of the fact.

Proposition 11.4 Let \((X, \mathcal{B}, \mu)\) be a measure space. \(\forall A, B \in \mathcal{B}\) such that \(A \subseteq B\), then, \(\mu(A) \leq \mu(B)\).
Proof. The set \( B \cap \tilde{A} \in \mathcal{B} \) since \( \mathcal{B} \) is a \( \sigma \)-algebra. Note that \( A \cap \left( B \cap \tilde{A} \right) = \emptyset \), then, \( \mu(B) = \mu \left( A \cup \left( B \cap \tilde{A} \right) \right) = \mu(A) + \mu \left( B \cap \tilde{A} \right) \geq \mu(A) \), by Fact 11.3. This completes the proof of the proposition. \( \square \)

Proposition 11.5 Let \((X, \mathcal{B}, \mu)\) be a measure space. A sequence \((E_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) is such that \( E_n \supseteq E_{n+1}, \ n = 1, 2, \ldots \). If \( \mu(E_1) < +\infty \), then, \( \lim_{n \to \infty} \mu \left( E_{n+1} \right) = \mu \left( \bigcap_{n=1}^{\infty} E_n \right) \).

Proof. Let \( E = \bigcap_{n=1}^{\infty} E_n \in \mathcal{B}. \) Then, \( \tilde{E} = \bigcup_{n=1}^{\infty} \tilde{E}_n \in \mathcal{B}. \) Note that

\[
\mu \left( E_1 \right) = \mu(E) + \sum_{n=1}^{\infty} \mu \left( E_n \cap \tilde{E}_{n+1} \right) \tag{11.1}
\]

On the other hand, note that \( E_n = E_{n+1} \cup \left( E_n \cap \tilde{E}_{n+1} \right), \ n = 1, 2, \ldots \), which implies that \( \mu(E_n) = \mu(E_{n+1}) + \mu \left( E_n \cap \tilde{E}_{n+1} \right) \), since the two sets on the right-hand-side of the above equality are disjoint, and all three numbers are finite by Proposition 11.4.

Then, substituting the above equality into (11.1) yields

\[
\mu \left( E_1 \right) = \mu(E) + \sum_{n=1}^{\infty} \left( \mu(E_n) - \mu \left( E_{n+1} \right) \right) = \mu(E) + \mu \left( E_1 \right) - \lim_{n \to \infty} \mu \left( E_{n+1} \right).
\]

Since \( \mu \left( E_1 \right) < +\infty \), then \( \mu(E) < +\infty \) by Proposition 11.4 and \( E \subseteq E_1 \). Therefore, we have \( \lim_{n \to \infty} \mu \left( E_n \right) = \mu(E) \).

This completes the proof of the proposition. \( \square \)

Proposition 11.6 Let \((X, \mathcal{B}, \mu)\) be a measure space. A sequence \((E_n)_{n=1}^{\infty} \subseteq \mathcal{B}. \) Then, \( \mu \left( \bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \mu \left( E_n \right). \)

Proof. Let \( G_1 = E_1, \ G_n = E_n \setminus \left( \bigcup_{i=1}^{n-1} E_i \right), \ n = 2, 3, \ldots \) Then, \((G_n)_{n=1}^{\infty} \) is a sequence of disjoint and measurable subsets of \( X. \) Therefore, \( \mu \left( \bigcup_{i=1}^{\infty} G_i \right) = \sum_{i=1}^{\infty} \mu \left( G_i \right). \) Note that \( \bigcup_{i=1}^{n} G_i = \bigcup_{i=1}^{n} E_i, \ n = 1, 2, \ldots \) Therefore, \( \bigcup_{i=1}^{\infty} G_i = \bigcup_{i=1}^{\infty} E_i. \) Thus,

\[
\mu \left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mu \left( G_i \right) \tag{11.2}
\]

By Proposition 11.4, \( G_n \subseteq E_n \) implies that \( \mu(G_n) \leq \mu(E_n), \ n = 1, 2, \ldots \). Applying these inequalities in the equality (11.2), yields \( \mu \left( \bigcup_{i=1}^{\infty} E_i \right) \leq \sum_{i=1}^{\infty} \mu \left( E_i \right). \)

This completes the proof of the proposition. \( \square \)

Proposition 11.7 Let \((X, \mathcal{B}, \mu)\) be a measure space and \((A_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) with \( A_i \subseteq A_{i+1}, \ \forall i \in \mathbb{N}. \) Then, \( \mu \left( \bigcup_{i=1}^{\infty} A_i \right) = \lim_{i \to \infty} \mu(A_i) \in \mathbb{R}. \)
11.1. MEASURE SPACES

Proof Let $A := \bigcup_{i=1}^{\infty} A_i \in B$. Note that $A = A_1 \cup (\bigcup_{i=1}^{\infty} (A_{i+1} \setminus A_i))$, where the sets in the union are pairwise disjoint. Then, by countable additivity of measure, $\mu(A) = \mu(A_1) + \sum_{i=1}^{\infty} \mu(A_{i+1} \setminus A_i) = \lim_{n \to \infty} (\mu(A_1) + \sum_{i=1}^{n} \mu(A_{i+1} \setminus A_i)) = \lim_{n \to \infty} \mu(A_{n+1}) = \lim_{i \to \infty} \mu(A_i)$, where the third equality follows from Fact 11.3. This completes the proof of the proposition.

Proposition 11.8 Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, $\forall E_1, E_2 \in \mathcal{B}$ with $\mu(E_1 \triangle E_2) = 0$, we have $\mu(E_1) = \mu(E_2)$.

Proof $0 = \mu(E_1 \triangle E_2) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1)$. Then, $\mu(E_1 \setminus E_2) = \mu(E_1 \setminus E_1) = 0$. Hence, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_2 \cap E_1) = \mu(E_2)$. This completes the proof of the proposition.

Lemma 11.9 (Borel-Cantelli) Let $(X, \mathcal{B}, \mu)$ be a measure space and $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$. Assume that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then, $\mu(G) := \mu(\bigcap_{n \in \mathbb{N}} \bigcup_{m=n}^{\infty} E_m) = 0$.

Proof $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N}$ such that $\sum_{n=N}^{\infty} \mu(E_n) < \epsilon$. Then, $0 \leq \mu(G) \leq \mu(\bigcup_{n=N}^{\infty} E_m) \leq \sum_{n=N}^{\infty} \mu(E_m) < \epsilon$, where the second inequality follows from Proposition 11.4; and the third inequality follows from Proposition 11.6. By the arbitrariness of $\epsilon$, we have $\mu(G) = 0$. This completes the proof of the lemma.

The set $\bigcap_{n \in \mathbb{N}} \bigcup_{m=n}^{\infty} E_m$ is equal to the set of elements that is in $E_n$, $n \in \mathbb{N}$, infinitely often. The book Williams (1991) uses the notation $\limsup_{n \in \mathbb{N}} E_n$ to denote this set. The book Williams (1991) also uses the notation $\liminf_{n \in \mathbb{N}} E_n$ to denote the set $\bigcup_{n \in \mathbb{N}} \bigcap_{m=n}^{\infty} E_m$, which is equal to the set of elements that is in all of $E_m$, $m \in \mathbb{N}$ with $n \leq m$, eventually. We will not formally use these notations. But, these sets are very important in the study of probability measure spaces.

Definition 11.10 Let $(X, \mathcal{B}, \mu)$ be a measure space. $\mu$ is said to be finite if $\mu(X) < +\infty$. $\mu$ is said to be $\sigma$-finite if there exists a sequence of measurable subsets of $X$, $(X_n)_{n=1}^{\infty}$, such that $\bigcap_{n=1}^{\infty} X_n = X$ and $\mu(X_n) < +\infty$, $n = 1, 2, \ldots$.

A set $E \subseteq X$ is said to be of finite measure, if $E \in \mathcal{B}$, and $\mu(E) < +\infty$. A set $E \subseteq X$ is said to be of $\sigma$-finite measure, if there exists a sequence of measurable subsets of $X$, $(E_n)_{n=1}^{\infty}$, such that $\bigcup_{n=1}^{\infty} E_n = E$ and $\mu(E_n) < +\infty$, $n = 1, 2, \ldots$.

Definition 11.11 Let $(X, \mathcal{B}, \mu)$ be a measure space. It is said to be complete if $\forall A, E \subseteq X$, with $A \subseteq E$, $E \in \mathcal{B}$, and $\mu(E) = 0$, then, $A \in \mathcal{B}$.

Proposition 11.12 Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, there is a unique complete measure space $(X, \mathcal{B}_0, \mu_0)$ such that
(i) $\mathcal{B} \subseteq \mathcal{B}_0$;

(ii) $\forall E \in \mathcal{B}, \mu(E) = \mu_0(E)$;

(iii) $\forall E \subseteq X, E \in \mathcal{B}_0$ if, and only if, $E = A \cup B$, where $A, B \subseteq X$ with $B \in \mathcal{B}$, and there exists a $C \in \mathcal{B}$, such that $A \subseteq C$ and $\mu(C) = 0$.

Then, $(X, \mathcal{B}_0, \mu_0)$ is said to be the completion of $(X, \mathcal{B}, \mu)$.

Furthermore, let $(X, \mathcal{B}_1, \mu_1)$ be another complete measure space that satisfies (i) and (ii), then, $\mathcal{B}_0 \subseteq \mathcal{B}_1$ and $\mu_0(E) = \mu_1(E)$, $\forall E \in \mathcal{B}_0$.

**Proof**

Define $\mathcal{B}_0$ to be the collection of subsets of $X$ as the following.

$$\mathcal{B}_0 = \{ E \subseteq X \mid \text{there exists } A, B, C \subseteq X \text{ such that } E = A \cup B, B, C \in \mathcal{B}, \mu(C) = 0, \text{ and } A \subseteq C \}$$

$\forall E \in \mathcal{B}$, let $A = \emptyset, B = E$, and $C = \emptyset$. Then, $E = A \cup B, B, C \in \mathcal{B}, \mu(C) = 0$, and $A \subseteq C$. Therefore, $E \in \mathcal{B}_0$. Hence, $\mathcal{B} \subseteq \mathcal{B}_0$.

The collection $\mathcal{B}_0$ form a $\sigma$-algebra on $X$, which is proved in the following.

$\forall E_1, E_2 \in \mathcal{B}_0$, there exist $A_1, B_1, C_1, A_2, B_2, C_2 \subseteq X$ such that $E_1 = A_1 \cup B_1, E_2 = A_2 \cup B_2, B_1, C_1, B_2, C_2 \in \mathcal{B}, \mu(C_1) = \mu(C_2) = 0$, and $A_1 \subseteq C_1, A_2 \subseteq C_2$. Then, $E_1 \cup E_2 = (A_1 \cup A_2) \cup (B_1 \cup B_2)$. The sets $B_1 \cup B_2, C_1 \cup C_2 \subseteq \mathcal{B}$, and $A_1 \cup A_2 \subseteq C_1 \cup C_2$. The sequence of inequalities holds, $0 \leq \mu(C_1 \cup C_2) \leq \mu(C_1) + \mu(C_2) \leq 0$, where the first one follows from Proposition 11.4, and the second one follows from Proposition 11.6. This further implies that $\mu(C_1 \cup C_2) = 0$. By the definition of $\mathcal{B}_0$, $E_1 \cup E_2 \in \mathcal{B}_0$.

Note that $\widehat{E}_1 = \widehat{B}_1 \cap \widehat{A}_1 = \widehat{B}_1 \cap \left( \left( \widehat{A}_1 \setminus \widehat{C}_1 \right) \cup \widehat{C}_1 \right)$, since $\widehat{C}_1 \subseteq \widehat{A}_1$.

This implies that $\widehat{E}_1 = \left( \widehat{B}_1 \cap \left( \widehat{A}_1 \setminus \widehat{C}_1 \right) \right) \cup \left( \widehat{B}_1 \cap \widehat{C}_1 \right)$. Since $\widehat{B}_1 \cap \widehat{C}_1 \in \mathcal{B}$ and $\widehat{B}_1 \cap \left( \widehat{A}_1 \setminus \widehat{C}_1 \right) = \widehat{B}_1 \cap \widehat{A}_1 \cap \widehat{C}_1 \subseteq C_1$ and $\mu(C_1) = 0$, then, $\widehat{E}_1 \in \mathcal{B}_0$.

Therefore, $\mathcal{B}_0$ form an algebra on the set $X$.

Consider an arbitrary sequence $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_0$. Then, there exists $A_n, B_n, C_n \subseteq X$, such that $E_n = A_n \cup B_n, B_n, C_n \in \mathcal{B}, A_n \subseteq C_n$, and $\mu(C_n) = 0, n = 1, 2, \ldots$. Then, $\bigcup_{n=1}^{\infty} E_n = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n)$. Since $\mathcal{B}$ is a $\sigma$-algebra, then, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$ and $\bigcup_{n=1}^{\infty} C_n \in \mathcal{B}$. Furthermore, $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n$ and $0 \leq \mu(\bigcup_{n=1}^{\infty} C_n) \leq \sum_{n=1}^{\infty} \mu(C_n) = 0$. Then, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_0$.

This proves that $\mathcal{B}_0$ is a $\sigma$-algebra.

Define a mapping $\mu_0 : \mathcal{B}_0 \to [0, \infty] \subseteq \mathbb{R}$ as following. $\forall E \in \mathcal{B}_0$, there exists $A, B, C \subseteq X$ such that $E = A \cup B, B, C \in \mathcal{B}, A \subseteq C$, and $\mu(C) = 0$. $\mu_0(E) := \mu(B)$.

The mapping $\mu_0$ is well defined because of the following. $\forall E \in \mathcal{B}_0$, let $A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X$ be such that,

$$E = A_1 \cup B_1 = A_2 \cup B_2$$
11.1. MEASURE SPACES

\[ B_1, C_1, B_2, C_2 \in \mathcal{B} \]
\[ \mu(C_1) = \mu(C_2) = 0 \quad A_1 \subseteq C_1 \quad A_2 \subseteq C_2 \]

Then, \( \mu(B_1) = \mu \left( \left( B_1 \cap \tilde{B}_2 \right) \cup (B_1 \cap B_2) \right) = \mu \left( B_1 \cap \tilde{B}_2 \right) + \mu \left( B_1 \cap B_2 \right) \leq \mu \left( B_1 \cap \tilde{B}_2 \right) + \mu(B_2) \), where the last inequality follows from Proposition 11.4.

Since \( A_1 \cup B_1 = A_2 \cup B_2 = E \) implies that \( \tilde{B}_2 \cap (A_1 \cup B_1) = \tilde{B}_2 \cap (A_2 \cup B_2) = \tilde{B}_2 \cap A_2 \subseteq C_2 \), then \( \tilde{B}_2 \cap B_1 \subseteq \left( \tilde{B}_2 \cap B_1 \right) \cup \left( \tilde{B}_2 \cap A_1 \right) \subseteq C_2 \). By Proposition 11.4, this implies that \( \mu \left( \tilde{B}_2 \cap B_1 \right) = 0 \). Therefore, \( \mu(B_1) \leq \mu(B_2) \).

By an argument similar to the above, it can be concluded that \( \mu(B_2) \leq \mu(B_1) \). Combining the two inequalities, yields \( \mu(B_2) = \mu(B_1) \).

This shows that \( \mu_0(E) \) is well defined.

\[ \forall E \in \mathcal{B} \subseteq \mathcal{B}_0, \ E = E \cup \emptyset, \text{ then, } \mu_0(E) = \mu(E). \] This proves that \( \mu_0 \) agrees with \( \mu \) on \( \mathcal{B} \).

In the following, it will be shown that the nonnegative, extended real-valued function \( \mu_0 \) defines a measure on the measurable space \((X, \mathcal{B}_0)\). Since \( \emptyset \in \mathcal{B} \), then \( \mu_0(\emptyset) = \mu(\emptyset) = 0 \).

Consider any sequence of subsets \( (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_0 \), which are pairwise disjoint. There exists \( A_n, B_n, C_n \subseteq X \), such that \( E_n = A_n \cup B_n, B_n, C_n \in \mathcal{B}, A_n \subseteq C_n \), and \( \mu(C_n) = 0, n = 1, 2, \ldots \). Since \( B_n \subseteq E_n, n = 1, 2, \ldots \), then \( (B_n)_{n=1}^{\infty} \) are pairwise disjoint. Hence, \( \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \).

Since \( \bigcup_{n=1}^{\infty} E_n = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n) \), \( \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n, 0 \leq \mu(\bigcup_{n=1}^{\infty} C_n) \leq \sum_{n=1}^{\infty} \mu(C_n) = 0 \), and \( \bigcup_{n=1}^{\infty} B_n, \bigcup_{n=1}^{\infty} C_n \in \mathcal{B} \) then, \( \mu_0(\bigcup_{n=1}^{\infty} E_n) = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \mu_0(E_n) \). This proves that \( \mu_0 \) is countably additive for disjoint \( \mathcal{B}_0 \)-measurable sets.

Hence, \( \mu_0 \) is a measure on \((X, \mathcal{B}_0)\).

In the following, it is proved that \((X, \mathcal{B}_0, \mu_0)\) is a complete measure space.

\[ \forall E_0 \subseteq X, \text{ such that there exists } E \in \mathcal{B}_0 \text{ with } E_0 \subseteq E \text{ and } \mu_0(E) = 0. \]

Then, there exists \( A, B, C \subseteq X \) such that \( E = A \cup B, B \subseteq C \), and \( \mu(C) = 0 \), by the definition of \( \mathcal{B}_0 \) and \( E \in \mathcal{B}_0 \). By the definition of \( \mu_0 \), we have \( \mu_0(E) = \mu(B) = 0 \). Hence, \( E_0 = E_0 \cup \emptyset \) and \( E_0 \subseteq A \cup B \subseteq C \cup B \).

Since \( \emptyset \subseteq C \cup B \subseteq \mathcal{B} \), and \( 0 \leq \mu(C \cup B) \leq \mu(C) + \mu(B) = 0 \), then \( E_0 \in \mathcal{B}_0 \).

This proves that \((X, \mathcal{B}_0, \mu_0)\) is complete.

Let \((X, \mathcal{B}_1, \mu_1)\) be any complete measure space that satisfies the (i) and (ii). It is shown in the following that \( \mathcal{B}_0 \subseteq \mathcal{B}_1 \) by an argument of contradiction. Suppose there exist a set \( E \in \mathcal{B}_0 \setminus \mathcal{B}_1 \), then \( E \notin \mathcal{B}_1 \). Since \( E \in \mathcal{B}_0 \), then, there exists \( A, B, C \subseteq X \), such that \( E = A \cup B, A \subseteq C, B, C \in \mathcal{B}, \text{ and } \mu(C) = 0 \). This implies \( B, C \in \mathcal{B}_1 \) and \( \mu_1(C) = \mu(C) = 0 \).

Hence, \( A \notin \mathcal{B}_1 \), otherwise, we have \( E \in \mathcal{B}_1 \). But, \( A \subseteq C, \text{ with } \mu_1(C) = 0 \), which contradicts the fact that \((X, \mathcal{B}_1, \mu_1)\) being complete.

Therefore, we have \( \mathcal{B}_0 \subseteq \mathcal{B}_1 \).
∀E ∈ B₀, there exists A, B, C ⊆ X, such that E = A ∪ B, A ⊆ C, B, C ∈ B, and µ(C) = 0. Then, µ₀(E) = µ(B) = µ₁(B) ≤ µ₁(E) ≤ µ₁(B ∪ C) ≤ µ₁(B) + µ₁(C) = µ(B), where Propositions 11.4 and 11.6 have been applied. Therefore, µ₀ and µ₁ agrees on the σ-algebra B₀.

Hence, (X, B₀, µ₀) is a complete measure space that satisfies (i), (ii), and (iii). It is the unique such space since B₀ is unique by (iii), then µ₀ is also unique since any complete measure space (X, B₀, µ₁) satisfying (i) and (ii) will be such that µ₁(E) = µ₀(E), ∀E ∈ B₀.

This completes the proof of the proposition.

Proposition 11.13 Let X := (X, B, µ) be a measure space, A ∈ B, B ⊆ B be a σ-algebra on X, and Bₐ := {C ⊆ A | C ∈ B}. Then, the following statements hold.

(i) (X, B, µ|B) and A := (A, Bₐ, µ|Bₐ) are measure spaces. The measure space A is said to be the subspace of the measure space X.

(ii) If, in addition, X is complete, then A is also complete.

(iii) Let X := (X, B, µ) be the completion of X as defined in Proposition 11.12, and A := (A, Bₐ, µ|A) be the measurable subspace of X. Then, A is the completion of A.

Proof (i) Since B is a σ-algebra on X, then (X, B) is a measurable space. ∀B ∈ B ⊆ B, 0 ≤ µ(B) ≤ +∞ is well-defined. Then, µ|B : B → [0, ∞] ⊆ ℝ₊. Clearly, µ|B (∅) = 0. ∀ pairwise disjoint sequence (Eᵢ)ᵢ=1∞ ⊆ B ⊆ B, we have E := \[ \bigcup_{i=1}^{∞} Eᵢ \] ∈ B and µ|B (E) = \[ \sum_{i=1}^{∞} µ(Eᵢ) \] = \[ \sum_{i=1}^{∞} |B(Eᵢ) \]. Hence, (X, B, µ|B) is a measure space.

Note that ∅ ∈ Bₐ and A ∈ Bₐ, ∀Eₐ ∈ Bₐ, then A ⊇ A \ Eₐ ∈ B and A \ Eₐ ∈ B. ∀(Eₐ)ᵢ=1∞ ⊆ Bₐ, ∀i ∈ N, A ⊇ Eₐᵢ ∈ B. Then, A ⊇ \[ \bigcup_{i=1}^{∞} Eₐᵢ \] ∈ B and \[ \bigcup_{i=1}^{∞} Eₐᵢ \] ∈ Bₐ. Hence, Bₐ is a σ-algebra on A. It is straightforward to show that µ|Bₐ is a measure on the measurable space (A, Bₐ). Hence (A, Bₐ, µ|Bₐ) is a measure space.

(ii) If, in addition, (X, B, µ) is complete, ∀Eₐ ⊆ A such that there exists B ∈ B with Eₐ ⊆ B and µ|Bₐ (B) = 0. Then, B ∈ B and µ(B) = 0. Then, Eₐ ∈ B by the completeness of (X, B, µ). Thus, Eₐ ∈ Bₐ. By the arbitrariness of Eₐ, (A, Bₐ, µ|Bₐ) is complete.

(iii) Note that Bₐ := {C ⊆ A | C ∈ B} and µₐ = µ|Bₐ. By (ii), A is a complete measure space. Since B ⊆ B, then Bₐ ⊆ Bₐ. ∀E ∈ Bₐ, µ|Bₐ (E) = µ(E)|B = µₐ (E)|B = µₐ (E). ∀E ⊆ A, assume that E = Eₐ ∪ E₉ with E₉, E₉ ∈ Bₐ, Eₐ ⊆ E₉, and µ|Bₐ (E₉) = 0. Then, µₐ (E₉) = 0 and E₉ ∈ Bₐ since A is complete. Then, E ∈ Bₐ. ∀E ⊆ A, assume that E ∈ Bₐ. Then, E ⊆ A and E ∈ B. By Proposition 11.12, E = Eₐ ∪ E₉ with E₉, E₉ ∈ Bₐ, Eₐ ⊆ E₉, and µ(E₉) = 0. Then, E₉, E₉ ∩ A ∈ Bₐ, E₉ ⊆ E₉ ∩ A, and 0 ≤ µ|Bₐ (E₉ ∩ A) = µ(E₉ ∩ A) ≤ µ(E₉) = 0. Hence, (i) — (iii) of Proposition 11.12 are satisfied and A is the completion of A.

This completes the proof of the proposition. □
11.2 Outer Measure and the Extension Theorem

**Definition 11.14** Let \( X \) be a set. An outer measure is a function \( \mu_o : X^2 \to [0, \infty] \subset \mathbb{R}_e \) satisfying

(i) \( \mu_o(\emptyset) = 0; \)

(ii) if \( A \subseteq B \subseteq X \), then \( \mu_o(A) \leq \mu_o(B) \); \hspace{1cm} (monotonicity)

(iii) if \( E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq X \), then \( \mu_o(E) \leq \sum_{i=1}^{\infty} \mu_o(E_i) \). \hspace{1cm} (countable subadditivity)

Note that (i) and (iii) above implies finite subadditivity.

**Definition 11.15** Let \( X \) be a set and \( \mu_o : X^2 \to [0, \infty] \subset \mathbb{R}_e \) be an outer measure. \( E \subseteq X \) is said to be measurable with respect to \( \mu_o \) if \( \forall A \subseteq X \), we have \( \mu_o(A) = \mu_o(A \cap E) + \mu_o(A \cap \bar{E}) \).

**Lemma 11.16** Let \( X \) be a set and \( \mu_o : X^2 \to [0, \infty] \subset \mathbb{R}_e \) be an outer measure. \( \forall E \subseteq X \) with \( \mu_o(E) = 0 \), then \( E \) is measurable with respect to \( \mu_o \).

**Proof** Fix any \( E \subseteq X \) with \( \mu_o(E) = 0 \). \( \forall A \subseteq X \), \( 0 \leq \mu_o(A \cap E) \leq \mu_o(E) = 0 \) and \( \mu_o(A \cap \bar{E}) \leq \mu_o(A) \). Then, we have \( \mu_o(A) \leq \mu_o(A \cap E) + \mu_o(A \cap \bar{E}) \leq \mu_o(A) \). Hence, \( E \) is measurable with respect to \( \mu_o \). This completes the proof of the lemma.

**Theorem 11.17** Let \( X \) be a set, \( \mu_o : X^2 \to [0, \infty] \subset \mathbb{R}_e \) be an outer measure, \( \mathcal{B} := \{ E \subseteq X \mid E \text{ is measurable with respect to } \mu_o \} \), and \( \bar{\mu}_o := \mu_o|_\mathcal{B} : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_e \). Then, \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \) and \( (X, \mathcal{B}, \bar{\mu}_o) \) is a complete measure space.

**Proof** Clearly \( \emptyset, X \subseteq \mathcal{B} \). \( \forall E \in \mathcal{B} \), we have \( \bar{E} \in \mathcal{B} \). \( \forall E_1, E_2 \in \mathcal{B} \), \( \forall A \subseteq X \), by the measurability of \( E_1 \) and \( E_2 \), we have \( \mu_o(A) = \mu_o(A \cap E_1) + \mu_o(A \cap E_2) + \mu_o(A \cap \bar{E}_1 \cap \bar{E}_2) + \mu_o(A \cap \bar{E}_2 \cap \bar{E}_1) + \mu_o(A \cap \bar{E}_1 \cap E_2) + \mu_o(A \cap \bar{E}_2 \cap E_1) \). Note that \( \mu_o(A \cap E_1 \cup E_2) = \mu_o(A \cap E_1) + \mu_o(A \cap E_2) + \mu_o(A \cap E_1 \cap E_2) + \mu_o(A \cap E_1 \cap \bar{E}_2) + \mu_o(A \cap \bar{E}_1 \cap E_2) + \mu_o(A \cap \bar{E}_1 \cap \bar{E}_2) \). Then, we have \( \mu_o(A) \geq \mu_o(A \cap (E_1 \cup E_2)) + \mu_o(A \cap \bar{E}_1 \cap \bar{E}_2) + \mu_o(A \cap \bar{E}_1 \cap \bar{E}_2) \). Hence, \( \mathcal{B} \) is an algebra on \( X \).

\( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \). Let \( G_n := \bigcup_{i=1}^{n} E_i \in \mathcal{B} \) and \( \bar{E}_1 := G_1 \in \mathcal{B} \), \( \bar{E}_{n+1} := G_{n+1} \setminus \bar{G}_n \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \bar{E}_1, \bar{E}_2, \ldots \) are pairwise disjoint. Let \( \bar{E} := \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} \bar{E}_i \), \( \forall A \subseteq X \), we have \( \mu_o(A) = \mu_o(A \cap \bar{G}_n) + \mu_o(A \cap \bar{E}_n) \). Then, \( \mu_o(A \cap G_n) = \mu_o(A \cap G_{n+1} \cap \bar{E}_n + 1) \). Hence, we have \( \mu_o(A \cap G_n) = \sum_{i=1}^{n} \mu_o(A \cap \bar{E}_i) \) and \( \mu_o(A) = \sum_{i=1}^{\infty} \mu_o(A \cap \bar{E}_i) \).
\( \forall n \in \mathbb{N} \). Therefore, \( \mu_o(A) \geq \sum_{i=1}^{\infty} \mu_o(A \cap E_i) + \mu_o(A \cap \overline{E}) \). Note that \( A \cap E = \bigcup_{i=1}^{\infty}(A \cap E_i) \), which implies that \( \mu_o(A \cap E) \leq \sum_{i=1}^{\infty} \mu_o(A \cap E_i) \), by countable subadditivity of outer measures. Then, we have \( \mu_o(A) \geq \sum_{i=1}^{\infty} \mu_o(A \cap E_i) + \mu_o(A \cap \overline{E}) \geq \mu_o(A \cap E) + \mu_o(A \cap \overline{E}) \geq \mu_o(A) \). Hence, by the arbitrariness of \( A, E \in \mathcal{B} \). This implies that \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \).

Set \( A = E \) in the above, we have \( \bar{\mu}(E) = \mu_o(E) = \sum_{i=1}^{\infty} \mu_o(E \cap E_i) = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \). Hence, \( (X, \mathcal{B}, \bar{\mu}) \) is a measure space.

Fix any \( A, E \subseteq X \) with \( A \subseteq E, E \in \mathcal{B} \), and \( \bar{\mu}(E) = 0 \). Then, \( 0 \leq \mu_o(A) \leq \mu_o(E) = \bar{\mu}(E) = 0 \). By Lemma 11.16, \( A \in \mathcal{B} \). Therefore, \( (X, \mathcal{B}, \bar{\mu}) \) is complete.

This completes the proof of the theorem. \( \square \)

**Definition 11.18** Let \( X \) be a set and \( \mathcal{A} \subseteq \mathcal{X} \) be an algebra on \( X \). A measure on algebra \( \mathcal{A} \) is a function \( \mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_{+} \) that satisfies:

(i) \( \mu(\emptyset) = 0 \);

(ii) \( \forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) with \( A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \) and \( A_1, A_2, \ldots \) being pairwise disjoint, then \( \mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \).

\( \mu \) is said to be finite if \( \mu(X) < +\infty \); and it is said to be \( \sigma \)-finite if \( \exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) such that \( X = \bigcup_{i=1}^{\infty} X_i \) and \( \mu(X_i) < \infty, \forall i \in \mathbb{N} \).

Clearly, a measure on an algebra is monotonic.

**Theorem 11.19 (Carathéodory Extension Theorem)** Let \( X \) be a set, \( \mathcal{A} \subseteq \mathcal{X} \) be an algebra on \( X \), and \( \mu : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R}_{+} \) be a measure on algebra \( \mathcal{A} \). Define \( \mu_o : \mathcal{X} \rightarrow [0, \infty] \subset \mathbb{R}_{+} \) by

\[
\mu_o(A) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, A \subseteq \bigcup_{i=1}^{\infty} A_i, A_i \neq \emptyset} \sum_{i=1}^{\infty} \mu(A_i), \quad \forall A \subseteq X
\]

Then, the following statements hold.

(i) \( \mu_o \) is an outer measure on \( X \) and is said to be the outer measure induced by \( \mu \).

(ii) Let \( \mathcal{B} := \{ E \subseteq X \mid E \text{ is measurable with respect to } \mu_o \} \) and \( \bar{\mu} := \mu_o|_\mathcal{B} : \mathcal{B} \rightarrow [0, \infty] \subset \mathbb{R}_{+} \). Then, \( (X, \mathcal{B}, \bar{\mu}) \) is a complete measure space and \( \bar{\mu}(A) = \mu(A), \forall A \in \mathcal{A} \). Hence, \( \bar{\mu} \) is an extension of \( \mu \) to the \( \sigma \)-algebra \( \mathcal{B} \supseteq \mathcal{A} \).

(iii) \( \bar{\mu} \) is finite if \( \mu \) is finite; and \( \bar{\mu} \) is \( \sigma \)-finite if \( \mu \) is \( \sigma \)-finite.

(iv) Let \( \mathcal{B}_o \) be the \( \sigma \)-algebra on \( X \) generated by \( \mathcal{A} \). If \( \mu \) is \( \sigma \)-finite, then \( \bar{\mu}|_{\mathcal{B}_o} \) is the unique measure on the measurable space \( (X, \mathcal{B}_o) \) that is an extension of \( \mu \).
(v) If $\mu$ is $\sigma$-finite, then the measure space $(X, \mathcal{B}, \tilde{\mu})$ is the completion of the measure space $(X, \mathcal{B}_0, \tilde{\mu}|_{\mathcal{B}_0})$.

**Proof** We need the following intermediate results.

**Claim 11.19.1** Let $A \in \mathcal{A}$, $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$, and $A \subseteq \bigcup_{i=1}^{\infty} A_i$. Then,

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

**Proof of claim:** Let $G_1 := A \cap A_1 \in \mathcal{A}$ and $G_{n+1} := A \cap A_{n+1} \cap \left(\bigcap_{i=1}^{n} A_i\right) \in \mathcal{A}$, $\forall n \in \mathbb{N}$. Then, $A = \bigcup_{i=1}^{\infty} G_i$ and $G_1, G_2, \ldots$ are pairwise disjoint. This implies that $\mu(A) = \sum_{i=1}^{\infty} \mu(G_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$, where the inequality follows from the monotonicity of $\mu$. This completes the proof of the claim.

Hence, $\forall A \in \mathcal{A}$, we have $\mu_o(A) = \mu(A)$.

**Claim 11.19.2** $\mu_o$ is an outer measure.

**Proof of claim:** Clearly, $\mu_o(\emptyset) = \mu(\emptyset) = 0$ since $\emptyset \in \mathcal{A}$. $\forall A_1 \subseteq A_2 \subseteq X$, by the definition of $\mu_o$, we have $\mu_o(A_1) \leq \mu_o(A_2)$. We need only to show the countable subadditivity for $\mu_o$. Fix any $E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq X$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu_o(E_i) < \infty$, $\forall i \in \mathbb{N}$; Case 2: $\exists i_0 \in \mathbb{N}$ such that $\mu_o(E_{i_0}) = \infty$.

Case 1: $\mu_o(E_i) < \infty$, $\forall i \in \mathbb{N}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists (A_{i,j})_{j=1}^{\infty} \subseteq \mathcal{A}$ such that $E_i \subseteq \bigcup_{j=1}^{\infty} A_{i,j}$ and $\mu_o(E_i) \leq \sum_{j=1}^{\infty} \mu(A_{i,j}) < \mu_o(E_i) + \epsilon/2^i$. Then, $E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} A_{i,j}$. By the definition of $\mu_o$, we have $\mu_o(E) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu(A_{i,j}) \leq \sum_{i=1}^{\infty} \mu_o(E_i) + \epsilon$.

In both cases, we have shown that $\mu_o(E) \leq \sum_{i=1}^{\infty} \mu_o(E_i)$. Hence, $\mu_o$ is an outer measure. This completes the proof of the claim.

Hence, the statement (i) holds.

(ii) By Theorem 11.17, $(X, \mathcal{B}, \tilde{\mu})$ is a complete measure space.

**Claim 11.19.3** $\mathcal{A} \subseteq \mathcal{B}$.

**Proof of claim:** Fix any $A \in \mathcal{A}$. $\forall \tilde{A} \subseteq X$, we will show that $\mu_o(\tilde{A}) = \mu_o(\tilde{A} \cap A) + \mu_o(\tilde{A} \cap \bar{A})$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mu_o(\tilde{A}) = \infty$; Case 2: $\mu_o(\tilde{A}) < \infty$. Case 1: $\mu_o(\tilde{A}) = \infty$. Then, $\mu_o(\tilde{A}) \geq \mu_o(\tilde{A} \cap A) + \mu_o(\tilde{A} \cap \bar{A}) \geq \mu_o(\tilde{A})$, where the second inequality follows from the countable subadditivity of an outer measure. This case is proved.

Case 2: $\mu_o(\tilde{A}) < \infty$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ such that $\tilde{A} \subseteq \bigcup_{i=1}^{\infty} A_i$ and $\mu_o(\tilde{A}) \leq \sum_{i=1}^{\infty} \mu(A_i) < \mu_o(\tilde{A}) + \epsilon$. By the countable additivity of $\mu$, we have $\mu(A_i) = \mu(A_i \cap \tilde{A}) + \mu(A_i \cap \bar{A})$, $\forall i \in \mathbb{N}$. Then,
CHAPTER 11. GENERAL MEASURE AND INTEGRATION

\( \mu_0(\bar{A}) + \epsilon > \sum_{i=1}^{\infty} (\mu(A_i \cap A) + \mu(A_i \cap \bar{A})) \geq \mu_0(\bar{A} \cap A) + \mu_0(\bar{A} \cap \bar{A}) \geq \mu_0(\bar{A}) \), where the second inequality follows from the definition of \( \mu_0 \); and the third inequality follows from the countable subadditivity of \( \mu_0 \). By the arbitrariness of \( \epsilon \), we have \( \mu_0(\bar{A}) = \mu_0(\bar{A} \cap A) + \mu_0(\bar{A} \cap \bar{A}) \). This case is proved.

Hence, in both cases, we have \( \mu_0(\bar{A}) = \mu_0(\bar{A} \cap A) + \mu_0(\bar{A} \cap \bar{A}) \). By the arbitrariness of \( \bar{A} \), we have \( \bar{A} \) is measurable with respect to \( \mu_0 \). Hence, \( A \in \mathcal{B} \) and \( A \subseteq \mathcal{B} \). This completes the proof of the claim. \( \square \)

\( \forall A \in \mathcal{A} \), we have \( \bar{\mu}(A) = \mu_0(A) \). Hence, the statement (ii) holds.

(iii) Note that \( X \in \mathcal{A} \). If \( \mu \) is finite, then \( \bar{\mu}(X) = \mu(X) < \infty \). Hence, \( \bar{\mu} \) is finite.

If \( \mu \) is \( \sigma \)-finite, then \( \exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) with \( X = \bigcup_{i=1}^{\infty} X_i \) such that \( \mu(X_i) < +\infty \), \( \forall i \in \mathbb{N} \). Then, \( \bar{\mu}(X_i) = \mu(X_i) < +\infty \) and \( (X_i)_{i=1}^{\infty} \subseteq \mathcal{B} \). Hence, \( \bar{\mu} \) is \( \sigma \)-finite.

(iv) Let \( \mu \) be \( \sigma \)-finite and \( \bar{\mu} : \mathcal{B}_a \to [0, \infty] \subset \mathbb{R} \) be any measure that is an extension of \( \mu \), that is, \( \bar{\mu}(A) = \mu(A), \forall A \in \mathcal{A} \). Clearly, \( \mathcal{B}_a \subseteq \mathcal{B} \). Then, by Proposition 11.13, \( \bar{\mu}|_{\mathcal{B}_a} \) is a measure on \( (X, \mathcal{B}_a) \) and it is an extension of \( \mu \). Let \( \mathcal{A}_\sigma := \{ E \subseteq X \mid \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni \exists E = \bigcup_{i=1}^{\infty} A_i \} \). Since \( \mathcal{B} \supseteq \mathcal{A} \) and is a \( \sigma \)-algebra on \( X \), then \( \mathcal{A}_\sigma \subseteq \mathcal{B}_a \subseteq \mathcal{B} \).

Claim 11.19.4 \( \forall E \in \mathcal{A}_\sigma \), \( \bar{\mu}(E) = \bar{\mu}(E) \).

Proof of claim: Fix any \( E \in \mathcal{A}_\sigma \). Then, \( \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni \exists E = \bigcup_{i=1}^{\infty} A_i \). Let \( A_1 := A_1 \) and \( A_{n+1} = (\bigcup_{i=1}^{n+1} A_i) \setminus (\bigcup_{i=1}^{n} A_i), \forall n \in \mathbb{N} \). Then, \( (A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, E = \bigcup_{i=1}^{\infty} A_i \), and \( A_1, A_2, \ldots \) are pairwise disjoint. Then, we have \( \bar{\mu}(E) = \sum_{i=1}^{\infty} \bar{\mu}(A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \bar{\mu}(A_i) = \bar{\mu}(E) \), where the first and last equality follows from the countable additivity of measures. This completes the proof of the claim. \( \square \)

Claim 11.19.5 \( \forall B \in \mathcal{B}_a \) with \( \bar{\mu}(B) < \infty \), we have \( \bar{\mu}(B) = \bar{\mu}(B) \).

Proof of claim: Fix any \( B \in \mathcal{B}_a \) with \( \bar{\mu}(B) < \infty \). Then, \( \mu_0(B) = \bar{\mu}(B) < \infty \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the definition of \( \mu_0 \), \( \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) such that \( B \subseteq \bigcup_{i=1}^{\infty} A_i =: E \in \mathcal{A}_\sigma \subseteq \mathcal{B}_a \) and \( \mu_0(B) \leq \sum_{i=1}^{\infty} \mu(A_i) < \bar{\mu}(B) + \epsilon \). Then, \( \bar{\mu}(B) \leq \bar{\mu}(E) = \bar{\mu}(E) \leq \sum_{i=1}^{\infty} \mu(A_i) < \bar{\mu}(B) + \epsilon \), where the first inequality follows from the monotonicity of measure, Proposition 11.4; the first equality follows from Claim 11.19.4; and the second inequality follows from the countable subadditivity of measure, Proposition 11.6. By the arbitrariness of \( \epsilon \), we have \( \bar{\mu}(B) \leq \bar{\mu}(B) \).

On the other hand, \( \bar{\mu}(E) = \bar{\mu}(B) + \bar{\mu}(E \setminus B) < \bar{\mu}(B) + \epsilon \). Since \( \bar{\mu}(B) < \infty \), then \( \bar{\mu}(E \setminus B) < \epsilon \). Note that \( E \setminus B \in \mathcal{B}_a \). This implies that \( \bar{\mu}(B) \leq \bar{\mu}(E) = \bar{\mu}(B) + \bar{\mu}(E \setminus B) \leq \bar{\mu}(B) + \bar{\mu}(E \setminus B) < \bar{\mu}(B) + \epsilon \), where the second inequality follows from the result of the previous paragraph. By the arbitrariness of \( \epsilon \), we have \( \bar{\mu}(B) \leq \bar{\mu}(B) \).

Hence, \( \bar{\mu}(B) = \bar{\mu}(B) \). This completes the proof of the claim. \( \square \)
Since \( \mu \) is \( \sigma \)-finite, then \( \exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) such that \( X = \bigcup_{i=1}^{\infty} X_i \) and \( \mu(X_i) < \infty, \forall i \in \mathbb{N} \). Without loss of generality, we may assume that \( X_1, X_2, \ldots \) are pairwise disjoint. \( \forall B \in \mathcal{B}_a \), we have \( B = \bigcup_{i=1}^{\infty} (B \cap X_i) \), \( B \cap X_1, B \cap X_2, \ldots \) are pairwise disjoint, and \( \bar{\mu}(B \cap X_i) \leq \bar{\mu}(X_i) < \infty \) with \( B \cap X_i \in \mathcal{B}_a, \forall i \in \mathbb{N} \). This implies that \( \bar{\mu}(B) = \sum_{i=1}^{\infty} \bar{\mu}(B \cap X_i) = \sum_{i=1}^{\infty} \bar{\mu}(B \cap X_i) = \bar{\mu}(B) \), where the first and last equality follows from the countable additivity of measures; and the second equality follows from Claim 11.19.5. Hence, \( \bar{\mu}|_{\mathcal{B}_a} = \bar{\mu} \). Therefore, \( \bar{\mu}|_{\mathcal{B}_a} \) is the unique measure on \( (X, \mathcal{B}_a) \) that is an extension of \( \mu \).

(v) Let \( \mu \) be \( \sigma \)-finite. Consider the measure space \( (X, \mathcal{B}_a, \bar{\mu}|_{\mathcal{B}_a}) \). By Proposition 11.12, this measure space admits the completion \( (X, \mathcal{B}, \bar{\mu}) \).

Since \( (X, \mathcal{B}, \bar{\mu}) \) is a complete measure space that agrees with \( \bar{\mu}|_{\mathcal{B}_a} \) on \( \mathcal{B}_a \), by Proposition 11.12, we have \( \bar{\mathcal{B}} \subseteq \mathcal{B} \) and \( \bar{\mu} = \bar{\mu}|_{\mathcal{B}_a} \).

Claim 11.19.6 \( \forall B \in \mathcal{B}, \exists U \in \mathcal{B}_a \) with \( B \subseteq U \) such that \( \bar{\mu}(U \setminus B) = 0 \).

Proof of claim: \( \forall B \in \mathcal{B}, \bar{\mu}(B) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, B \subseteq \bigcup_{i=1}^{\infty} A_i} \sum_{i=1}^{\infty} \mu(A_i). \)

We first prove the special case \( \bar{\mu}(B) < +\infty. \) \( \forall i \in \mathbb{N}, \exists (U_{i,j})_{j=1}^{\infty} \subseteq \mathcal{A} \) with \( B \subseteq \bigcup_{j=1}^{\infty} U_{i,j} =: U_i \in \mathcal{B}_a \) such that \( \bar{\mu}(B) \leq \bar{\mu}(U_i) = \bar{\mu}(U_i) \leq \sum_{j=1}^{\infty} \mu(U_{i,j}) < \bar{\mu}(B) + 2^{-i} < +\infty. \) Then, \( \bar{\mu}(U_i) = \bar{\mu}(B) + \bar{\mu}(U_i \setminus B) < \bar{\mu}(B) + 2^{-i} < +\infty. \) This implies that \( \bar{\mu}(U_i \setminus B) < 2^{-i}. \) Let \( U := \bigcap_{i=1}^{\infty} U_i \in \mathcal{B}_a \).

Clearly, \( B \subseteq U \) and \( \bar{\mu}(U \setminus B) \leq \bar{\mu}(U_i \setminus B) < 2^{-i}, \forall i \in \mathbb{N}. \) Hence, \( \bar{\mu}(U \setminus B) = 0. \)

Now, we prove the general case \( \bar{\mu}(B) \in [0, \infty] \subseteq \mathbb{R}. \) \( B = \bigcup_{i=1}^{\infty} (X_i \cap B) =: \bigcup_{i=1}^{\infty} B_i, \forall i \in \mathbb{N}, \bar{\mu}(B_i) \leq \bar{\mu}(X_i) = \mu(X_i) < +\infty. \) By the special case, \( \exists U_i \in \mathcal{B}_a \) with \( B_i \subseteq U_i \) such that \( \bar{\mu}(U_i \setminus B_i) = 0. \) Then, \( B \subseteq \bigcup_{i=1}^{\infty} U_i =: U \in \mathcal{B}_a \) and \( 0 \leq \bar{\mu}(U \setminus B) \leq \sum_{i=1}^{\infty} \bar{\mu}(U_i \setminus B_i) = 0, \) where the second inequality follows from the countable subadditivity of measure. This completes the proof of the claim. \( \Box \)

\( \forall B \in \mathcal{B}, \) by Claim 11.19.6, \( \exists U, \bar{U} \in \mathcal{B}_a \) with \( B \subseteq U \) and \( \bar{B} \subseteq \bar{U} \) such that \( \bar{\mu}(U \setminus B) = \bar{\mu}(\bar{U} \setminus \bar{B}) = 0. \) Let \( F := \bar{U} \in \mathcal{B}_a. \) Then, \( F \subseteq B \) and \( \bar{\mu}(B \setminus F) = \bar{\mu}(\bar{U} \setminus \bar{B}) = 0. \) Then, \( B = F \cup (B \setminus F), F \setminus F \subseteq U \setminus F, F \setminus F \in \mathcal{B}_a \) and \( \bar{\mu}(U \setminus F) = \bar{\mu}|_{\mathcal{B}_a}(U \setminus F) = \bar{\mu}(U \setminus F) = \bar{\mu}(U \setminus B) + \bar{\mu}(B \setminus F) = 0. \) This shows that \( B \in \mathcal{B}, \) by the completeness of \( (X, \mathcal{B}, \bar{\mu}). \) By the arbitrariness of \( B, \) we have \( \mathcal{B} \subseteq \mathcal{B}. \) Hence, \( B = \mathcal{B} \) and \( \bar{\mu} = \bar{\mu}. \) Hence \( (X, \mathcal{B}, \bar{\mu}) \) is the completion of \( (X, \mathcal{B}_a, \bar{\mu}|_{\mathcal{B}_a}). \)

This completes the proof of the theorem. \( \Box \)

Definition 11.20 An interval \( I \) in \( \mathbb{R} \) is any of the following sets: (i) \( \emptyset; \) (ii) \( (a, b) \subseteq \mathbb{R} \) with \( a, b \in \mathbb{R} \) and \( a < b; \) (iii) \( [a, b) \subseteq \mathbb{R} \) with \( a \in \mathbb{R}, b \in \mathbb{R}, \) and \( a < b; \) (iv) \( (a, b] \subseteq \mathbb{R} \) with \( a \in \mathbb{R}, b \in \mathbb{R}, \) and \( a < b; \) and (v) \( \mathbb{R} \subseteq \mathbb{R} \) with \( a, b \in \mathbb{R} \) and \( a \leq b. \)

Example 11.21 Let \( X = \mathbb{R}, \)

\( \mathcal{A} := \{ E \subseteq \mathbb{R} | E \) is the union of finitely many intervals in \( \mathbb{R} \}. \)
It is clear that \( A \) is an algebra on \( \mathbb{R} \). \( \forall E \in A, E \) may be written as the union of finitely many pairwise disjoint intervals: \( E = \bigcup_{i=1}^{n} I_i \), for some \( n \in \mathbb{N} \). Define \( \mu(E) = \sum_{i=1}^{n} \mu(I_i) \), where

\[
\mu(I) = \begin{cases} 
0 & \text{if } I = \emptyset \\
b - a & \text{if } I \text{ is any other types of interval}
\end{cases}
\]

Clearly, \( \mu(\emptyset) = 0; \mu(I_1 \cup I_2) = \mu(I_1) + \mu(I_2), \forall \) intervals \( I_1, I_2 \subseteq \mathbb{R} \) with \( I_1 \cap I_2 = \emptyset \). Then, \( \mu : A \to [0, \infty] \subset \mathbb{R} \) is well-defined and \( \mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2), \forall E_1, E_2 \in A \) with \( E_1 \cap E_2 = \emptyset \). Hence, \( \mu \) is finitely additive.

Fix any interval \( I = \bigcup_{i=1}^{\infty} E_i \), where \( (E_i)_{i=1}^{\infty} \subseteq A \) and \( E_1, E_2, \ldots \) are pairwise disjoint. We will show that \( \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \). \( \forall n \in \mathbb{N}, I \supseteq \bigcup_{i=1}^{n} E_i \). Then, by finite additivity of \( \mu, \mu(I) \geq \sum_{i=1}^{n} \mu(E_i) \). Hence, \( \mu(I) \geq \sum_{i=1}^{\infty} \mu(E_i) \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( \sum_{i=1}^{\infty} \mu(E_i) = \infty \); Case 2: \( I = \emptyset \); Case 3: \( I \neq \emptyset \) and \( \sum_{i=1}^{\infty} \mu(E_i) < \infty \). Then, \( \infty > \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \). Hence, \( \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \).

Case 2: \( I = \emptyset \). Then, \( 0 = \mu(I) \leq \sum_{i=1}^{\infty} \mu(E_i) \). Hence, we have \( 0 = \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \). Case 3: \( I \neq \emptyset \) and \( \sum_{i=1}^{\infty} \mu(E_i) < \infty \). \( \forall i \in \mathbb{N}, E_i = \bigcup_{j=1}^{n_i} I_{i,j} \), where \( n_i \in \mathbb{N} \) and \( I_{i,1}, \ldots, I_{i,n_i} \) are pairwise disjoint intervals. Clearly, \( \mu(E_i) = \sum_{j=1}^{n_i} \mu(I_{i,j}) \). Then, \( I = \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} I_{i,j} \). Then, we have \( \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j}) < \infty \). Fix any closed interval \( [a, b] := \bar{I} \subseteq I \), where \( a, b \in \mathbb{R} \) and \( a \leq b \). Such \( a, b \in \mathbb{R} \) exists since \( I \neq \emptyset \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall i \in \mathbb{N}, \forall j \in \{1, \ldots, n_i\} \), we may enlarge \( I_{i,j} \) to an open interval \( I_{i,j} \supseteq I_{i,j} \) with \( \mu(I_{i,j}) \leq \mu(I_{i,j}) + \frac{2^{-i,j}}{n_i} \).

Thus, we have \( \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \). Hence, in all three cases, we have \( \mu(I) = \sum_{i=1}^{\infty} \mu(E_i) \).

\( \forall E \in A, \forall (E_i)_{i=1}^{\infty} \subseteq A \) with \( E = \bigcup_{i=1}^{n} I_i \), \( E_1, E_2, \ldots \) being pairwise disjoint, we have \( E = \bigcup_{i=1}^{n} I_i \), where \( n \in \mathbb{N} \) and \( I_1, \ldots, I_n \) are pairwise disjoint intervals. Then, \( I_j = \bigcup_{i=1}^{n} (I_j \cap E_i), \forall j \in \{1, \ldots, n\} \). By the above argument, we have \( \mu(I_j) = \sum_{i=1}^{n} \mu(I_j \cap E_i) \). Then, \( \mu(E) = \sum_{j=1}^{n} \mu(I_j) = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \mu(I_j \cap E_i) = \sum_{j=1}^{\infty} \sum_{i=1}^{n} \mu(I_j \cap E_i) = \sum_{j=1}^{\infty} \mu(E_i) \), where the last equality follows from the finite additivity of \( \mu \).

Hence, \( \mu \) is a measure on algebra \( A \) and is \( \sigma \)-finite. By Carathéodory Extension Theorem 11.19, \( \mu \) extends uniquely to a \( \sigma \)-finite complete measure space \( (X, B, \bar{\mu}) \). This measure space is called the Lebesgue measure space. Since \( B \) is a \( \sigma \)-algebra on \( \mathbb{R} \) and the topology \( O_R \) on \( \mathbb{R} \) is second countable, then \( B \supseteq O_R \). We will denote the Lebesgue measure space by \( (\mathbb{R}, B_L, \mu_L) \), where \( B_L \) is the collection of Lebesgue measurable sets in \( \mathbb{R} \) and \( \mu_L \) is the Lebesgue measure.
11.2. OUTER MEASURE AND THE EXTENSION THEOREM

Let \( \mu_{\text{Lo}} \) be the Lebesgue outer measure induced by \( \mu \) as defined in Carathéodory Extension Theorem 11.19. We claim that

\[
\mu_{\text{Lo}}(E) = \inf_{\{I_i\}_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu(I_i); \quad \forall E \subseteq \mathbb{R} \quad (11.3)
\]

To prove this, we note that \( \mu_{\text{Lo}}(E) = \inf_{\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{A}, E \subseteq \bigcup_{i=1}^{\infty} E_i} \sum_{i=1}^{\infty} \mu(E_i) \).

Then, clearly, we have

\[
\mu_{\text{Lo}}(E) \leq \inf_{\{I_i\}_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu(I_i)
\]

On the other hand, \( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) with \( E \subseteq \bigcup_{i=1}^{\infty} E_i, \forall i \in \mathbb{N}, E_i = \bigcup_{j=1}^{n_i} I_{i,j} \) where \( n_i \in \mathbb{N} \) and \( I_{i,1}, \ldots, I_{i,n_i} \) are pairwise disjoint intervals, and \( \mu(E_i) = \sum_{j=1}^{n_i} \mu(I_{i,j}) \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \forall j \in \{1, \ldots, n_i\} \), we may enlarge \( I_{i,j} \) to an open interval \( \bar{I}_{i,j} \supseteq I_{i,j} \) with \( \mu(\bar{I}_{i,j}) \leq \mu(I_{i,j}) + \frac{\epsilon}{n_i} \).

Then, we have \( E \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} I_{i,j} \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} \bar{I}_{i,j} \) and \( \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j}) \geq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} (\mu(\bar{I}_{i,j}) - \frac{\epsilon}{n_i}) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(\bar{I}_{i,j}) - \epsilon \geq -\epsilon + \inf_{\{I_i\}_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu(I_i) \).

This implies that \( \mu_{\text{Lo}}(E) \geq -\epsilon + \inf_{\{I_i\}_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu(I_i) \).

Hence, by the arbitrariness of \( \epsilon \), we have

\[
\mu_{\text{Lo}}(E) \geq \inf_{\{I_i\}_{i=1}^{\infty} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{\infty} I_i} \sum_{i=1}^{\infty} \mu(I_i)
\]

Therefore, (11.3) is true.

By Carathéodory Extension Theorem 11.19, \( \mu_{\text{L}} = \mu_{\text{Lo}|_{\mathcal{B}_{\text{L}}}} \) and \( (\mathbb{R}, \mathcal{B}_{\text{L}}, \mu_{\text{L}}) \) is the completion of \( (\mathbb{R}, \mathcal{B}_\sigma, \mu|_{\mathbb{R}}) \), where \( \mathcal{B}_\sigma \) is the \( \sigma \)-algebra on \( \mathbb{R} \) generated by \( \mathcal{A} \). Clearly, \( \mathcal{O}_{\mathbb{R}} \subseteq \mathcal{B}_\sigma \) and \( \mathcal{B}_\sigma \) is the \( \sigma \)-algebra on \( \mathbb{R} \) generated by \( \mathcal{O}_{\mathbb{R}} \).

**Definition 11.22** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space. The collection of Borel sets on \( \mathcal{X} \) is the smallest \( \sigma \)-algebra on \( X \) that contains \( \mathcal{O} \), which will be denoted by \( \mathcal{B}_\mathcal{B}(\mathcal{X}) \). The smallest algebra on \( X \) that contains \( \mathcal{O} \) will be denoted by \( \mathcal{A}(\mathcal{X}) \).

Clearly, we have \( \mathcal{B}_\mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}_{\text{L}} \).

**Proposition 11.23** Let \( E \subseteq \mathbb{R} \). Then, the following statements are equivalent.

(i) \( E \in \mathcal{B}_{\text{L}} \).

(ii) \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \mathcal{O} \in \mathcal{O}_{\mathbb{R}} \text{ with } E \subseteq O \text{ such that } \mu_{\text{Lo}}(O \setminus E) < \epsilon \).
(iii) $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists F \subseteq \mathbb{R}$ with $\tilde{F} \in \mathcal{O}_\mathbb{R}$ and $F \subseteq E$ such that

$$\mu_{\mathcal{O}_\mathbb{R}}(E \setminus F) < \epsilon.$$ 

If $\mu_{\mathcal{O}_\mathbb{R}}(E) < +\infty$. Then, (i)–(iii) are also equivalent to

(iv) $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists U \in \mathcal{O}_\mathbb{R}$ which is finite union of open intervals, such that $\mu_{\mathcal{O}_\mathbb{R}}(U \triangle E) < \epsilon$.

Hence, the Lebesgue measure space $(\mathbb{R}, \mathcal{B}_L, \mu_L)$ is the completion of the Borel measure space $(\mathbb{R}, \mathcal{B}_B(\mathbb{R}), \mu_B := \mu_{\mathcal{B}_B(\mathbb{R})})$.

**Proof** Fix $E \subseteq \mathbb{R}$.

(i) $\Rightarrow$ (ii). First, we show the special case that $\mu_{\mathcal{O}_\mathbb{R}}(E) < +\infty$. Let $E \in \mathcal{B}_L$. By Example 11.21,

$$\mu_{\mathcal{O}_\mathbb{R}}(E) = \mu_L(E) = \inf_{\{I_i\}_{i=1}^\infty, \text{open intervals}, \bigcup_{i=1}^\infty I_i \subseteq E} \sum_{i=1}^\infty \mu_L(I_i)$$

Then, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists$ a sequence of open intervals $(I_i)_{i=1}^\infty \subseteq \mathcal{O}_\mathbb{R}$ with $E \subseteq \bigcup_{i=1}^\infty I_i =: O \in \mathcal{O}_\mathbb{R}$ such that $\mu_L(O) \leq \mu_L(E) \leq \sum_{i=1}^\infty \mu_L(I_i) < \mu_L(E) + \epsilon < +\infty$. Then, $\epsilon > \mu_L(O) - \mu_L(E) = \mu_L(O \setminus E) = \mu_{\mathcal{O}_\mathbb{R}}(O \setminus E)$, where the first equality follows from the countable additivity of $\mu_L$. Hence, (ii) holds.

Next, we show the general case $\mu_{\mathcal{O}_\mathbb{R}}(E) \in [0, \infty] \subseteq \mathbb{R}$. Let $E \in \mathcal{B}_L$ and $I_1 := (-1, 1) \subseteq \mathbb{R}$, $I_n := (-n, -n + 1] \cup (n - 1, n] \subseteq \mathbb{R}$, $\forall n = 2, 3, \ldots$. Clearly, $E = \bigcup_{i=1}^\infty (I_i \cap E) =: \bigcup_{i=1}^\infty E_i$. Fix any $\epsilon \in (0, \infty) \subseteq \mathbb{R}$. $\forall i \in \mathbb{N}$, $I_i \in \mathcal{B}_L$ and $E_i \in \mathcal{B}_L$ with $\mu_L(E_i) \leq \mu_L(I_i) = 2$. By the special case we have just shown, $\exists O_i \in \mathcal{O}_\mathbb{R} \subseteq \mathcal{B}_L$ with $E_i \subseteq O_i$, such that $\mu_{\mathcal{O}_\mathbb{R}}(O_i \setminus E_i) = \mu_L(O_i \setminus E_i) < 2^{-i} \epsilon$. Let $O := \bigcup_{i=1}^\infty O_i \in \mathcal{O}_\mathbb{R}$. Then, $\mu_{\mathcal{O}_\mathbb{R}}(O \setminus E) = \mu_L(O \setminus E) = \mu_L((\bigcup_{i=1}^\infty O_i) \cap \bigcup_{i=1}^\infty E_i) = \mu_L((\bigcup_{i=1}^\infty O_i \cap (\bigcap_{j=1}^\infty \overline{E_j})) \cap (\bigcup_{i=1}^\infty E_i)) \leq \sum_{i=1}^\infty \mu_L(O_i \cap \bigcap_{j=1}^\infty \overline{E_j}) \leq \sum_{i=1}^\infty \mu_L(O_i \cap E_i) = \sum_{i=1}^\infty \mu_L(O_i \setminus E_i) < \epsilon$. Hence (ii) holds.

(ii) $\Rightarrow$ (i). Fix any $A \subseteq \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists O \in \mathcal{O}_\mathbb{R}$ with $E \subseteq O$ such that $\mu_{\mathcal{O}_\mathbb{R}}(O \setminus E) < \epsilon$. Note that $A \cap \overline{E} = (A \cap \overline{E} \cap O) \cup (A \cap \overline{E} \cap \overline{O}) \subseteq (O \setminus E) \cup (A \cap \overline{O})$ and $A \cap \overline{E} \subseteq A \cap O$. By Definition 11.14, we have $\mu_{\mathcal{O}_\mathbb{R}}(A \cap \overline{E}) \leq \mu_{\mathcal{O}_\mathbb{R}}(A \cap E) + \mu_{\mathcal{O}_\mathbb{R}}(A \cap \overline{O}) \leq \mu_{\mathcal{O}_\mathbb{R}}(A \cap O) + \mu_{\mathcal{O}_\mathbb{R}}(O \setminus E) + \mu_{\mathcal{O}_\mathbb{R}}(A \cap \overline{O}) \leq \mu_{\mathcal{O}_\mathbb{R}}(A) + \epsilon$, where the third inequality follows from the fact that $O \in \mathcal{O}_\mathbb{R} \subseteq \mathcal{B}_L$ and Definition 11.15. By the arbitrariness of $\epsilon$, we have $\mu_{\mathcal{O}_\mathbb{R}}(A) = \mu_{\mathcal{O}_\mathbb{R}}(A \cap E) + \mu_{\mathcal{O}_\mathbb{R}}(A \cap \overline{E})$. By the arbitrariness of $A$, we have $E \in \mathcal{B}_L$.

(i) $\Leftrightarrow$ (iii). $E \in \mathcal{B}_L \iff \overline{E} \in \mathcal{B}_L \iff \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists O \in \mathcal{O}_\mathbb{R}$ with $\overline{E} \subseteq O$ such that $\mu_{\mathcal{O}_\mathbb{R}}(O \setminus \overline{E}) < \epsilon$, where the last $\Leftrightarrow$ follows from (i) $\Leftrightarrow$ (ii). Let $F := \overline{O}$, we have $F \subseteq E$ and $\mu_{\mathcal{O}_\mathbb{R}}(E \setminus F) = \mu_{\mathcal{O}_\mathbb{R}}(O \setminus \overline{E}) < \epsilon$.

Next, we show (i) – (iii) are equivalent to (iv) under the additional assumption that $\mu_{\mathcal{O}_\mathbb{R}}(E) < +\infty$. 


(i) $\Rightarrow$ (iv). Let $E \in \mathcal{B}_L$. \(\forall \varepsilon \in (0, \infty) \subset \mathbb{R}\), by Example 11.21, \(\exists\) a sequence of open intervals \((I_i)_{i=1}^\infty\) with \(E \subseteq \bigcup_{i=1}^\infty I_i =: O \in \mathcal{O}_\mathbb{R}\) such that 
\[
\mu_{\text{lo}}(E) = \mu_L(E) \leq \sum_{i=1}^\infty \mu_L(I_i) < \mu_L(E) + \varepsilon/2 < +\infty.
\]
Then, \(\exists N \in \mathbb{N}\) such that \(\sum_{i=N+1}^\infty \mu_L(I_i) < \varepsilon/2\). Let \(U := \bigcup_{i=1}^N I_i \in \mathcal{O}_\mathbb{R}\). We have \(\mu_{\text{lo}}(U \Delta E) = \mu_L(U \setminus E) + \mu_L(E \setminus U) \leq \mu_L(O \setminus E) + \mu_L(E \cap \bar{U})\). Note that 
\[
\mu_L(O) = \mu_L(E) + \mu_L(O \setminus E) \leq \sum_{i=1}^\infty \mu_L(I_i) \leq \mu_L(E) + \varepsilon/2 < +\infty.
\]
and \(E \cap \bar{U} = (E \cap \bar{U} \cap O) \cup (E \cap \bar{U} \cap \bar{O}) = E \cap \bar{U} \cap O \subseteq O \setminus U \subseteq \bigcup_{i=N+1}^\infty I_i\). Then, 
\[
\mu_{\text{lo}}(U \Delta E) \leq \varepsilon/2 + \mu_L(\bigcup_{i=N+1}^\infty I_i) \leq \varepsilon/2 + \sum_{i=N+1}^\infty \mu_L(I_i) < \varepsilon.
\]
Hence, (iv) holds.

(iv) $\Rightarrow$ (i). Fix any set \(A \subseteq \mathbb{R}\). By (iv), \(\forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \ \exists U \in \mathcal{O}_\mathbb{R}\) which is finite union of open intervals, such that \(\mu_{\text{lo}}(U \Delta E) < \varepsilon/2\). Note that 
\[
A \cap E = (A \cap U \cap E) \cup (A \cap U \cap \bar{E}) \subseteq (A \cap U) \cup (E \setminus U) \text{ and } A \cap \bar{E} = (A \cap \bar{E} \cap U) \cup (A \cap \bar{E} \cap \bar{U}) \subseteq (U \setminus E) \cup (A \setminus U).
\]
By Definition 11.14, 
\[
\mu_{\text{lo}}(A) \leq \mu_{\text{lo}}(A \cap E) + \mu_{\text{lo}}(A \cap \bar{E}) \leq \mu_{\text{lo}}((A \cap U) \cup (E \setminus U)) + \mu_{\text{lo}}((U \setminus E) \cup (A \setminus U)) \leq \mu_{\text{lo}}(A \cap U) + \mu_{\text{lo}}(E \setminus U) + \mu_{\text{lo}}(U \setminus E) + \mu_{\text{lo}}(A \setminus U).
\]
By the arbitrariness of \(\varepsilon\), we have \(\mu_{\text{lo}}(A) = \mu_{\text{lo}}(A \cap E) + \mu_{\text{lo}}(A \cap \bar{E})\). By the arbitrariness of \(A\), we have \(E \in \mathcal{B}_L\). Hence (i) holds.

Finally, by Example 11.21, the Lebesgue measure space is the completion of the Borel measure space \((\mathbb{R}, \mathcal{B}_\mathbb{R}(\mathbb{R}), \mu_{\text{lo}}|_{\mathcal{B}_\mathbb{R}(\mathbb{R})})\). This completes the proof of the proposition.

\[\square\]

**Proposition 11.24** Let \(X_i := (X_i, O_i)\) be a second countable topological space, \(i := 1, 2\), \(X := X_1 \times X_2 =: (X_1 \times X_2, O)\) be the product topological space, \(E := \{B_1 \times B_2 \subseteq X_1 \times X_2 \mid B_1 \in \mathcal{B}_\mathbb{R}(X_1), B_2 \in \mathcal{B}_\mathbb{R}(X_2)\}, \ i := 1, 2\), and \(B\) be the \(\sigma\)-algebra on \(X_1 \times X_2\) generated by \(E\). Then, \(\mathcal{B} = \mathcal{B}_\mathbb{R}(X)\).

**Proof** Let \(O_{B_i} \subseteq O_i\) be a countable basis for \(X_i\), \(i := 1, 2\). Without loss of generality, assume that \(X_i \in \mathcal{O}_{B_i}, \ i := 1, 2\). By Proposition 3.28, \(X\) is second countable with a countable basis \(\mathcal{O}_B := \{O_{B_1} \times O_{B_2} \subseteq X_1 \times X_2 \mid O_{B_1} \in \mathcal{O}_{B_1}, i := 1, 2\}\). Clearly, \(O_{B_1} \subseteq O_i \subseteq \mathcal{B}_\mathbb{R}(X_i), i := 1, 2\). Then, \(\mathcal{O}_B \subseteq E\). \(\forall O \in O, O = \bigcup_{i=1}^\infty O_i\), where \(O_i \in \mathcal{O}_B, \forall i \in \mathbb{N}\). Hence, \(O \in B\). This implies that \(O \subseteq B\). Since \(B\) is a \(\sigma\)-algebra, then \(O \subseteq \mathcal{B}_\mathbb{R}(X) \subseteq B\).

On the other hand, note that \(\mathcal{B}_\mathbb{R}(X)\) is the \(\sigma\)-algebra generated by \(O\). Then, \(\forall O_1 \in O_1, \forall O_2 \in O_2, \text{ we have } O_1 \times O_2 \in O \subseteq \mathcal{B}_\mathbb{R}(X)\).

**Claim 11.24.1** \(\forall O_1 \in O_1, \forall O_2 \in \mathcal{B}_\mathbb{R}(X_2), \text{ we have } O_1 \times O_2 \in \mathcal{B}_\mathbb{R}(X)\).

**Proof of claim:** Fix any \(O_1 \in O_1\). Define \(F := \{E \subseteq X_2 \mid O_1 \times E \in \mathcal{B}_\mathbb{R}(X)\}\). Clearly, \(O_2 \subseteq F\). Then, \(\emptyset, X_2 \in O_2 \subseteq F\). \(\forall E \in F\), we have \(O_1 \times E \in \mathcal{B}_\mathbb{R}(X)\). Note that \(O_1 \times X_2 \in \mathcal{B}_\mathbb{R}(X)\). This leads to \(O_1 \times (X_2 \setminus E) = (O_1 \times X_2) \setminus (O_1 \times E) \in \mathcal{B}_\mathbb{R}(X)\). Thus, \(X_2 \setminus E \in F\). \(\forall (E_i)_{i=1}^\infty \subseteq F\), we have \(O_1 \times E_i \in \mathcal{B}_\mathbb{R}(X), \forall i \in \mathbb{N}\). This implies that
Let \( E \subseteq B_B(\mathcal{X}) \). Then, \( O_2 \subseteq B_B(\mathcal{X}_2) \subseteq \mathcal{F} \). Hence, \( \forall b_2 \in B_B(\mathcal{X}_2) \), we have \( b_2 \in \mathcal{F} \) and \( O_1 \times b_2 \in B_B(\mathcal{X}) \). This completes the proof of the claim.

**Claim 11.24.2** \( \forall b_1 \in B_B(\mathcal{X}_1), \forall b_2 \in B_B(\mathcal{X}_2) \), we have \( b_1 \times b_2 \in B_B(\mathcal{X}) \).

**Proof of claim:** Fix any \( b_2 \in B_B(\mathcal{X}_2) \). Define \( \mathcal{F} := \{ E \subseteq \mathcal{X} \mid E \times b_2 \in B_B(\mathcal{X}) \} \). By Claim 11.24.1, \( O_1 \subseteq \mathcal{F} \). Then, \( \emptyset, X_1 \in O_1 \subseteq \mathcal{F} \). \( \forall E \in \mathcal{F} \), we have \( E \times b_2 \in B_B(\mathcal{X}) \). Note that \( X_1 \times b_2 \in B_B(\mathcal{X}) \). This leads to \( (X_1 \setminus E) \times b_2 = (X_1 \times b_2) \setminus (E \times b_2) \in B_B(\mathcal{X}) \). Thus, \( X_1 \setminus E \in \mathcal{F} \). \( \forall (E_i)_{i=1}^\infty \subseteq \mathcal{F} \), we have \( E_i \times b_2 \in B_B(\mathcal{X}) \). This implies that \( (\bigcup_{i=1}^\infty E_i) \times b_2 = \bigcup_{i=1}^\infty (E_i \times b_2) \in B_B(\mathcal{X}) \) and \( \bigcup_{i=1}^\infty E_i \in \mathcal{F} \). The above shows that \( \mathcal{F} \) is a \( \sigma \)-algebra on \( \mathcal{X}_1 \). Then, \( O_1 \subseteq B_B(\mathcal{X}_1) \subseteq \mathcal{F} \). Hence, \( \forall b_1 \in B_B(\mathcal{X}_1) \), we have \( b_1 \in \mathcal{F} \) and \( b_1 \times b_2 \in B_B(\mathcal{X}) \). This completes the proof of the claim.

By Claim 11.24.2, \( \mathcal{F} \subseteq B_B(\mathcal{X}) \). By the definition of \( \mathcal{B} \), we have \( \mathcal{B} \subseteq B_B(\mathcal{X}) \). Hence, \( \mathcal{B} = B_B(\mathcal{X}) \). This completes the proof of the proposition.

**Proposition 11.25** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( E \in B_B(\mathcal{X}) \), \( \mathcal{E} := (E, \mathcal{O}_E) \) be the topological space with the subset topology of \( \mathcal{X} \), and \( B_E := \{ C \subseteq E \mid C \in B_B(\mathcal{X}) \} \). Then, \( B_E = B_B(\mathcal{E}) \).

**Proof** Clearly \( \emptyset, E \in B_E \). \( \forall A \in B_E \), we have \( A \subseteq E \) and \( A \in B_B(\mathcal{X}) \). Then, \( E \supseteq E \setminus A \in B_B(\mathcal{X}) \) and \( E \setminus A \in B_E \). \( \forall (A_i)_{i=1}^\infty \subseteq B_E \), we have \( E \supseteq \bigcup_{i=1}^\infty A_i \in B_B(\mathcal{X}) \) and \( \bigcup_{i=1}^\infty A_i \in B_E \). This shows that \( B_E \) is a \( \sigma \)-algebra on \( E \).

\( \forall O_E \in \mathcal{O}_E \), \( \exists O \in \mathcal{O} \subseteq B_B(\mathcal{X}) \) such that \( O_E = O \cap E \). Then, \( O_E \in B_E \) and \( O_E \subseteq B_E \). Hence, \( \mathcal{O}_E \subseteq B_B(\mathcal{E}) \subseteq B_E \).

On the other hand, let \( \mathcal{E} := (X \setminus E, \mathcal{O}_{X \setminus E}) \) be the topological space with the subset topology of \( \mathcal{X} \).

**Claim 11.25.1** \( \forall C \in B_B(\mathcal{X}) \), we have \( C_1 := C \cap E \in B_B(\mathcal{E}) \) and \( C_2 := C \cap (X \setminus E) \in B_B\left(\mathcal{E}\right)\).

**Proof of claim:** Define \( \mathcal{F} := \left\{ A_1 \cup A_2 \subseteq X \mid A_1 \in B_B(\mathcal{E}), A_2 \in B_B\left(\mathcal{E}\right)\right\} \). \( \forall O \in \mathcal{O} \), \( O \cap E \in \mathcal{O}_E \subseteq B_B(\mathcal{E}) \) and \( O \cap (X \setminus E) \in \mathcal{O}_{X \setminus E} \subseteq B_B\left(\mathcal{E}\right) \). Then, \( O = (O \cap E) \cup (O \cap (X \setminus E)) \in \mathcal{F} \) and \( O \subseteq \mathcal{F} \).

Clearly, \( \emptyset, E \in B_B(\mathcal{E}) \) and \( \emptyset, X \setminus E \in B_B\left(\mathcal{E}\right) \). Then, \( \emptyset = \emptyset \cup \emptyset \in \mathcal{F} \) and \( X = E \cup (X \setminus E) \in \mathcal{F} \). \( \forall A \in \mathcal{F} \), \( \exists A_1 \in B_B(\mathcal{E}) \) and \( \exists A_2 \in B_B\left(\mathcal{E}\right) \) such that \( A = A_1 \cup A_2 \). Then, \( A_1 = A \cap E \) and \( A_2 = A \cap (X \setminus E) \). Note that
11.2. OUTER MEASURE AND THE EXTENSION THEOREM

Let \( X := (X, \mathcal{O}) \) be a topological space and \((X, \mathcal{B}, \mu)\) be a measure space on the same set \( X \). The triple \( \mathcal{X} := (X, \mathcal{B}, \mu) \) is said to be a topological measure space if \( \mathcal{B} = \mathcal{B}_\mathcal{B}(\mathcal{X}) \) and \( \forall E \in \mathcal{B}_\mathcal{B}(\mathcal{X}), \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists U \in \mathcal{O} \text{ with } E \subset U \text{ such that } \mu(U \setminus E) < \varepsilon. \) We will say that \( \mathcal{X} \) is finite or \( \sigma \)-finite if the underlying measure space is so. We will say that \( \mathcal{X} \) is Tychonoff, Hausdorff, regular, completely regular, or normal if \( \mathcal{X} \) is so. We will say that \( \mathcal{X} \) is first countable or second countable if \( \mathcal{X} \) is so. We will say that \( \mathcal{X} \) is separable, second category everywhere, connected, or locally connected if \( \mathcal{X} \) is so. We will say that \( \mathcal{X} \) is compact, countably compact, sequentially compact, locally compact, \( \sigma \)-compact, or paracompact if \( \mathcal{X} \) is so. We will say that \( \mathcal{X} \) is locally finite if \( \forall \text{compact } K \subset \mathcal{X}, \mu(K) < \infty. \)

Let \( X := (X, \rho) \) be a metric space, \( \mathcal{O} \) be the natural topology on \( X \) generated by the metric \( \rho \), and \((X, \mathcal{B}, \mu)\) be a measure space on the same set \( X \). The triple \( \mathcal{X} := (X, \mathcal{B}, \mu) \) is said to be a metric measure space if \((X, \mathcal{O}), \mathcal{B}, \mu) \) is a topological measure space. \( \mathcal{X} \) is said to be a complete metric measure space, if \( \mathcal{X} \) is a complete metric space. \( \mathcal{X} \) is said to be totally bounded if \( \mathcal{X} \) is so.

Let \( \mathcal{X} := (X, \mathcal{K}, \| \cdot \|) \) be a normed linear space over the field \( \mathbb{K} \), \( \mathcal{O} \) be the natural topology on \( \mathcal{X} \) generated by the norm \( \| \cdot \| \), and \((X, \mathcal{B}, \mu)\) be a measure space on the same set \( X \). The triple \( \mathcal{X} := (X, \mathcal{B}, \mu) \) is said to be a normed linear measure space if \((X, \mathcal{O}), \mathcal{B}, \mu) \) is a topological measure space. \( \mathcal{X} \) is said to be a complete normed linear measure space. Furthermore, if \( \mathcal{X} \) is a Banach space, then \( \mathcal{X} \) is said to be a Banach measure space. Depending on whether \( \mathbb{K} = \mathbb{R} \) or \( \mathbb{K} = \mathbb{C} \), we will say that \( \mathcal{X} \) is a real or complex Banach measure space.

Proposition 11.27 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \((X, \mathcal{B}, \mu)\) be a measure space on the same set \( X \), and \( \mathcal{B} = \mathcal{B}_\mathcal{B}(\mathcal{X}) \). The triple \( \mathcal{X} := (X, \mathcal{B}, \mu) \) is topological measure space if, and only if, \( \forall E \in \mathcal{B}_\mathcal{B}(\mathcal{X}), \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists X \setminus F \in \mathcal{O} \text{ with } F \subset E \text{ such that } \mu(E \setminus F) < \varepsilon. \)

Proof “Necessity” \( \forall E \in \mathcal{B}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, X \setminus E \in \mathcal{B}. \) By \( \mathcal{X} \) being a topological measure space, \( \exists U \in \mathcal{O} \text{ with } X \setminus E \subset U \text{ such that } \mu(U \setminus (X \setminus E)) \subset \mathbb{R}. \) Then, \( X \setminus U \subset E \text{ and } \mu(E \setminus (X \setminus U)) = \mu(U \setminus (X \setminus E)) < \varepsilon. \) Hence, \( F := X \setminus U \) is the set we seek.
“Sufficiency” \( \forall E \in \mathcal{B}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \ X \setminus E \in \mathcal{B}. \) By the assumption, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq X \setminus E \) such that \( \mu((X \setminus E) \setminus F) < \varepsilon. \) Let \( U := X \setminus F. \) Then, \( E \subseteq U \) and \( \mu(U \setminus E) = \mu((X \setminus F) \setminus E) = \mu((X \setminus E) \setminus F) < \varepsilon. \) Hence, \( X \) is a topological measure space. \( \square \)

**Example 11.28** The real line \( \mathbb{R} \) is a real Banach space with norm defined to be the absolute value. Then, by Proposition 11.23 and Example 11.21, \( \mathbb{R} := ((\mathbb{R}, \mathbb{R}, \cdot)), \mathcal{B}_\mathbb{R} (\mathbb{R}), \mu_\mathbb{R} \) is a \( \sigma \)-finite real Banach measure space. \( \diamond \)

**Proposition 11.29** Let \( \mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu) \) be a topological measure space, where \( \mathcal{X} := (X, \mathcal{O}) \) is a topological space, \( E \in \mathcal{B}, \) \( \mathcal{E} := (E, \mathcal{O}_E) \) be the topological space with the subset topology of \( \mathcal{X}, \) and \( (E, \mathcal{B}_E, \mu_E) \) be the measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13. Then, \( \mathcal{E} := (\mathcal{E}, \mathcal{B}_E, \mu_E) \) is a topological measure space and is said to be the topological measure subspace of \( \mathcal{X}. \)

**Proof** By Definition 11.26, \( \mathcal{B} = \mathcal{B}_\mathcal{X} (\mathcal{X}). \) By Proposition 11.13, \( \mathcal{B}_E = \{ C \subseteq E \mid C \in \mathcal{B} \} \) and \( \mu_E = \mu|_{\mathcal{B}_E}. \) By Proposition 11.25, we have \( \mathcal{B}_E = \mathcal{B}_\mathcal{X} (\mathcal{E}). \)

\( \forall A \in \mathcal{B}_E, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \ E \supseteq A \in \mathcal{B}. \) By \( \mathcal{X} \) being a topological measure space, \( \exists U \in \mathcal{O} \) with \( A \subseteq U \) such that \( \mu(U \setminus A) < \varepsilon. \) Let \( U := U \cap E \in \mathcal{O}_E \subseteq \mathcal{B}_E. \) Clearly, \( A \subseteq U \) and \( 0 \leq \mu_E(U \setminus A) = \mu(U \cap E \setminus A) \leq \mu(U \setminus A) < \varepsilon. \) Hence, \( \mathcal{E} \) is a topological measure space. This completes the proof of the proposition. \( \square \)

**Definition 11.30** Let \( X \) be a set and \( \mathcal{C} \subseteq \mathcal{X}^2 \) be a nonempty collection of subsets of \( X. \) \( \mathcal{C} \) is said to be a semialgebra on \( X \) if \( \forall C_1, C_2 \in \mathcal{C}, C_1 \cap C_2 \in \mathcal{C} \) and \( C_1 \) is a finite disjoint union of sets in \( \mathcal{C}. \)

**Proposition 11.31** Let \( X \) be a set and \( \mathcal{C} \subseteq \mathcal{X}^2 \) be a semialgebra on \( X. \) \( \mathcal{A} := \{ A \subseteq X \mid A \) is the finite disjoint union of sets in \( \mathcal{C} \}. \) Then, \( \mathcal{A} \) is the algebra on \( X \) generated by \( \mathcal{C}. \)

**Proof** Clearly \( \emptyset \in \mathcal{A}. \) Fix \( C_1 \in \mathcal{C} \neq \emptyset. \) Then, \( \widetilde{C_1} = \bigcup_{i=2}^{n} C_i, \) where \( n \in \mathbb{N}, \) \( (C_i)_{i=2}^{n} \subseteq \mathcal{C} \) is pairwise disjoint. Then, \( X = C_1 \cup \widetilde{C_1} = \bigcup_{i=1}^{n} C_i, \) where the sets in the union are pairwise disjoint. Hence, \( X \in \mathcal{A}. \)

\( \forall A_1, A_2 \in \mathcal{A}, \) then \( A_i = \bigcup_{j=1}^{n_i} C_{i,j,i} \), where \( n_i \in \mathbb{Z}_+ \) and \( (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C} \) is pairwise disjoint, \( i = 1, 2. \) Note that \( \widetilde{A_1} = \bigcap_{j=1}^{n_1} \widetilde{C_{1,j,i}}, \forall j \in \{1, \ldots, n_1\}, \) \( \widetilde{C_{1,j,1}} = \bigcup_{i=1}^{m_j} C_{1,j,1,i}, \) where \( m_j \in \mathbb{Z}_+ \) and \( (C_{1,j,1})_{i=1}^{m_j} \subseteq \mathcal{C} \) is pairwise disjoint. Then,

\[
\widetilde{A_1} = \bigcap_{j=1}^{n_1} \bigcup_{i=1}^{m_j} C_{1,j,1,i} = \bigcup_{i=1}^{m_1} \bigcap_{j=1}^{n_1} \bigcup_{i=1}^{m_j} C_{1,j,1,i} = \bigcup_{i=1}^{m_2} \bigcap_{j=1}^{n_1} \bigcup_{i=1}^{m_j} C_{1,j,1,i}
\]
When \( n_1 = 0 \), clearly \( \bar{A}_1 = X \in \mathcal{A} \). When \( n_1 \geq 1 \), \( \bigcap_{l=1}^{n_1} C_{i_l,j_l} \in \mathcal{C} \) for any admissible choice of indices \( l_1, \ldots, l_{n_1} \). For different admissible choices of indices \( l_1, \ldots, l_{n_1} \) and \( 
abla_1, \ldots, \nabla_{n_1} \), \( (\bigcap_{l=1}^{n_1} C_{i_l,j_l}) \cap (\bigcap_{l=1}^{n_1} C_{j_l,j_l}) = \emptyset \). Hence, \( \bar{A}_1 \in \mathcal{A} \). Note that \( A_1 \cap A_2 = \bigcup_{j=1}^{n_1} \bigcup_{l=1}^{n_2} (C_{i_1,j_1} \cap C_{j_2,j_2}) \). \( C_{i_1,j_1} \cap C_{j_2,j_2} \in \mathcal{C} \) for any admissible choice of indices \( j_1, j_2 \). For different admissible choices of indices \( j_1, j_2 \) and \( j_1, j_2 \), \( (C_{i_1,j_1} \cap C_{j_2,j_2}) \cap (C_{j_1,j_1} \cap C_{j_2,j_2}) = \emptyset \). Hence, \( A_1 \cap A_2 \in \mathcal{A} \).

Therefore, \( \mathcal{A} \) is an algebra on \( X \). Clearly, \( \mathcal{A} \) is the smallest algebra on \( X \) containing \( \mathcal{C} \). Hence, \( \mathcal{A} \) is the algebra on \( X \) generated by \( \mathcal{C} \). This completes the proof of the proposition.

**Proposition 11.32** Let \( X \) be a set, \( \mathcal{C} \) be a semialgebra on \( X \), and \( \mu : \mathcal{C} \to [0, +\infty] \subset \mathbb{R} \). Assume that

(i) \( \forall C \in \mathcal{C}, \forall n \in \mathbb{Z}_+ \), \( \forall \) pairwise disjoint \( (C_i)_{i=1}^n \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^n C_i \), then \( \mu(C) = \sum_{i=1}^n \mu(C_i) \);

(ii) \( \forall C \in \mathcal{C}, \forall \) pairwise disjoint \( (C_i)_{i=1}^\infty \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^\infty C_i \), then \( \mu(C) \leq \sum_{i=1}^{\infty} \mu(C_i) \).

Then, \( \mu \) admits a unique extension to a measure \( \bar{\mu} \) on the algebra on \( X \), \( \mathcal{A} \), generated by \( \mathcal{C} \).

**Proof** By Proposition 11.31, \( \mathcal{A} = \{ A \subseteq X \mid A \) is the finite disjoint union of sets in \( \mathcal{C} \} \). Define \( \bar{\mu} : \mathcal{A} \to [0, +\infty] \subset \mathbb{R} \) by \( \bar{\mu}(A) = \sum_{i=1}^n \mu(C_i) \), \( \forall A = \bigcup_{i=1}^n C_i \in \mathcal{A} \), where \( n \in \mathbb{Z}_+ \) and \( (C_i)_{i=1}^n \subseteq \mathcal{C} \) is pairwise disjoint. By (i), \( \bar{\mu} \) is well-defined. Clearly \( \bar{\mu}(\emptyset) = 0 \). \( \forall A_1, A_2 \in \mathcal{A} \) with \( A_1 \cap A_2 = \emptyset \), \( A_l = \bigcup_{i=1}^{n_l} C_{l,i}, \) where \( n_l \in \mathbb{Z}_+ \) and \( (C_{l,i})_{i=1}^{n_l} \subseteq \mathcal{C} \) is pairwise disjoint, \( l = 1, 2 \), \( A_1 \cup A_2 = \bigcup_{l=1}^2 \bigcup_{i=1}^{n_l} C_{l,i} \) where \( C_{l,i}, l = 1, 2, i = 1, \ldots, n_l \) are pairwise disjoint. Then, \( \bar{\mu}(A_1 \cup A_2) = \sum_{l=1}^2 \sum_{i=1}^{n_l} \mu(C_{l,i}) = \bar{\mu}(A_1) + \bar{\mu}(A_2) \).

Hence, \( \bar{\mu} \) is finitely additive.

\( \forall (A_i)_{i=1}^\infty \subseteq \mathcal{A} \) with \( A = \bigcup_{i=1}^\infty A_i \in \mathcal{A} \) and \( A_1, A_2, \ldots \) being pairwise disjoint. By finite additivity \( \bar{\mu}(A) = \bar{\mu}(\bigcup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \bar{\mu}(A_i), \forall \mu \in \mathbb{N} \). Then, \( \bar{\mu}(A) \geq \sum_{i=1}^{\infty} \mu(A_i) \). On the other hand, since \( A \in \mathcal{A} \), then \( A = \bigcup_{i=1}^{n_i} C_i, \) where \( n_i \in \mathbb{Z}_+ \) and \( (C_i)_{i=1}^{n_i} \subseteq \mathcal{C} \) is pairwise disjoint. \( \forall i \in \mathbb{N} \), \( A_i = \bigcup_{l=1}^{n_i} C_{i,l} \), where \( n_i \in \mathbb{Z}_+ \) and \( (C_{i,l})_{l=1}^{n_i} \subseteq \mathcal{C} \) is pairwise disjoint. Then, \( \forall i \in \{1, \ldots, n\}, \bar{\mu}(A_i) = \bar{\mu}(A_1 \cap (\bigcup_{j=1}^\infty A_j)) = \sum_{l=1}^{\infty} \sum_{j=1}^{n_i} \mu(C_{j,l} \cap C_i) \).

By (ii), we have \( \mu(C_i) \leq \sum_{j=1}^{\infty} \sum_{i=1}^{n_i} \mu(C_{j,l} \cap C_i) \). Therefore, \( \bar{\mu}(A) = \sum_{i=1}^{n_i} \mu(C_i) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \sum_{i=1}^{\infty} \mu(C_{j,l} \cap C_i) = \sum_{l=1}^{\infty} \sum_{j=1}^{n_i} \sum_{i=1}^{\infty} \mu(C_{j,l} \cap C_i) \).
Hence, \( Y \) is measurable if, and only if, \( \forall \alpha \in \mathbb{R}, the\ set\ \{ x \in D \mid f(x) < \alpha \} \in \mathcal{B} \).

(ii) \( \forall \alpha \in \mathbb{R}, the\ set\ \{ x \in D \mid f(x) \geq \alpha \} \in \mathcal{B} \).

(iii) \( \forall \alpha \in \mathbb{R}, the\ set\ \{ x \in D \mid f(x) \leq \alpha \} \in \mathcal{B} \).

(iv) \( \forall \alpha \in \mathbb{R}, the\ set\ \{ x \in D \mid f(x) \leq \alpha \} \in \mathcal{B} \).

(v) \( \forall \alpha \in \mathbb{R}, the\ set\ \{ x \in D \mid f(x) > \alpha \} \in \mathcal{B} \).
11.3. MEASURABLE FUNCTIONS

These statements imply

\[(vi) \forall \alpha \in \mathbb{R}, \text{the set } \{x \in D \mid f(x) = \alpha \} \in \mathcal{B}.\]

**Proof** The proof is straightforward, and is therefore omitted. \qed

**Proposition 11.36** Let \(X := (X, \mathcal{B})\) be a measurable space, \(D \in \mathcal{B}\), \(f_i : D \to [0, \infty] \subset \mathbb{R}_e\) be \(\mathcal{B}\)-measurable, \(i = 1, 2\), and \(c \in (0, \infty) \subset \mathbb{R}\). Then, \(cf_1\) and \(f_1 + f_2\) are \(\mathcal{B}\)-measurable.

**Proof** \(\forall \alpha \in \mathbb{R}, \{x \in D \mid cf_1(x) < \alpha\} = \{x \in D \mid f_1(x) < \alpha/c\} \in \mathcal{B}\) by Proposition 11.35. Then, by Proposition 11.35, \(cf_1\) is \(\mathcal{B}\)-measurable. \(\{x \in D \mid f_1(x) + f_2(x) < \alpha\} = \bigcup_{r \in \mathbb{Q}}(\{x \in D \mid f_1(x) < r\} \cap \{x \in D \mid f_2(x) < \alpha - r\}) \in \mathcal{B}\) by Proposition 11.35. Then, by Proposition 11.35, \(f_1 + f_2\) is \(\mathcal{B}\)-measurable. This completes the proof of the proposition. \qed

**Proposition 11.37** Let \(X := (X, \mathcal{B}, \mu)\) be a topological measure space, \(\gamma\) be a topological space, \(D \in \mathcal{B}\), and \(f : D \to \gamma\) be continuous. Then, \(f\) is \(\mathcal{B}\)-measurable.

**Proof** The proof is straightforward, and is therefore omitted. \qed

**Proposition 11.38** Let \((X, \mathcal{B})\) be a measurable space, \(\gamma\) and \(Z\) be topological spaces, \(D \in \mathcal{B}\), \(f : D \to \gamma\) be \(\mathcal{B}\)-measurable, and \(g : \gamma \to Z\) be continuous. Then, \(h := g \circ f : D \to Z\) is \(\mathcal{B}\)-measurable.

**Proof** \(\forall O \in \mathcal{O}_Z\), \(g_{\text{inv}}(O) \in \mathcal{O}_\gamma\) since \(g\) is continuous. \(h_{\text{inv}}(O) = f_{\text{inv}}(g_{\text{inv}}(O)) \in \mathcal{B}\). Hence, \(h\) is \(\mathcal{B}\)-measurable. \qed

**Proposition 11.39** Let \((X, \mathcal{B})\) be a measurable space, \(\gamma_i := (Y_i, \mathcal{O}_i)\) be a second countable topological space, \(D \in \mathcal{B}\), \(f_i : D \to \gamma_i\), \(\forall i \in N \subseteq \mathbb{N}\), \(N\) is a countable index set, \(\gamma := (\gamma, \mathcal{O}) := \prod_{i \in \mathbb{N}} \gamma_i\) be the product topological space, and \(f : D \to \gamma\) be given by \(\pi_i(f(x)) = f_i(x)\), \(\forall x \in D, \forall i \in N\). Then, \(f\) is \(\mathcal{B}\)-measurable if, and only if, \(f_i\) is \(\mathcal{B}\)-measurable, \(\forall i \in N\).

**Proof** “Sufficiency” Let \(f_i\)'s be \(\mathcal{B}\)-measurable. \(\forall i \in N\), let \(B_{Y_i}\) be a countable basis for \(\gamma_i\). Without loss of generality, assume \(\emptyset, Y_i \in B_{Y_i}\). Then, \(\gamma\) is second countable with a countable basis \(B_{Y}\) as defined in Proposition 3.28. \(\forall O \in \mathcal{O}, O = \bigcup_{i \in \mathbb{N}} O_{B_i}\) where \(O_{B_i} \in B_{Y_i}, \forall i \in \mathbb{N}\). \(\forall i \in \mathbb{N}\), \(O_{B_i} = \prod_{i \in \mathbb{N}} B_{Y_i}\), where \(B_{Y_i} \in B_{Y_i}, \forall i \in N, B_{Y_i} = Y_i, \forall i \in N \setminus N_i\), and \(N_i \subseteq N\) is a finite set. Then, \(f_{\text{inv}}(O_{B_i}) = \bigcap_{i \in \mathbb{N}} f_{i \text{inv}}(B_{Y_i}) \in \mathcal{B}\) since \(f_i\)'s are \(\mathcal{B}\)-measurable and \(\mathcal{B}\) is a \(\sigma\)-algebra. Then, \(f_{\text{inv}}(O) = f_{\text{inv}}(\bigcup_{i \in N} O_{B_i}) = \bigcup_{i \in N} f_{\text{inv}}(O_{B_i}) \in \mathcal{B}\). Hence, \(f\) is \(\mathcal{B}\)-measurable.

“Necessity” Let \(f\) be \(\mathcal{B}\)-measurable. \(\forall i_0 \in N, \forall O_{i_0} \in \mathcal{O}_{i_0}\), let \(O := \prod_{i \in \mathbb{N}} O_i \in \mathcal{O}\), where \(O_i = Y_i, \forall i \in N\) with \(i \neq i_0\). By the measurability of \(f\), we have \(f_{i_0 \text{inv}}(O_{i_0}) = f_{\text{inv}}(O) \in \mathcal{B}\). Hence, \(f_{i_0}\) is \(\mathcal{B}\)-measurable.

This completes the proof of the proposition. \qed
Lemma 11.43 Let \( D \) be a separable metric space, \( f_n : D \to \mathbb{R} \) be \( B \)-measurable, \( \forall n \in \mathbb{N} \), \( f_S := \limsup_{n \in \mathbb{N}} f_n : D \to \mathbb{R} \) defined by \( f_S(x) = \limsup_{n \in \mathbb{N}} f_n(x), \forall x \in D \), and \( f_I := \liminf_{n \in \mathbb{N}} f_n : D \to \mathbb{R} \) defined by \( f_I(x) = \liminf_{n \in \mathbb{N}} f_n(x), \forall x \in D \). Then, \( \forall n \in \mathbb{Z}_+, \bigcup_{i=1}^{n} f_i, \sup_{i \in \mathbb{N}} f_i, \inf_{i \in \mathbb{N}} f_i, \limsup_{n \in \mathbb{N}} f_n = f_S \), and \( \liminf_{n \in \mathbb{N}} f_n = f_I \) are \( B \)-measurable.

Proof Assume that \( \forall \alpha \in \mathbb{R}, \{ x \in D \mid \bigcup_{i=1}^{n} f_i(x) > \alpha \} = \bigcup_{i=1}^{n} \{ x \in D \mid f_i(x) > \alpha \} \in B \). Hence, by Proposition 11.35, \( \bigcup_{i=1}^{n} f_i \) is \( B \)-measurable.

Proposition 11.44 Let \( (X, \mathcal{B}) \) be a measurable space, \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) be such that \( \bigcup_{i=1}^{\infty} E_i = X \), \( \mathcal{Y} := (Y, \mathcal{O}) \) be a topological space, \( D \in \mathcal{B}, \mathcal{B}_{E_i} := \{ B \in \mathcal{B} \mid B \subseteq E_i \}, \forall i \in \mathbb{N}, \) and \( f : D \to \mathcal{Y} \). Then, \( f|_{E_i} \) is \( \mathcal{B}_{E_i} \)-measurable, \( \forall i \in \mathbb{N}, \) if, and only if, \( f|_{E_i} \) is \( \mathcal{B}_{E_i} \)-measurable, \( \forall i \in \mathbb{N}, \) if, and only if, \( f \) is \( \mathcal{B} \)-measurable.

Proof (i) \( \Rightarrow \) (ii). Assume that \( f|_{E_i} \) is \( \mathcal{B}_{E_i} \)-measurable, \( \forall i \in \mathbb{N}, \forall O \in \mathcal{O} \), by the assumption, we have \( (f|_{E_i})_{inv}(O) \in \mathcal{B}_{E_i} \subseteq \mathcal{B}, \forall i \in \mathbb{N} \). Then, \( f|_{E_i} \) is \( \mathcal{B} \)-measurable.

(ii) \( \Rightarrow \) (iii). Assume that \( f|_{E_i} \) is \( \mathcal{B} \)-measurable, \( \forall i \in \mathbb{N}, \forall O \in \mathcal{O} \), by the assumption, we have \( (f|_{E_i})_{inv}(O) \in \mathcal{B}, \forall i \in \mathbb{N} \). Then, \( f_{inv}(O) = \bigcup_{i=1}^{\infty} (f|_{E_i})_{inv}(O) \in \mathcal{B} \). Hence, \( f \) is \( \mathcal{B} \)-measurable.

(iii) \( \Rightarrow \) (i). Assume that \( f \) is \( \mathcal{B} \)-measurable. \( \forall O \in \mathcal{O} \), by the assumption, \( f_{inv}(O) \in \mathcal{B} \). Then, \( (f|_{E_i})_{inv}(O) = f_{inv}(O) \cap E_i \in \mathcal{B}_{E_i}, \forall i \in \mathbb{N} \). Hence, \( f|_{E_i} \) is \( \mathcal{B}_{E_i} \)-measurable, \( \forall i \in \mathbb{N} \). This completes the proof of the proposition. \( \square \)

Definition 11.42 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space. A property \( P \) is said to hold almost everywhere in \( \mathcal{X} \) (abbreviated \( a.e. \)) if the set of points where it fails to hold or does not make sense belongs to \( \mathcal{B} \) and have measure \( 0 \). We will write \( P \) \( a.e. \) in \( \mathcal{X} \) or \( P(x) \) \( a.e. \) \( x \in \mathcal{X} \).

Lemma 11.43 Let \( \mathcal{X} := (X, \mathcal{B}) \) be a measurable space, \( \mathcal{Y} := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, \) and \( f : D_1 \to \mathcal{Y} \) and \( g : D_2 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Then, the function \( h : D_1 \cap D_2 \to \mathbb{R}, \) defined by \( h(x) = \rho(f(x), g(x)), \forall x \in D_1 \cap D_2 \), is \( \mathcal{B} \)-measurable.

Proof By Proposition 4.4, \( \mathcal{Y} \) is second countable. By Propositions 11.39 and 11.41, the function \( h_1 : D_1 \cap D_2 \to \mathcal{Y} \times \mathcal{Y}, \) defined by \( h_1(x) = (f(x), g(x)), \forall x \in D_1 \cap D_2 \), is \( \mathcal{B} \)-measurable. By Propositions 11.38 and 4.30, the function \( h \) is \( \mathcal{B} \)-measurable. \( \square \)
Lemma 11.44 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_i \in \mathcal{B}$, and $f_i : D_i \to \mathcal{Y}$ be $\mathcal{B}$-measurable, $i = 1, 2, 3$. Assume that $f_1 = f_2$ a.e. in $\mathcal{X}$ and $f_2 = f_3$ a.e. in $\mathcal{X}$. Then, $f_1 = f_3$ a.e. in $\mathcal{X}$.

Proof Let $E_1 := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x) \}$, $E_2 := \tilde{D}_2 \cup D_3 \cup \{ x \in D_2 \cap D_3 \mid f_2(x) \neq f_3(x) \}$, and $E_3 := \tilde{D}_1 \cup \tilde{D}_3 \cup \{ x \in D_1 \cap D_3 \mid f_3(x) \neq f_1(x) \}$. By the assumption, we have $E_1, E_2 \in \mathcal{B}$ and $\mu(E_1) = \mu(E_2) = 0$. Clearly, $E_3 \subseteq E_1 \cup E_2$. Note that $E_3 = \tilde{D}_1 \cup \tilde{D}_3 \cup \{ x \in D_1 \cap D_3 \mid \rho(f_1(x), f_3(x)) > 0 \} \in \mathcal{B}$ by Lemma 11.43. Then, we have $0 \leq \mu(E_3) \leq \mu(E_1) + \mu(E_2) = 0$. Hence, $f_1 = f_3$ a.e. in $\mathcal{X}$. This completes the proof of the lemma.

Lemma 11.45 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}_i := (Y_i, \mathcal{O}_i)$ be a second countable metrizable topological space, $\forall i \in N \subseteq \mathbb{N}$, where $N$ is a countable index set, $\mathcal{Y} := \prod_{i \in N} \mathcal{Y}_i$ be the product space, $D_1, D_2 \in \mathcal{B}$, $f_i : D_i \to \mathcal{Y}_i$ and $g_i : D_i \to \mathcal{Y}_i$ be $\mathcal{B}$-measurable, $\forall i \in N$, $f : D_1 \to \mathcal{Y}$ be defined by $\pi_i(f(x)) = f_i(x)$, $\forall i \in N$, $\forall x \in D_1$, and $g : D_2 \to \mathcal{Y}$ be defined by $\pi_i(g(x)) = g_i(x)$, $\forall i \in N$, $\forall x \in D_2$. Then, $f = g$ a.e. in $\mathcal{X}$ if, and only if, $f_i = g_i$ a.e. in $\mathcal{X}$, $\forall i \in N$.

Proof Define $E_i := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid f_i(x) \neq g_i(x) \}$, $\forall i \in N$, and $E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid f(x) \neq g(x) \}$.

"Sufficiency" Let $f_i = g_i$ a.e. in $\mathcal{X}$, $\forall i \in N$. Then, $E_i \in \mathcal{B}$ and $\mu(E_i) = 0$, $\forall i \in N$. Note that $E = \bigcup_{i \in N} E_i \in \mathcal{B}$ and $\mu(E) = 0$. Hence, $f = g$ a.e. in $\mathcal{X}$.

"Necessity" Let $f = g$ a.e. in $\mathcal{X}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. $\forall i \in N$, $E_i \subseteq E$. Since $\mathcal{Y}_i$ is a second countable metrizable topological space, let $\rho_i : Y_i \times Y_i \to [0, \infty) \subseteq \mathbb{R}$ be the metric on $\mathcal{Y}_i$ whose natural topology is $\mathcal{O}_i$. Then, $(Y_i, \rho_i)$ is a separable metric space, by Proposition 4.4. By Lemma 11.43, we have $E_i = \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid \rho_i(f_i(x), g_i(x)) > 0 \} \in \mathcal{B}$. Then, $0 \leq \mu(E_i) \leq \mu(E) = 0$. Hence, $f_i = g_i$ a.e. in $\mathcal{X}$.

This completes the proof of the lemma.

Lemma 11.46 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a topological space, $\mathcal{Z} := (Z, \rho)$ be a separable metric space, $D_1, D_2 \in \mathcal{B}$, $f_i : D_i \to \mathcal{Y}$ be $\mathcal{B}$-measurable, $i = 1, 2$, $g : \mathcal{Y} \to \mathcal{Z}$ be continuous. Assume that $f_1 = f_2$ a.e. in $\mathcal{X}$. Then, $g \circ f_1 = g \circ f_2$ a.e. in $\mathcal{X}$.

Proof By the assumption, $E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x) \} \in \mathcal{B}$ and $\mu(E) = 0$. By Proposition 11.38, we have $g \circ f_1$ is $\mathcal{B}$-measurable, $i = 1, 2$. By Lemma 11.43, $E_1 \supseteq E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid g(f_1(x)) \neq g(f_2(x)) \} = \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid \rho(g(f_1(x)), g(f_2(x))) > 0 \} \in \mathcal{B}$. Then, $0 \leq \mu(E) \leq \mu(E) = 0$. Hence, $g \circ f_1 = g \circ f_2$ a.e. in $\mathcal{X}$. This completes the proof of the lemma.
Proposition 11.47 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a complete measure space, $\mathcal{Y}$ be a topological space, $D_1, D_2 \in \mathcal{X}^2$, $f : D_1 \to \mathcal{Y}$ be $\mathcal{B}$-measurable, $g : D_2 \to \mathcal{Y}$, and $g = f$ a.e. in $\mathcal{X}$. Then, $g$ is $\mathcal{B}$-measurable.

Proof Let $E := \overline{D_1} \cup \overline{D_2} \cup \{ x \in D_1 \cap D_2 \mid f(x) \neq g(x) \}$. Then, $E \in \mathcal{B}$ and $\mu(E) = 0$. \(\forall O \in \mathcal{O}_Y, \ g_{\text{inv}}(O) = g_{\text{inv}}(O) \cap \big( (g_{\text{inv}}(O) \cap E) \cup (g_{\text{inv}}(O) \cap (X \setminus E)) \big) \). Note that $X \setminus E, f_{\text{inv}}(O) \in \mathcal{B}$. By the completeness of $\mathcal{X}$, $E \supseteq g_{\text{inv}}(O) \cap E \in \mathcal{B}$. Then, $g_{\text{inv}}(O) \in \mathcal{B}$. Hence, $g$ is $\mathcal{B}$-measurable.

Proposition 11.48 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $D \in \mathcal{B}, f_n : D \to \mathcal{Y}$ be $\mathcal{B}$-measurable, \(\forall n \in \mathbb{N}\), and $f : D \to \mathcal{Y}$. Assume that $\lim_{n \to \infty} f_n(x) = f(x)$, \(\forall x \in D\). Then, $f$ is $\mathcal{B}$-measurable.

Proof Fix any open set $O \subseteq \mathcal{Y}$. \(\forall n \in \mathbb{N}\), define $O_n := \{ y \in O \mid B_Y(y, 1/n) \subseteq O \}$. Then, $O_n = O \cap \big( \bigcap_{y \in \mathcal{Y} \setminus O} (Y \setminus B_Y(y, 1/n)) \big)$. Note that $\bigcap_{y \in \mathcal{Y} \setminus O} (Y \setminus B_Y(y, 1/n))$ is a closed set in $\mathcal{Y}$. Then, $O_n \in \mathcal{B}_B(\mathcal{Y})$.

Claim 11.48.1 $f_{\text{inv}}(O) = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_k) = : U$.

Proof of claim: $\forall x \in f_{\text{inv}}(O) \subseteq D$, then $y := f(x) \in O$. Since $O$ is open, then $\exists L \in \mathbb{N}$ such that $B_Y(y, 1/L) \subseteq O$. Then, $\forall l \geq L + 1, \forall \tilde{y} \in B_Y(y, 1/l), \exists \tilde{y} \in B_Y(\tilde{y}, 1/l)$, we have $\rho(y, \tilde{y}) \leq \rho(y, \tilde{y}) + \rho(\tilde{y}, \tilde{y}) < 1$. This implies that $\tilde{y} \in B_Y(y, 1/L) \subseteq O$. Then, $B_Y(\tilde{y}, 1/l) \subseteq O$. Thus, $\tilde{y} \in O_1$ and $B_Y\left(y, \frac{1}{L(L+1)}\right) \subseteq O_1$. Since $\lim_{n \to \infty} f_n(x) = f(x) = y$, then $\exists N \in \mathbb{N}$ such that $\forall n \geq N, \rho(f_n(x), y) < 1$. Then, $f_n(x) \in B_Y\left(y, \frac{1}{L(L+1)}\right) \subseteq O_1$ and $x \in f_{n\text{inv}}(O)$. By arbitrariness of $n, x \in \bigcap_{n=1}^{\infty} f_{n\text{inv}}(O_1) \subseteq \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_k)$ by arbitrariness of $l$, we have $x \in \bigcap_{k=L+1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_k) \subseteq U$. Hence, by the arbitrariness of $x, f_{\text{inv}}(O) \subseteq U$.

On the other hand, \(\forall x \in U, \exists L \in \mathbb{N}\) such that $x \in \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j\text{inv}}(O_L)$. Then, $\exists N \in \mathbb{N}$ such that $\forall n \geq N, x \in f_{n\text{inv}}(O_L)$ and $f_n(x) \in O_L$. Since $\lim_{n \to \infty} f_n(x) = f(x)$, then $\exists N \in \mathbb{N}$ with $\tilde{N} \geq N$ such that $\rho(f_{\tilde{N}}(x), f(x)) < 1$. Then, $f(x) \in B_Y(f_{\tilde{N}}(x), 1/L)$. Note that $f_{\tilde{N}}(x) \in O_L$ and $B_Y(f_{\tilde{N}}(x), 1/L) \subseteq O$. Then, $f(x) \in O$ and $x \in f_{\text{inv}}(O)$. By arbitrariness of $x, U \subseteq f_{\text{inv}}(O)$.

Hence, $f_{\text{inv}}(O) = U$. This completes the proof of the claim.

Proposition 11.49 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a complete measure space, $\mathcal{Y} := (Y, \rho)$ be a metric space, $D_1, D_2 \in \mathcal{X}^2$, $f_n : D_1 \to \mathcal{Y}$ be $\mathcal{B}$-measurable,
11.3. MEASURABLE FUNCTIONS

∀n ∈ N, and f : D_2 → Y. Assume that lim_{n∈N} f_n = f a.e. in X. Then, f is B-measurable.

**Proof** Let E := \{ x ∈ D_1 ∩ D_2 \mid (f_n(x))_{n=1}^{∞} \} does not converge to f(x) \} \cup \widetilde{D}_1 \cup \widetilde{D}_2. By the assumption, E \in B and μ(E) = 0. Let g := f\rceil_{X \setminus E} and g_n := f_n\rceil_{X \setminus E}, ∀n ∈ N. Then, lim_{n∈N} g_n(x) = g(x), ∀x ∈ X \setminus E \in B. By Proposition 11.48, g is B-measurable. ∀ open set O ⊆ Y, f_{inv}(O) = g_{inv}(O) \cup (f_{inv}(O) \cap E). By the completeness of X, we have E ⊇ f_{inv}(O) \cap E \in B. Then, f_{inv}(O) \in B. Hence, f is B-measurable. This completes the proof of the proposition. □

**Proposition 11.50** Let (X, B) be a measurable space, Y := (Y, ρ) be a separable metric space, D_1, D_2 ∈ B, f_n : D_1 → Y be B-measurable, ∀n ∈ N, and f : D_2 → Y be B-measurable. Then, the set

E := \{ x ∈ D_1 ∩ D_2 \mid (f_n(x))_{n=1}^{∞} \} \in B

**Proof** We will show that

\[ E = \bigcap_{m=1}^{∞} \bigcup_{i=1}^{∞} \bigcap_{k=1}^{∞} E_{k,m} =: \tilde{E} \]

where E_{k,m} := \{ x ∈ D_1 ∩ D_2 \mid ρ(f_k(x), f(x)) < 1/m \} \in B, ∀m ∈ N, ∀k ∈ N, by Lemma 11.43. This further implies that \tilde{E} ∈ B.

∀x ∈ E, ∀m ∈ N, ∃k_0 ∈ N, ∀k ≥ k_0, x ∈ D_1 ∩ D_2 and ρ(f_k(x), f(x)) < 1/m. Then, x ∈ E_{k,m} and x ∈ \bigcap_{k=k_0}^{∞} E_{k,m} ≤ \bigcup_{i=1}^{∞} \bigcap_{k=1}^{∞} E_{k,m}. By the arbitrariness of m, we have x ∈ \tilde{E}. This implies that E ⊆ \tilde{E}. On the other hand, ∀x ∈ \tilde{E}, ∀m ∈ N, x ∈ \bigcup_{i=1}^{∞} \bigcap_{k=1}^{∞} E_{k,m}. Then, ∃k_0 ∈ N, such that x ∈ \bigcap_{k=k_0}^{∞} E_{k,m}. ∀k ∈ N with k ≥ k_0, x ∈ E_{k,m}. Then, x ∈ D_1 ∩ D_2 and ρ(f_k(x), f(x)) < 1/m. This leads to lim_{n∈N} f_n(x) = f(x) and x ∈ E. Then, we have \tilde{E} ⊆ E. Hence, E = \tilde{E} ∈ B. This completes the proof of the proposition. □

**Proposition 11.51** Let X := (X, B, μ) be a measure space, Y := (Y, ρ) be a separable complete metric space, D ∈ B, f_i : D → Y be B-measurable, ∀i ∈ N, E := \{ x ∈ D \mid (f_i(x))_{i=1}^{∞} \} converges in Y \} and f : E → Y defined by f(x) = lim_{i∈N} f_i(x), ∀x ∈ E. Then, E = \bigcap_{j=1}^{∞} \bigcup_{m=1}^{∞} \bigcap_{i=m}^{∞} \bigcap_{n=m}^{∞} E_{j,l,n} =: \tilde{E} \in B, where E_{j,l,n} := \{ x ∈ D \mid ρ(f_i(x), f_n(x)) < 1/j \}, ∀j, l, n ∈ N, and f is B-measurable.

**Proof** ∀j, l, n ∈ N, by Lemma 11.43 and the separability of Y, E_{j,l,n} ∈ B. Then, \tilde{E} ∈ B.

∀x ∈ \tilde{E}, ∀j ∈ N, ∃m_0 ∈ N, ∀l, n ∈ N with l ≥ m_0 and n ≥ m_0, we have x ∈ D and ρ(f_l(x), f_n(x)) < 1/j. Then, (f_i(x))_{i=1}^{∞} ⊆ Y is a Cauchy sequence. By the completeness of Y, (f_i(x))_{i=1}^{∞} converges in Y. Hence, x ∈ E. By the arbitrariness of x, we have \tilde{E} ⊆ E.
Proposition 11.52 Let \( a.e. \) in \( \bigcup E \) Proposition 3.67, we have \( E \) arbitrari-ness of \( \forall \) be as defined above, \( E \) the proof of the proposition.

Let \( \rho \) \( \rho \) the "Sufficiency" Assume \( \lim_{n \to \infty} f_n(x) = f(x) \), \( \forall x \in E \). By Proposition 11.41, \( f_i|_E \) is \( \mathcal{B} \)-measurable, \( \forall i \in \mathbb{N} \). By Proposition 11.48, \( f \) is \( \mathcal{B} \)-measurable.

This completes the proof of the proposition. \( \square \)

Proposition 11.53 Let \( X := (Y, \mathcal{B}, \mu) \) be a measure space, \( Z := (Z, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, f_n : D_1 \to Y \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, f : D_2 \to Y \) be \( \mathcal{B} \)-measurable, and \( g : Y \to Z \) be continuous. Assume that \( \lim_{n \to \infty} f_n = f \) \( \mu \)-a.e. in \( X \). Then, \( \lim_{n \to \infty} g \circ f_n = g \circ f \) \( \mu \)-a.e. in \( X \).

Proof Let \( E := \{ x \in D_1 \cap D_2 \mid \lim_{n \to \infty} f_n(x) = f(x) \} \) and \( \bar{E} := \{ x \in D_1 \cap D_2 \mid \lim_{n \to \infty} g(f_n(x)) = g(f(x)) \} \). By Proposition 3.66, we have \( E \subseteq \bar{E} \). By the assumption, \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). By Proposition 11.38, \( g \circ f_n \) and \( g \circ f \) are \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). By Proposition 11.50, \( \bar{E} \in \mathcal{B} \). Then, \( 0 \leq \mu(\bar{E}) \leq \mu(E) = 0 \). Hence, \( \lim_{n \to \infty} g \circ f_n = g \circ f \) \( \mu \)-a.e. in \( X \). This completes the proof of the proposition. \( \square \)
11.3. MEASURABLE FUNCTIONS

Proposition 11.54 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} \) be a separable metric space, \( D_i \in \mathcal{B}, \ i = 1, 2, 3, 4, \ f_n : D_1 \to \mathcal{Y} \) and \( g_n : D_2 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( f : D_3 \to \mathcal{Y} \) and \( g : D_4 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \to \infty} f_n = f \ \text{a.e. in} \ \mathcal{X}, \ f_n = g_n \ \text{a.e. in} \ \mathcal{X}, \ \forall i \in \mathbb{N}, \) and \( f = g \ \text{a.e. in} \ \mathcal{X}. \) Then, \( \lim_{n \to \infty} g_n = g \ \text{a.e. in} \ \mathcal{X}. \) This completes the proof of the proposition.

Proof By the assumption, we have \( E_1 := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_3 \mid (f_n(x))^\infty_{n=1} \) does not converge to \( f(x) \}\) \( \in \mathcal{B}; \ E_n := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid f_n(x) \neq g_n(x) \} \in \mathcal{B}, \ \forall n \in \mathbb{N}; \ E := \tilde{D}_3 \cup \tilde{D}_4 \cup \{ x \in D_3 \cap D_4 \mid f(x) \neq g(x) \} \in \mathcal{B}; \) and \( \mu(E_1) = \mu(E_n) = \mu(E) = 0, \ \forall n \in \mathbb{N}. \) Then, \( E := E_1 \cup (\bigcup_{n=1}^{\infty} E_n) \cup \tilde{E} \in \mathcal{B} \) and \( \mu(E) = 0. \) Let \( E_2 := \tilde{D}_2 \cup D_1 \cup \{ x \in D_2 \cap D_4 \mid (g_n(x))^\infty_{n=1} \) does not converge to \( g(x) \}\). By Proposition 11.50, \( E_2 \in \mathcal{B}. \) Clearly, \( \tilde{E} \subseteq \tilde{E}_2 \) and \( E_2 \subseteq E. \) Then, \( 0 \leq \mu(E_2) \leq \mu(E) = 0. \) This shows that \( \lim_{n \to \infty} g_n = g \ \text{a.e. in} \ \mathcal{X}. \) This completes the proof of the proposition.

Theorem 11.55 (Egoroff’s Theorem) Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, \ f_n : D_1 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, \) and \( f : D_2 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \to \infty} f_n = f \ \text{a.e. in} \ \mathcal{X}. \) Then, \( \forall \eta \in (0, \infty) \subseteq \mathbb{R}, \exists A \in \mathcal{B} \) with \( \mu(A) < \eta \) and \( A \supseteq \tilde{D}_1 \cup \tilde{D}_2 \) such that \( \left( f_n|_{\mathcal{X}\setminus A} \right)^\infty_{n=1} \) converges uniformly to \( f|_{\mathcal{X}\setminus A}. \)

Proof Let \( E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid (f_n(x))^\infty_{n=1} \) does not converge to \( f(x) \}\). By the assumption, \( E \in \mathcal{B} \) and \( \mu(E) = 0. \)

\( \forall \eta \in (0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N}, \forall m \in \mathbb{N}, \) let

\[
G_{m,n} := \{ x \in X \setminus E \mid \rho(f_m(x), f(x)) \geq 2^{-n} \}
\]

By Lemma 11.43, \( G_{m,n} \in \mathcal{B}. \) Let

\[
H_{m,n} := \bigcup_{j=m}^{\infty} G_{j,n} = \{ x \in X \setminus E \mid \rho(f_i(x), f(x)) \geq 2^{-n} \ \text{for some} \ i \geq m \}
\]

Then, \( H_{m+1,n} \subseteq H_{m,n} \subseteq \mathcal{B}. \) \( \forall x \in X \setminus E, \) we have \( \lim_{i \to \infty} f_j(x) = f(x), \) then \( \exists m_x \in \mathbb{N} \) such that \( x \notin H_{m_x,n}. \) Hence, \( \bigcap_{m=1}^{\infty} H_{m,n} = \emptyset. \) Note that \( \mu(H_{1,n}) \leq \mu(X) < \infty. \) By Proposition 11.5, \( \lim_{m \to \infty} \mu(H_{m,n}) = \mu(\bigcap_{m=1}^{\infty} H_{m,n}) = 0. \) Then, \( \exists m_n \in \mathbb{N} \) such that \( \mu(H_{m,n}) < 2^{-n} \eta. \) Let \( A_n := H_{m,n} \in \mathcal{B}. \) Then, \( \mu(A_n) < 2^{-n} \eta \) and \( \forall i \geq m_n, \forall x \in (X \setminus E) \setminus A_n, \) we have \( \rho(f_i(x), f(x)) < 2^{-n}. \)

Let \( A := E \cup (\bigcup_{n=1}^{\infty} A_n) \in \mathcal{B}. \) Clearly, \( A \supseteq \tilde{D}_1 \cup \tilde{D}_2. \) By countable subadditivity of measure, we have \( \mu(A) \leq \mu(E) + \sum_{n=1}^{\infty} \mu(A_n) < \eta. \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N} \) with \( 0 < 2^{-n_0} \leq \epsilon. \forall j \in \mathbb{N} \) with \( m_{n_0} \leq j, \forall x \in X \setminus A \subseteq (X \setminus E) \setminus A_n, \) we have \( \rho(f_j(x), f(x)) < 2^{-n_0} \leq \epsilon. \) Hence,
\[
\left(f_n|_{X \setminus A}\right)_{n=1}^{\infty} \text{ converges uniformly to } f|_{X \setminus A}. \text{ This completes the proof of the proposition.}
\]

**Definition 11.56** Let \( X := (\mathcal{X}, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} := (Y, \rho) \) be a separable metric space, \( D_n \in \mathcal{B}, f_n : D_n \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, D \in \mathcal{B}, \) and \( f : D \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. We will say that \( (f_n)_{n=1}^{\infty} \) converges to \( f \) in measure in \( \mathcal{X} \) if \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \mathcal{N} \in \mathbb{N} \) such that \( \forall n \in \mathbb{N} \) with \( n \geq \mathcal{N} \), we have \( \mu(A_{n, \varepsilon}) := \mu(D_n \cup \overline{D} \cup \{ x \in D_n \cap D \mid \rho(f_n(x), f(x)) \geq \varepsilon \}) < \varepsilon \). In this case, we will write \( \lim_{n \to \infty} f_n = f \) in measure in \( \mathcal{X} \).

Note that in the above definition, \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N}, A_{n, \varepsilon} \in \mathcal{B} \) by Lemma 11.43. Hence, the definition is well-defined.

**Proposition 11.57** Let \( X := (\mathcal{X}, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, f_n : D_1 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, \) and \( f : D_2 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \to \infty} f_n = f \) in measure in \( \mathcal{X} \). Then, there exists a subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) such that \( \lim_{k \to \infty} f_{n_k} = f \) a.e. in \( \mathcal{X} \).

**Proof** Define \( A_{n,k} := \overline{D}_1 \cup \overline{D}_2 \cup \{ x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \geq 2^{-k} \}, \forall n, k \in \mathbb{N} \). By Lemma 11.43, \( A_{n,k} \in \mathcal{B} \). By the assumption, \( \exists n_1 \in \mathbb{N} \) such that \( \forall n \in \mathbb{N} \) with \( n \geq n_1 \), we have \( \mu(A_{n,1}) < 2^{-1} \). \( \forall k \in \mathbb{N} \) with \( k > 1 \), by the assumption, \( \exists n_k \in \mathbb{N} \) with \( n_k > n_{k-1} \) such that \( \forall n \in \mathbb{N} \) with \( n \geq n_k \), we have \( \mu(A_{n,k}) < 2^{-k} \). This defines a subsequence \( (f_{n_k})_{k=1}^{\infty} \).

Let \( A := \bigcap_{i=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n_k,k} \in \mathcal{B} \). Then, \( \mu(A) \leq \mu(\bigcup_{k=1}^{\infty} A_{n_k,k}) \leq \sum_{k=1}^{\infty} \mu(A_{n_k,k}) < 2^{-i+1}, \forall i \in \mathbb{N} \). Hence, \( \mu(A) = 0 \). \( \forall x \in \mathcal{X} \setminus A = \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n_k,k}, \forall i_0 \in \mathbb{N} \) such that \( \forall k \in \mathbb{N} \) with \( k \geq i_0, x \in A_{n_k,k} \). Then, we have \( x \in D_1 \cap D_2 \) and \( \rho(f_{n_k}(x), f(x)) < 2^{-k} \). Hence, \( \lim_{k \to \infty} f_{n_k}(x) = f(x) \). Then, \( E := \overline{D}_1 \cup \overline{D}_2 \cup \{ x \in D_1 \cap D_2 \mid (f_{n_k}(x))_{k=1}^{\infty} \text{ does not converge to } f(x) \} \subseteq A \). By Proposition 11.50, we have \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). Therefore, \( \lim_{k \to \infty} f_{n_k} = f \) a.e. in \( \mathcal{X} \). This completes the proof of the proposition.

**Proposition 11.58** Let \( X := (\mathcal{X}, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, f_n : D_1 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, \) and \( f : D_2 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \to \infty} f_n = f \) a.e. in \( \mathcal{X} \). Then, \( \lim_{n \to \infty} f_n = f \) in measure in \( \mathcal{X} \).

**Proof** \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R} \), by Egoroff’s Theorem 11.55, \( \exists A \in \mathcal{B} \) with \( \mu(A) < \varepsilon \) and \( A \supseteq \overline{D}_1 \cup \overline{D}_2 \) such that \( \left(f_n|_{X \setminus A}\right)_{n=1}^{\infty} \) converges uniformly to \( f|_{X \setminus A} \). Then, \( \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n \geq n_0 \), we have \( \rho(f_n(x), f(x)) < \varepsilon \), \( \forall x \in X \setminus A \). Then, \( A_{n,\varepsilon} := \overline{D}_1 \cup \overline{D}_2 \cup \{ x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \geq \varepsilon \} \subseteq A \). By Lemma 11.43, \( A_{n,\varepsilon} \in \mathcal{B} \) and \( \mu(A_{n,\varepsilon}) \leq \mu(A) < \varepsilon \). This shows that \( \lim_{n \to \infty} f_n = f \) in measure in \( \mathcal{X} \). This completes the proof of the proposition.
11.3. MEASURABLE FUNCTIONS

Proposition 11.59 Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y := (Y, \rho) \) be a separable metric space, \( D_n \in \mathcal{B} \), \( f_n : D_n \rightarrow Y \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), \( D \in \mathcal{B} \), and \( f : D \rightarrow Y \) be \( \mathcal{B} \)-measurable. Then, \( \lim_{n \in \mathbb{N}} f_n = f \) in measure in \( X \) if, and only if, every subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) admits a subsequence \( (f_{n_{k_j}})_{j=1}^{\infty} \) that converges to \( f \) in measure.

Proof

"Necessity" Let \( (f_n)_{n=1}^{\infty} \) converges to \( f \) in measure and \( (f_{n_k})_{k=1}^{\infty} \) be a subsequence of \( (f_n)_{n=1}^{\infty} \). Clearly, \( (f_{n_k})_{k=1}^{\infty} \) converges to \( f \) in measure, and it is a subsequence of itself. Hence, the result holds.

"Sufficiency" Assume that every subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) admits a subsequence \( (f_{n_{k_j}})_{j=1}^{\infty} \) that converges to \( f \) in measure. We will prove the result using an argument of contradiction. Suppose \( (f_n)_{n=1}^{\infty} \) does not converge to \( f \) in measure. Then, \( \exists \epsilon_0 \in (0, \infty) \subset \mathbb{R} \), \( \forall n_0 \in \mathbb{N} \), \( \exists n \in \mathbb{N} \) with \( n \geq n_0 \) such that \( \mu(A_n) := \mu(D_n \cap \{ x \in D_n \cap D : |f_n(x) - f(x)| \geq \epsilon_0 \}) \geq \epsilon_0 \), where \( A_n \in \mathcal{B} \) by Lemma 11.43. Then, \( \exists n_1 \in \mathbb{N} \) such that \( \mu(A_{n_1}) \geq \epsilon_0 \). \( \forall k \in \mathbb{N} \), \( \exists n_{k+1} \in \mathbb{N} \) with \( n_{k+1} > n_k \) such that \( \mu(A_{n_{k+1}}) \geq \epsilon_0 \). This defines a subsequence \( (f_{n_k})_{k=1}^{\infty} \), which satisfies that \( \forall k \in \mathbb{N} \), \( \mu(A_{n_k}) \geq \epsilon_0 \). Clearly, there does not exist a subsequence of \( (f_{n_k})_{k=1}^{\infty} \) that converges to \( f \) in measure. This contradicts with the assumption. Therefore, \( (f_n)_{n=1}^{\infty} \) converges to \( f \) in measure.

This completes the proof of the proposition.

Proposition 11.60 Let \( X := (X, \mathcal{B}, \mu) \) be a finite measure space, \( Y := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B} \), \( f_n : D_1 \rightarrow Y \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( f : D_2 \rightarrow Y \) be \( \mathcal{B} \)-measurable. Then, \( \lim_{n \in \mathbb{N}} f_n = f \) in measure in \( X \) if, and only if, every subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) admits a subsequence \( (f_{n_{k_j}})_{j=1}^{\infty} \) that converges to \( f \) a.e. in \( X \).

Proof

"Necessity" Let \( (f_n)_{n=1}^{\infty} \) converge to \( f \) in measure. Let \( (f_{n_k})_{k=1}^{\infty} \) be a subsequence of \( (f_n)_{n=1}^{\infty} \). Clearly, \( (f_{n_k})_{k=1}^{\infty} \) converges to \( f \) in measure. By Proposition 11.57, there exists a subsequence \( (f_{n_{k_j}})_{j=1}^{\infty} \) of \( (f_{n_k})_{k=1}^{\infty} \) that converges to \( f \) a.e. in \( X \).

"Sufficiency" Assume that every subsequence \( (f_{n_k})_{k=1}^{\infty} \) of \( (f_n)_{n=1}^{\infty} \) admits a subsequence \( (f_{n_{k_j}})_{j=1}^{\infty} \) that converges to \( f \) a.e. in \( X \). By Proposition 11.58, \( (f_{n_{k_j}})_{j=1}^{\infty} \) converges to \( f \) in measure. By Proposition 11.59, we have \( (f_n)_{n=1}^{\infty} \) converges to \( f \) in measure in \( X \).

This completes the proof of the proposition.

Definition 11.61 Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y := (Y, \rho) \) be a separable metric space, \( D_n \in \mathcal{B} \), and \( f_n : D_n \rightarrow Y \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). We will say that \( (f_n)_{n=1}^{\infty} \) is Cauchy in measure in \( X \) if \( \forall \epsilon \in \)
(0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall n, m \in \mathbb{N} with n \geq n_0 and m \geq n_0, we have 
\mu(D_n \cup D_m \cup \{ x \in D_n \cap D_m \mid \rho(f_n(x), f_m(x)) \geq \epsilon \}) < \epsilon.

Clearly, a sequence of measurable functions that converges in measure to a measurable function is Cauchy in measure. The converse is established in the following proposition.

**Proposition 11.62** Let \( X := (X, B, \mu) \) be a measure space, \( Y := (Y, \rho) \) be a separable complete metric space, \( D_n \in B \), and \( f_n : D_n \to Y \) be \( B \)-measurable, \( \forall n \in \mathbb{N} \). Assume that \( (f_n)_{n=1}^\infty \) is Cauchy in measure in \( X \). Then, \( \exists f : D \to Y \), where \( D := \bigcup_{n=1}^\infty D_n \in B \) such that \( f \) is \( B \)-measurable and \( \lim_{n \to \infty} f_n = f \) in measure in \( X \).

**Proof** Let \( n_0 := 0 \). \( \forall i \in \mathbb{N} \), by the assumption, \( \exists n_i \in \mathbb{N} \) with \( n_i > n_{i-1} \) such that \( \forall n, m \in \mathbb{N} with n \geq n_i \) and \( m \geq n_i \), we have \( \mu(D_n \cup D_m \cup \{ x \in D_n \cap D_m \mid \rho(f_n(x), f_m(x)) \geq 2^{-i} \}) < 2^{-i} \). This defines a subsequence \( (f_{n_i})_{i=1}^\infty \) of \( (f_n)_{n=1}^\infty \). \( \forall i \in \mathbb{N} \), let \( A_i := D_{n_i} \cup D_{n_{i+1}} \cup \{ x \in D_{n_i} \cap D_{n_{i+1}} \mid \rho(f_{n_i}(x), f_{n_{i+1}}(x)) \geq 2^{-i} \} \). Then, \( A_i \in B \) and \( \mu(A_i) < 2^{-i} \). Let \( A := \bigcap_{i=1}^\infty \bigcup_{j=i}^\infty A_j \in B \). Then, \( \mu(A) \leq \mu(\bigcup_{j=i}^\infty A_j) \leq \sum_{j=i}^\infty \mu(A_j) < 2^{-i+1}, \forall i \in \mathbb{N} \). Hence, \( \mu(A) = 0 \). \( \forall x \in X \setminus A = \bigcup_{i=1}^\infty \bigcap_{j=i}^\infty A_j \subseteq D, \exists i_0 \in \mathbb{N} \) such that \( x \in A_i \) and \( \rho(f_{n_i}(x), f_{n_{i+1}}(x)) < 2^{-i}, \forall i \in \mathbb{N} \) with \( i \geq i_0 \). Then, \( (f_{n_i}(x))_{i=i_0}^\infty \subseteq Y \) is a Cauchy sequence. By the completeness of \( Y \), \( \exists f(x) \in Y \) such that \( \lim_{i \to \infty} f_{n_i}(x) = f(x) \). \( \forall x \in A \cap D \), we assign \( f(x) := f_{n}(x) \in Y \), where \( n \in \mathbb{N} \) and \( x \in D_n \subseteq D \) are fixed independent of \( x \). This defines a function \( f : D \to Y \) that satisfies \( \lim_{i \to \infty} f_{n_i}(x) = f(x), \forall x \in X \setminus A \).

\( \forall i \in \mathbb{N} \), define \( \tilde{f}_{n_i} : D \to Y \) by \( \tilde{f}_{n_i}(x) := \begin{cases} f_{n_i}(x) & x \in D_{n_i} \\ f(x) & x \in D \setminus D_{n_i} \end{cases}, \forall x \in D \).

By Proposition 11.41, \( \tilde{f}_{n_i} \) is \( B \)-measurable. Clearly, \( \lim_{i \to \infty} \tilde{f}_{n_i}(x) = f(x), \forall x \in D \). By Proposition 11.48, \( f \) is \( B \)-measurable.

\( \forall i \in \mathbb{N} \), \( \forall x \in \bigcup_{j=i}^\infty A_j = \bigcap_{j=i}^\infty A_j \subseteq X \setminus A \), we have \( \rho(f_{n_i}(x), f_{n_{i+1}}(x)) < 2^{-i}, \forall j \geq i \). Then, by Propositions 4.30, 3.66, and 3.67, \( \rho(f_{n_i}(x), f(x)) = \lim_{i \to \infty} \rho(f_{n_i}(x), f_{n_i}(x)) \leq \lim_{i \to \infty} \sum_{j=i}^{i-1} \rho(f_{n_j}(x), f_{n_{j+1}}(x)) < 2^{-i+1} \). Then, \( E_i := D_{n_i} \cup \{ x \in D_{n_i} \mid \rho(f_{n_i}(x), f(x)) \geq 2^{-i+1} \} \subseteq \bigcup_{j=i}^\infty A_j \). By Lemma 11.43, \( E_i \in B \) and \( \mu(E_i) \leq \sum_{j=i}^\infty \mu(A_j) < 2^{-i+1} \).

\( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists i_0 \in \mathbb{N} \) such that \( 2^{-i_0+1} < \epsilon/2 \). \( \forall n \in \mathbb{N} \) with \( n \geq n_{i_0} \), we have \( \mu(D_n \cup \{ x \in D_n \mid \rho(f_n(x), f(x)) \geq \epsilon \}) \leq \mu(D_n \cup D_{n_0} \cup \{ x \in D_n \cap D_{n_0} \mid \rho(f_n(x), f_{n_0}(x)) + \rho(f_{n_0}(x), f(x)) \geq \epsilon \}) \leq \mu(D_n \cup D_{n_0} \cup \{ x \in D_n \cap D_{n_0} \mid \rho(f_n(x), f_{n_0}(x)) \geq \epsilon/2 \}) + \mu(D_{n_0} \cup \{ x \in D_{n_0} \mid \rho(f_{n_0}(x), f(x)) \geq \epsilon/2 \}) < 2^{-i_0} + 2^{-i_0+1} < \epsilon \). Hence, \( \lim_{n \to \infty} f_n = f \) in measure in \( X \). This completes the proof of the proposition. \( \Box \)
11.3. MEASURABLE FUNCTIONS

**Definition 11.63** Let \( X \) be a set and \( A \subseteq X \). The indicator function \( \chi_{A,X} : X \to \{0,1\} \subseteq \mathbb{R} \) is defined by \( \chi_{A,X}(x) = \begin{cases} 1 & x \in A \\ 0 & x \in X \setminus A \end{cases}, \forall x \in X \).

**Definition 11.64** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} \) be a normed linear space. \( \phi : X \to \mathcal{Y} \) is said to be a simple function if \( \exists n \in \mathbb{Z}_+ \), \( \exists y_1, \ldots, y_n \in \mathcal{Y} \), and \( \exists A_1, \ldots, A_n \in \mathcal{B} \) with \( \mu(A_i) < +\infty, i = 1, \ldots, n, \) such that \( \phi(x) = \sum_{i=1}^{n} y_i \chi_{A_i}(x), \forall x \in X \). We will say that a simple function \( \phi \) is in canonical representation if \( y_1, \ldots, y_n \) are distinct and none equals to \( \vartheta_y \), and \( A_1, \ldots, A_n \) are nonempty and pairwise disjoint.

Clearly, every simple function admits a unique canonical representation.

**Proposition 11.65** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} := (\mathcal{Y}, \mathcal{K}, \| \cdot \|) \) be a separable normed linear space, and \( f : X \to U \subseteq \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Then,

(i) \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \) a simple function \( \phi_\epsilon : X \to U \) such that \( \mu(\{ x \in X : \| f(x) - \phi_\epsilon(x) \| \geq \epsilon \}) < \epsilon \);

(ii) there exists a sequence of simple functions \( (\varphi_n)_{n=1}^\infty, \varphi_n : X \to U, \forall n \in \mathbb{N}, \) that converges to \( f \) in measure.

**Proof** (i) By Proposition 4.38, \( U \), considered as a subspace of the metric space \( \mathcal{Y} \), is separable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( U = \emptyset \); Case 2: \( U \neq \emptyset \). Case 1: \( U = \emptyset \). Then, \( X = \emptyset, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \) a simple function \( \phi_\epsilon : X \to U \) to be the simple function with \( \epsilon = 0 \) as in Definition 11.64. Then, \( \mu(\{ x \in X : \| f(x) - \phi_\epsilon(x) \| \geq \epsilon \}) = \mu(\emptyset) = 0 < \epsilon \). Hence, (i) holds. Case 2: \( U \neq \emptyset \). Then, \( \exists (y_i)_{i=1}^\infty \subseteq U \) such that \( \forall y \in U, \forall \delta \in (0, \infty) \subseteq \mathbb{R}, \exists i \in \mathbb{N}, \exists \delta \in \mathbb{R} : \| y - y_i \| < \delta \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists U \subseteq \bigcup_{i=1}^\infty B_y(y, \epsilon) \in : = \bigcup_{i=1}^\infty V_i. \) Define \( V_i := V_i \in B_y(y_i) \), \( V_{i+1} := V_{i+1} \setminus (\bigcup_{j=1}^i V_j) \in B_y(y), \forall i \in \mathbb{N} \). Clearly, \( U \subseteq \bigcup_{i=1}^\infty V_i \) and \( V_1, V_2, \ldots \) are pairwise disjoint. Then, by Propositions 2.5 and 11.34, \( X = f^{-1}(U) = \bigcup_{i=1}^\infty f^{-1}(V_i) \) and \( (f^{-1}(V_i))_{i=1}^\infty \subseteq \mathcal{B} \) and are pairwise disjoint. Then, \( \mu(X) = \sum_{i=1}^\infty \mu(f^{-1}(V_i)) < +\infty \). This implies that \( \exists n_0 \in \mathbb{N} \) such that \( \sum_{i=n_0+1}^\infty \mu(f^{-1}(V_i)) < \epsilon \). Let \( A_i := f^{-1}(V_i) \in \mathcal{B}, i = 1, \ldots, n_0 \), and \( A_{n_0+1} := X \setminus (\bigcup_{i=1}^{n_0} A_i) \in \mathcal{B} \). Define the simple function \( \phi_\epsilon : X \to U \) by \( \phi_\epsilon(x) = \sum_{i=1}^{n_0} y_i \chi_{A_i}(x), \forall x \in X \). \( \forall x \in X, \exists i_x \in \{1, \ldots, n_0 \} \) such that \( x \in A_{i_x} \). Then, \( \phi_\epsilon(x) = y_{i_x} \in U \). \( \forall x \in X \setminus A_{n_0+1}, i_x \in \{1, \ldots, n_0 \} \) and \( x \in A_{i_x} = f^{-1}(V_{i_x}) \subseteq f^{-1}(V_{i_x}). \) Then, \( \phi_\epsilon(x) = y_{i_x} \) and \( f(x) \in V_{i_x} = B_y(y_{i_x}, \epsilon) \). Hence, \( \| f(x) - \phi_\epsilon(x) \| < \epsilon \). Then, \( A := \{ x \in X : \| f(x) - \phi_\epsilon(x) \| \geq \epsilon \} \subseteq A_{n_0+1} = \bigcup_{i=n_0+1}^\infty f^{-1}(V_i) \). By Lemma 11.43, \( A \in \mathcal{B} \). Then, \( \mu(A) \leq \sum_{i=n_0+1}^\infty \mu(f^{-1}(V_i)) < \epsilon \). Hence, (i) holds. In both cases, we have proved (i).

(ii) \( \forall n \in \mathbb{N}, \) let \( \varphi_n := \phi_{2^{-n}} \). Then, \( (\varphi_n)_{n=1}^\infty \) converges to \( f \) in measure. This completes the proof of the proposition. □
Proposition 11.66 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathcal{Y} \) be a separable normed linear space, \( U \subseteq \mathcal{Y} \) be a conic segment, and \( f : X \to U \subseteq \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Then, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^{\infty}, \varphi_n : X \to U, \forall n \in \mathbb{N} \), such that \( \lim_{n \to \infty} \varphi_n = f \) a.e. in \( \mathcal{X} \) and \( \| \varphi_n(x) \| \leq \| f(x) \|, \forall x \in X, \forall n \in \mathbb{N} \).

Proof We will first consider the special case \( \mathcal{Y} = \mathbb{R} \) and \( \mathcal{X} \) is finite. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by Proposition 11.5, \( \exists M_\epsilon \in \mathbb{N} \) such that \( \mu(\{ x \in X \mid |f(x)| \geq M_\epsilon \epsilon \}) =: \mu(A_\epsilon) < \epsilon \). Let \( I_i := \{-(M_i - i + 1)\epsilon, -(M_i - i)\epsilon \} \subset \mathbb{R}, i = 1, \ldots, 2M_i \) and \( A_i := \text{fix}(I_i) \in \mathcal{B} \). Note that \( \mu(A_i) \leq \mu(X) < +\infty, i = 1, \ldots, 2M_i \) and \( A_i \)'s are pairwise disjoint. Define \( \phi_i : X \to U, |\phi_i(x)| \leq |f(x)|, \forall x \in X, \) and \( \mu(\{ x \in X \mid |f(x) - \phi_i(x)| > \epsilon \}) \leq \mu(A_i) < \epsilon \). Then, the sequence of simple functions \( (\phi_{2^n})_{n=1}^{\infty} \) converges to \( f \) in measure in \( \mathcal{X} \). By Proposition 11.57, there exists a subsequence \( (\varphi_k)_{k=1}^{\infty} = (\phi_{2^{-n_k}})_{k=1}^{\infty} \) such that \( \lim_{k \to \infty} \varphi_k = f \) a.e. in \( \mathcal{X} \). The result holds in this special case.

Next, we consider the special case \( \mathcal{X} \) is finite. By Proposition 11.65, there exists a sequence of simple functions \( (\psi_1)_{i=1}^{\infty}, \psi_i : X \to U, \forall i \in \mathbb{N} \), that converges to \( f \) in measure. By Proposition 11.57, there exists a subsequence \( (\psi_{n_i})_{i=1}^{\infty} = (\psi_{i_1})_{n_1=1}^{\infty} \) such that \( \lim_{n_i \to \infty} \psi_{n_i} = f \) a.e. in \( \mathcal{X} \). By Propositions 7.21 and 11.38, \( \mathcal{P} \circ f \) is \( \mathcal{B} \)-measurable, where \( \mathcal{P} \circ f : X \to [0, \infty) \subset \mathbb{R} \) is defined by \( \mathcal{P} \circ f(x) = \| f(x) \|, \forall x \in X \).\(^1\) By the previous special case, there exists a sequence of simple functions \( (\phi_{n_k})_{k=1}^{\infty}, \phi_n : X \to [0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N} \), such that \( \lim_{n \to \infty} \phi_n = \mathcal{P} \circ f \) a.e. in \( \mathcal{X} \) and \( 0 \leq \phi_n(x) \leq \mathcal{P} \circ f(x), \forall x \in X, \forall n \in \mathbb{N} \). Let \( E_1 := \{ x \in X \mid (\psi_{n_1}(x))_{n_1=1}^{\infty} \) does not converge to \( f(x) \} \) and \( E_2 := \{ x \in X \mid (\phi_n(x))_{n=1}^{\infty} \) does not converge to \( \| f(x) \| \} \). Then, \( E_1, E_2 \in \mathcal{B} \) and \( \mu(E_1) = \mu(E_2) = 0 \). Fix any \( n \in \mathbb{N} \). Let \( \psi_n \) admit the canonical representation \( \psi_n = \sum_{j=1}^{n} y_j \chi_{A_j}, x \), where \( n \in \mathbb{Z}_+, y_1, \ldots, y_n \in U \) are distinct and none equals to \( \psi_y, A_1, \ldots, A_n \in \mathcal{B} \) are pairwise disjoint, noneempty, and of finite measure. Let \( y_{n+1} := \psi_y \in U \) and \( A_{n+1} := X \setminus (\bigcup_{j=1}^{n} A_j) \in \mathcal{B} \).

Then, \( \psi_n = \sum_{j=1}^{n+1} y_j \chi_{A_j}, x \). Let \( \phi_n \) admit the canonical representation \( \sum_{j=1}^{n} a_j \chi_{A_j}, x \), where \( n \in \mathbb{Z}_+, a_1, \ldots, a_n \in [0, \infty) \subset \mathbb{R} \) are distinct and none equals to \( 0, A_1, \ldots, A_n \in \mathcal{B} \) are pairwise disjoint, noneempty, and of finite measure. Let \( a_{n+1} := 0 \) and \( A_{n+1} := X \setminus (\bigcup_{j=1}^{n} A_j) \in \mathcal{B} \).

Then, \( \phi_n = \sum_{j=1}^{n+1} a_j \chi_{A_j}, x \). Define a simple function \( \varphi_n : X \to U \) by \( \varphi_n(x) = \begin{cases} y_j & \text{if } y_j \leq a_l, \forall x \in A_j \cap \tilde{A}_l, j = 1, \ldots, n + 1, \lfloor l = 1, \ldots, n + 1 \rfloor. \end{cases} \)

Clearly, \( \varphi_n(x) \in U, \forall x \in X \). Alternatively, we

\(^1\)This notation will be used throughout the rest of the notes. With this notation, we distinguish that \( \| f \| \) is a nonnegative real number and \( \mathcal{P} \circ f \) is a nonnegative real valued function.
Proposition 11.67 completes the proof of the proposition.

Then, we have \( \|\phi_n(x)\| \leq \phi_n(x) \leq \mathcal{P} \circ f(x), \forall x \in X, \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n(x) = f(x), \forall x \in X \setminus (E_1 \cup E_2) \). Let \( E := \{ x \in X \mid (\phi_n(x))_{n=1}^{\infty} \) does not converge to \( f(x) \} \). Then, \( E \subseteq E_1 \cup E_2 \). By Proposition 11.50, \( E \in \mathcal{B} \). This implies that \( 0 \leq \mu(E) \leq \mu(E_1) + \mu(E_2) = 0 \). Therefore, \( \lim_{n \to \infty} \phi_n = f \) a.e. in \( X \). Hence, the result holds in this special case.

Finally, we consider the general case. Since \( X \) is \( \sigma \)-finite, then \( \exists (X_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{i=1}^{\infty} X_i \) and \( \mu(X_i) < +\infty, \forall i \in \mathbb{N} \). Without loss of generality, we may assume that \( X_i \subseteq X_{i+1}, \forall i \in \mathbb{N} \). Fix any \( i \in \mathbb{N} \), let \( X_i := (X_i, \mathcal{B}_i, \mu_i) \) be the finite measure subspace of \( X \) as defined in Proposition 11.13. Then, \( f|_{X_i} : X_i \to \mathcal{B} \)-measurable. By the second special case, there exists a sequence of simple functions \( (\phi_{i,n})_{n=1}^{\infty}, \phi_i : X_i \to \mathcal{B}, \forall n \in \mathbb{N}, \) such that \( \|\phi_{i,n}(x)\| \leq \|f(x)\| \), \( \forall x \in X_i, \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_{i,n} = f|_{X_i} \) a.e. in \( X_i \). By Proposition 11.58, \( \lim_{n \to \infty} \phi_{i,n} = f|_{X_i} \) in measure in \( X_i \). Then, \( \exists n_i \in \mathbb{N} \) such that \( \mu_i(E_{i,n}) := \mu_i(\{x \in X_i \mid \|f|_{X_i}(x) - \phi_{i,n}(x)\| \geq 2^{-i}\}) < 2^{-i} \). Let \( \phi_{i,n_i} \) admit the canonical representation \( \phi_{i,n_i} = \sum_{j=1}^{n_i} y_{i,j} \chi_{A_{i,j},X_i}, \) where \( y_{i,j} \in \mathbb{Z}_+, y_{1,1}, \ldots, y_{i,n_i} \in U \) are distinct and none equals to \( \vartheta_j \), and \( A_{i,1}, \ldots, A_{i,n_i} \in \mathcal{B}_i \subseteq \mathcal{B} \) are pairwise disjoint, nonempty, and of finite measure. Then, we may define a simple function \( \phi_i : X \to U \cup \{ \vartheta_j \} = U \) by \( \phi_i = \sum_{j=1}^{n_i} y_{i,j} \chi_{A_{i,j},X}. \) Note that \( \phi_i(x) = \phi_{i,n_i}(x) \) and \( \|\phi_i(x)\| \leq \|f(x)\| \), \( \forall x \in X_i, \forall i \in \mathbb{N} \) such that \( x \in X_{i_0} \). Then, \( x \in X_{j_i}, \forall j \in \mathbb{N} \) with \( j \geq i_0 \). \( \exists i_1 \in \mathbb{N} \) such that \( x \in X_{j_i} \setminus E_{j_i}, \forall j \in \mathbb{N} \) with \( j \geq i_1 \). Let \( k_0 = \max(i_0,i_1) \). \( \forall k \geq k_0, x \in X_{k} \setminus E_{k} \) and \( \|\phi_k(x) - f(x)\| < 2^{-k} \). This implies that \( \lim_{k \to \infty} \phi_k(x) = f(x) \). Then, \( E := \{ x \in X \mid (\phi_k(x))_{k=1}^{\infty} \) does not converge to \( f(x) \} \subseteq E \). By Proposition 11.50, we have \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). Hence, \( \lim_{n \to \infty} \phi_k = f \) a.e. in \( X \). This completes the proof of the proposition. \( \square \)

Proposition 11.67 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space and \( f : X \to [0, \infty] \subseteq \mathbb{R} \) be \( \mathcal{B} \)-measurable. Then, there exists a sequence of simple functions \( (\phi_n)_{n=1}^{\infty}, \phi_n : X \to [0, \infty] \subseteq \mathbb{R}, \forall n \in \mathbb{N}, \) such that \( 0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq f(x), \forall x \in X, \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n(x) = f(x), \forall x \in X \).

Proof \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mu(X_n) < \infty, \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( X_n \subseteq X_{n+1}, \forall n \in \mathbb{N} \). Fix any \( n \in \mathbb{N} \). Let \( I_{i,n} := ((i-1)/n, i/n], E_{i,n} := I_{i,n}(E_i) \cap X_n \in \mathcal{B}, i = 1, \ldots, n^2, \) and \( E_{i,n+1,n} := X_n \setminus (\bigcup_{i=1}^{n^2} E_{i,n}). \) Note that \( E_{i,n} \)'s are
pairwise disjoint and \( \mu(E_i) \leq \mu(X_n) < \infty \), \( \forall i \in \{1, \ldots, n^2 + 1\} \). Define \( \phi_n := \sum_{i=1}^{n^2+1} (i-1)/n \chi_{E_i,n} \). Then, \( \phi_n : X \to [0, \infty) \subset \mathbb{R} \), \( 0 \leq \phi_n(x) \leq f(x) \), \( \forall x \in X \), \( \phi_n \) is a simple function, and \( 0 \leq f(x) - \phi_n(x) \leq 1/n \), \( \forall x \in f\text{inv}(\{0\}) \cap X_n \). Define \( \varphi_n := \bigvee_{i=1}^n \phi_i \). Then, \( \varphi_n : X \to [0, \infty) \subset \mathbb{R} \), \( 0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x) \), \( \forall x \in X \), \( \varphi_n \) is a simple function, and \( 0 \leq f(x) - \varphi_n(x) \leq 1/n \), \( \forall x \in f\text{inv}(\{0\}) \cap X_n \).

Fix any \( x \in X \). If \( f(x) = \infty \), then \( \exists n_0 \in \mathbb{N} \) such that \( x \in X_{n_0} \), \( \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( x \in X_n \), \( x \in E_{n^2+1,n} \), and \( \phi_n(x) = n \). This further implies that \( \varphi_n(x) = n \), \( \forall n \geq n_0 \). Hence, \( \infty = f(x) = \lim_{n \to \infty} \varphi_n(x) \). On the other hand, if \( f(x) < \infty \), then \( \exists n_0 \in \mathbb{N} \) such that \( x \in X_{n_0} \) and \( 0 \leq f(x) \leq 0 \). Then, \( \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( |f(x) - \varphi_n(x)| \leq 1/n \). This implies that \( \lim_{n \to \infty} \varphi_n(x) = f(x) \).

Hence, we have \( f(x) = \lim_{n \to \infty} \varphi_n(x) \), \( \forall x \in X \). This completes the proof of the proposition. \( \Box \)

**Lemma 11.68** Let \( X := (X,B,\mu) \) be a \( \sigma \)-finite measure space, \( Y \) be a normed linear space over \( \mathbb{K} \), \( U \subseteq Y \) be a \( \sigma \)-compact conic segment, and \( f : X \to U \) be \( B \)-measurable. Then, there exists a sequence of simple functions \( \{\varphi_n\}_{n=1}^{\infty} \), \( \varphi_n : X \to U \), \( \forall n \in \mathbb{N} \) such that \( \lim_{n \to \infty} \varphi_n(x) = f(x) \), \( \forall x \in X \), and \( \|\varphi_n(x)\| \leq \|f(x)\| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \).

**Proof** Since \( U \) is \( \sigma \)-compact, then \( U = \bigcup_{n=1}^{\infty} K_n \), where \( K_n \) is compact, \( \forall n \in \mathbb{N} \). Since \( X \) is \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^{\infty} \subseteq B \) such that \( X = \bigcup_{n=1}^{\infty} X_n \), and \( \mu(X_n) < +\infty \), \( \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( K_n \subseteq K_{n+1} \) and \( X_n \subseteq X_{n+1} \), \( \forall n \in \mathbb{N} \).

Fix any \( n \in \mathbb{N} \), \( K_n \subseteq \bigcup_{y \in K_n} B_y(y,1/n) \). By the compactness of \( K_n \), \( \exists D_n \subseteq K_n \) where \( D_n \) is a finite set, such that \( K_n \subseteq \bigcup_{y \in D_n} B_y(y,1/n) \). Without loss of generality, assume that \( D_n = \{y_1, \ldots, y_{m_n}\} \), where \( m_n \in \mathbb{Z}^+ \). Let \( E_{n,1} := B_{y_1}(y_1,1/n) \), and \( E_{n,i} := B_{y_i}(y_{i-1,1/n}) \setminus \bigcup_{j=1}^{i-1} E_{n,j} \), \( i = 2, \ldots, m_n \). Then, \( K_n \subseteq \bigcup_{i=1}^{m_n} E_{n,i} \) and the sets in the union are pairwise disjoint. Let \( A_{n,i} := f\text{inv}(E_{n,i}) \cap X_n \in B \), \( i = 1, \ldots, m_n \), since \( E_{n,i} \in \mathcal{B}(Y) \) and \( f \) is \( B \)-measurable. Note that \( \mu(A_{n,i}) \leq \mu(X_n) < +\infty \), \( i = 1, \ldots, m_n \), and \( \{A_{n,i}\}_{i=1}^{m_n} \) is pairwise disjoint. Define \( \psi_n : X \to K_n \subseteq U \) by \( \psi_n := \sum_{i=1}^{m_n} y_{i-1,1/n} \chi_{A_{n,i}} \). Clearly, \( \psi_n \) is a simple function. \( \forall x \in X \), \( \exists N \in \mathbb{N} \) such that \( x \in X_N \) and \( f(x) \in K_N \). Then, \( \|f(x) - \psi_n(x)\| < 1/n \), \( \forall n \in \mathbb{N} \) with \( n \geq N \). Hence, \( \lim_{n \to \infty} \psi_n(x) = f(x) \), \( \forall x \in X \).

By Propositions 7.21 and 11.38, \( \mathcal{P} \circ f : X \to [0, \infty) \subset \mathbb{R} \) is \( B \)-measurable. By Proposition 11.67, there exists a sequence of simple functions \( \{\phi_n\}_{n=1}^{\infty} \) such that \( 0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq \mathcal{P} \circ f(x) = \|f(x)\| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n(x) = \mathcal{P} \circ f(x), \forall x \in X \).

Fix any \( n \in \mathbb{N} \), let \( \phi_n \) admit the canonical representation \( \phi_n = \sum_{i=1}^{m_n} a_j \chi_{A_{n,j}} \), where \( m_n \in \mathbb{Z}^+ \), \( a_1, \ldots, a_{m_n} \in [0, \infty) \subset \mathbb{R} \) are distinct and none equals to 0, \( A_{n,1}, \ldots, A_{n,m_n} \in \mathcal{B} \) are pairwise disjoint, nonempty,
and of finite measure. Define $\varphi_n : X \to U$ by

$$
\varphi_n(x) = \begin{cases} 
\frac{y_{n,i}}{\|y_{n,i}\|} & \text{if } x \in A_{n,i} \cap \tilde{A}_{n,j}, \|y_{n,i}\| \leq a_j, i = 1, \ldots, m_n, j = 1, \ldots, m_n \\
\frac{y_{n,i}}{\|y_{n,i}\|} & \text{if } x \in A_{n,i} \cap \tilde{A}_{n,j}, \|y_{n,i}\| > a_j, i = 1, \ldots, m_n, j = 1, \ldots, m_n \\
\varnothing & x \in X \setminus \left(\bigcup_{i=1}^{m_n} A_{n,i} \cap \bigcup_{j=1}^{\tilde{m_n}} \tilde{A}_{n,j}\right)
\end{cases}
$$

Clearly $\varphi_n(x) \in U$, $\forall x \in X$, since $U$ is a conic segment. An equivalent way to express $\varphi_n$ is that $\varphi_n(x) = \begin{cases} 
\psi_n(x) & \text{if } \|\psi_n(x)\| \leq \phi_n(x) \\
\frac{\phi_n(x)}{\|\psi_n(x)\|} & \text{if } \|\psi_n(x)\| \geq \phi_n(x)
\end{cases}$, $\forall x \in X$. Then, we have $\|\varphi_n(x)\| \leq \phi_n(x) \leq P \circ f(x)$, $\forall x \in X$. Furthermore, $\lim_{n \to \infty} \varphi_n(x) = f(x)$, $\forall x \in X$.

This completes the proof of the lemma. $\square$

Thus, we have generalized Littlewood’s three principles: “every Lebesgue measurable set in $\mathbb{R}$ is almost a finite union of open intervals” (Proposition 11.23); “every measurable function taking value in a separable normed linear space is almost a simple function” (Propositions 11.65 and 11.66); “every convergent sequence of measurable functions taking value in a separable metric space is nearly uniformly convergent” (Egoroff’s Theorem 11.55).

11.4 Integration

Proposition 11.69 Let $X$ be a normed linear space, $\mathcal{B}_B(X)$ be the collection of Borel sets on $X$. A representation of $X$ is the collection $R := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}$, where $\Lambda$ is a finite index set, $(U_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}_B(X)$ are nonempty and pairwise disjoint, $\bigcup_{\alpha \in \Lambda} U_\alpha = X$, $x_\alpha \in U_\alpha$, and $\|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1$, $\forall \alpha \in \Lambda$. The set of all representations of $X$ is denoted $\mathcal{R}(X)$. Introduce a relation $\preceq$ on $\mathcal{R}(X)$ by $\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}, R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(X)$, we will say $R_1 \preceq R_2$ if, $\forall \alpha \in \Lambda, \exists \beta \in \Gamma$ such that $x_\alpha = y_\beta$, and, $\forall \beta \in \Gamma, \exists \alpha \in \Lambda$ such that $V_\beta \subseteq U_\alpha$. Then, $\mathcal{J}(X) := (\mathcal{R}(X), \preceq)$ is a directed system and $\preceq$ is an antisymmetric partial ordering on $\mathcal{R}(X)$. $\mathcal{J}(X)$ is said to be the integration system on $X$.

Proof Clearly, $R_0 := \{(\varnothing, X)\} \in \mathcal{R}(X) \neq \emptyset$. Clearly, $\preceq$ is reflexive and transitive. $\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}, R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(X)$ with $R_1 \preceq R_2$ and $R_2 \preceq R_1$, $\forall \alpha \in \Lambda, \exists \beta \in \Gamma$ such that $U_\alpha \subseteq V_\beta$, by $R_2 \preceq R_1$. By $R_1 \preceq R_2$, $\exists \tilde{\alpha} \in \Lambda$ such that $V_{\tilde{\beta}} \subseteq U_{\tilde{\alpha}}$. Then, we have $U_{\tilde{\alpha}} \subseteq V_{\tilde{\beta}} \subseteq U_{\tilde{\alpha}}$. By $R_1 \in \mathcal{R}(X)$, $(U_\alpha)_{\alpha \in \Lambda}$ are nonempty and pairwise disjoint. Then, we must have $\alpha = \tilde{\alpha}$ and $U_\alpha = V_{\tilde{\beta}}$. By $R_1 \preceq R_2$, $\exists \beta \in \Gamma$ such that $x_\alpha = y_\beta$. Then, $U_\alpha = V_{\beta} \ni x_\alpha = y_\beta \in V_{\beta}$. By $R_2 \in \mathcal{R}(X)$, $(V_\beta)_{\beta \in \Gamma}$ are nonempty and pairwise disjoint. Then, we must have $\beta = \beta$ and $x_\alpha = y_\beta$. Then, $(x_\alpha, U_\alpha) \in R_2$. Hence, by the arbitrariness of $\alpha,$
we have $R_1 \subseteq R_2$. By an argument that is similar to the above, we have $R_2 \subseteq R_1$. Therefore, $R_1 = R_2$. This shows that $\preceq$ is an antisymmetric partial ordering on $\mathcal{R}(\mathcal{X})$.

$\forall R_1 := \{(x_\alpha, U_\alpha) \mid \alpha \in \Lambda\}$, $R_2 := \{(y_\beta, V_\beta) \mid \beta \in \Gamma\} \in \mathcal{R}(\mathcal{X})$, define $R_3 \in \mathcal{R}(\mathcal{X})$ as follows. $\forall \alpha \in \Lambda$, $\forall \beta \in \Gamma$, let $W_{\alpha,\beta} := U_\alpha \cap V_\beta \in \mathcal{B}(\mathcal{X})$.

We will distinguish six exhaustive and mutually exclusive cases.

Case 1: $W_{\alpha,\beta} = \emptyset$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta}^{(2)} = \emptyset$.

Case 2: $W_{\alpha,\beta} \neq \emptyset$, $x_\alpha \notin W_{\alpha,\beta}$, and $y_\beta \notin W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} \in W_{\alpha,\beta}^{(1)}$ such that $\|z_{\alpha,\beta}^{(1)}\| < \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$, and $W_{\alpha,\beta}^{(2)} = \emptyset$.

Case 3: $x_\alpha \in W_{\alpha,\beta}$ and $y_\beta \notin W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 4: $x_\alpha \notin W_{\alpha,\beta}$ and $y_\beta \in W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = y_\beta \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|y_\beta\| < \inf_{x \in V_\beta} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 5: $x_\alpha = y_\beta \in W_{\alpha,\beta}$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \in \mathcal{B}(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, and $W_{\alpha,\beta}^{(2)} = \emptyset$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$.

Case 6: $x_\alpha \in W_{\alpha,\beta}$, $y_\beta \in W_{\alpha,\beta}$, and $x_\alpha \neq y_\beta$. Let $\delta = \|x_\alpha - y_\beta\| > 0$. Define $W_{\alpha,\beta}^{(1)} = W_{\alpha,\beta} \cap \mathcal{B}(\mathcal{X}, \delta/2) \in \mathcal{B}(\mathcal{X})$, $z_{\alpha,\beta}^{(1)} = x_\alpha \in W_{\alpha,\beta}^{(1)}$, $W_{\alpha,\beta}^{(2)} = W_{\alpha,\beta} \cap \mathcal{B}(\mathcal{X}, \delta/2)^{\complement} \in \mathcal{B}(\mathcal{X})$, and $z_{\alpha,\beta}^{(2)} = y_\beta \in W_{\alpha,\beta}^{(2)}$. Then, $\|z_{\alpha,\beta}^{(1)}\| = \|x_\alpha\| < \inf_{x \in U_\alpha} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(1)}} \|x\| + 1$ and $\|z_{\alpha,\beta}^{(2)}\| = \|y_\beta\| < \inf_{x \in V_\beta} \|x\| + 1 \leq \inf_{x \in W_{\alpha,\beta}^{(2)}} \|x\| + 1$.

Define $R_3 := \left\{(z_{\alpha,\beta}^{(i)}, W_{\alpha,\beta}^{(i)}) \mid \alpha \in \Lambda, \beta \in \Gamma, i = 1, 2, W_{\alpha,\beta}^{(i)} \neq \emptyset\right\}$.

Clearly, $R_3 \in \mathcal{R}(\mathcal{X})$, $R_1 \subseteq R_3$, and $R_2 \subseteq R_3$. Therefore, $\mathcal{J}(\mathcal{X})$ is a directed system. This completes the proof of the proposition.

**Definition 11.70** Let $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a normed linear space with the integration system $\mathcal{J}(\mathcal{Y}) := (\mathcal{Y}(\mathcal{Y}), \preceq)$, and $f : \mathcal{X} \to \mathcal{Y}$ be $\mathcal{B}$-measurable. $\forall R = \{(y_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{R}(\mathcal{Y})$, define $F_R := \sum_{\alpha \in \Lambda} y_\alpha \mu(f^{-1}(U_\alpha)) \in \mathcal{Y}$. This defines a net $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$. $f$ is said to be integrable if the net admits a limit in $\mathcal{Y}$. In this case, $\lim_{R \in \mathcal{J}(\mathcal{Y})} F_R \in \mathcal{Y}$ is said to be the integral of $f$ over $\mathcal{X}$ and denoted by $\int_X f \, d\mu$ or $\int_X f \, d\mu(x)$. When $\mathcal{Y} = \mathbb{R}$, we will denote $\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{R \in \mathcal{J}(\mathcal{Y})} F_R \in \mathbb{R}$ whenever the limit exists, which is said to be the integral of $f$ over $\mathcal{X}$.

**Definition 11.71** Let $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a measure space with $\mu(\mathcal{X}) = \infty$, $\mathcal{Y}$ be a normed linear space, and $f : \mathcal{X} \to \mathcal{Y}$ be $\mathcal{B}$-measurable. Define the set $\mathcal{M}(\mathcal{X}) := \{ A \in \mathcal{B} \mid \mu(A) < +\infty\}$. Clearly, $\mathcal{M}(\mathcal{X}) := (\mathcal{M}(\mathcal{X}), \subseteq)$ is
11.4. INTEGRATION

a directed system. ∀A ∈ M(X), let \((A, B_A, \mu_A)\) be the finite measure subspace of \(X\) as defined in Proposition 11.13. Define a net \((F_A)_{A ∈ M(X)}\) by

\[
F_A := \int_X f \; d\mu_A ∈ Y, \forall A ∈ M(X).
\]

\(f\) is said to be integrable if the net is well defined and admits a limit in \(Y\). In this case, \(\lim_{A ∈ M(X)} F_A ∈ Y\) is said to be the integral of \(f\) over \(X\) and denoted by \(\int_X f \; d\mu\) or \(\int_X f(x) \; d\mu(x)\).

When \(Y = \mathbb{R}\), we will allow the net \((F_A)_{A ∈ M(X)} \subseteq \mathbb{R}_e\) and denote \(\int_X f \; d\mu := \int_X f(x) \; d\mu(x) := \lim_{A ∈ M(X)} F_A ∈ \mathbb{R}_e\) whenever the limit exists, which is said to be the integral of \(f\) over \(X\).

**Proposition 11.72** Let \(X := (X, B, \mu)\) be a finite measure space, \(Y\) be a normed linear space, \(f : X → Y\) be \(B\)-measurable, and \(X := (\hat{X}, \hat{B}, \hat{\mu})\) be a finite measure space that is an extension of \(X\), i.e., \(B ⊆ \hat{B}\) and \(\mu = \hat{\mu}|_B\).

Then, \(\int_X f \; d\mu = \int_{\hat{X}} \int_{\hat{X}} f \; d\hat{\mu}\), whenever one of the integrals exists.

**Proof** Let \(J(Y)\) be the integration system on \(Y\) as defined in Proposition 11.69, \((F_R)_{R ∈ J(Y)}\) be the net for \(\int_X f \; d\mu\), and \((\hat{F}_R)_{\hat{R} ∈ J(\hat{Y})}\) be the net for \(\int_{\hat{X}} f \; d\hat{\mu}\), as defined in Definition 11.70. ∀\(R := \{(y_α, U_α) \mid α ∈ Λ\} ∈ J(Y), \forall α ∈ Λ, we have \(f_{\text{inv}}(U_α) ∈ B \subseteq \hat{B}\) since \(f\) is \(B\)-measurable, and \(\mu(f_{\text{inv}}(U_α)) = \hat{\mu}(f_{\text{inv}}(U_α)) ∈ \mathbb{R}_e\) since \(\mu = \hat{\mu}|_B\). Then, \(F_R = \sum_{α ∈ Λ} y_α \mu(f_{\text{inv}}(U_α)) = \sum_{α ∈ Λ} y_α \hat{\mu}(f_{\text{inv}}(U_α)) = \hat{F}_R ∈ \hat{Y}\). Hence, \(\int_X f \; d\mu = \lim_{R ∈ J(Y)} F_R = \lim_{\hat{R} ∈ J(\hat{Y})} \hat{F}_R = \int_{\hat{X}} f \; d\hat{\mu}\), whenever one of the integrals exists. This completes the proof of the proposition. □

**Lemma 11.73** Let \(X := (X, B, \mu)\) be a finite measure space, \(Y\) be a normed linear space, \(f : X → Y\) be \(B\)-measurable, \(\bar{X} ∈ B\), and \(X := (\bar{X}, \hat{B}, \hat{\mu})\) be the finite measure subspace of \(X\) as defined in Proposition 11.13. Assume that \(\mu(X \setminus \bar{X}) = 0\). Then, \(\int_X f \; d\mu = \int_{\bar{X}} f \; d\mu\), whenever one of the integrals exists.

**Proof** Let \(J(Y)\) be the integration system on \(Y\) as defined in Proposition 11.69, \((F_R)_{R ∈ J(Y)}\) be the net for \(\int_X f \; d\mu\), and \((\hat{F}_R)_{\hat{R} ∈ J(\hat{Y})}\) be the net for \(\int_{\bar{X}} f \; d\hat{\mu}\), as defined in Definition 11.70. ∀\(R := \{(y_α, U_α) \mid α ∈ Λ\} ∈ J(Y), \forall α ∈ Λ, we have \(f_{\text{inv}}(U_α) \cap \bar{X} = f|_{\bar{X}} \text{inv}(U_α) ∈ \hat{B} \subseteq \hat{B}\) since \(f\) is \(B\)-measurable, and \(0 ≤ \mu(f_{\text{inv}}(U_α) \cap (X \setminus \bar{X})) ≤ \mu(X \setminus \bar{X}) = 0\). Then, \(F_R = \sum_{α ∈ Λ} y_α \mu(f_{\text{inv}}(U_α)) = \sum_{α ∈ Λ} y_α \hat{\mu}(f|_{\bar{X}} \text{inv}(U_α)) = \hat{F}_R ∈ \hat{Y}\). Hence, \(\int_X f \; d\mu = \lim_{R ∈ J(Y)} F_R = \lim_{\hat{R} ∈ J(\hat{Y})} \hat{F}_R = \int_{\bar{X}} f \; d\hat{\mu}\), whenever one of the integrals exists. This completes the proof of the proposition. □

**Proposition 11.74** Let \(X := (X, B, \mu)\) be a measure space with \(\mu(X) = +∞\), \(Y\) be a normed linear space, \(f : X → Y\) be \(B\)-measurable, and \(X := (X, B, \mu)\) be the completion of \(X\). Then, \(\int_X f \; d\mu = \int_X f \; d\hat{\mu}\), whenever one of the integrals exists.
Proposition 3.70, Let \( \int_X f \, d\mu \) be the net for \( \int_X f \, d\bar{\mu} \) as defined in Definition 11.71. Clearly, \( \mathcal{M}(\mathcal{X}) \subseteq \mathcal{M}(\mathcal{\bar{X}}) \), since \( \mathcal{X} \) is the completion of \( \mathcal{\bar{X}} \). We will distinguish two exhaustive cases: Case 1: \( \int_X f \, d\mu \) exists; Case 2: \( \int_X f \, d\bar{\mu} \) exists.

Case 1: \( \int_X f \, d\mu \) exists. Then, the net \( (\bar{F}_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \) is well defined. 

\[ \forall A \in \mathcal{M}(\mathcal{\bar{X}}), \quad A \in \mathcal{M}(\mathcal{\bar{X}}) \] 

By Proposition 11.72, \( F_A = \int_A f|_A \, d\mu_A = \int_A f|_A \, d\bar{\mu}_A = F_A \), where \( (A, B_A, \mu_A) \) is the finite measure subspace of \( \mathcal{X} \) and \( (A, B_A, \bar{\mu}_A) \) is the finite complete measure subspace of \( \mathcal{X} \). Then, \( (F_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \) is well defined. \( \forall A \in \mathcal{M}(\mathcal{\bar{X}}) \), we have \( A \in \mathcal{B} \) with \( \bar{\mu}(A) < +\infty \). By Proposition 11.12 and its proof, \( \exists A, B, \bar{\mathcal{C}} \subseteq \mathcal{X} \) with \( A = \bar{\mathcal{C}} \cup \bar{\mathcal{B}} \), \( B, \bar{\mathcal{C}} \subseteq \mathcal{B} \), \( A \subseteq \bar{\mathcal{C}} \), \( \mu(\bar{\mathcal{C}}) = 0 \), and \( \bar{\mu}(A) = \mu(\bar{\mathcal{B}}) \). Then, \( A \subseteq B \cup \bar{\mathcal{C}} =: A \in \mathcal{B} \) and \( \mu(A) \leq \mu(\bar{\mathcal{B}}) + \mu(\bar{\mathcal{C}}) = \mu(\bar{\mathcal{B}}) = \bar{\mu}(A) < +\infty \). This shows that \( A \in \mathcal{M}(\mathcal{\bar{X}}) \) and \( A \subseteq A \). Hence, \( (F_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \) is a subnet of \( (F_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \).

By Proposition 3.70, \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{\bar{X}})} F_A = \lim_{\bar{\mu} \in \mathcal{\bar{X}}(\mathcal{\bar{X}})} \bar{F}_A = \int_X f \, d\bar{\mu} \).

Case 2: \( \int_X f \, d\bar{\mu} \) exists. Then, the net \( (F_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \) is well defined. 

\[ \forall \bar{A} \in \mathcal{M}(\mathcal{\bar{X}}), \quad \bar{A} \in \mathcal{M}(\mathcal{\bar{X}}) \] 

By Proposition 11.72 and its proof, \( \exists A, B, \bar{\mathcal{C}} \subseteq \mathcal{X} \) with \( A = \bar{\mathcal{C}} \cup \bar{\mathcal{B}} \), \( B, \bar{\mathcal{C}} \subseteq \mathcal{B} \), \( A \subseteq \bar{\mathcal{C}} \), \( \mu(\bar{\mathcal{C}}) = 0 \), and \( \mu(A) = \mu(\bar{\mathcal{B}}) \). Let \( \bar{B} := A \in \mathcal{B} \). Then, \( 0 \leq \bar{\mu}(A \setminus A) = \bar{\mu}(A \setminus \bar{\mathcal{B}}) \leq \bar{\mu}(\bar{\mathcal{C}}) = \mu(\bar{\mathcal{C}}) = 0 \) and \( \mu(A) = \mu(A) < +\infty \). This shows that \( A \in \mathcal{M}(\mathcal{\bar{X}}) \) and \( A \subseteq A \). Let \( \bar{A} := (A, B_A, \mu_A) \) be the finite measure subspace of \( \mathcal{X} \), \( \bar{A} := (A, B_A, \bar{\mu}_A) \) be the finite complete measure subspace of \( \mathcal{X} \). By Lemma 11.73, we have \( \bar{F}_A = \int_A f|_A \, d\bar{\mu}_A = \int_A f|_A \, d\mu_A \), whenever one of them exists. Since \( \mathcal{X} \) is the completion of \( \mathcal{\bar{X}} \), then \( \bar{A} \) is an extension of \( A \). By Proposition 11.72, we have \( \int_A f|_A \, d\mu_A = \int_A f|_A \, d\bar{\mu}_A = F_A \). Hence, \( \bar{F}_A = F_A \in U \). Then, the net \( (\bar{F}_A)_{A \in \mathcal{M}(\mathcal{\bar{X}})} \) is well defined.

Fix any open set \( U \) \((U \subseteq \mathbb{R} \) or \( U \subseteq \mathbb{R} \) if \( U \neq \mathbb{R} \)) with \( \int_X f \, d\mu \in U \). Since \( \lim_{A \in \mathcal{M}(\mathcal{\bar{X}})} F_A = \int_X f \, d\mu \in U \), then \( \exists A_0 \in \mathcal{M}(\mathcal{\bar{X}}) \) such that \( \forall A \in \mathcal{M}(\mathcal{\bar{X}}) \) with \( A_0 \subseteq A \), we have \( F_A \in U \). Note that \( A_0 \in \mathcal{B} \subseteq \mathcal{B} \) and \( \mu(A_0) = \bar{\mu}(A_0) < +\infty \). Then, \( A_0 \in \mathcal{M}(\mathcal{\bar{X}}) \). \( \forall \bar{A} \in \mathcal{M}(\mathcal{\bar{X}}) \) with \( A_0 \subseteq A \), then \( \bar{A} \in \mathcal{B} \) and \( \mu(\bar{A}) = \mu(A) < +\infty \). By Proposition 11.12 and its proof, \( \exists A, B, \bar{\mathcal{C}} \subseteq \mathcal{X} \) with \( A = \bar{\mathcal{C}} \cup \bar{\mathcal{B}} \), \( B, \bar{\mathcal{C}} \subseteq \mathcal{B} \), \( A \subseteq \bar{\mathcal{C}} \), \( \mu(\bar{\mathcal{C}}) = 0 \), and \( \mu(A) = \mu(\bar{\mathcal{B}}) \). Let \( A := A_0 \cup B \subseteq \mathcal{B} \). Then, \( 0 \leq \mu(A \setminus A_0) = \mu(A \setminus \bar{\mathcal{B}}) \leq \mu(\bar{\mathcal{C}}) = \mu(\bar{\mathcal{C}}) = 0 \), and \( \mu(A) = \mu(B) + \mu(A_0) = \mu(A) + \mu(A) < +\infty \). This shows that \( A \in \mathcal{M}(\mathcal{\bar{X}}) \) and \( A_0 \subseteq A \subseteq A \). Let \( \bar{A} := (A, B_A, \mu_A) \) be the finite measure subspace of \( \mathcal{X} \), \( \bar{A} := (A, B_A, \bar{\mu}_A) \) be the finite complete measure subspace of \( \mathcal{X} \). By Lemma 11.73, we have \( \bar{F}_A = \int_A f|_A \, d\bar{\mu}_A = \int_A f|_A \, d\mu_A \), whenever one of them exists. Since \( \mathcal{X} \) is the completion of \( \mathcal{\bar{X}} \), then \( \bar{A} \) is an extension of \( A \). By Proposition 11.72, we have \( \int_A f|_A \, d\mu_A = \int_A f|_A \, d\bar{\mu}_A = F_A \). Hence, \( \bar{F}_A = F_A \in U \). Therefore, we have \( \int_X f \, d\bar{\mu} = \lim_{\bar{\mu} \in \mathcal{\bar{X}}(\mathcal{\bar{X}})} \bar{F}_A = \int_X f \, d\mu \).
11.4. INTEGRATION

Hence, in both cases, we have $\int_X f \, d\mu = \int_X f \, d\tilde{\mu}$. This completes the proof of the proposition. \hfill \Box

Proposition 11.75 Let $X := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a normed linear space, $\phi : X \to \mathcal{Y}$ be a simple function in canonical representation, i.e., $\exists n \in \mathbb{Z}_+$, $\exists y_1, \ldots, y_n \in \mathcal{Y}$, which are distinct and none equals to $\vartheta_y$, $\exists A_1, \ldots, A_n \in \mathcal{B}$, which are nonempty, pairwise disjoint, and of finite measure, such that $\phi(x) = \sum_{i=1}^n y_i \chi_{A_i}(x)$, $\forall x \in X$. Then, $\int_X \phi \, d\mu = \sum_{i=1}^n y_i \mu(A_i) =: I \in \mathcal{Y}$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. Let $y_{n+1} := \vartheta_y$ and $A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B}$. Let $\epsilon_0 := \min\{1, \min_{1 \leq i < j \leq n+1} \|y_i - y_j\|\} > 0$, by the assumption. Let $(\Phi_R)_{R \in \mathcal{Y}(\mathcal{Y})}$ be the net as defined in Definition 11.70 for the simple function $\phi$.

$\forall \epsilon \in (0, \epsilon_0)/2 \subset \mathbb{R}$, let $U_i := B_{\mathcal{Y}}(y_i, \epsilon) \subset \mathcal{B}(\mathcal{Y})$, $i = 1, \ldots, n+1$, which are clearly pairwise disjoint and nonempty. We will distinguish two exhaustive and mutually exclusive subcases: Case 1a: $\mathcal{Y} = \bigcup_{i=1}^{n+1} U_i$; Case 1b: $\mathcal{Y} \supset \bigcup_{i=1}^{n+1} U_i$.

Case 1a: $\mathcal{Y} = \bigcup_{i=1}^{n+1} U_i$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{Y} = B_{\mathcal{Y}}(\vartheta_y, \delta)$. Then, $\mathcal{Y}$ must be the trivial normed linear space, i.e., $\mathcal{Y}$ is a singleton set containing $\vartheta_y$. There is a single representation of $\mathcal{Y}$, which is $R_0 := \{(\vartheta_y, \mathcal{Y})\}$. Then, $\Phi_{R_0} = \vartheta_y$. Clearly, we must have $n = 0$, since $y_1, \ldots, y_n \in \mathcal{Y}$ are distinct and none equals to $\vartheta_y$. Then, $I = \vartheta_y$. Hence, $\forall R \in \mathcal{Y}(\mathcal{Y})$ with $R_0 \preceq R$, $\|\Phi_R - I\| = \|\Phi_{R_0} - I\| = 0$.

Case 1b: $\mathcal{Y} \supset \bigcup_{i=1}^{n+1} U_i$. Let $U_{n+2} = \bigcap_{i=1}^{n+1} \tilde{U}_i \in \mathcal{B}(\mathcal{Y})$, which is nonempty. Let $y_{n+2} \in U_{n+2}$ be such that $\|y_{n+2}\| = \inf_{y \in U_{n+2}} \|y\| + 1$. Let $R_0 := \{(y_i, U_i) : i = 1, \ldots, n+2\}$. It is easy to check that $R_0 \in \mathcal{Y}(\mathcal{Y})$.

$\forall R \in \mathcal{Y}(\mathcal{Y})$ with $R_0 \preceq R$, then $\tilde{R} = \{(y_\alpha, \tilde{U}_\alpha) : \alpha \in \Lambda\}$, where $\Lambda$ is a finite index set, $(\tilde{U}_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}(\mathcal{Y})$ are nonempty and pairwise disjoint, $\bigcap_{\alpha \in \Lambda} \tilde{U}_\alpha = \vartheta_y$, $\forall \tilde{y}_\alpha \in \tilde{U}_\alpha$, and $\|\tilde{y}_\alpha\| = \inf_{y \in U_\alpha} \|y\| + 1$, $\forall \alpha \in \Lambda$, $\forall \alpha \in \{1, \ldots, n+2\}$, and $\forall \alpha \in \Lambda$, $\alpha = \tilde{y}_\alpha$. Again by $R_0 \preceq \tilde{R}$, $\forall \alpha \in \Lambda$ such that $y_\alpha \in \tilde{U}_\alpha$, $\exists! \alpha \in \Lambda$ such that $y_\alpha \in \tilde{U}_\alpha$. Since $\left(\tilde{U}_\alpha\right)_{\alpha \in \Lambda} \subseteq \mathcal{B}(\mathcal{Y})$ are nonempty and pairwise disjoint and $y_\alpha \in \tilde{U}_\alpha$, $\forall \alpha \in \Lambda$, then $\alpha = \tilde{y}_\alpha$. By $R_0 \preceq \tilde{R}$, we have $\tilde{U}_\alpha \subseteq U_\alpha$. Hence, $\alpha_1, \ldots, \alpha_{n+2}$ are distinct. Let $\Lambda := \{\alpha_1, \ldots, \alpha_{n+2}\}$.

$\forall \alpha \in \Lambda \setminus \tilde{\Lambda}$, $\phi_{\text{inv}}(U_\alpha) = \emptyset$. Then, $\Phi_{\tilde{R}} = \sum_{\alpha \in \Lambda} \sum_{i=1}^{n+2} y_i \mu(\phi_{\text{inv}}(U_\alpha)) = \sum_{i=1}^{n+2} y_i \mu(\phi_{\text{inv}}(\tilde{U}_\alpha))$. Note that $\forall i \in \{1, \ldots, n+1\}$, $\phi_{\text{inv}}(U_{\alpha_{n+1}}) = A_i$ and $\mu(A_i) \leq \mu(X) < +\infty$. Furthermore, $\phi_{\text{inv}}(\tilde{U}_{\alpha_{n+2}}) = \emptyset$. Hence, $\Phi_{\tilde{R}} = \sum_{i=1}^{n+1} y_i \mu(A_i) = I$. Then, we have $\|\Phi_{\tilde{R}} - I\| = 0$.

In both subcases, we have $\exists R_0 \in \mathcal{Y}(\mathcal{Y})$, $\forall R \in \mathcal{Y}(\mathcal{Y})$ with $R_0 \preceq R$, we have $\|\Phi_{\tilde{R}} - I\| = 0 < \epsilon$. Hence, $\int_X \phi \, d\mu = \lim_{R \in \mathcal{Y}(\mathcal{Y})} \Phi_{\tilde{R}} = I$. \hfill \Box
Case 2: \( \mu(X) = +\infty \). \( \forall A \in \mathcal{M}(X) \), \( \mu(A) < +\infty \). \( \phi|_A : A \rightarrow \mathbb{Y} \) is a simple function. By Case 1, \( \Phi_A = \int_{A} \phi|_A \, d\mu_A \in \mathbb{Y} \) is well-defined, where \( (A, \mathcal{B}_A, \mu_A) \) is the finite measure subspace of \( X \). Hence, the net \( \Phi_{A*} \in \mathcal{M}(X) \) is well-defined. Take \( A_0 := \bigcup_{i=1}^{n} A_i \in \mathcal{B} \). Then, \( \mu(A_0) = \sum_{i=1}^{n} \mu(A_i) < +\infty \). Therefore, \( A_0 \in \mathcal{M}(X) \). \( \forall A \in \mathcal{M}(X) \) with \( A_0 \subseteq A \), by Case 1, \( \Phi_A = \sum_{i=1}^{n} \phi_i \mu(A_i) = I \). Then, \( \int_{X} \phi \, d\mu = \lim_{A \in \mathcal{M}(X)} \Phi_A = I \).

This completes the proof of the proposition. \( \square \)

Clearly, the above result holds for simple functions given in any form, not necessarily in canonical representation. It also shows that integral of simple functions are linear, that is, \( \forall \) simple functions \( \phi_1 \) and \( \phi_2 \), \( \forall c \in \mathbb{K} \), we have \( I_{X} (\phi_1 + \phi_2) \, d\mu = I_{X} \phi_1 \, d\mu + I_{X} \phi_2 \, d\mu \) and \( I_{X} (c \phi_1) \, d\mu = c \, I_{X} \phi_1 \, d\mu \).

**Lemma 11.76** Let \( X := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathbb{Y} \) be a separable Banach space, \( \phi_n : X \rightarrow \mathbb{Y} \) be a simple function in canonical representation, \( \forall n \in \mathbb{N} \), and \( f : X \rightarrow \mathbb{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \in \mathbb{N}} \phi_n = f \) a.e. in \( X \) and \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| \phi_n(x) \| \leq M \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \). Then, \( f \) is integrable over \( X \) and \( I_{X} f \, d\mu = \lim_{n \in \mathbb{N}} I_{X} \phi_n \, d\mu \in \mathbb{Y} \).

**Proof** Let \( E := \{ x \in X \mid (\phi_n(x))_{n=1}^{\infty} \) does not converge to \( f(x) \} \). Then, \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). Let \( (F_R)_{R \subseteq \mathbb{Y}} \) be the net for \( I_{X} f \, d\mu \) as defined in Definition 11.70.

\( \forall \varepsilon \in (0, 1) \subset \mathbb{R} \), by Egoroff’s Theorem 11.55, \( \exists E_1 \in \mathcal{B} \) with \( \mu(E_1) < \frac{\varepsilon}{4M+2} \) such that \( (\phi_n|_{X \setminus E_1})_{n=1}^{\infty} \) converges uniformly to \( f|_{X \setminus E_1} \). Clearly, \( E \subseteq E_1 \). Then, \( \exists n_0 \in \mathbb{N} \), \( n \in \mathbb{N} \) with \( n \geq n_0 \), \( \forall x \in X \setminus E_1 \), \( \| \phi_n(x) - f(x) \| < \frac{\varepsilon}{4(4M+2)} := \varepsilon \). Fix any \( n \geq n_0 \). Let \( \phi_n \) admit the canonical representation \( \phi_n = \sum_{i=1}^{n} y_i \chi_{A_i} \), where \( y_i \in \mathbb{Z} \), \( y_1, \ldots, y_{\bar{n}} \in \mathbb{Y} \) are distinct and none equals to \( \vartheta_y \), and \( A_1, \ldots, A_{\bar{n}} \in \mathcal{B} \) are nonempty, pairwise disjoint, and of finite measure. Let \( y_{\bar{n}+1} := \vartheta_y \) and \( A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_i) \in \mathcal{B} \). Then, \( X = \bigcup_{i=1}^{\bar{n}+1} A_i \) and the sets in the union are pairwise disjoint and of finite measure. Define \( V_i := E_{B_y(y_i, \varepsilon)} \), \( i = 1, \ldots, \bar{n} + 1 \). Define \( V_{\bar{n}+2} := E \setminus (\bigcup_{j=1}^{\bar{n}+1} V_j) \). Clearly, we may form a representation \( \hat{R} := \{ (\hat{y}_i, V_i) \mid i = 1, \ldots, \bar{n} + 2, V_i \neq \emptyset \} \in \mathcal{J}(\mathbb{Y}) \), where \( \hat{y}_i \in V_i \), \( i = 1, \ldots, \bar{n} + 2 \), are any vectors that satisfy the assumption of Proposition 11.69. \( \forall R \in \mathcal{J}(\mathbb{Y}) \) with \( \hat{R} \preceq R \). Let \( R := \{ (\hat{y}_\gamma, \hat{U}_\gamma) \mid \gamma \in \Gamma \} \).

\( \forall \gamma \in \Gamma \), by \( \hat{R} \preceq R \), \( \exists i_\gamma \in \{1, \ldots, \bar{n} + 2\} \) such that \( \hat{U}_\gamma \subseteq V_{i_\gamma} \). Define \( \hat{\Gamma} := \{ \gamma \in \Gamma \mid i_\gamma = \bar{n} + 2 \} \). Let \( A_{\gamma,i} := A_i \cap f_{inv}(\hat{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}, \) \( i = 1, \ldots, \bar{n} + 1 \).

**Claim 11.76.1** \( \forall \gamma \in \hat{\Gamma} \), \( f_{inv}(\hat{U}_\gamma) \subseteq E_1 \) and \( \| \hat{y}_\gamma \mu(f_{inv}(\hat{U}_\gamma) \cap E_1) \| \leq (M + 1)\mu(f_{inv}(\hat{U}_\gamma)) \).
Proof of claim: Fix any $\gamma \in \Gamma$. $\hat{U}_\gamma \subseteq V_{\bar{n}+2} = \gamma \setminus (\bigcup_{j=\bar{n}+1}^{\bar{n}+2} \bar{V}_j)$. Then, $B \ni f^{-1}_n(\hat{U}_\gamma) \subseteq E_1$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $f^{-1}_n(\hat{U}_\gamma) \subseteq E$; Case 2: $f^{-1}_n(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$. 

Case 1: $f^{-1}_n(\hat{U}_\gamma) \subseteq E$. Then, $\mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1) = 0$ and $\|\hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1)\| = 0 = (M + 1)\mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1)$. Case 2: $f^{-1}_n(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$. Then, $\exists x \in E_1 \setminus E$ such that $f(x) \in \hat{U}_\gamma$. Note that $\lim_{m \in \mathbb{N}} \phi_m(x) = f(x)$ and $\|\phi_m(x)\| \leq M$, $\forall m \in \mathbb{N}$. Then, by Propositions 7.21 and 3.66, we have $\|f(x)\| = \lim_{m \in \mathbb{N}} \|\phi_m(x)\| \leq M$. Then, by $R \in \mathcal{J}(\gamma)$, $\|\hat{y}_\gamma\| < \inf_{y \in \hat{U}_\gamma} \|y\| + 1 \leq M + 1$. Therefore, $\|\hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1)\| \leq (M + 1)\mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1)$. This completes the proof of the claim.

$\forall \gamma \in \Gamma \setminus \Gamma, \hat{U}_\gamma \subseteq V_{\bar{n}}, \text{ where } i_\gamma \in \{1, \ldots, \bar{n} + 1\}$. $\forall i \in \{1, \ldots, \bar{n} + 1\}$ with $A_{i_\gamma,i} \neq \emptyset$, then $\exists x \in A_{i_\gamma,i}$, $\phi_n(x) = y_i$ and $f(x) \in B_y(y_i, \epsilon) \cap \hat{U}_\gamma \subseteq B_y(y_i, \epsilon) \cap V_{i_\gamma} \subseteq B_y(y_i, \epsilon) \cap \bar{V}_{i_\gamma} \subseteq B_y(y_i, \epsilon) \cap B_y(y_{i_\gamma}, \epsilon)$. Note that $\hat{y}_\gamma \in \hat{U}_\gamma \subseteq B_y(y_{i_\gamma}, \epsilon)$. Therefore, we have $\|\hat{y}_\gamma - y_i\| \leq \|\hat{y}_\gamma - y_{i_\gamma}\| + \|y_{i_\gamma} - f(x)\| + \|f(x) - y_i\| < 3\epsilon$.

Note that $F_R = \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma))$ and $\int_X \phi_n \, d\mu = \sum_{i=1}^{\bar{n}} y_i \mu(A_i)$, by Proposition 11.75. Then,

\[
\| F_R - \int_X \phi_n \, d\mu \| = \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} y_i \mu(A_i) \right\|
\]
\[
= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \left( \mu(f^{-1}_n(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1) \right) \right\|
\]
\[
= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \mu(A_{i_\gamma,i}) + \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i) \right\|
\]
\[
= \left\| \sum_{\gamma \in \Gamma} \left( \sum_{i=1}^{\bar{n}+1} \mu(A_{i_\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i) \right) \right\|
\]
\[
= \left\| \sum_{\gamma \in \Gamma} \left( \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{i_\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f^{-1}_n(\hat{U}_\gamma) \cap E_1) \right) \right\|
\]
\[
- \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \\
= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) \\
+ \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\| \\
= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \\
- \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\| \\
\leq \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \| \hat{y}_\gamma - y_i \| \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \| \hat{y}_\gamma \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \| \\
+ \sum_{\gamma \in \Gamma} \| \hat{y}_\gamma \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \| + \sum_{i=1}^{\bar{n}+1} \| y_i \| \mu(A_i \cap E_1) \\
\leq 3\bar{\epsilon} \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \\
+ \sum_{\gamma \in \Gamma} (M + 1) \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) + \sum_{i=1}^{\bar{n}+1} M \mu(A_i \cap E_1) \\
< \epsilon/2 + \sum_{\gamma \in \Gamma} (M + 1) \mu(\hat{f}_{\text{inv}}(\hat{U}_\gamma) \cap E_1) + \sum_{i=1}^{\bar{n}+1} M \mu(A_i \cap E_1) \\
= \epsilon/2 + (2M + 1) \mu(E_1) < \epsilon
\]

where the seventh equality follows from the fact that \( \forall \gamma \in \tilde{\Gamma}, A_{\gamma,i} \subseteq f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1) = \emptyset; \) and the third inequality follows from the fact that \( \forall \gamma \in \Gamma \setminus \tilde{\Gamma}, \hat{y}_\gamma \in \hat{U}_\gamma \subseteq V_i \subseteq \tilde{V}_i = B_{\gamma} (y_{i_1}, \bar{\epsilon}) \) and \( \| \hat{y}_\gamma \| \leq \| y_{i_1} \| + \| \hat{y}_\gamma - y_{i_1} \| < M + \bar{\epsilon} < M + 1. \)

In summary, we have shown that \( \forall \epsilon \in (0, 1) \subseteq \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n \geq n_0, \exists R \in \mathcal{I}(\mathbb{Y}) \) such that \( \forall R \in \mathcal{I}(\mathbb{Y}) \) with \( R \leq R, \) we have \( \| F_{R} - \int_X \phi_n \, d\mu \| < \epsilon. \) Then, the net \( (F_{R})_{R \in \mathcal{I}(\mathbb{Y})} \) is a Cauchy net. By Proposi-
11.4. INTEGRATION

Theorem 11.77 (Bounded Convergence Theorem) Let \((X, \mathcal{B}, \mu) =: \mathcal{X}\) be a finite measure space, \(\mathcal{Y}\) be a separable Banach space, \(f_n : X \to \mathcal{Y}\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), and \(f : X \to \mathcal{Y}\) be \(\mathcal{B}\)-measurable. Assume that \(\lim_{n \to \infty} f_n = f\) a.e. in \(\mathcal{X}\) and \(\exists M \in [0, \infty) \subset \mathbb{R}\) such that \(\| f_n(x) \| \leq M, \forall x \in X, \forall n \in \mathbb{N}\). Then, \(f\) is integrable over \(X\) and \(\int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathcal{Y}\).

Proof
Let \(U := \overline{\mathcal{B}}(\phi, M)\), which is a conic segment. Then, \(\forall n \in \mathbb{N}, f_n : X \to U \subseteq \mathcal{Y}\). By Proposition 11.66, there exists a sequence of simple functions \((\phi_{n,i})_{i=1}^{\infty}\), \(\phi_{n,i} : X \to U\), \(\forall i \in \mathbb{N}\), that converges to \(f_n\), a.e. in \(\mathcal{X}\) and \(\| \phi_{n,i}(x) \| \leq \| f_n(x) \| \leq M, \forall x \in X, \forall i \in \mathbb{N}\). By Lemma 11.76, \(\int_X f_n \, d\mu = \lim_{i \to \infty} \int_X \phi_{n,i} \, d\mu \in \mathcal{Y}\). By Proposition 11.58, we have \(\lim_{i \to \infty} \phi_{n,i} = f_n\) in measure in \(\mathcal{X}\). Then, \(\exists i_n \in \mathbb{N}\) such that \(\| \int_X f_n \, d\mu - \int_X \phi_{n,i_n} \, d\mu \| < 2^{-n}\) and \(\mu(E_n) := \mu(\{ x \in X \mid \| f_n(x) - \phi_{n,i_n}(x) \| \geq 2^{-n}\}) < 2^{-n}\). Denote \(\phi_n := \phi_{n,i_n}\).

By Proposition 11.58, \(\lim_{n \to \infty} f_n = f\) in measure in \(\mathcal{X}\). \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}\) such that \(\forall n \in \mathbb{N}\) with \(n \geq n_0\), we have \(2^{-n} < \epsilon/2\) and \(\mu(E_n) := \mu(\{ x \in X \mid \| f_n(x) - f(x) \| \geq \epsilon/2\}) < \epsilon/2\). Note that, by Lemma 11.43, \(\mathcal{B} \ni F_n := \{ x \in X \mid \| \phi_n(x) - f(x) \| \geq \epsilon \} \subseteq \{ x \in X \mid \| \phi_n(x) - f_n(x) \| + \| f_n(x) - f(x) \| \geq \epsilon \} \subseteq E_n \cup \overline{E_n}\). Then, we have \(\mu(F_n) \leq \mu(E_n) + \mu(\overline{E_n}) < \epsilon\). This implies that \(\lim_{n \to \infty} \phi_n = f\) in measure in \(\mathcal{X}\). By Proposition 11.57, there exists a subsequence \((\phi_{n_k})_{k=1}^{\infty}\) of \((\phi_n)_{n=1}^{\infty}\) that converges to \(f\) a.e. in \(\mathcal{X}\). By Lemma 11.76, \(f\) is integrable over \(X\) and \(\int_X f \, d\mu = \lim_{k \to \infty} \int_X \phi_{n_k} \, d\mu \in \mathcal{Y}\). By the fact that \(\| \int_X f_n \, d\mu - \int_X \phi_{n_k} \, d\mu \| < 2^{-n_k}, \forall k \in \mathbb{N}\), then we have \(\lim_{k \to \infty} \int_X f_n \, d\mu = \lim_{k \to \infty} \int_X \phi_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\). So far, we have shown that for any uniformly bounded \(\mathcal{B}\)-measurable sequence of functions \((f_n)_{n=1}^{\infty}\) that converges to \(f\) a.e. in \(\mathcal{X}\), there exists a subsequence \((f_{n_k})_{k=1}^{\infty}\) such that \(\lim_{k \to \infty} \int_X f_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\). This then holds for any subsequence of \((f_n)_{n=1}^{\infty}\). By Proposition 3.71, we have \(\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\) and \(f\) is integrable over \(X\).

This completes the proof of the theorem.

□

Proposition 11.78 Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a measure space and \(f : X \to [0, \infty) \subset \mathbb{R}\) be \(\mathcal{B}\)-measurable. Then,

\[
\int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu =: S \in \mathbb{R}_e
\]

where the supremum is over all simple functions \(\phi : X \to [0, \infty) \subset \mathbb{R}\) that satisfy \(0 \leq \phi(x) \leq f(x), \forall x \in X\).
CHAPTER 11. GENERAL MEASURE AND INTEGRATION

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \).

Case 1: \( \mu(X) < +\infty \). Let \( (F_R)_{R \in \mathcal{J}(\mathbb{R})} \) be the net as defined in Definition 11.70 for \( \int_X f \, d\mu \). Clearly, \( S \geq 0 \). Fix any simple function \( \phi : X \to [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi(x) \leq f(x) \), \( \forall x \in X \). Without loss of generality, we may assume that \( \phi \) is in canonical representation. Then, \( \exists n \in \mathbb{Z}_+ \), \( \exists y_1, \ldots, y_n \in [0, \infty) \subset \mathbb{R} \), which are distinct and none equals to 0, and \( \exists A_1, \ldots, A_n \in \mathcal{B} \), which are nonempty, pairwise disjoint, and \( \mu(A_i) < \infty \), \( i = 1, \ldots, n \), such that \( \phi = \sum_{i=1}^n y_i \chi_{A_i} \). Let \( y_{n+1} := 0 \) and \( A_{n+1} := X \setminus \bigcup_{i=1}^n A_i \in \mathcal{B} \). Then, \( X = \bigcup_{i=1}^{n+1} A_i \) and the sets in the union are pairwise disjoint and of finite measure. Without loss of generality, assume that \( 0 = y_{n+1} < y_n < \cdots < y_1 < +\infty \). Let \( I_1 := [y_1, \infty) \subset \mathbb{R} \), \( I_i := [y_i, y_{i-1}) \subset \mathbb{R} \), \( i = 2, \ldots, n + 1 \), \( I_{n+2} := (-\infty, 0) \subset \mathbb{R} \), and \( y_{n+2} := -1 \). Then, \( R_0 := \{ (y_i, I_i) \mid i = 1, \ldots, n + 2 \} \in \mathcal{J}(\mathbb{R}) \). \( \forall R \in \mathcal{J}(\mathbb{R}) \) with \( R_0 \preceq R \), let \( R = \{ (y_{\alpha}, \U_\alpha) \mid \alpha \in \Lambda \} \). \( \forall \alpha \in \Lambda \), by \( R_0 \preceq R \), \( \exists \alpha_i \in \{ 1, \ldots, n + 2 \} \) such that \( \U_\alpha \subseteq I_{\alpha_i} \). Define \( \Lambda := \{ \alpha \in \Lambda \mid i_\alpha = n + 2 \} \). Note that \( \forall \alpha \in \Lambda \setminus \Lambda, \forall i \in \{ 1, \ldots, n + 1 \} \), if \( E_{i,\alpha} \neq \emptyset \), then \( \exists x \in E_{i,\alpha} \), \( y_i = \phi(x) \leq f(x) \in \U_\alpha \subseteq I_{i_\alpha} \), which further implies that \( y_i \leq y_{i_\alpha} \). Note that \( y_{\alpha_i} \in \U_\alpha \subseteq I_{i_\alpha} \) and \( y_{\alpha_i} \geq y_i \). Then, \( \bar{y}_\alpha \mu(E_{i,\alpha}) \geq y_i \mu(E_{i,\alpha}) \). If \( E_{i,\alpha} = \emptyset \), clearly \( 0 = \bar{y}_\alpha \mu(E_{i,\alpha}) = y_i \mu(E_{i,\alpha}) = 0 \). Therefore, we have \( \bar{y}_\alpha \mu(E_{i,\alpha}) \geq y_i \mu(E_{i,\alpha}) \), \( \forall \alpha \in \Lambda \setminus \Lambda, \forall i \in \{ 1, \ldots, n + 1 \} \). This leads to

\[
F_R = \sum_{\alpha \in \Lambda \setminus \Lambda} y_{\alpha_i} \mu(E_{i,\alpha}) \geq \sum_{\alpha \in \Lambda \setminus \Lambda} y_i \mu(E_{i,\alpha}) = \sum_{i=1}^{n+1} y_i \mu(A_i) = \int_X \phi \, d\mu
\]

where the last equality follows from Proposition 11.75. In summary, we have shown that \( \forall R \in \mathcal{J}(\mathbb{R}) \) with \( R_0 \preceq R \), we have \( F_R \geq \int_X \phi \, d\mu \). Then, \( \liminf_{R \in \mathcal{J}(\mathbb{R})} F_R \geq \int_X \phi \, d\mu \). By the arbitrariness of \( \phi \), we have \( \liminf_{R \in \mathcal{J}(\mathbb{R})} F_R \geq S \).

We will show that \( \limsup_{R \in \mathcal{J}(\mathbb{R})} F_R \leq S \) by an argument of contradiction. Suppose \( \bar{S} := \limsup_{R \in \mathcal{J}(\mathbb{R})} F_R > S \). Then, \( \exists \epsilon_0 \in (0, \infty) \subset \mathbb{R} \), \( \forall R_0 \in \mathcal{J}(\mathbb{R}) \), \( \exists R \in \mathcal{J}(\mathbb{R}) \) with \( R_0 \preceq R \), such that \( F_R < S + \epsilon_0 \). Let \( \epsilon_1 := \min\{ 1, \frac{\epsilon_0}{2 + 2\mu(X)} \} > 0 \). By Proposition 11.5, \( \exists E_1 \in \mathcal{B} \) with \( \mu(E_1) < \epsilon_1 \).
and \( \exists M \in \mathbb{N} \) such that \( 0 \leq f(x) \leq M - 1, \forall x \in X \setminus E_1 \). Then, \( \exists n \in \mathbb{N} \) such that \( M/n < \epsilon_1 \). Let \( U_i := [(i-1)M/n, iM/n) \subset \mathbb{R} \), \( y_i := (i-1)M/n \), \( i = 1, \ldots, n, U_{n+1} := [M, \infty) \subset \mathbb{R} \), \( y_{n+1} := M, U_{n+2} := (-\infty, 0) \subset \mathbb{R} \), and \( y_{n+2} = -1 \). Then, \( R_0 := \{(y_i, U_i) \mid i = 1, \ldots, n + 2\} \subset J(\mathbb{R}) \). Then, \( \exists R \in J(\mathbb{R}) \) with \( R_0 \leq R \) such that \( F_R > S + \epsilon_0 \). Let \( R = \{(\tilde{y}, \tilde{U}) \mid \alpha \in \Lambda\} \). \( \forall \alpha \in \Lambda \), \( \exists \ i_{\alpha} \in \{1, \ldots, n + 2\} \) such that \( \tilde{U} \alpha \subseteq U_{i_{\alpha}} \).

Define \( \Lambda := \{\alpha \in \Lambda \mid i_{\alpha} = n + 2\} \), \( \Lambda_1 := \{\alpha \in \Lambda \mid i_{\alpha} = 1, \ldots, n\} \), and \( \Lambda_2 := \{\alpha \in \Lambda \mid i_{\alpha} = n + 1\} \). Clearly, \( \Lambda = \Lambda_1 \cup \Lambda_2 \cup \Lambda \) and the sets in the union are pairwise disjoint. \( \forall \alpha \in \Lambda \), \( \tilde{U} \alpha \subseteq U_{n+2} = (-\infty, 0) \subset \mathbb{R} \). Then, \( f_{\text{inv}}(\tilde{U} \alpha) = 0 \). \( \forall \alpha \in \Lambda_1 \), note that \( \tilde{y}_{i_{\alpha}} \in U_{i_{\alpha}} \setminus U_i \), then \( y_{i_{\alpha}} \leq \tilde{y}_{i_{\alpha}} < y_{i_{\alpha} + 1} \). Let \( \tilde{y}_{i_{\alpha}} := \inf_{y \in \tilde{U} \alpha} y \in U_{i_{\alpha}} \subseteq U_{i_{\alpha} - 1} \). Then, \( y_{i_{\alpha}} \leq \tilde{y}_{i_{\alpha}} \leq y_{i_{\alpha} + 1} \). Let \( A_\alpha := f_{\text{inv}}(\tilde{U} \alpha) \). \( \forall \alpha \in \Lambda_2 \), note that \( \tilde{y}_{i_{\alpha}} \in U_{i_{\alpha}} \subseteq U_{n+1} = [M, \infty) \subset \mathbb{R} \) and \( \tilde{y}_{i_{\alpha}} = \lfloor \tilde{y}_{i_{\alpha}} \rfloor + 1 = \inf_{y \in \tilde{U} \alpha} \lfloor y \rfloor + 1 \). Let \( \tilde{y}_{i_{\alpha}} := \inf_{y \in \tilde{U} \alpha} y \).

Then, \( \tilde{y}_{i_{\alpha}} \leq \tilde{y}_{i_{\alpha}} < \tilde{y}_{i_{\alpha} + 1} \). Let \( A_\alpha := f_{\text{inv}}(\tilde{U} \alpha) \). Since \( \tilde{U} \alpha \subseteq U_{n+1} \), then \( f_{\text{inv}}(\tilde{U} \alpha) \subseteq E_1 \).

Define a simple function \( \phi : X \to [0, \infty) \subset \mathbb{R} \) by \( \phi := \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_{i_{\alpha}} \chi_{\tilde{A}_{\alpha}, x} \).

Clearly, \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). Then, by Proposition 11.75,

\[
\int_X \phi \, d\mu = \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_{i_{\alpha}} \mu(A_{\alpha}) = \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_{i_{\alpha}} \mu(f_{\text{inv}}(\tilde{U} \alpha)) - F_R + F_R \\
\geq S + \epsilon_0 + \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} (\tilde{y}_{i_{\alpha}} - \tilde{y}_{i_{\alpha}}) \mu(f_{\text{inv}}(\tilde{U} \alpha)) - F_R \\
= S + \epsilon_0 + \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} (\tilde{y}_{i_{\alpha}} - \tilde{y}_{i_{\alpha}}) \mu(f_{\text{inv}}(\tilde{U} \alpha)) \\
\geq S + \epsilon_0 - M/n \sum_{\alpha \in \Lambda_1} \mu(f_{\text{inv}}(\tilde{U} \alpha)) - \sum_{\alpha \in \Lambda_2} \mu(f_{\text{inv}}(\tilde{U} \alpha)) \\
\geq S + \epsilon_0 - \epsilon_1 \mu(X) - \mu(E_1) > S
\]

This contradicts with the definition of \( S \). Therefore, \( \limsup_{R \in J(\mathbb{R})} F_R \leq S \).

Then, we have \( S \leq \liminf_{R \in J(\mathbb{R})} F_R \leq \limsup_{R \in J(\mathbb{R})} F_R \leq S \). By Proposition 3.83, \( S = \lim_{R \in J(\mathbb{R})} F_R = \int_X f \, d\mu \).

Case 2: \( \mu(X) = +\infty \). Let \( (F_A)_{A \in \mathcal{M}(X)} \) be the net as defined in Definition 11.71. \( \forall A \in \mathcal{M}(X), A \in \mathcal{B} \) and \( \mu(A) < \infty \). Let \( (A, B_A, \mu_A) \) be the measure subspace of the measure space \( \mathcal{X} \) as defined in Proposition 11.13. By Case 1, \( (F_A)_{A \in \mathcal{M}(X)} \subseteq \mathcal{R}_e \) is well-defined.

Fix any simple function \( \phi : X \to [0, \infty) \subset \mathbb{R} \) with \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). Then, \( \exists n \in \mathbb{Z}_+ \), \( \exists y_1, \ldots, y_n \in [0, \infty) \subset \mathbb{R} \), and \( \exists A_1, \ldots, A_n \in \mathcal{B} \), which are of finite measure, such that \( \phi = \sum_{i=1}^n y_i \chi_{A_i, x} \). Take \( A_0 = \).
\[ \bigcup_{i=1}^{n} A_i \in \mathcal{B}. \] Then, \( \mu(A_0) \leq \sum_{i=1}^{n} \mu(A_i) < +\infty. \) Then, \( A_0 \in \mathcal{M}(\mathcal{X}) \).

\( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_0 \subseteq A \), we have, \( F_A = \sup_{0 \leq \phi \leq f_A} \int_A \phi \, d\mu_A \geq \int_A \phi \, d\mu_A = \sum_{i=1}^{n} y_i \mu(A_i) = \int_X \phi \, d\mu \), where the last two equalities follow from Proposition 11.75. Then, \( \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq \int_X \phi \, d\mu \). By the arbitrariness of \( \phi \), we have \( \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S \).

On the other hand, \( \sup_{\phi \in \mathcal{M}(\mathcal{X})} F_A =: S \geq \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S. \) Suppose \( \tilde{S} > S \). Then, \( \exists A \in \mathcal{M}(\mathcal{X}) \) such that \( S < F_A = \int_A f \, d\mu_A = \sup_{0 \leq \phi \leq f_A} \int_A \phi \, d\mu_A \). This implies that \( \exists \) a simple function \( \phi : A \to [0, \infty) \subset \mathbb{R} \) with \( 0 \leq \phi(x) \leq f(x), \forall x \in A \), such that \( \int_A \phi \, d\mu_A > S \).

We may extend this simple function on \( A \) to a simple function on \( X \) that is zero on \( X \setminus A \). Let the extended simple function be \( \tilde{\phi} : X \to [0, \infty) \subset \mathbb{R} \). Then, \( \forall x \in X \), and by Proposition 11.75, \( \int_X \tilde{\phi} \, d\mu = \int_A \phi \, d\mu_A > S \). This contradicts with definition of \( S \). Therefore, we have \( \tilde{S} \leq S \).

In summary, we have \( \exists S : \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \limsup_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \sup_{\phi \in \mathcal{M}(\mathcal{X})} F_A = \tilde{S} \leq S. \) Then, by Proposition 3.83, we have \( \int_X f \, d\mu = \liminf_{A \in \mathcal{M}(\mathcal{X})} F_A = S \in \mathbb{R}. \)

This completes the proof of the proposition.

**Definition 11.79** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( f : X \to [0, \infty] \subset \mathbb{R}_e \) be \( \mathcal{B} \)-measurable. We will define the integral of \( f \) over \( \mathcal{X} \) by \( \int_X f \, d\mu := \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu =: S \in [0, \infty] \subset \mathbb{R}_e \), where the supremum is over all simple functions \( \phi : X \to [0, \infty] \subset \mathbb{R} \) that satisfy \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). We will say that \( f \) is integrable if \( S \in \mathbb{R} \).

This definition generalizes the definition of integral to nonnegative extended real-valued functions. By Proposition 11.78, it is consistent with our earlier definitions of integral. The reason that we introduce this generalization is that this leads to considerable simplification in the integrals of nonnegative functions over product measure spaces.

**Theorem 11.80** (Fatou’s Lemma) Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( f_n : X \to [0, \infty] \subset \mathbb{R}_e \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( f : X \to [0, \infty] \subset \mathbb{R}_e \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \in \mathbb{N}} f_n = f \) a.e. in \( \mathcal{X} \). Then, \( \int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu \).

**Proof** By Definition 11.79, \( \int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu \), where the supremum is over all simple function \( \phi : X \to [0, \infty] \subset \mathbb{R} \) that satisfy \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). Fix any such simple function \( \phi \). Let \( E := \{ x \in X \mid \phi(x) > 0 \} \). We must have \( E \in \mathcal{B} \) and \( \mu(E) < +\infty \). Let \( M \in [0, \infty] \subset \mathbb{R} \) be such that \( 0 \leq \phi(x) \leq M, \forall x \in X \). Define \( g_n := f_n \land \phi, \forall n \in \mathbb{N} \). By Proposition 11.40, \( g_n \)’s are \( \mathcal{B} \)-measurable. Since \( \lim_{n \in \mathbb{N}} f_n = f \) a.e. in \( \mathcal{X} \) and \( 0 \leq \phi(x) \leq f(x), \forall x \in X \), then \( \lim_{n \in \mathbb{N}} g_n = \phi \) a.e. in \( \mathcal{X} \). By Bounded Convergence Theorem 11.77 and Proposition 11.75, we have \( \int_X \phi \, d\mu = \int_E \phi \, d\mu_E = \lim_{n \in \mathbb{N}} \int_E g_n \, d\mu_E \),
where \((E, \mathcal{B}_E, \mu_E)\) is the finite measure subspace of \(X\) as defined in Proposition 11.13. By Definition 11.79, we have \(\int_E g_n \, d\mu_E \leq \int_X g_n \, d\mu \leq \int_X f_n \, d\mu\). Then, \(\int_X \phi \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu\). By the arbitrariness of \(\phi\), we have \(\int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu\). This completes the proof of the theorem. 

\(\square\)

**Theorem 11.81 (Monotone Convergence Theorem)** Let \(X := (X, \mathcal{B}, \mu)\) be a measure space, \(f_n : X \to [0, \infty] \subset \mathbb{R}_e\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), and \(f : X \to [0, \infty] \subset \mathbb{R}_e\) be \(\mathcal{B}\)-measurable. Assume that \(\lim_{n \in \mathbb{N}} f_n = f\) a.e. in \(X\) and \(f_n(x) \leq f_{n+1}(x), \forall x \in X, \forall n \in \mathbb{N}\). Then, \(\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu\).

**Proof** By the assumption, \(f_n \leq f\) a.e. in \(X\), \(\forall n \in \mathbb{N}\). By Definition 11.79 and the fact that any two simple functions that equal to each other almost everywhere have the same integral, we have \(\mathbb{R}_e \ni \int_X f \, d\mu \geq \int_X f_n \, d\mu \in \mathbb{R}_e, \forall n \in \mathbb{N}\). By Fatou’s Lemma 11.80, we have \(\int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu \leq \limsup_{n \in \mathbb{N}} \int_X f_n \, d\mu \leq \int_X f \, d\mu\). Hence, by Proposition 3.83, we have \(\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathbb{R}_e\). This completes the proof of the proposition. 

\(\square\)

**Proposition 11.82** Let \(X := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite measure space and \(g : X \to [0, \infty] \subset \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that \(\int_X g \, d\mu =: \lambda < \infty\).

Then, define \(f : X \to [0, \infty] \subset \mathbb{R}\) by \(f(x) = \begin{cases} g(x) & \text{if } g(x) < \infty \\ 0 & \text{if } g(x) = \infty \end{cases}\), \(\forall x \in X\).

Then, \(f\) is \(\mathcal{B}\)-measurable, \(f = g\) a.e. in \(X\) and \(\int_X f \, d\mu = \int_X g \, d\mu \in \mathbb{R}\).

**Proof** Let \(E := \{ x \in X \mid g(x) < \infty \} \in \mathcal{B}\). We will show that \(\mu(X \setminus E) = 0\) by an argument of contradiction. Suppose \(\mu(X \setminus E) > 0\). Since \(X\) is \(\sigma\)-finite, \(\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) such that \(X = \bigcup_{n=1}^{\infty} X_n\) and \(\mu(X_n) < +\infty, \forall n \in \mathbb{N}\). Without loss of generality, we may assume \(X_n \subseteq X_{n+1}\), \(\forall n \in \mathbb{N}\). Let \(B_n := X_n \setminus E \in \mathcal{B}, \forall n \in \mathbb{N}\). By Proposition 11.7, we have \(\lim_{n \in \mathbb{N}} \mu(B_n) = \mu(X \setminus E) > 0\). Then, \(\exists n \in \mathbb{N}\) such that \(\mu(B_n) \in (0, +\infty) \subset \mathbb{R}\). Clearly \(\lambda \in [0, \infty) \subset \mathbb{R}\). Let \(M := \frac{\lambda}{\mu(B_n)} \in (0, \infty) \subset \mathbb{R}\). Let \(\tilde{g} := g^\wedge M\). By Proposition 11.40, \(\tilde{g}\) is \(\mathcal{B}\)-measurable. Then, \(\tilde{g} : X \to [0, M] \subset \mathbb{R}\) and \(\lambda = \int_X g \, d\mu \geq \int_X \tilde{g} \, d\mu \geq \mu(B_n) = \lambda + 1 > \lambda\), where the first inequality follows from Definition 11.79; and the second inequality follows from Definition 11.79 and Proposition 11.75. This is a contradiction. Hence, \(\mu(X \setminus E) = 0\).

Clearly, \(f : X \to [0, \infty] \subset \mathbb{R}\) satisfies \(f(x) = \begin{cases} g(x) & \text{if } x \in E \\ 0 & \text{if } x \in X \setminus E \end{cases}\). By Proposition 11.41, \(f\) is \(\mathcal{B}\)-measurable and \(f = g\) a.e. in \(X\). By Monotone Convergence Theorem 11.81, \(\int_X f \, d\mu = \int_X g \, d\mu = \lambda \in \mathbb{R}\). This completes the proof of the proposition. 

\(\square\)

The above proposition says that for a measurable extended real-valued function, if its integral is finite then it can be modified to be a real-valued function by replacing its \(\infty\) value by zero on a set of measure zero, and the integral remains the same for modified and the unmodified function.
Proposition 11.83 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $f_i : X \to [0, \infty] \subseteq \mathbb{R}_+$ be $\mathcal{B}$-measurable, $i = 1, 2$, $g_n : X \to [0, \infty] \subseteq \mathbb{R}_+$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, and $g : X \to [0, \infty] \subseteq \mathbb{R}_+$ be $\mathcal{B}$-measurable. Then, the following statements hold.

(i) $\forall c \in (0, \infty) \subseteq \mathbb{R}$, $\int_X (cg) \, d\mu = c \int_X g \, d\mu; \forall c \in [0, \infty) \subseteq \mathbb{R}$, if $g : X \to [0, \infty)$ and $\int_X g \, d\mu < \infty$, then $\int_X (cg) \, d\mu = c \int_X g \, d\mu$.

(ii) If $f_1 \leq f_2$ a.e. in $\mathcal{X}$, then $\int_X f_1 \, d\mu \leq \int_X f_2 \, d\mu$.

(iii) $\forall E \in \mathcal{B}$, we have $\int_E g \, d\mu_E = \int_X (g|_{\mathcal{X}(E)}) \, d\mu = \int_E g \, d\mu \leq \int_X g \, d\mu$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the measure subspace of $\mathcal{X}$.

(iv) $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathbb{R}_+$.

(v) $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \cap E_2 = \emptyset$, we have $\int_{E_1 \cup E_2} g \, d\mu = \int_{E_1} g \, d\mu + \int_{E_2} g \, d\mu \in \mathbb{R}_+$.

(vi) $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$, we have $\int_{\bigcup_{n=1}^\infty E_n} g \, d\mu = \sum_{n=1}^\infty \int_{E_n} g \, d\mu \in \mathbb{R}_+$.

(vii) If $g$ and $g_n$'s are integrable over $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$, and $\lim_{n \in \mathbb{N}} \int_X g_n \, d\mu = \int_X g \, d\mu \in \mathbb{R}$. Then, $\forall E \in \mathcal{B}$, we have $\int_E g \, d\mu = \lim_{n \in \mathbb{N}} \int_{E \cap X} \phi \, d\mu$.

Proof (i) and (ii) follows directly from Propositions 11.75 and 3.81 and Definition 11.79.

(iii) Note that $g|_{\mathcal{E}}$ is $\mathcal{B}_E$-measurable by Proposition 11.41. By Propositions 7.23, 11.38, and 11.39, $g|_{\mathcal{X}(E)}$ is $\mathcal{B}$-measurable. Thus, the result follows directly from Definition 11.79 and Proposition 11.75.

(iv) By Propositions 7.23, 11.38, and 11.39, $f_1 + f_2$ is $\mathcal{B}$-measurable. By Definition 11.79, we have $\int_X f_1 \, d\mu = \sup_{0 \leq \phi \leq f_1} \int_X \phi \, d\mu$, $\int_X f_2 \, d\mu = \sup_{0 \leq \phi \leq f_2} \int_X \phi \, d\mu$, and $\int_X (f_1 + f_2) \, d\mu = \sup_{0 \leq \phi \leq f_1 + f_2} \int_X \phi \, d\mu$, where the suprema are over all simple functions $\phi : X \to [0, \infty) \subseteq \mathbb{R}$ satisfying the stated inequalities. Any simple function $\phi_1 : X \to [0, \infty) \subseteq \mathbb{R}$ and $\phi_2 : X \to [0, \infty) \subseteq \mathbb{R}$ satisfying $0 \leq \phi_1 \leq f_1$ and $0 \leq \phi_2 \leq f_2$, we have $\phi_1 + \phi_2$ is a simple function satisfying $0 \leq \phi_1 + \phi_2 \leq f_1 + f_2$. Then, $\int_X \phi_1 \, d\mu + \int_X \phi_2 \, d\mu = \int_X \phi_1 \, d\mu \leq \int_X (f_1 + f_2) \, d\mu$ and $\mu(E) < +\infty$ since $\phi$ is a simple function. Let $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ be the finite measure subspace of $\mathcal{X}$ as
11.4. INTEGRATION

defined in Proposition 11.13. By Proposition 11.66, there exists sequences of simple functions \((\phi_{1,n})_{n=1}^\infty\) and \((\phi_{2,n})_{n=1}^\infty\), \(\phi_{1,n} : E \to [0, M] \subseteq \mathbb{R}\), \(\forall i \in \{1, 2\}, \forall n \in \mathbb{N}\), that converges to \(f_1|_E\) and \(f_2|_E\) a.e. in \(\mathcal{E}\), respectively. By Bounded Convergence Theorem 11.77, we have \(\int_E f_1|_E d\mu_E = \lim_{n \to \infty} \int_E \phi_{1,n} d\mu_E\) and \(\int_E f_2|_E d\mu_E = \lim_{n \to \infty} \int_E \phi_{2,n} d\mu_E\). By Propositions 11.52, 7.23, 11.53, \(\lim_{n \to \infty} \int_E (\phi_{1,n} + \phi_{2,n}) = \int_E f_1|_E + \int_E f_2|_E = \phi|_E\) a.e. in \(\mathcal{E}\). Again, by Bounded Convergence Theorem 11.77 and Proposition 11.75, we have \(\int_E \phi|_E d\mu_E = \lim_{n \to \infty} \int_E \phi_{1,n} d\mu_E + \int_E \phi_{2,n} d\mu_E = (\int_E f_1|_E d\mu_E + \int_E f_2|_E d\mu_E),\) By Proposition 11.75, (ii), and (iii), we have \(\int_X \phi|_E d\mu = \int_X \phi|_E d\mu_E = \int_X f_1|_E d\mu + \int_X f_2|_E d\mu\). By the arbitrariness of \(\phi\), we have \(\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu\). Therefore, we have \(\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu\).

(v) By (iii) and (iv), we have \(\int_{E_1 \cup E_2} g d\mu = \int_{E_1} (g\chi_{E_1, X}) d\mu = \int_X (g\chi_{E_1, X}) d\mu + \int_X (g\chi_{E_2, X}) d\mu = \int_X g d\mu + \int_{E_2} g d\mu \subseteq \mathbb{R}_e\).

(vi) Fix any pairwise disjoint \((E_n)_{n=1}^\infty \subseteq \mathcal{B}\). Let \(E := \bigcup_{n=1}^\infty E_n \subseteq \mathcal{B}\) and \(E_n := \bigcup_{i=1}^n E_i \subseteq \mathcal{B}, \forall n \in \mathbb{N}\). By Propositions 7.23, 11.38, and 11.39, \(g\chi_{E_n, X}\) and \(g\chi_{E, X}\) are \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\). Clearly, \(\lim_{n \to \infty} g(x)\chi_{E_n, X}(x) = g(x)\chi_{E, X}(x), \forall x \in X\), and \(0 \leq g(x)\chi_{E_n, X}(x) \leq g(x)\chi_{E_n, X}(x), \forall x \in X, \forall n \in \mathbb{N}\). By Monotone Convergence Theorem 11.81, \(\int_E g d\mu = \int_X (g\chi_{E_n, X}) d\mu = \lim_{n \to \infty} \int_X (g\chi_{E_n, X}) d\mu = \lim_{n \to \infty} \sum_{i=1}^n \int_{E_i} g d\mu = \sum_{i=1}^\infty \int_{E_i} g d\mu,\) where the fourth equality follows from (v).

(vii) By (iii), we have \(0 \leq \int_{E_1 \setminus E} g d\mu \leq \int_X (g\chi_{E_1, X}) d\mu < +\infty, 0 \leq \int_{X \setminus E_1} g d\mu \leq \int_X g d\mu < +\infty, 0 \leq \int_{X \setminus E_1} g d\mu \leq \int_X g d\mu < +\infty, \forall n \in \mathbb{N}\). By Fatou’s Lemma 11.80, \(\int_{E_1 \setminus E} g d\mu \leq \liminf_{n \to \infty} \int_{E_1} g d\mu \leq \liminf_{n \to \infty} \int_X g d\mu\). By a similar argument, \(\int_{X \setminus E_1} g d\mu \leq \liminf_{n \to \infty} \int_{X \setminus E_1} g d\mu\). By (v) and Proposition 3.83, we have \(\int_E g d\mu = \limsup_{n \to \infty} \int_{E_1} g d\mu \leq \limsup_{n \to \infty} \int_{E_1 \setminus E} g d\mu \leq \limsup_{n \to \infty} \left( \int_{E_1} g d\mu - \int_{X \setminus E_1} g d\mu \right) = \limsup_{n \to \infty} \int_{X \setminus E_1} g d\mu \leq \int_X g d\mu - \int_{X \setminus E_1} g d\mu = \int_E g d\mu\). By Proposition 3.83, we have \(\int_E g d\mu = \lim_{n \to \infty} \int_{E_n} g d\mu \subseteq \mathbb{R}\).

This completes the proof of the proposition. 

\begin{proof}
\end{proof}

Proposition 11.84 Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a measure space and \(f : X \to [0, \infty] \subseteq \mathbb{R}_e\) be \(\mathcal{B}\)-measurable. Assume that \(f\) is integrable over \(\mathcal{X}\), that is, \(\int_X f d\mu \in \mathbb{R}\). Then, \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists \delta \in (0, \infty) \subseteq \mathbb{R}\) such that \(\forall A \subseteq \mathcal{B}\) with \(\mu(A) < \delta\), we have \(\int_A f_A d\mu_A < \epsilon\), where \(A := (A, \mathcal{B}_A, \mu_A)\) is the measure subspace of \(\mathcal{X}\) as defined in Proposition 11.13.

\begin{proof}
\end{proof}
Proposition 11.40, \( f_n \) is \( \mathcal{B} \)-measurable. By Propositions 7.23, 11.38, and 11.39, \( f - f_n \) is \( \mathcal{B} \)-measurable. By Monotone Convergence Theorem 11.81, \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N} \) such that \( \left| \int_X f \, d\mu - \int_X f_n \, d\mu \right| < \epsilon/2 \) Choose \( \delta := \frac{\epsilon}{2n_0} \in (0, \infty) \subset \mathbb{R} \). \( \forall A \in \mathcal{B} \) with \( \mu(A) < \delta \), by Proposition 11.83 and Definition 11.79, we have

\[
\int_A f |_A \, d\mu_A = \int_A (f - f_n) |_A \, d\mu_A + \int_A f_n |_A \, d\mu_A
\]

\[
\leq \int_X (f - f_n) \, d\mu + n_0 \mu(A) = \int_X f \, d\mu - \int_X f_n \, d\mu + n_0 \mu(A)
\]

\[
< \epsilon/2 + \epsilon/2 = \epsilon
\]

This completes the proof of the proposition. \( \square \)

11.5 General Convergence Theorems

Lemma 11.85 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( f_n : X \to [0, \infty] \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable, \( f : X \to [0, \infty] \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \to \infty} f_n = f \) a.e. in \( \mathcal{X} \) and \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathbb{R} \). Then, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R} \) and \( \forall n_0 \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \) with \( \mu(A) < \delta \), \( \forall n \in \mathbb{N} \) with \( n \geq n_0 \), we have \( 0 \leq \int_A f |_A \, d\mu_A < \epsilon \) and \( 0 \leq \int_A f_n |_A \, d\mu_A < \epsilon \), where \( (A, \mathcal{B}_A, \mu_A) \) is the finite measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13.

Proof Let \( \bar{f}_n := f_n \wedge f, \forall n \in \mathbb{N} \). By Proposition 11.40, \( \bar{f}_n \) is \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). Then, \( 0 \leq \bar{f}_n(x) \leq f(x), \forall x \in X, \forall n \in \mathbb{N}, \) and \( \int_X \bar{f}_n \, d\mu \leq \int_X f \, d\mu, \forall n \in \mathbb{N}, \) by Definition 11.79. By Proposition 11.50, \( \lim_{n \to \infty} \bar{f}_n = f \) a.e. in \( \mathcal{X} \). By Fatou’s Lemma 11.80, \( \int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu \leq \limsup_{n \to \infty} \int_X f_n \, d\mu \leq \int_X f \, d\mu, \) where the last two inequalities follows from Proposition 3.83 and Definition 3.82, respectively. By Proposition 3.83, \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathbb{R} \).

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by Proposition 11.84, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) and \( \exists n_0 \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \) with \( \mu(A) < \delta \), \( \forall n \in \mathbb{N} \) with \( n \geq n_0 \), we have \( 0 \leq \int_A f |_A \, d\mu_A < \epsilon/3, \) \( \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu | < \epsilon/3, \) and \( \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu | < \epsilon/3. \) Then,

\[
0 \leq \int_A f_n |_A \, d\mu_A = \int_A (f_n - \bar{f}_n) |_A \, d\mu_A + \int_A \bar{f}_n |_A \, d\mu_A
\]

\[
\leq \int_A (f_n - \bar{f}_n) |_A \, d\mu_A + \int_A f |_A \, d\mu_A + \int_X (f_n - \bar{f}_n) \, d\mu + \epsilon/3
\]

\[
= \int_X f_n \, d\mu - \int_X \bar{f}_n \, d\mu + \epsilon/3
\]

\[
\leq \epsilon/3 + \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu \}

| \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu | < \epsilon
where the first inequality follows from Definition 11.79; the first equality, the second inequality, and the third inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.83 and the fact that \( \int_X f_\alpha \, d\mu < +\infty \). This completes the proof of the lemma.

**Lemma 11.86** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} \) be a Banach space, \( \mathcal{W} \) be a separable subspace of \( \mathcal{Y} \), \( \phi_\alpha : X \to \mathcal{W} \) be a simple function, \( \forall n \in \mathbb{N}, f : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable, \( g_\alpha : X \to [0, \infty) \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( g : X \to [0, \infty) \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable. Assume that

(i) \( \lim_{n \in \mathbb{N}} \phi_\alpha = f \) a.e. in \( \mathcal{X} \) and \( \lim_{n \in \mathbb{N}} g_\alpha = g \) a.e. in \( \mathcal{X} \);

(ii) \( \| \phi_\alpha(x) \| \leq g_\alpha(x), \forall x \in X, \forall n \in \mathbb{N} \);

(iii) \( g_\alpha \)'s and \( g \) are integrable over \( X \) and \( \lim_{n \in \mathbb{N}} \int_X g_\alpha \, d\mu = \int_X g \, d\mu \in \mathbb{R} \).

Then, \( f \) is integrable over \( X \) and \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_\alpha \, d\mu \in \mathbb{R} \).

**Proof** Let \( E := \{ x \in X \mid (\phi_\alpha(x))_{n=1}^\infty \) does not converge to \( f(x) \) or \( (g_\alpha(x))_{n=1}^\infty \) does not converge to \( g(x) \} \). Then, by (i), \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). Let \( (F_\alpha)_{R \in (y)} \) be the net for \( \int_X f \, d\mu \) as defined in Definition 11.70.

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by (i), (iii), and Lemma 11.85, \( \exists \delta \in (0, \epsilon/6) \subset \mathbb{R} \) and \( \exists n_0 \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \) with \( \mu(A) < \delta \), \( \forall n \in \mathbb{N} \) with \( n \geq n_0 \), we have \( 0 \leq \int_A \phi_A \, d\mu_A < \epsilon/6 \) and \( 0 \leq \int_A g_n \, d\mu_A < \epsilon/6 \), where \( (A, \mathcal{B}_A, \mu_A) \) is the finite measure subspace of \( \mathcal{X} \). By Egoroff's Theorem 11.55, \( \exists E_1 \in \mathcal{B} \) with \( \mu(E_1) < \delta \) such that \( \left( \phi_\alpha\mid_{X \setminus E_1} \right)_{n=1}^\infty \) converges uniformly to \( f|_{X \setminus E_1} \).

Let \( E_1 := E \cup E_1 \). Then, \( E_1 \in \mathcal{B} \) with \( \mu(E_1) = \mu(E_1) < \delta \) such that \( \left( \phi_\alpha\mid_{X \setminus E_1} \right)_{n=1}^\infty \) converges uniformly to \( f|_{X \setminus E_1} \). Then, \( \forall n \in \mathbb{N} \) with \( n_1 \geq n_0, \forall x \in X \setminus E_1, \| \phi_\alpha(x) - f(x) \| < \frac{\epsilon}{6\mu(X)+\epsilon} \). Fix any \( n \geq n_1 \). Let \( \phi_n \) admit the canonical representation \( \phi_n = \sum_{i=1}^n w_i \chi_{A_i \setminus X} \), where \( \tilde{w}_i \in \mathbb{Z}_+ \), \( w_1, \ldots, w_n \in \mathbb{W} \) are distinct and none equals to \( \tilde{v}_y \), and \( A_1, \ldots, A_n \in \mathcal{B} \) are nonempty, pairwise disjoint, and of finite measure. Let \( \tilde{w}_{n+1} := \tilde{v}_y \) and \( \tilde{A}_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B} \).

Then, \( X = \bigcup_{i=1}^{n+1} A_i \) and the sets in the union are pairwise disjoint and of finite measure. Define \( V_i := B_{B}(w_i, \epsilon), \) \( i = 1, \ldots, n + 1 \). Define \( V_i := V_i \cap \mathcal{B}(y), V_{i+1} := V_i + (\bigcup_{j=1}^i V_j) \in \mathcal{B}(y), \) \( i = 1, \ldots, n \).

Let \( V_{n+2} := y \setminus (\bigcup_{j=1}^n V_j) \in \mathcal{B}(y) \). Clearly, we may form a representation \( \tilde{R} = \{ (\tilde{y}_i, V_i) \mid i = 1, \ldots, n + 2, V_i \neq \emptyset \} \in \mathcal{I}(y) \), where \( \tilde{y}_i \in V_i, i = 1, \ldots, n + 2 \), are any vectors that satisfy the assumption of Proposition 11.69. \( \forall R \in \mathcal{I}(y) \) with \( \tilde{R} \preceq R \). Let \( R := \{ (\tilde{y}_i, U_i) \mid \gamma \in \Gamma \} \).

\( \forall \gamma \in \Gamma \), by \( \tilde{R} \preceq R \), \( \exists i_\gamma \in \{1, \ldots, n + 2 \} \) such that \( U_\gamma \subseteq V_{i_\gamma} \). Define \( \Gamma := \{ \gamma \in \Gamma \mid i_\gamma = n + 2 \} \). Let \( A_{\gamma,i} := A_i \cap f_{\mathcal{B}}(U_\gamma) \cap (X \setminus E_1) \in \mathcal{B}, i = 1, \ldots, n + 1 \). Let \( \tilde{A}_{\gamma} := f_{\mathcal{B}}(U_\gamma) \cap (E_1 \setminus E) \in \mathcal{B} \).
Claim 11.86.1 \( \forall \gamma \in \bar{\Gamma}, f_{\text{inv}}(\bar{U}_{\gamma}) \subseteq E_1; \forall \gamma \in \Gamma_s := \{ \gamma \in \Gamma \mid \bar{A}_\gamma \neq \emptyset \}, \| \hat{\gamma}_* \mu(\bar{A}_\gamma) \| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma); \text{ and } \| \sum_{\gamma \in \Gamma} \hat{\gamma}_* \mu( f_{\text{inv}}(\bar{U}_{\gamma}) \cap E_1) \| < \epsilon/3. \)

Proof of claim: Fix any \( \gamma \in \bar{\Gamma}, \bar{U}_{\gamma} \subseteq V_{n+2} = \bigcup_{j=1}^{n+1} V_j \). Then, \( \mathcal{B} \ni f_{\text{inv}}(\bar{U}_{\gamma}) \subseteq E_1 \).

\( \forall \gamma \in \Gamma_s, \bar{A}_\gamma = f_{\text{inv}}(\bar{U}_{\gamma}) \cap (E_1 \setminus E) \neq \emptyset. \forall x \in \bar{A}_\gamma, \) we have \( f(x) \in \bar{U}_{\gamma}, \lim_{n \in \mathbb{N}} \phi_m(x) = f(x), \| \phi_m(x) \| \leq g_m(x), \forall m \in \mathbb{N}, \) and \( \lim_{m \in \mathbb{N}} g_m(x) = g(x). \) Then, by Propositions 4.30, 3.66, and 3.67, we have \( \| f(x) \| = \lim_{m \in \mathbb{N}} \| \phi_m(x) \| \leq \lim_{m \in \mathbb{N}} g_m(x) = g(x). \) Then, by \( R \in \mathcal{F}(Y), \| \hat{y}_* \| < \inf_{\gamma \in \bar{U}_{\gamma}} \| y \| + 1 \leq \| f(x) \| + 1 \leq g(x) + 1. \) By the arbitrariness of \( x, \) we have \( \| \hat{y}_* \| \leq \inf_{x \in \bar{A}_\gamma} g(x) + 1. \) Therefore, \( \| \hat{y}_* \mu(\bar{A}_\gamma) \| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma). \)

Now that
\[
\| \sum_{\gamma \in \Gamma} \hat{y}_* \mu( f_{\text{inv}}(\bar{U}_{\gamma}) \cap E_1) \| = \| \sum_{\gamma \in \Gamma} \hat{y}_* \mu(\bar{A}_\gamma) + \mu( f_{\text{inv}}(\bar{U}_{\gamma}) \cap E_1 \cap E) \|
\]
\[
= \| \sum_{\gamma \in \Gamma} \hat{y}_* \mu(\bar{A}_\gamma) \| = \| \sum_{\gamma \in \Gamma} \hat{y}_* \mu(\bar{A}_\gamma) \| \leq \| \sum_{\gamma \in \Gamma} \hat{y}_* \mu(\bar{A}_\gamma) \|
\]
\[
\leq \sum_{\gamma \in \Gamma} (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma) = \sum_{\gamma \in \Gamma} \mu(\bar{A}_\gamma) + \sum_{\gamma \in \Gamma} \inf_{x \in \bar{A}_\gamma} g(x) \mu(\bar{A}_\gamma)
\]

The simple function \( \psi : X \to [0, \infty) \subseteq \mathbb{R}, \) defined by \( \psi(x) = \sum_{\gamma \in \Gamma_s} \phi_{x \in \bar{A}_\gamma} \chi_{\bar{A}_\gamma}(x), \forall x \in X, \) satisfies \( 0 \leq \psi(x) \leq g(x), \forall x \in X \) and \( \psi(x) = 0, \forall x \in X \setminus E_1. \) Then, \( \sum_{\gamma \in \Gamma_s} \phi_{x \in \bar{A}_\gamma} \chi_{\bar{A}_\gamma}(x) \mu(\bar{A}_\gamma) = \int_X \psi \, d\mu = \int_{E_1} \psi \, d\mu_{E_1} < \epsilon/6, \) where the equalities follow from Proposition 11.75; and the first inequality follows from Proposition 11.83. Then, \( \| \sum_{\gamma \in \Gamma} \hat{y}_* \mu( f_{\text{inv}}(\bar{U}_{\gamma}) \cap E_1) \| < \mu(E_1) + \epsilon/6 < \delta + \epsilon/6 \leq \epsilon/3. \)

This completes the proof of the claim. \( \square \)

Claim 11.86.2 \( \| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \| < \epsilon/6. \)

Proof of claim: \( \forall \gamma \in \Gamma, \forall x \in A_i, \| \phi_n(x) \| \leq g_n(x), \forall x \in X, \) then \( \| w_i \| \leq g_n(x), \forall x \in A_i, i = 1, \ldots, \bar{n} + 1. \) Then, \( \| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \| \leq \sum_{i=1}^{\bar{n}+1} \| w_i \| \mu(A_i \cap E_1) \leq \int_{E_1} g_n \, d\mu_{E_1} < \epsilon/6, \) where the second inequality follows from Definition 11.79. \( \square \)

Claim 11.86.3 \( \forall \gamma \in \Gamma \setminus \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}, \) we have \( \| \hat{y}_* - w_i \| \mu(A_{\gamma,i}) \leq 3\epsilon \mu(A_{\gamma,i}). \)

Proof of claim: \( \forall \gamma \in \Gamma \setminus \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}, \gamma \in \{1, \ldots, \bar{n} + 1\} \) and \( \hat{U}_\gamma \subseteq V_{\bar{n}+1}. \) The result is trivial if \( A_{\gamma,i} = \emptyset. \) If \( A_{\gamma,i} \neq \emptyset, \) then \( \exists x \in A_{\gamma,i} = A_i \cap \)
Proposition 11.75. Then, \( \hat{\gamma} \in U_y \subseteq B_y \left( w_i, \hat{\epsilon} \right) \cap \hat{U}_y \subseteq B_y \left( w_i, \epsilon \right) \cap \hat{V}_y \subseteq B_y \left( w_i, \epsilon \right) \cap B_y \left( w_i, \hat{\epsilon} \right) \). Note that \( \hat{\gamma} \in \hat{U}_y \subseteq B_y \left( w_i, \epsilon \right) \). Therefore, we have \( \| \hat{\gamma} - w_i \| + \| w_i - f(x) \| + \| f(x) - w_i \| < 3\hat{\epsilon} \). Hence, the result holds.

Note that \( F_R = \sum_{\gamma \in \Gamma} \hat{\gamma} \mu(f_{\text{inv}}(\hat{U}_\gamma)) \) and \( \int_X \phi_\mu = \sum_{i=1}^{n} w_i \mu(A_i) \), by Proposition 11.75. Then,

\[
\left\| F_R - \int_X \phi_\mu \, d\mu \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{\gamma} \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{n} w_i \mu(A_i) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma} \left( \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right)
- \sum_{i=1}^{n} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma} \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{n+1} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma} \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{n+1} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\|
\]

\[
< \left\| \sum_{\gamma \in \Gamma} \hat{\gamma} \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) \right) \right\| + \frac{\epsilon}{3} + \frac{\epsilon}{6}
\]

\[
= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \Gamma'} \hat{\gamma} \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) \right) \right\|
\]

\[
= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \Gamma'} \hat{\gamma} \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) \right) \right\| \leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \Gamma'} \sum_{i=1}^{n+1} \| \hat{\gamma} - w_i \| \mu(A_{\gamma,i})
\]

\[
\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \Gamma'} \sum_{i=1}^{n+1} 3\epsilon \mu(A_{\gamma,i}) \leq \frac{\epsilon}{2} + \frac{3\mu(X)\epsilon}{6\mu(X) + 1} < \epsilon
\]
where the first inequality follows from Claims 11.86.1 and 11.86.2; and the seventh equality follows from the fact that $A_{\gamma,i} = \emptyset$, $\forall \gamma \in \Gamma$, $\forall i \in \{1, \ldots, \bar{n} + 1\}$; and the third inequality follows from Claim 11.86.3.

In summary, we have shown that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_1 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\exists R \in \mathcal{I}(Y)$ such that $\forall R \in \mathcal{I}(Y)$ with $R \leq R$, we have $\|F_R - \int_X \phi_n \, d\mu\| < \epsilon$. Then, the net $(F_R)_{R \in \mathcal{I}(Y)}$ is a Cauchy net. By Proposition 4.44, the net admits a limit $\int_X f \, d\mu \in \mathcal{Y}$. By Propositions 4.30, 3.66, and 3.67, we have $\|\int_X f \, d\mu - \int_X \phi_n \, d\mu\| = \lim_{R \in \mathcal{I}(Y)} \|F_R - \int_X \phi_n \, d\mu\| \leq \epsilon$, $\forall n \geq n_1$. Hence, we have $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu \in \mathcal{Y}$ and $f$ is integrable over $X$. This completes the proof of the lemma.

Lemma 11.87 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, $f : X \to U \subseteq \mathcal{W}$ be $\mathcal{B}$-measurable, and $g : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable. Assume that $\|f(x)\| \leq g(x)$, $\forall x \in X$, and $g$ is integrable over $X$. Then, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \to U$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$, $\|\varphi_k(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and $f$ is integrable over $X$ with $\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k \, d\mu \in \mathcal{Y}$.

Furthermore, if $\mathcal{Y}$ admits a positive cone $P$ and $U \subseteq P \subseteq \mathcal{Y}$, then $\int_X f \, d\mu \in P$.

Proof By Proposition 11.66, there exists a sequence of simple functions $(\varphi_k)_{k=1}^\infty$, $\varphi_k : X \to U$, $\forall k \in \mathbb{N}$, such that $\|\varphi_k(x)\| \leq \|f(x)\| \leq g(x)$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $\mathcal{X}$. By Lemma 11.86, $f$ is integrable over $X$ and $\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k \, d\mu \in \mathcal{Y}$.

Furthermore, if $\mathcal{Y}$ admits a positive cone $P$ and $U \subseteq P \subseteq \mathcal{Y}$, then, by Propositions 8.41 and 11.75, $\varphi_k : X \to U \subseteq P \subseteq \mathcal{Y}$, $\forall k \in \mathbb{N}$, and $\int_X \varphi_k \, d\mu \in P$. Then, by Proposition 4.13, $\int_X f \, d\mu \in P = \mathcal{P}$, since $P$ is closed. This completes the proof of the lemma.

Theorem 11.88 (Lebesgue Dominated Convergence) Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $f_n : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $g_n : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, and $g : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable. Assume that

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|f_n(x)\| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$;

(iii) $g_n$'s and $g$ are integrable over $X$ and $\int_X g \, d\mu = \lim_{n \in \mathbb{N}} \int_X g_n \, d\mu \in \mathbb{R}$.

Then, $f$ is integrable over $X$ and $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{Y}$.

Proof $\forall n \in \mathbb{N}$, by Lemma 11.87, there exists a sequence of simple functions $(\varphi_{n,k})_{k=1}^\infty$, $\varphi_{n,k} : X \to \mathcal{W}$, $\forall k \in \mathbb{N}$, such that $\lim_{k \in \mathbb{N}} \varphi_{n,k} = f_n$ a.e. in $\mathcal{X}$ and $\lim_{k \in \mathbb{N}} \varphi_{n,k} = f$ a.e. in $\mathcal{X}$.
11.5. GENERAL CONVERGENCE THEOREMS

Proposition 11.89 Let $X := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $f_i : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $g_i : X \to [0, \infty) \subset \mathbb{R}$ be integrable over $\mathcal{X}$, $i = 1, 2$. Assume that $\| f_i(x) \| \leq g_i(x)$, $\forall x \in X, i = 1, 2$. Then, the following statements hold.

(i) $f_1, f_2$, and $f_1 + f_2$ are integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu \in \mathcal{Y}$.

(ii) $\forall A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $Af_1$ is integrable over $\mathcal{X}$ and $\int_X (Af_1) d\mu = A \int_X f_1 d\mu \in \mathcal{Z}$.

(iii) $\forall c \in \mathbb{K}, cf_1$ is integrable over $\mathcal{X}$ and $\int_X (cf_1) d\mu = c \int_X f_1 d\mu \in \mathcal{Y}$.

(iv) $\forall E \in \mathcal{B}$, $f_1|_E$ is integrable over $E$ and $\int_X (f_1 \chi_{E,X}) d\mu = \int_E f_1 d\mu \in \mathcal{Y}$, where $E := (E, \mathcal{B}_E, \mu_E)$ is the finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. We will henceforth denote $\int_E f_1 d\mu$ by $\int_E f_1 d\mu$.

(v) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 d\mu = \int_X f_2 d\mu \in \mathcal{Y}$.

(vi) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, $\sum_{i=1}^\infty \int_{E_i} f_1 d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 d\mu \in \mathcal{Y}$.

(vii) $0 \leq \| \int_X f_1 d\mu \| \leq \int_X P \circ f_1 d\mu \leq \int_X g_1 d\mu < +\infty$, where $P \circ f_1 : X \to [0, \infty) \subset \mathbb{R}$ is as defined on page 384.
(viii) If $Y$ admits a positive cone $P$ and $f_1 \leq f_2$ a.e. in $\mathcal{X}$, then $\int_X f_1 \, d\mu \leq \int_X f_2 \, d\mu$.

**Proof**

By Lemma 11.87, there exist sequences of simple functions $(\phi_{i,n})_{n=1}^\infty$, $\phi_{i,n} : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$, $i = 1, 2$ such that $i = 1, 2$, $\lim_{n \to \infty} \phi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\|\phi_{i,n}(x)\| \leq \|f_i(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $f_i$ is integrable over $\mathcal{X}$, and $\int_X f_i \, d\mu = \lim_{n \to \infty} \int_X \phi_{i,n} \, d\mu \in Y$.

(i) By Propositions 7.23, 11.38, and 11.39, $f_1 + f_2$, $g_1 + g_2$, and $\phi_{i,n} + \phi_{i,2,n}$ are $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. By Proposition 11.83, $0 \leq \int_X (g_1 + g_2) \, d\mu = \int_X g_1 \, d\mu + \int_X g_2 \, d\mu < +\infty$. Note that, by Propositions 7.23, 11.52, and 11.53, $\lim_{n \to \infty} (\phi_{1,n} + \phi_{2,n}) = f_1 + f_2$ a.e. in $\mathcal{X}$ and $\|\phi_{1,n}(x) + \phi_{2,n}(x)\| \leq \|\phi_{1,n}(x)\| + \|\phi_{2,n}(x)\| \leq g_1(x) + g_2(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Clearly, $\phi_{1,n} + \phi_{2,n} : X \to \mathcal{W}$ and $f_1 + f_2 : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.88, $f_1 + f_2$ is integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{1,n} + \phi_{2,n}) \, d\mu = \lim_{n \to \infty} (\int_X \phi_{1,n} \, d\mu + \int_X \phi_{2,n} \, d\mu) = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in Y$, where the second equality follows from Proposition 11.75, and the third equality follows from Propositions 3.66, 3.67, and 7.23.

(ii) $\forall A \in \mathcal{B}(Y, Z)$, by Proposition 11.83 and the fact that $\int_X g_1 \, d\mu \in \mathbb{R}$, we have $\int_X (\|A\| g_1) \, d\mu = \|A\| \int_X g_1 \, d\mu \in \mathbb{R}$. By Propositions 7.62 and 11.38, $A_{f_1}$ and $A\phi_{1,n}$ are $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. Clearly, $A\phi_{1,n} : X \to \mathcal{W}$ and $A_{f_1} : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$, where $\mathcal{W} := \{z \in \mathbb{Z} \mid z = Av, w \in \mathbb{W}\}$ is a separable subspace of $\mathbb{Z}$. By Propositions 11.52 and 7.62, $\lim_{n \to \infty} A\phi_{1,n} = A_{f_1}$ a.e. in $\mathcal{X}$. By Proposition 11.75, we have $\int_X (A\phi_{1,n}) \, d\mu = A \int_X \phi_{1,n} \, d\mu \in \mathbb{Z}$. Note that $\|A\phi_{1,n}(x)\| \leq \|A\| \cdot \|\phi_{1,n}(x)\| \leq A\|g_1(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.88, $A_{f_1}$ is integrable over $\mathcal{X}$ and $\int_X (A_{f_1}) \, d\mu = \lim_{n \to \infty} \int_X (A\phi_{1,n}) \, d\mu = \lim_{n \to \infty} A \int_X \phi_{1,n} \, d\mu = A \int_X f_1 \, d\mu \in \mathbb{Z}$, where the last equality follows from Propositions 3.66 and 7.62.

(iii) This follows immediately from (ii).

(iv) $(\phi_{i,n} | E)_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1 | E$ a.e. in $E$. Note that $\|\phi_{i,n} | E (x)\| \leq g_1 | E (x)$, $\forall x \in E$, $\forall n \in \mathbb{N}$. By Proposition 11.83, $0 \leq \int_E g_1 | E \, d\mu = \int_X g_1 \, d\mu < +\infty$. By Lebesgue Dominated Convergence Theorem 11.88, we have $\int_E f_1 | E \, d\mu = \lim_{n \to \infty} \int_E \phi_{i,n} | E \, d\mu \in Y$. Note that $(\phi_{i,n} \chi_{E,X})_{n=1}^\infty$ is a sequence of simple functions that converges to $f_1 \chi_{E,X}$ a.e. in $\mathcal{X}$, and $\|\phi_{i,n}(x) \chi_{E,X}(x)\| \leq g_1(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.38, and 11.39, $\phi_{i,n} \chi_{E,X}$ and $f_1 \chi_{E,X}$ are $\mathcal{B}$-measurable. Again by Lebesgue Dominated Convergence Theorem 11.88, we have $\int_X (f_1 \chi_{E,X}) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{i,n} \chi_{E,X}) \, d\mu = \lim_{n \to \infty} \int_X \phi_{i,n} | E \, d\mu = \int_E f_1 | E \, d\mu \in \mathbb{Z}$, where the second equality follows from Proposition 11.75.

(v) If $f_1 = f_2$ a.e. in $\mathcal{X}$, then $\lim_{n \to \infty} \phi_{1,n} = f_2$ a.e. in $\mathcal{X}$ by Proposition 11.54. By Lebesgue Dominated Convergence Theorem 11.88, we have $\int_X f_2 \, d\mu = \lim_{n \to \infty} \int_X \phi_{1,n} \, d\mu = \int_X f_1 \, d\mu$. 

408 CHAPTER 11. GENERAL MEASURE AND INTEGRATION
Proposition 11.41, \( E_i \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \) and \( \tilde{E}_n := \bigcup_{i=1}^{n} E_i \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \lim_{n \to \infty} f_1(x) \chi_{\tilde{E}_n} = f_1(x) \chi_{E_n} \), \( \forall x \in X \). Then, we have \( \sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \int_{X} (f_1 \chi_{E_i} \chi_{X}) \, d\mu = \lim_{n \to \infty} \int_{X} (f_1 \chi_{E_n} \chi_{X}) \, d\mu = \lim_{n \to \infty} \int_{X} f_1 \, d\mu = \int_{X} f_1 \, d\mu \), where the first inequality follows from (iv); the second equality follows from \( \text{(i)} \); the third inequality follows from Proposition 11.83.

Theorem 11.91 (Lebesgue Dominated Convergence) Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a normed linear space, and \( f : X \to Y \) be \( \mathcal{B} \)-measurable. \( f \) is said to be absolutely integrable if \( \mathcal{P} \circ f \) is integrable over \( X \), that is, \( 0 \leq \int_{X} \mathcal{P} \circ f \, d\mu < +\infty \).

Definition 11.90 Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a Banach space, \( W \) be a separable subspace of \( Y \), \( f_n : X \to W \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), \( f : X \to W \) be \( \mathcal{B} \)-measurable, \( g_n : X \to [0, \infty) \subseteq \mathbb{R} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( g : X \to [0, \infty) \subseteq \mathbb{R} \) be \( \mathcal{B} \)-measurable. Assume that

\[
\begin{align*}
(i) \quad & \lim_{n \to \infty} f_n = f \text{ a.e. in } X, \lim_{n \to \infty} g_n = g \text{ a.e. in } X; \\
(ii) \quad & \| f_n(x) \| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}; \\
(iii) \quad & g_n \text{'s and } g \text{ are integrable over } X \text{ and } \int_{X} g \, d\mu = \lim_{n \to \infty} \int_{X} g_n \, d\mu \in \mathbb{R}.
\end{align*}
\]
Then, \( f \) is integrable over \( \mathcal{X} \) and \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathbb{Y} \). Furthermore, \( f \) is absolutely integrable and \( 0 \leq \int_X (P \circ f) \, d\mu = \lim_{n \in \mathbb{N}} \int_X (P \circ f_n) \, d\mu \leq \int_X g \, d\mu < \infty \).

**Proof** We will show that \( f \) is integrable over \( \mathcal{X} \) and \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathbb{Y} \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \).

Case 1: \( \mu(X) < +\infty \). The result follows immediately from Lebesgue Dominated Convergence Theorem 11.88.

Case 2: \( \mu(X) = +\infty \). Let \((F_A)_{A \in \mathcal{M}(\mathcal{X})}\) and \((F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}\) be the nets for the functions \( f \) and \( f_n \) as defined in Definition 11.71, respectively, \( \forall n \in \mathbb{N} \). \( \forall A \in \mathcal{M}(\mathcal{X}) \), we have \( A \in \mathcal{B} \) and \( \mu(A) < \infty \). Let \( A := (A, B_A, \mu_A) \) be the finite measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13. By Proposition 11.83, \( g|_A \) and \( g_n|_A \)'s are integrable over \( A \), \( \lim_{n \in \mathbb{N}} g_n|_A = g|_A \) a.e. in \( A \), and \( \int_A g|_A \, d\mu_A = \lim_{n \in \mathbb{N}} \int_A g_n|_A \, d\mu_A \). By Lebesgue Dominated Convergence Theorem 11.88, we have \( f|_A \) and \( f_n|_A \)'s are integrable over \( A \) and \( \int_A f|_A \, d\mu_A = \lim_{n \in \mathbb{N}} \int_A f_n|_A \, d\mu_A \in \mathbb{Y} \). Note that \( F_A = \int_A f|_A \, d\mu_A \in \mathbb{Y} \) and \( F_{n,A} = \int_A f_n|_A \, d\mu_A \in \mathbb{Y} \), \( \forall n \in \mathbb{N} \). Then, the nets \((F_A)_{A \in \mathcal{M}(\mathcal{X})}\) and \((F_{n,A})_{A \in \mathcal{M}(\mathcal{X})}\)'s are well defined.

\( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), by Definition 11.79, \( \exists \) a simple function \( \phi : X \to [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi(x) \leq g(x) \), \( \forall x \in X \), and \( \int_X g \, d\mu - \varepsilon/5 < \int_X \phi \, d\mu \leq \int_X g \, d\mu < +\infty \). Let \( \mathcal{A}_0 := \{ x \in X \mid \phi(x) > 0 \} \). Then, \( \mathcal{A}_0 \in \mathcal{B} \), \( \mu(\mathcal{A}_0) = +\infty \), and \( \mathcal{A}_0 \in \mathcal{M}(\mathcal{X}) \). \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A \subseteq \mathcal{A}_0 \), we have \( 0 \leq \int_{\mathcal{A}_0 \setminus A} g|_A \, d\mu_A = \int_A g|_A \, d\mu_A - \int_{\mathcal{A}_0 \setminus A} g|_A \, d\mu_A \leq \int_X g \, d\mu - \int_{\mathcal{A}_0 \setminus A} \phi|_A \, d\mu_A = \int_X g \, d\mu - \int_X \phi \, d\mu < \varepsilon/5 \), where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that \( \| F_A - F_{A_0} \| = \left\| \int_A f|_A \, d\mu_A - \int_{A_0} f|_{A_0} \, d\mu_{A_0} \right\| = \left\| \int_{A_0 \setminus A} f|_A \, d\mu_A \right\| = \lim_{n \in \mathbb{N}} \left\| \int_{A_0 \setminus A} f_n|_{A_0} \, d\mu_{A_0} \right\| \leq \lim_{n \in \mathbb{N}} \int_{A_0 \setminus A} g_n|_{A_0} \, d\mu_{A_0} = \int_X g \, d\mu - \int_X \phi \, d\mu < \varepsilon/5 \), where the second equality follows form Proposition 11.89; the third equality follows from the preceding paragraph; the first inequality follows from Proposition 11.89; and the fourth equality follows from the preceding paragraph. Then, the net \((F_A)_{A \in \mathcal{M}(\mathcal{X})}\) is a Cauchy net, which admits a limit, by Proposition 4.44, \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathbb{Y} \). Hence, \( f \) is integrable over \( \mathcal{X} \).

\( \forall n \in \mathbb{N} \), by Definition 11.79, \( \exists \) a simple function \( \phi_n : X \to [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi_n(x) \leq g_n(x) \), \( \forall x \in X \), and \( \int_X g_n \, d\mu - \varepsilon/5 < \int_X \phi_n \, d\mu \leq \int_X g_n \, d\mu < +\infty \). Let \( \mathcal{A}_n := \{ x \in X \mid \phi_n(x) > 0 \} \). Then, \( \mathcal{A}_n \in \mathcal{B} \), \( \mu(\mathcal{A}_n) < +\infty \), and \( \mathcal{A}_n \in \mathcal{M}(\mathcal{X}) \). \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A \subseteq \mathcal{A}_n \), we have \( 0 \leq \int_{\mathcal{A}_n \setminus A} g_n|_A \, d\mu_A = \int_A g_n|_A \, d\mu_A - \int_{\mathcal{A}_n \setminus A} g_n|_A \, d\mu_A \leq \int_X g_n \, d\mu - \int_{\mathcal{A}_n \setminus A} \phi_n|_A \, d\mu_A = \int_X g_n \, d\mu - \int_X \phi_n \, d\mu < \varepsilon/5 \), where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that \( \| F_{n,A} - F_{n,A_0} \| = \lim_{n \in \mathbb{N}} \left\| \int_{A_0 \setminus A} f_n|_{A_0} \, d\mu_{A_0} \right\| \leq \lim_{n \in \mathbb{N}} \int_{A_0 \setminus A} g_n|_{A_0} \, d\mu_{A_0} = \int_X g_n \, d\mu - \int_X \phi_n \, d\mu < \varepsilon/5 \), where the second equality follows form Proposition 11.89; the third equality follows from the preceding paragraph; the first inequality follows from Proposition 11.89; and the fourth equality follows from the preceding paragraph. Then, the net \((F_A)_{A \in \mathcal{M}(\mathcal{X})}\) is a Cauchy net, which admits a limit, by Proposition 4.44, \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A \in \mathbb{Y} \). Hence, \( f \) is integrable over \( \mathcal{X} \).
11.5. GENERAL CONVERGENCE THEOREMS

Let $\mathcal{M}(\mathcal{X})$ be a Banach space over $\mathbb{K}$, $\mathbb{W}$ be a separable subspace of $\mathbb{Y}$, and $\mathbb{Z}$ be a Banach space over $\mathbb{K}$, $f_i : X \to \mathbb{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.

(i) $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in \mathbb{Y}$, $i = 1, 2$.

(ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathbb{Y}$.

Proposition 11.92 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $f_i : X \to \mathbb{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.

(i) $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in \mathcal{Y}$, $i = 1, 2$.

(ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathcal{Y}$.
(iii) \( \forall A \in \mathcal{B}(\mathbb{Y}, \mathbb{Z}) \), \( Af_1 \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (Af_1) \, d\mu = A \int_{\mathcal{X}} f_1 \, d\mu \in \mathbb{Z} \).

(iv) \( \forall c \in \mathbb{K}, cf_1 \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (cf_1) \, d\mu = c \int_{\mathcal{X}} f_1 \, d\mu \in \mathbb{Y} \).

(v) \( \forall H \in \mathcal{B}, f_1|_H \) is absolutely integrable over \( \mathcal{H} \) and \( \int_{\mathcal{X}} (f_1|_H) \, d\mu = \int_H f_1|_H \, d\mu_H \in \mathbb{Y} \), where \( \mathcal{H} := (H, \mathcal{B}_H, \mu_H) \) is the measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13. We will henceforth denote \( \int_H f_1|_H \, d\mu_H \) by \( \int_H f_1 \, d\mu \).

(vi) If \( f_1 = f_2 \) a.e. in \( \mathcal{X} \) then \( \int_{\mathcal{X}} f_1 \, d\mu = \int_{\mathcal{X}} f_2 \, d\mu \in \mathbb{Y} \).

(vii) \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}, \sum_{i=1}^{\infty} f_1 \, d\mu = \int_{\bigcup_{i=1}^{\infty} E_i} f_1 \, d\mu \in \mathbb{Y} \).

(viii) \( 0 \leq \left\| \int_{\mathcal{X}} f_1 \, d\mu \right\| \leq \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mu < +\infty \).

(ix) If \( \mathbb{Y} \) admits a positive cone \( \mathcal{P} \) and \( f_1 \leq f_2 \) a.e. in \( \mathcal{X} \), then \( \int_{\mathcal{X}} f_1 \, d\mu \leq \int_{\mathcal{X}} f_2 \, d\mu \).

**Proof**

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mu(\mathcal{X}) < +\infty \); Case 2: \( \mu(\mathcal{X}) = +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their proof.

Case 2: \( \mu(\mathcal{X}) = +\infty \). \( f_1, f_2, \mathcal{P} \circ f_1, \) and \( \mathcal{P} \circ f_2 \) are \( \mathcal{B} \)-measurable by Propositions 7.21 and 11.38.

(i) By Lebesgue Dominated Convergence Theorem 11.91, we have that \( f_1 \) is integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} f_1 \, d\mu \in \mathbb{Y}, i = 1, 2 \). Let \( (F_{i,E})_{E \in \mathcal{M}(\mathcal{X})} \) be the net for \( \int_{\mathcal{X}} f_1 \, d\mu \) as defined in Definition 11.71, \( i = 1, 2 \). \( \forall E \in \mathcal{M}(\mathcal{X}), F_{i,E} = \int_E f_1 \, d\mu_E \), \( i = 1, 2 \). Then, by Definition 11.71, we have the net \( (F_{i,E})_{E \in \mathcal{M}(\mathcal{X})} \) is well-defined, \( i = 1, 2 \), and \( \lim_{E \in \mathcal{M}(\mathcal{X})} F_{i,E} = \int_{\mathcal{X}} f_1 \, d\mu \in \mathbb{Y}, i = 1, 2 \).

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, \( f_1 + f_2 \) and \( \mathcal{P} \circ (f_1 + f_2) \) are \( \mathcal{B} \)-measurable. Note that \( \mathcal{P} \circ (f_1 + f_2)(x) = \left\| f_1(x) + f_2(x) \right\| \leq \left\| f_1(x) \right\| + \left\| f_2(x) \right\| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x), \forall x \in \mathcal{X} \). By Proposition 11.83, \( 0 \leq \int_{\mathcal{X}} \mathcal{P} \circ (f_1 + f_2) \, d\mu \leq \int_{\mathcal{X}} \mathcal{P} \circ (f_1 + f_2) \, d\mu = \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mu + \int_{\mathcal{X}} \mathcal{P} \circ f_2 \, d\mu < +\infty \).

Hence, \( f_1 + f_2 \) is absolutely integrable over \( \mathcal{X} \). Let \( (F_{E})_{E \in \mathcal{M}(\mathcal{X})} \) be the net for \( \int_{\mathcal{X}} (f_1 + f_2) \, d\mu \) as defined in Definition 11.71. \( \forall E \in \mathcal{M}(\mathcal{X}), F_{E} = \int_{E} (f_1 + f_2) \, d\mu_E = \int_{E} f_1 \, d\mu_E + \int_{E} f_2 \, d\mu_E = F_{1,E} + F_{2,E}, \) where the second equality follows from Proposition 11.89. Then, \( \int_{\mathcal{X}} (f_1 + f_2) \, d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} F_{E} = \lim_{E \in \mathcal{M}(\mathcal{X})}(F_{1,E} + F_{2,E}) = \int_{\mathcal{X}} f_1 \, d\mu + \int_{\mathcal{X}} f_2 \, d\mu \in \mathbb{Y}, \) where the third equality follows from Propositions 3.66, 3.67, and 7.23. Hence, \( f_1 + f_2 \) is integrable over \( \mathcal{X} \).

(iii) By Propositions 11.38, 7.21, and 7.62, \( Af_1 \) and \( \mathcal{P} \circ (Af_1) \) are \( \mathcal{B} \)-measurable. Note that \( \mathcal{P} \circ (Af_1)(x) = \left\| Af_1(x) \right\| \leq \left\| A \right\| \left\| f_1(x) \right\| = \left\| A \right\| \mathcal{P} \circ f_1(x), \forall x \in \mathcal{X} \). By Proposition 11.83 and the fact that \( \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mu \in \mathbb{R} \),
we have $0 \leq \int_X |P \circ (Af_i)| \, d\mu \leq \int_X (\|P \circ f_i\| \, d\mu = \|P \circ f_i\| \, d\mu < +\infty$. Hence, $Af_i$ is absolutely integrable over $\mathcal{X}$. Let $(F_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (Af_i) \, d\mu$ as defined in Definition 11.71. $\forall E \in \mathcal{M}(\mathcal{X})$, $F_E = \int_E (Af_i) \, d\mu = A \int_E f_i \, d\mu = AF_{1,E}$, where the second equality follows from Proposition 11.89. Then, by Propositions 3.66 and 7.62, we have $\int_X (Af_i) \, d\mu = \lim_{E \in \mathcal{X}(\mathcal{X})} F_E = \lim_{E \in \mathcal{M}(\mathcal{X})} AF_{1,E} = A \int_X f_i \, d\mu \in \mathbb{Z}$.

Hence, $Af_i$ is integrable over $\mathcal{X}$.

(iv) This follows immediately from (iii).

(v) Fix any $H \in \mathcal{B}$. By Proposition 11.41, $f_i|_H$ and $P \circ (f_i|_H) = (P \circ f_i)|_H$ are $\mathcal{B}$-measurable. By Proposition 11.83, $f_i|_H$ is absolutely integrable over $\mathcal{H}$. Then, by (i), $f_i|_H$ is integrable over $\mathcal{H}$. Again, by (i), $f_1|_H \chi_{\mathcal{X}}$ is integrable over $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $\|\int_A (f_1 \chi_{\mathcal{X}}) \|_A \, d\mu_A = \int_X (f_1 \chi_{\mathcal{X}}) \, d\mu < \epsilon$. We will distinguish two exhaustive and mutually exclusive subcases: Case 2a: $\mu(H) = +\infty$; Case 2b: $\mu(H) < +\infty$.

Case 2a: $\mu(H) = +\infty$. $\exists E_0 \in \mathcal{M}(\mathcal{H})$ such that $\forall E \in \mathcal{M}(\mathcal{H})$ with $E_0 \subseteq E$, we have $\|\int_E f_i \, d\mu_E - \int_H f_i \, d\mu_E\| < \epsilon/2$. Let $A := A_0 \cup E_0 \in \mathcal{M}(\mathcal{X})$ and $E_1 := A \cap H \in \mathcal{M}(\mathcal{H})$. By Proposition 11.89, $\int_{A_1} (f_1 \chi_{\mathcal{X}}) \, d\mu_{A_1} = \int_H \int_{E_1} f_i \, d\mu_{E_1}$. Then, we have $\|\int_X (f_1 \chi_{\mathcal{X}}) \, d\mu - \int_H f_i \, d\mu\| < \epsilon/2 + \epsilon/2 = \epsilon$. By the arbitrariness of $\epsilon$, we have $\int_X (f_1 \chi_{\mathcal{X}}) \, d\mu = \int_H f_i \, d\mu \in \mathbb{Y}$.

Case 2b: $\mu(H) < +\infty$. Let $A := A_0 \cup H \in \mathcal{M}(\mathcal{X})$. By Proposition 11.89, $\int_{A_1} (f_1 \chi_{\mathcal{X}}) \, d\mu_{A_1} = \int_H f_i \, d\mu_H$. Then, we have $\|\int_X (f_1 \chi_{\mathcal{X}}) \, d\mu - \int_H f_i \, d\mu_H\| = \int_X (f_1 \chi_{\mathcal{X}}) \, d\mu - \int_{A_1} (f_1 \chi_{\mathcal{X}}) \, d\mu_{A_1} < \epsilon/2$. By the arbitrariness of $\epsilon$, we have $\int_X (f_1 \chi_{\mathcal{X}}) \, d\mu = \int_H f_i \, d\mu_H \in \mathbb{Y}$.

Hence, in both subcases, $\int_X (f_1 \chi_{\mathcal{X}}) \, d\mu = \int_H f_i \, d\mu_H \in \mathbb{Y}$.

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.91.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}$ and $E_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \to \infty} \int_{E_n} f_i \, d\mu - \int_X f_i \, d\mu \leq +\infty$, where the second inequality follows from Proposition 11.83. Then, $0 \leq \|\int_X f_i \, d\mu\| = \lim_{E \in \mathcal{M}(\mathcal{X})} \|F_{1,E}\| = \|\int_X f_i \, d\mu\| < +\infty$, where the equality follows from Propositions 3.66 and 7.21.
Proposition 11.93 Let $\mathcal{X} = (X, \mathcal{B}, \mu)$ be a measure space, $Y_i$ be a Banach space over $\mathbb{K}$, $W_i \subseteq Y_i$ be a separable subspace, $f_i : X \to W_i$, $i = 1, 2$, and $h : X \to W_1 \times W_2$ be defined by $h(x) = (f_1(x), f_2(x))$, $\forall x \in X$. Then, $h$ is absolutely integrable over $\mathcal{X}$ if, and only if, $f_1$ and $f_2$ are absolutely integrable over $\mathcal{X}$. In this case, we have $\int_X h \, d\mu = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in Y_1 \times Y_2$.

Proof Let $f_1$ and $f_2$ be absolutely integrable over $\mathcal{X}$. By Proposition 11.39, $h$ is $\mathcal{B}$-measurable. By Propositions 7.21 and 11.38, $P \circ h$ is $\mathcal{B}$-measurable. $\forall x \in X$, $P \circ h(x) = \|(f_1(x), f_2(x))\| = (\|f_1(x)\|^2 + \|f_2(x)\|^2)^{1/2} \leq P \circ f_1(x) + P \circ f_2(x)$. By Proposition 11.83, $0 \leq \int_X (P \circ h) \, d\mu \leq \int_X (P \circ f_1 + P \circ f_2) \, d\mu = \int_X (P \circ f_1) \, d\mu + \int_X (P \circ f_2) \, d\mu < \infty$. Hence, $h$ is absolutely integrable over $\mathcal{X}$. By Propositions 4.31 and 7.22, $Y_1 \times Y_2$ is a Banach space over $\mathbb{K}$. By Propositions 7.22, 3.28, and 4.4, $W_1 \times W_2 \subseteq Y_1 \times Y_2$ is a separable subspace. By Proposition 11.92, $h$, $f_1$, and $f_2$ are integrable over $\mathcal{X}$ and $\int_X h \, d\mu \in Y_1 \times Y_2$.

On the other hand, let $h$ be absolutely integrable over $\mathcal{X}$. Then, by Proposition 11.39, $f_1$ and $f_2$ are $\mathcal{B}$-measurable. Fix any $i \in \{1, 2\}$. By Propositions 7.21 and 11.38, $P \circ f_i$ is $\mathcal{B}$-measurable. $\forall x \in X$, $P \circ f_i(x) \leq P \circ h(x)$. Then, by Proposition 11.83, $0 \leq \int_X P \circ f_i \, d\mu \leq \int_X P \circ h \, d\mu < \infty$. Hence, $f_i$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in Y_i$.

Hence, $h$ is absolutely integrable over $\mathcal{X}$ if, and only if, $f_1$ and $f_2$ are absolutely integrable over $\mathcal{X}$.

Let $h$, $f_1$, and $f_2$ be absolutely integrable over $\mathcal{X}$. We will show that $\int_X h \, d\mu = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in Y_1 \times Y_2$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < \infty$; Case 2: $\mu(X) = \infty$.

Case 1: $\mu(X) < \infty$. By Lemma 11.87, $\forall i \in \{1, 2\}$, $\exists$ sequence of simple functions $(\varphi_{i,n})_{n=1}^\infty$ sequence of simple functions $(\varphi_{i,n})_{n=1}^\infty$ for $i \in \{1, 2\}$, $\exists$ sequence of simple functions $(\varphi_{i,n})_{n=1}^\infty$, $\forall n \in \mathbb{N}$, such that $\lim_{n \to \infty} \varphi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\|\varphi_{i,n}(x)\| \leq \|f_i(x)\|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\int_X f_i \, d\mu = \lim_{n \to \infty} \int_X \varphi_{i,n} \, d\mu \in Y_i$. By Proposition 11.53, we have $\lim_{n \to \infty} \varphi_{i,n} = h$ a.e. in $\mathcal{X}$, where $\psi_n : X \to W_1 \times W_2$ is defined by $\psi_n(x) = (\varphi_{1,n}(x), \varphi_{2,n}(x))$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Clearly, $\psi_n$ is a simple function, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, $\int_X \psi_n \, d\mu = \lim_{n \to \infty} \int_X \varphi_{1,n} \, d\mu + \int_X \varphi_{2,n} \, d\mu = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in Y_1 \times Y_2$, where the second equality follows from Proposition 11.75; and the third equality follows from Proposition 3.67.
Case 2: $\mu(X) = \infty$. Let $(H_E)_{E \in M(X)}$ and $(F_i,E)_{E \in M(X)}$ be the nets for $\int_{X} h \, d\mu$ and $\int_{X} f \, d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2$. \\
\forall E \in M(X)$, by Case 1,
\[ H_E = \int_{E} h \, d\mu_{E} = (\int_{E} f_1 \, d\mu_{E}, \int_{E} f_2 \, d\mu_{E}) = (F_1,E, F_2,E) \]
where $\mathcal{E} := (E, \mathcal{B}_{E}, \mu_{E})$ is the finite measure subspace of $\mathcal{X}$. Since $h$, $f_1$, and $f_2$ are integrable over $\mathcal{X}$, we have $\int_{X} h \, d\mu = \lim_{E \in M(X)} H_E = \lim_{E \in M(X)} (F_1,E, F_2,E) = (\int_{X} f_1 \, d\mu, \int_{X} f_2 \, d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$, where the last equality follows from Proposition 3.67.
Hence, in both cases, we have $\int_{X} h \, d\mu = (\int_{X} f_1 \, d\mu, \int_{X} f_2 \, d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$.
This completes the proof of the proposition.

**Proposition 11.94** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, and $f : X \to U$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}$, $\varphi_n : X \to U$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$, $\lim_{n \in \mathbb{N}} \int_{X} \varphi_n \, d\mu = \int_{X} f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_{X} \mathcal{P} \circ \varphi_n \, d\mu = \int_{X} \mathcal{P} \circ f \, d\mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_{X} \mathcal{P} \circ (\varphi_n - f) \, d\mu = 0$.

**Proof**  By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}$, $\varphi_n : X \to U$, $\forall n \in \mathbb{N}$, such that $\|\varphi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.91, $\lim_{n \in \mathbb{N}} \int_{X} \varphi_n \, d\mu = \int_{X} f \, d\mu \in \mathcal{Y}$. Note that $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \varphi_n = \mathcal{P} \circ f$ a.e. in $\mathcal{X}$, by Propositions 7.21 and 11.52. Again by Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{n \in \mathbb{N}} \int_{X} (\mathcal{P} \circ \varphi_n) \, d\mu = \int_{X} (\mathcal{P} \circ f) \, d\mu \in [0, \infty) \subset \mathbb{R}$. By Lemma 11.43, $\mathcal{P} \circ (\varphi_n - f)$ is $\mathcal{B}$-measurable. By Propositions 7.21, 7.23, 11.53, and 11.52, we have $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\varphi_n - f) = 0$ a.e. in $\mathcal{X}$. Note that $0 \leq \mathcal{P} \circ (\varphi_n - f)(x) = \|\varphi_n(x) - f(x)\| \leq \|\varphi_n(x)\| + \|f(x)\| \leq 2\mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have $\lim_{n \in \mathbb{N}} \int_{X} \mathcal{P} \circ (\varphi_n - f) \, d\mu = 0$. This completes the proof of the proposition.

**Proposition 11.95** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $U \subseteq \mathcal{W}$ be a conic segment, and $f : X \to U$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\phi_n)_{n=1}^{\infty}$, $\phi_n : X \to U$, $\forall n \in \mathbb{N}$, such that $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \int_{X} \phi_n \, d\mu = \int_{X} f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_{X} \mathcal{P} \circ \phi_n \, d\mu = \int_{X} \mathcal{P} \circ f \, d\mu \in [0, \infty) \subset \mathbb{R}$, and $\lim_{n \in \mathbb{N}} \int_{X} \mathcal{P} \circ (\phi_n - f) \, d\mu = 0$.

**Proof**  We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) < +\infty$; Case 2: $\mu(X) = +\infty$.

Case 1: $\mu(X) < +\infty$. Then, $\mathcal{X}$ is a finite measure space. The result follows immediately from Proposition 11.94.
Case 2: \( \mu(X) = +\infty \). Since \( \int_X \mathcal{P} \circ f \, d\mu < +\infty \), then, by Definition 11.71, \( \forall n \in \mathbb{N} \), \( \exists E_n \in \mathcal{B} \) with \( \mu(E_n) < +\infty \) such that \( \int_X \mathcal{P} \circ f \, d\mu < 2^{-n-1} \). By Case 1, \( \exists \) a simple function \( \bar{\phi}_n : E_n \rightarrow U \) such that \( \| \bar{\phi}_n(x) \| \leq \mathcal{P} \circ f(x) \), \( \forall x \in E_n \), \( \left\| \int_{E_n} \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu \right\| < 2^{-n-1} \). Let \( \mathcal{P} \circ \bar{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu < 2^{-n-1} \), and \( 0 \leq \int_{E_n} \mathcal{P} \circ \bar{\phi}_n \, d\mu_{E_n} < 2^{-n-1} \), where \( (E_n, \mathcal{B}_{E_n}, \mu_{E_n}) \) is the finite measure subspace of \( X \). This simple function \( \bar{\phi}_n \) may be extended to a simple function \( \phi_n : X \rightarrow U \) such that \( \phi_n|_{E_n} = \bar{\phi}_n \) and \( \phi_n(x) = \bar{\phi}_y \), \( \forall x \in E \), and \( \forall x \in X \setminus E_n \), and \( \| \phi(x) \| \leq \mathcal{P} \circ f(x) \), \( \forall x \in X \), and \( \left\| \int_X \phi_n \, d\mu - \int_X f \, d\mu \right\| = \left\| \int_X (\phi_n \chi_{E_n} - f) \, d\mu + \int_X f \, d\mu - \int_X \mathcal{P} \circ f \, d\mu \right\| < 2^{-n-1} \). Let \( \mathcal{P} \circ f \, d\mu = \int_X f \, d\mu + \mu\left( \bigcup_{n=1}^{\infty} E_n \right) \). Finally, we have \( 0 \leq \int_X \mathcal{P} \circ f \, d\mu < 2^{-n-1} < 2^{-n} \), where the second inequality follows from Proposition 11.92. Hence, \( (\phi_n)_{n=1}^{\infty} \) is the sequence we seek.

This completes the proof of the proposition.

\( \square \)

**Proposition 11.96** Let \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, and \( f_1, f_2 : X \rightarrow \mathbb{R} \) and \( f_2, f_2 : X \rightarrow \mathbb{R} \) be \( \mathcal{B} \)-measurable. Assume that \( \int_E f_1 \, d\mu \leq \int_X \mathcal{P} \circ f_1 \, d\mu \) for all \( \mathcal{P} \in \mathcal{P} \) with \( \mathcal{P} \circ f_1 \, d\mu < \infty \), \( i = 1, 2 \). Then, \( f_1 \leq f_2 \) a.e. in \( X \).

**Proof** By Propositions 7.21, 7.23, 11.38, and 11.39, \( f_1, f_2, \mathcal{P} \circ f_1, \) and \( \mathcal{P} \circ f_2 \) are \( \mathcal{B} \)-measurable. Let \( E := \{ x \in X \mid f_2(x) - f_1(x) < 0 \} \). Suppose \( \mu(E) > 0 \). \( E = \bigcup_{n=1}^{\infty} E_n := \bigcup_{n=1}^{\infty} \{ x \in X \mid f_2(x) - f_1(x) \leq -1/n, \| f_1(x) \| \leq n, \| f_2(x) \| \leq n \} \). By Proposition 11.7, \( \mu(E_n) = \lim_{n \rightarrow \infty} \mu(E_n) \) and \( \exists n \in \mathbb{N} \) such that \( \mu(E_n) > 0 \). Since \( X \) is \( \sigma \)-finite, then \( \exists (X_m)_{m=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{m=1}^{\infty} X_m \) and \( \mu(X_m) < +\infty \), \( \forall m \in \mathbb{N} \). Without loss of generality, we may assume that \( X_m \subseteq X_{m+1} \), \( \forall m \in \mathbb{N} \). Then, by Proposition 11.7, \( \mu(E_n) = \lim_{m \rightarrow \infty} \mu(E_n \cap X_m) \). Then, \( \exists m \in \mathbb{N} \) such that \( \mu(E_m) = \mu(E_m \cap X_m) < +\infty \), \( \forall m \in \mathbb{N} \). By Definition 11.79, \( \int_{E} \mathcal{P} \circ f \, d\mu \leq n\mu(E) \leq \infty \), \( i = 1, 2 \). Then, by Propositions 11.75 and 11.92, \( \mathcal{P} \) \( \exists \int_{E} f_1 \, d\mu - \int_X f_2 \, d\mu - \mathcal{P}(E)/n = \int_{E} (f_1 - f_2 - 1/n) \, d\mu > 0 \). This contradicts the assumption. Hence, \( \mu(E) = 0 \) and \( f_1 \leq f_2 \) a.e. in \( X \).

\( \square \)
Proposition 11.97 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( f_i : X \to [0, \infty) \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable, \( i = 1, 2 \). Assume that \( f_1 \leq f_2 \) a.e. in \( \mathcal{X} \), \( \mu(\{ x \in X \mid f_1(x) < f_2(x) \}) > 0 \), and \( 0 \leq \int_X f_1 \, d\mu < +\infty \). Then, \( \int_X f_1 \, d\mu \leq \int_X f_2 \, d\mu \).

**Proof** By Proposition 11.83, \( \int_X f_1 \, d\mu \leq \int_X f_2 \, d\mu \). Suppose \( \int_X f_2 \, d\mu \leq \int_X f_1 \, d\mu \). Then, \( f_2 \) and \( f_1 \) are absolutely integrable over \( \mathcal{X} \). Let \( g := f_2 - f_1 \). Then, by Propositions 7.23, 11.38, and 11.39, we have \( g : X \to \mathbb{R} \) is \( \mathcal{B} \)-measurable and \( g \geq 0 \) a.e. in \( \mathcal{X} \). By Proposition 11.92, \( g \) is absolutely integrable over \( \mathcal{X} \) and \( \int_X g \, d\mu = 0 \). For every \( E \in \mathcal{B} \), by Proposition 7.16, \( \bar{\Omega} \) is a closed convex set. Let \( \Omega \) be a convex functional. Assume that \( f(X) \) is absolutely integrable over \( \mathcal{X} \) and the epigraph \( [G, \Omega] \) is closed. Then, \( \int_X g \, d\mu = 0 \). By Proposition 11.96, we have \( g = 0 \) a.e. in \( \mathcal{X} \). This contradicts the fact that \( g = 0 \) a.e. in \( \mathcal{X} \). Therefore, we must have \( \int_X f_1 \, d\mu < \int_X f_2 \, d\mu \). This completes the proof of the proposition. \( \square \)

Theorem 11.98 (Jensen’s Inequality) Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space with \( \mu(X) = 1 \), \( \mathcal{Y} \) be a real Banach space, \( \mathcal{W} \) be a separable subspace of \( \mathcal{Y} \), \( \Omega \subseteq \mathcal{Y} \) be a nonempty closed convex set, \( f : X \to \Omega \cap \mathcal{W} \) be absolutely integrable over \( \mathcal{X} \), and \( G : \Omega \to \mathbb{R} \) be a convex functional. Assume that \( G \circ f \) is absolutely integrable over \( \mathcal{X} \) and the epigraph \( [G, \Omega] \) is closed. Then, \( \int_X f \, d\mu \in \Omega \) and \( G(\int_X f \, d\mu) \leq \int_X (G \circ f) \, d\mu \in \mathbb{R} \).

**Proof** By Proposition 11.92, \( y_0 := \int_X f \, d\mu \in \mathcal{Y} \). We will first show that \( y_0 \) is a convex functional. Note that \( \mu(X) = 1 \) implies that \( \mathcal{X} \neq \emptyset \). Then, \( \Omega \cap \mathcal{W} \neq \emptyset \). By Lemma 11.87, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^\infty \), \( \varphi_n : \mathcal{X} \to \Omega \cap \mathcal{W} \), \( \forall n \in \mathbb{N} \), such that \( \lim_{n \to \infty} \varphi_n = f \) a.e. in \( \mathcal{X} \), \( \| \varphi_n(x) \| \leq \mathcal{P} \circ f(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), and \( y_0 = \int_X f \, d\mu = \lim_{n \to \infty} \int_X \varphi_n \, d\mu \in \mathcal{Y} \). Fix any \( n \in \mathbb{N} \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^n y_i \chi_{A_i} \). Let \( y_{n+1} := \varphi_n \in \Omega \cap \mathcal{W} \) and \( A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B} \). Then, by Proposition 7.16, \( \int_X \varphi_n \, d\mu = \sum_{i=1}^n y_i \mu(A_i) \). Note that \( \mu(X) = \sum_{i=1}^{n+1} \mu(A_i) \) and the summands are nonnegative. Since \( \Omega \) is convex, then \( \int_X \varphi_n \, d\mu \in \Omega \). Since \( \Omega \) is closed, then, by Proposition 4.13, \( y_0 = \lim_{n \to \infty} \int_X \varphi_n \, d\mu \in \mathcal{Y} \).

Case 2: \( \varphi \notin \Omega \). Note that \( \mu(X) = 1 \) implies that \( \mathcal{X} \neq \emptyset \). Then, \( \Omega \cap \mathcal{W} \neq \emptyset \). Let \( \bar{y} \in \Omega \cap \mathcal{W} \) and \( \bar{y} := \bar{y} - \bar{g} \). Then, \( \varphi \notin \Omega \) and, by Proposition 7.16, \( \bar{y} \) is a closed convex set. Let \( \bar{f} := f \circ \bar{y} \). Then, we have \( \bar{f} : X \to \bar{\Omega} \cap \mathcal{W} \). By Propositions 7.18, 7.39, and 7.43, \( \bar{f} \) is \( \mathcal{B} \)-measurable. Note that \( \mathcal{P} \circ f(x) = \| f(x) \| \leq \mathcal{P} \circ f(x) + \| \bar{y} \| \), \( \forall x \in X \). Then, by Propositions 11.83 and 11.75, \( \int_X (\mathcal{P} \circ f) \, d\mu \leq \int_X \| \bar{y} \| \, d\mu = \int_X \| \bar{y} \| \, d\mu = \int_X \| \bar{y} \| \, d\mu + \int_X \| \bar{y} \| \, d\mu = \int_X \| \bar{y} \| \, d\mu + \int_X \| \bar{y} \| \, d\mu < +\infty \). Hence, \( \bar{f} \) is absolutely integrable over \( \mathcal{X} \). By Case 1, we have \( \int_X f \, d\mu \in \Omega \). By Propositions 11.92 and 11.75, we have \( \int_X f \, d\mu = \int_X \bar{f} \, d\mu + \bar{y} \in \Omega \).
Hence, in both cases, we have $y_0 \in \Omega$. By Proposition 8.32, we have $G(y_n) = \sup_{y_n, \in \Omega_{\text{conj}}} (\langle y_n, y_0 \rangle - G_{\text{conj}}(y_n))$, where $G_{\text{conj}} : \Omega_{\text{conj}} \to \mathbb{R}$ is the conjugate functional to $G$. For $\epsilon \in (0, \infty) \subset \mathbb{R}$, there exists $y_{se} \in \Omega_{\text{conj}}$ such that $G(y_0) - \epsilon \leq \langle y_{se}, y_0 \rangle - G_{\text{conj}}(y_{se}) \leq \langle y_{se}, y_0 \rangle - \langle y_{se}, y \rangle + G(y) = G(y) - \langle y_{se}, y - y_0 \rangle$, $\forall y \in \Omega$. Then, $\forall x \in X$, we have $G(y_0) - \epsilon \leq G(f(x)) - \langle y_{se}, f(x) - y_0 \rangle$. By Proposition 11.92, we have

$$G(y_0) - \epsilon = \int_X (G(y_0) - \epsilon) \, d\mu \leq \int_X (G \circ f - \langle y_{se}, f - y_0 \rangle) \, d\mu$$

$$= \int_X (G \circ f) \, d\mu - \int_X \langle y_{se}, f \rangle \, d\mu + \int_X \langle y_{se}, y_0 \rangle \, d\mu$$

$$= \int_X (G \circ f) \, d\mu - \langle y_{se}, \int_X f \, d\mu \rangle + \langle y_{se}, y_0 \rangle = \int_X (G \circ f) \, d\mu \in \mathbb{R}$$

By the arbitrariness of $\epsilon$, we have $\mathbb{R} \ni G(\int_X f \, d\mu) \leq \int_X (G \circ f) \, d\mu \in \mathbb{R}$. This completes the proof of the theorem. \qed

### 11.6 Banach Space Valued Measures

**Definition 11.99** Let $(X, \mathcal{B})$ be a measurable space and $\mathcal{Y}$ be a normed linear space. A $\mathcal{Y}$-valued pre-measure $\mu$ on $(X, \mathcal{B})$ is a function $\mu : \mathcal{B} \to \mathcal{Y}$ such that

(i) $\mu(\emptyset) = \emptyset$;

(ii) $\forall (E_i)^\infty_{i=1} \subseteq \mathcal{B}$, which is pairwise disjoint, we have $\sum_{i=1}^\infty \| \mu(E_i) \| < \infty$ and $\mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i)$.

Then, the triple $(X, \mathcal{B}, \mu)$ is said to be a $\mathcal{Y}$-valued pre-measure.

**Proposition 11.100** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued pre-measure space, where $\mathcal{Y}$ is a normed linear space. Then, $\mathcal{P} \circ \mu$ defines a measure on the measurable space $(X, \mathcal{B})$.

**Proof** Clearly, $\mathcal{P} \circ \mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)^\infty_{i=1} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^\infty E_i$, $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(E_i) > 0$, $\forall 0 \leq t_i < \mathcal{P} \circ \mu(E_i)$, $\exists n_i \in \mathbb{Z}^+$ and $\exists$ pairwise disjoint $(E_{i,j})^n_{j=1} \subseteq \mathcal{B}$ with $E_i = \bigcup_{j=1}^{n_i} E_{i,j}$ such that $t_i < \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| \leq \mathcal{P} \circ \mu(E_i)$. $\forall i \in \mathbb{N}$ with $\mathcal{P} \circ \mu(E_i) = 0$, let $t_i = 0$, $n_i = 1$, $E_{i,1} = E_i$. Then, $0 = t_i \leq \| \mu(E_{i,1}) \| \leq \mathcal{P} \circ \mu(E_i) = 0$.

**Claim 11.100.1** $0 \leq t := \sum_{i=1}^\infty t_i \leq \mathcal{P} \circ \mu(E)$. 


11.6. BANACH SPACE VALUED MEASURES

Proof of claim: Clearly, \( t \geq 0 \) since \( t_i \geq 0, \forall i \in \mathbb{N} \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( t < +\infty \); Case 2: \( t = +\infty \).

Case 1: \( t < +\infty \). Then, \( t \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| \geq 0 \). Therefore, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists N \in \mathbb{N} \) such that \( t - \epsilon < \sum_{i=1}^{N} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| \leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| + \| \mu(E \setminus \left( \bigcup_{i=1}^{N} \bigcup_{j=1}^{n_i} E_{i,j} \right) \) \| \leq \mathcal{P} \circ \mu(E) \). By the arbitrariness of \( \epsilon \), we have \( t \leq \mathcal{P} \circ \mu(E) \).

Case 2: \( t = +\infty \). Then \( t = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| = \infty. \forall M \in [0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N} \) such that

\[
M \leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \|
\]

\[
\leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \| \mu(E_{i,j}) \| + \| \mu(E \setminus \left( \bigcup_{i=1}^{N} \bigcup_{j=1}^{n_i} E_{i,j} \right) \) \| \leq \mathcal{P} \circ \mu(E)
\]

This implies that \( \mathcal{P} \circ \mu(E) = +\infty = t \). This completes the proof of the claim. \( \square \)

Hence, \( \sum_{i=1}^{\infty} t_i \leq \mathcal{P} \circ \mu(E) \). By the arbitrariness of \( t_i \)'s, we have \( \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i) \leq \mathcal{P} \circ \mu(E) \).

On the other hand, \( \forall n \in \mathbb{Z}_+, \forall \) pairwise disjoint \( (A_j)_{j=1}^{n} \subseteq \mathcal{B} \) with \( E = \bigcup_{j=1}^{n} A_j \), we have

\[
\sum_{j=1}^{n} \| \mu(A_j) \| = \sum_{j=1}^{n} \| \mu(\bigcup_{i=1}^{\infty} (E_i \cap A_j)) \| = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(E_i \cap A_j) \|
\]

\[
\leq \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(E_i \cap A_j) \| = \sum_{i=1}^{\infty} \sum_{j=1}^{n} \| \mu(E_i \cap A_j) \| \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)
\]

By the arbitrariness of \( n \) and \( (A_j)_{j=1}^{n} \), we have \( \mathcal{P} \circ \mu(E) \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i) \).

Hence, \( \mathcal{P} \circ \mu(E) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i) \). This implies that \( \mathcal{P} \circ \mu \) is a measure on \( (X, \mathcal{B}) \). This completes the proof of the proposition. \( \square \)

Definition 11.101 Let \( X := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued pre-measure space, where \( \mathcal{Y} \) is a normed linear space. We will say \( X \) is finite if \( \mathcal{P} \circ \mu(X) < +\infty \). We will say a property holds almost everywhere in \( X \) if it holds almost everywhere in \( \hat{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu) \). We will say that \( X \) is complete if the measure space \( \hat{X} \) is complete.

Fact 11.102 Let \( (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued pre-measure space, where \( \mathcal{Y} \) is a normed linear space. Then, \( \mu \) is finitely additive, i.e., \( \mu(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mu(E_i) \), where \( n \in \mathbb{Z}_+ \) and \( (E_i)_{i=1}^{n} \subseteq \mathcal{B} \) is pairwise disjoint.

Proof Let \( E_{n+i} = \emptyset, \forall i \in \mathbb{N} \). Clearly, \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) is pairwise disjoint. Therefore, \( \mu(\bigcup_{i=1}^{n} E_i) = \mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{n} \mu(E_i) \). This completes the proof of the fact. \( \square \)
Proposition 11.103 Let \((X, \mathcal{B}, \mu)\) be a \(Y\)-valued pre-measure space, where \(Y\) is a normed linear space, and \((E_n)_{n=1}^\infty \subseteq \mathcal{B}\) be such that \(E_n \supseteq E_{n+1}\), \(\forall n \in \mathbb{N}\). Then, \(\lim_{n \to \infty} \mu(E_n) = \mu(\bigcap_{n=1}^\infty E_n) \in Y\).

Proof Let \(E := \bigcap_{n=1}^\infty E_n \subseteq \mathcal{B}\). Note that \(E_1 = E \cup (\bigcup_{n=1}^\infty (E_n \setminus E_{n+1}))\), and the sets in the union are pairwise disjoint and measurable. Therefore, \(\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n \setminus E_{n+1}) + \|\mu(E)\| + \sum_{n=1}^\infty \|\mu(E_n \setminus E_{n+1})\| < +\infty\). Note that \(E_n = E_{n+1} \cup (E_n \setminus E_{n+1})\), \(n = 1, 2, \ldots\), which implies that \(\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \setminus E_{n+1})\), by Fact 11.102. This leads to \(\mu(E_1) = \mu(E) + \sum_{n=1}^\infty (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \mu(E_1) - \lim_{n \to \infty} \mu(E_n) \in Y\). Hence, we have \(\lim_{n \to \infty} \mu(E_n) = \mu(E) \in Y\). This completes the proof of the proposition. \(\Box\)

Proposition 11.104 Let \((X, \mathcal{B}, \mu)\) be a \(Y\)-valued pre-measure space, where \(Y\) is a normed linear space, and \((A_i)_{i=1}^\infty \subseteq \mathcal{B}\) with \(A_i \subseteq A_{i+1}\), \(\forall i \in \mathbb{N}\). Then, \(\mu(\bigcup_{i=1}^\infty A_i) = \lim_{i \to \infty} \mu(A_i) \in Y\).

Proof Let \(A := \bigcup_{i=1}^\infty A_i \subseteq \mathcal{B}\). Note that \(A = A_1 \cup (\bigcup_{i=1}^\infty (A_{i+1} \setminus A_i))\), where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, \(\mu(A) = \mu(A_1) + \sum_{i=1}^\infty \mu(A_{i+1} \setminus A_i) = \lim_{i \to \infty} (\mu(A_1) + \sum_{i=1}^\infty \mu(A_{i+1} \setminus A_i)) = \lim_{i \to \infty} \mu(A_{i+1}) = \lim_{i \to \infty} \mu(A_i) \in Y\), where the third equality follows from Fact 11.102. This completes the proof of the proposition. \(\Box\)

Proposition 11.105 Let \((X, \mathcal{B}, \mu)\) be a \(Y\)-valued pre-measure space, where \(Y\) is a normed linear space. Then, \(\forall E_1, E_2 \in \mathcal{B}\) with \(\mathcal{P} \circ \mu(E_1 \triangle E_2) = 0\), we have \(\mu(E_1) = \mu(E_2) \in Y\).

Proof By Proposition 11.100, \(0 = \mathcal{P} \circ \mu(E_1 \triangle E_2) = \mathcal{P} \circ \mu(E_1 \setminus E_2) + \mathcal{P} \circ \mu(\overline{E_2 \setminus E_1})\). Then, \(\mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(\overline{E_2 \setminus E_1}) = 0\) and \(\mu(E_1 \setminus E_2) = \mu(\overline{E_2 \setminus E_1})\). Hence, we have \(\mu(E_1) = \mu(E_1 \triangle E_2) + \mu(\overline{E_2 \setminus E_1}) = \mu(E_1 \triangle E_2) = \mu(E_2)\), where the first and last equality follows from Fact 11.102. This completes the proof of the proposition. \(\Box\)

Proposition 11.106 Let \((X, \mathcal{B}, \mu)\) be a \(Y\)-valued pre-measure space, where \(Y\) is a normed linear space. Then, there is a unique complete \(Y\)-valued pre-measure space \((X, \mathcal{B}_0, \mu_0)\) such that

(i) \(\mathcal{B} \subseteq \mathcal{B}_0\);

(ii) \(\forall E \in \mathcal{B}, \mu(E) = \mu_0(E)\) and \(\mathcal{P} \circ \mu(E) = \mathcal{P} \circ \mu_0(E)\);

(iii) \(\forall E \subseteq X, E \in \mathcal{B}_0\) if, and only if, \(E = A \cup B\), where \(A, B \subseteq X\) with \(B \in \mathcal{B}\), and there exists a \(C \in \mathcal{B}\), such that \(A \subseteq C\) and \(\mathcal{P} \circ \mu(C) = 0\).

Then, \((X, \mathcal{B}_0, \mu_0)\) is said to be the completion of \((X, \mathcal{B}, \mu)\). Then, \((X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)\) is the completion of \((X, \mathcal{B}, \mathcal{P} \circ \mu)\).

Furthermore, let \((X, \mathcal{B}_1, \mu_1)\) be another complete \(Y\)-valued pre-measure space that satisfies (i) and (ii), then, \(\mathcal{B}_0 \subseteq \mathcal{B}_1\), \(\mu_0(E) = \mu_1(E)\), and \(\mathcal{P} \circ \mu_0(E) = \mathcal{P} \circ \mu_1(E)\), \(\forall E \in \mathcal{B}_0\).
Proof  
By Proposition 11.100, \((X, B, \mathcal{P} \circ \mu)\) is a measure space. By Proposition 11.12, it admits the completion \((X, B_0, \nu_0)\), where \(B_0 \subseteq B\) satisfies (i) and (iii) and \(\nu_0(E) = \mathcal{P} \circ \mu(E), \forall E \in B\). Define \(\mu_0 : B_0 \to \mathcal{Y}\) as following, \(\forall E \in B_0\), there exists \(A, B, C \subseteq X\) such that \(E = A \cup B, B, C \in B, A \subseteq C, \) and \(\mathcal{P} \circ \mu(C) = 0\). \(\mu_0(E) := \mu(B) \in \mathcal{Y}\).

The mapping \(\mu_0\) is well defined because of the following. \(\forall E \in B_0\), let \(A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X\) be such that

\[
E = A_1 \cup B_1 = A_2 \cup B_2
\]

\[
B_1, C_1, B_2, C_2 \in B
\]

Then, \(B \ni B_1 \setminus B_2 \subseteq E \setminus B_2 = A_2 \setminus B_2 \subseteq C_2\) and \(B \ni B_2 \setminus B_1 \subseteq E \setminus B_1 = A_1 \setminus B_1 \subseteq C_1\). This implies, by Proposition 11.4, \(0 \leq \mathcal{P} \circ \mu(B_1 \Delta B_2) = \mathcal{P} \circ \mu(B_1 \setminus B_2) + \mathcal{P} \circ \mu(B_2 \setminus B_1) \leq \mathcal{P} \circ \mu(C_2) + \mathcal{P} \circ \mu(C_1) = 0\). By Proposition 11.105, we have \(\mu(B_1) = \mu(B_2)\). This shows that \(\mu_0\) is well defined.

\(\forall E \in B \subseteq B_0\), \(E = E \cup \emptyset\), then, \(\mu_0(E) = \mu(E)\). This proves that \(\mu_0\) agrees with \(\mu\) on \(B\).

Next, we will show that \(\mu_0\) is a \(\mathcal{Y}\)-valued pre-measure on the measurable space \((X, B_0)\). Since \(\emptyset \in B\), then \(\mu_0(\emptyset) = \mu(\emptyset) = \emptyset \mathcal{Y}\). Consider any sequence of pairwise disjoint subsets \((E_n)_{n=1}^{\infty} \subseteq B_0, \forall n \in \mathbb{N}\), there exists \(A_n, B_n, C_n \subseteq X\) such that \(E_n = A_n \cup B_n, B_n, C_n \in B, A_n \subseteq C_n, \) and \(\mathcal{P} \circ \mu(C_n) = 0\). Since \(B_n \subseteq E_n, \forall n \in \mathbb{N}\), then, \((B_n)_{n=1}^{\infty} \subseteq B\) is pairwise disjoint. Hence, \(\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)\) and \(\sum_{n=1}^{\infty} \|\mu(B_n)\| < +\infty\).

Since \(\bigcup_{n=1}^{\infty} E_n = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n), \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n, 0 \leq \mathcal{P} \circ \mu(\bigcup_{n=1}^{\infty} C_n) \leq \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(C_n) = 0, \) and \(\sum_{n=1}^{\infty} B_n, \bigcup_{n=1}^{\infty} C_n \in B\), then, \(\mu_0(\bigcup_{n=1}^{\infty} E_n) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \|\mu(B_n)\| < +\infty\). This proves that \(\mu_0\) is countably additive. Hence, \((X, B_0, \mu_0)\) is a \(\mathcal{Y}\)-valued pre-measure space.

Next, we will show that \(\mathcal{P} \circ \mu_0 = \nu_0\). \(\forall E \in B_0, \forall n \in \mathbb{N}_+, \forall \) pairwise disjoint \((E_i)_{i=1}^{n} \subseteq B_0\) with \(E = \bigcup_{i=1}^{n} E_i, \forall i \in \{1, \ldots, n\}, E_i = A_i \cup B_i\) such that \(A_i \subseteq C_i, B_i, C_i \in B, \) and \(\mathcal{P} \circ \mu(C_i) = 0\). Then, \((B_i)_{i=1}^{n} \subseteq B\) is pairwise disjoint, \(E = (\bigcup_{i=1}^{n} A_i) \cup (\bigcup_{i=1}^{n} B_i) =: A \cup B, A \subseteq \bigcup_{i=1}^{n} C_i =: C \in B, B \in B, \) and \(0 \leq \mathcal{P} \circ \mu(C) \leq \sum_{i=1}^{n} \mathcal{P} \circ \mu(C_i) = 0\), where the last inequality follows from Proposition 11.6. By Proposition 11.12 and its proof, \(\nu_0(E) = \mathcal{P} \circ \mu(B) = \sum_{i=1}^{n} \mathcal{P} \circ \mu(B_i) \geq \sum_{i=1}^{n} \|\mu(B_i)\| = \sum_{i=1}^{n} \|\mu_0(E_i)\|\). Hence, \(\nu_0(E) \geq \mathcal{P} \circ \mu_0(E)\).

On the other hand, \(\nu_0(E) = \mathcal{P} \circ \mu(B) \leq \mathcal{P} \circ \mu_0(B) \leq \mathcal{P} \circ \mu_0(E)\), where the first inequality follows from Definition 11.99 and the fact that \(\mu = \mu_0|_{B_0}\); and the second inequality follows from Proposition 11.4. Therefore, \(\mathcal{P} \circ \mu_0 = \nu_0\). \(\forall E \in B, \mathcal{P} \circ \mu_0(E) = \nu_0(E) = \mathcal{P} \circ \mu_0(E)\). This shows that (ii) is satisfied. This proves that \((X, B_0, \mu_0)\) is a complete \(\mathcal{Y}\)-valued pre-measure space satisfying (i), (ii), and (iii) and \((X, B_0, \mathcal{P} \circ \mu_0)\) is the completion of \((X, B, \mathcal{P} \circ \mu)\).
Let \((X, \mathcal{B}_1, \mu_1)\) be any complete \(\mathcal{Y}\)-valued pre-measure space that satisfies the (i) and (ii). By Proposition 11.100, \((X, \mathcal{B}_1, \mathcal{P} \circ \mu_1)\) is a complete measure space that extends \((X, \mathcal{B}, \mathcal{P} \circ \mu)\). By Proposition 11.12 and the fact that \((X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)\) is the completion of \((X, \mathcal{B}, \mathcal{P} \circ \mu)\), then \(\mathcal{B}_0 \subseteq \mathcal{B}_1\) and \((\mathcal{P} \circ \mu_1)|_{\mathcal{B}_0} = \mathcal{P} \circ \mu_0\). \(\forall E \in \mathcal{B}_0\), there exists \(A, B, C \subseteq X\) such that \(E = A \cup B, A \subseteq C, B, C \in \mathcal{B}\), and \(\mathcal{P} \circ \mu(C) = 0\). Then, \(\mathcal{P} \circ \mu_1(C) = 0, E = B \cup (E \setminus B), \mathcal{B}_0 \ni E \setminus B = A \setminus B \subseteq C,\) and 
\(0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0\). This leads to \(\mu_1(E \setminus B) = \emptyset\) and \(\mu_1(E) = \mu_1(B) + \mu_1(E \setminus B) = \mu(B) = \mu_0(E)\). Therefore, \(\mu_0 = \mu_1|_{\mathcal{B}_0}\).

Therefore, \((X, \mathcal{B}_0, \mu_0)\) is a complete \(\mathcal{Y}\)-valued pre-measure space that satisfies (i), (ii), and (iii). It is the unique such space since \(\mathcal{B}_0\) is unique by (iii), and \(\mu_0\) is also unique since any complete \(\mathcal{Y}\)-valued pre-measure space \((X, \mathcal{B}_0, \mu_1)\) satisfying (i) and (ii) will be such that \(\mu_1 = \mu_0\) and \(\mathcal{P} \circ \mu_0 = \mathcal{P} \circ \mu_1\).

This completes the proof of the proposition. \(\square\)

**Proposition 11.107** Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued pre-measure space, where \(\mathcal{Y}\) is a normed linear space, \(A \in \mathcal{B}, \mathcal{B} \subseteq \mathcal{B}\) be a \(\sigma\)-algebra on \(X\), and \(\mathcal{B}_A := \{C \subseteq A \mid C \in \mathcal{B}\}\). Then, the following statements hold.

(i) \((X, \mathcal{B}, \mu|_{\mathcal{B}})\) and \(\mathcal{A} := (A, \mathcal{B}_A, \mu_A := \mu|_{\mathcal{B}_A})\) are \(\mathcal{Y}\)-valued pre-measure spaces, and \((\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ (\mu|_{\mathcal{B}_A}) = \mathcal{P} \circ \mu_A\). The \(\mathcal{Y}\)-valued pre-measure space \(\mathcal{A}\) is said to be the subspace of the \(\mathcal{Y}\)-valued pre-measure space \(\mathcal{X}\).

(ii) If, in addition, \(\mathcal{X}\) is complete, then \(\mathcal{A}\) is also complete.

(iii) Let \(\check{\mathcal{X}} := (X, \check{\mathcal{B}}, \check{\mu})\) be the completion of \(\mathcal{X}\) as defined in Proposition 11.106, and \(\check{\mathcal{A}} := (A, \check{\mathcal{B}}_A, \check{\mu}_A)\) be the \(\mathcal{Y}\)-valued pre-measure subspace of \(\check{\mathcal{X}}\). Then, \(\check{\mathcal{A}}\) is the completion of \(\mathcal{A}\).

**Proof**

(i) Since \(\mathcal{B}\) is a \(\sigma\)-algebra on \(X\), then \((X, \mathcal{B})\) is a measurable space. \(\forall B \in \mathcal{B} \subseteq \mathcal{B}, \mu|_{\mathcal{B}}(B) = \mu(B) \in \mathcal{Y}\). Then, \(\mu|_{\mathcal{B}} : \mathcal{B} \rightarrow \mathcal{Y}\). Clearly, \(\mu|_{\mathcal{B}}(\emptyset) = \mu(\emptyset) = \emptyset\). \(\forall\) pairwise disjoint sequence \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \subseteq \mathcal{B}\), we have \(E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}, \sum_{i=1}^{\infty} \| \mu|_{\mathcal{B}}(E_i) \| = \sum_{i=1}^{\infty} |\mu(E_i)| < +\infty\), and \(\mu|_{\mathcal{B}}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \| \mu|_{\mathcal{B}}(E_i) \|. \) Hence, \((X, \mathcal{B}, \mu|_{\mathcal{B}})\) is a \(\mathcal{Y}\)-valued pre-measure space.

Clearly, \(\mathcal{B}_A\) is a \(\sigma\)-algebra on \(A\) as we have shown in Proposition 11.13. \(\emptyset \in \mathcal{B}_A\) and \(\mu|_{\mathcal{B}_A}(\emptyset) = \mu(\emptyset) = \emptyset\). \(\forall\) pairwise disjoint sequence \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B}_A \subseteq \mathcal{B}\), we have \(E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_A, \sum_{i=1}^{\infty} \| \mu|_{\mathcal{B}_A}(E_i) \| = \sum_{i=1}^{\infty} |\mu(E_i)| < +\infty\), and \(\mu|_{\mathcal{B}_A}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \| \mu|_{\mathcal{B}_A}(E_i) \|. \) Hence \(\mathcal{A} := (A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})\) is a \(\mathcal{Y}\)-valued pre-measure space. By Definition 11.99, \((\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ (\mu|_{\mathcal{B}_A}) = \mathcal{P} \circ \mu_A\).

(ii) If, in addition, \(\mathcal{X}\) is complete, then \((X, \mathcal{B}, \mathcal{P} \circ \mu)\) is a complete measure space. \(\forall E_A \subseteq A\) such that there exists \(B \in \mathcal{B}_A\) with \(E_A \subseteq B\) and \(\mathcal{P} \circ (\mu|_{\mathcal{B}_A})(B) = 0\), we have \(B \in \mathcal{B}\) and \(\mathcal{P} \circ \mu(B) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(B) = \)
of Proposition 11.106 are satisfied and \( \bar{\mu} \) is complete. Hence, \( A \) is complete.

(iii) Note that \( \mathcal{B}_A = \{ C \subseteq A \mid C \in \mathcal{B} \} \), \( \bar{\mu}_A = \bar{\mu}|_{\mathcal{B}_A} \), and \( \mathcal{P} \circ \mu_A = (\mathcal{P} \circ \bar{\mu})|_{\mathcal{B}_A} \). By (ii), \( \bar{\mathcal{A}} \) is a complete \( \mathcal{Y} \)-valued measure space. Since \( \bar{\mu} \) is complete, then \( \mathcal{B}_A \subseteq \bar{\mathcal{B}}_A \). \( \forall E \in \mathcal{B}_A \), \( \mu_A(E) = \mu|_{\mathcal{B}_A}(E) = \mu(E) = \bar{\mu}(E) \) and, by (i), \( \mathcal{P} \circ \mu_A(E) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(E) = \mathcal{P} \circ \mu(E) = \mathcal{P} \circ \bar{\mu}(E) = (\mathcal{P} \circ \bar{\mu})|_{\mathcal{B}_A}(E) = \mathcal{P} \circ \bar{\mu}_A(E) \), where the third equality follows from Proposition 11.106. \( \forall E \subseteq A \), assume that \( E = E_A \cup E_B \) with \( E_B, E_C \subseteq \mathcal{B}_A \), \( E_A \subseteq E_C \), and \( \mathcal{P} \circ \mu_A(E_C) = 0 \). Then, \( \mathcal{P} \circ \bar{\mu}_A(E_C) = 0 \) and \( E_A \subseteq \mathcal{B}_A \) since \( \bar{\mathcal{A}} \) is complete. Then, \( E \in \bar{\mathcal{B}}_A \). \( \forall E \subseteq A \), assume that \( E \in \mathcal{B}_A \). Then, \( E \subseteq A \) and \( E \in \bar{\mathcal{B}} \). By Proposition 11.106, \( E = E_A \cup E_B \) with \( E_B, E_C \subseteq \mathcal{B}_A \), \( E_A \subseteq E_C \), and \( \mathcal{P} \circ \mu(E_C) = 0 \). Then, \( E_B, E_C \cap A \subseteq \mathcal{B}_A \), \( E_A \subseteq E_C \cap A \), and \( 0 \leq \mathcal{P} \circ \mu_A(E_C \cap A) = \mathcal{P} \circ \mu(E_C \cap A) \leq \mathcal{P} \circ \mu(E_C) = 0 \), where the last inequality follows from Proposition 11.4. Hence, (i) — (iii) of Proposition 11.106 are satisfied and \( \bar{\mathcal{A}} \) is the completion of \( \mathcal{A} \).

This completes the proof of the proposition. \( \square \)

**Definition 11.108** Let \((X, \mathcal{B}, \nu)\) be a measure space and \( \mathcal{Y} \) be a normed linear space. A \( \mathcal{Y} \)-valued measure \( \mu \) on \((X, \mathcal{B})\) is a function from \( \mathcal{B} \) to \( \mathcal{Y} \) that satisfies

(i) \( \mu(\emptyset) = 0 \);

(ii) \( \forall E \in \mathcal{B} \) with \( \nu(E) = +\infty \), \( \mu(E) \) is undefined;

(iii) \( \forall E \in \mathcal{B} \) with \( \nu(E) < +\infty \), \( \mu(E) \in \mathcal{Y} \) and \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{\infty} E_i \), we have \( \sum_{i=1}^{\infty} \| \mu(E_i) \| < +\infty \) and \( \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \); (countable additivity)

(iv) \( \forall E \in \mathcal{B} \) with \( \nu(E) < +\infty \), we have

\[
\nu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \bigcup_{i=1}^{n} E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(E_i) \|
\]

Then, \((X, \mathcal{B}, \mu)\) is said to be a \( \mathcal{Y} \)-valued measure space; and \( \nu \) is said to be the total variation of \( \mu \), denoted by \( \mathcal{P} \circ \mu \).

**Definition 11.109** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space. We will say \( \mathcal{X} \) is finite if \( \mathcal{P} \circ \mu(X) < +\infty \). We will say \( \mathcal{X} \) is \( \sigma \)-finite if \( \mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu) \) is \( \sigma \)-finite. We will say a property holds almost everywhere in \( \mathcal{X} \) if it holds almost everywhere in \( \bar{\mathcal{X}} \). We will say that \( \mathcal{X} \) is complete if the measure space \( \bar{\mathcal{X}} \) is complete. We will say that a sequence of measurable functions on \( \mathcal{X} \) converges in measure in \( \mathcal{X} \) if it converges in measure in \( \bar{\mathcal{X}} \).
Clearly, \((X, \mathcal{B}, \mu)\) is a finite \(\mathcal{Y}\)-valued measure space if, and only if, it is a finite \(\mathcal{Y}\)-valued pre-measure space. In this case, the definitions of total variations of \(\mu\) coincide considering \(\mu\) as a \(\mathcal{Y}\)-valued measure or as \(\mathcal{Y}\)-valued pre-measure. Hence, a \(\mathcal{Y}\)-valued pre-measure space becomes a finite \(\mathcal{Y}\)-valued measure space when we show that its total variation is finite.

**Fact 11.110** Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space. Then, \(\mu\) is finitely additive, i.e., \(\forall n \in \mathbb{Z}_+\) and \(\forall\) pairwise disjoint \((E_i)_{i=1}^n \subseteq \mathcal{B}\), \(\mu\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i)\), whenever the right-hand-side or the left-hand-side makes sense.

**Proof** Let \(E := \bigcup_{i=1}^n E_i \in \mathcal{B}\). Clearly, \(\mathcal{P} \circ \mu(E) = \sum_{i=1}^n \mathcal{P} \circ \mu(E_i)\). Then, \(E \in \text{dom}(\mu)\) whenever \((E_i)_{i=1}^n \subseteq \text{dom}(\mu)\). Let \(E_{n+i} = \emptyset\), \(\forall i \in \mathbb{N}\). Clearly, \((E_i)_{i=1}^n \subseteq \mathcal{B}\) is pairwise disjoint. Therefore, \(\mu(E) = \mu\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i) = \sum_{i=1}^n \mu(E_i) \in \mathcal{Y}\), when \(E \in \text{dom}(\mu)\). This completes the proof of the fact. \(\square\)

**Proposition 11.111** Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space, and \((E_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) be such that \(\mathcal{P} \circ \mu(E_1) < +\infty\) and \(E_n \supseteq E_{n+1}, \forall n \in \mathbb{N}\). Then, \(\lim_{n \to \infty} \mu(E_n) = \mu(\bigcap_{n=1}^{\infty} E_n) \in \mathcal{Y}\).

**Proof** Clearly, \(E_1 \in \text{dom}(\mu)\) and then \(E_n \in \text{dom}(\mu), \forall n \in \mathbb{N}\). Let \(E := \bigcap_{n=1}^{\infty} E_n \in \mathcal{B}\). Note that \(E_1 = E \cap \left(\bigcup_{n=1}^{\infty} \left(E_n \setminus E_{n+1}\right)\right)\), and the sets in the union are pairwise disjoint and measurable. Therefore, \(\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}\) and \(\|\mu(E)\| + \sum_{n=1}^{\infty} \|\mu(E_n \setminus E_{n+1})\| < \infty\). Note that \(E_n = E_{n+1} \cup \left(E_n \setminus E_{n+1}\right)\), \(n = 1, 2, \ldots\), which implies that \(\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}\), by Fact 11.110. This leads to \(\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} \mu(E_n \setminus E_{n+1})\). \(\mu(E) = \mu(E_1) = \lim_{n \to \infty} \mu(E_n) \in \mathcal{Y}\). Hence, we have \(\lim_{n \to \infty} \mu(E_n) = \mu(E) \in \mathcal{Y}\). This completes the proof of the proposition. \(\square\)

**Proposition 11.112** Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space, and \((A_i)_{i=1}^{\infty} \subseteq \mathcal{B}\) be such that \(A_i \subseteq A_{i+1}, \forall i \in \mathbb{N}\), and \(A := \bigcup_{i=1}^{\infty} A_i \in \text{dom}(\mu)\). Then, \(\mu(A) = \lim_{n \to \infty} \mu(A_i) \in \mathcal{Y}\).

**Proof** \(A \in \text{dom}(\mu)\) implies that \(\mathcal{P} \circ \mu(A) < +\infty\). Then, \(\mathcal{P} \circ \mu(A_i) \leq \mathcal{P} \circ \mu(A) < +\infty\) and \(A_i \in \text{dom}(\mu), \forall i \in \mathbb{N}\). Note that \(A = A_1 \cup \left(\bigcup_{i=1}^{\infty} (A_{i+1} \setminus A_i)\right)\), where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, \(\mu(A) = \mu(A_1) + \sum_{i=1}^{\infty} \mu(A_{i+1} \setminus A_i) = \lim_{n \to \infty} (\mu(A_1) + \sum_{i=1}^{n} \mu(A_{i+1} \setminus A_i)) = \lim_{n \to \infty} \mu(A_{n+1}) = \lim_{i \to \infty} \mu(A_i) \in \mathcal{Y}\), where the third equality follows from Fact 11.110. This completes the proof of the proposition. \(\square\)

**Proposition 11.113** Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space. Then, \(\forall E_1, E_2 \in \mathcal{B}\) with \(\mathcal{P} \circ \mu(E_1 \Delta E_2) = 0\), we have \(\mu(E_1) = \mu(E_2)\) whenever one of them makes sense.
Proposition 11.114 Let \((X, B, \mu)\) be a \(Y\)-valued measure space, where \(Y\) is a normed linear space. Then, there is a unique complete \(Y\)-valued measure space \((X, B_0, \mu_0)\) such that

(i) \(B \subseteq B_0\);

(ii) \((\mathcal{P} \circ \mu)|_B = \mathcal{P} \circ \mu\) and \(\mu_0|_B = \mu\);

(iii) \(\forall E \in X, E \in B_0\) if, and only if, \(E = A \cup B\), where \(A, B \subseteq X\) with \(B \in B\), and there exists a \(C \in B\), such that \(A \subseteq C\) and \(\mathcal{P} \circ \mu(C) = 0\).

Then, \((X, B_0, \mu_0)\) is said to be the completion of \((X, B, \mu)\). Then, \((X, B_0, \mathcal{P} \circ \mu)\) is the completion of \((X, B, \mathcal{P} \circ \mu)\).

Furthermore, let \((X, B_1, \mu_1)\) be another complete \(Y\)-valued measure space that satisfies (i) and (ii), then \(B_0 \subseteq B_1\), \(\mathcal{P} \circ \mu_0 = (\mathcal{P} \circ \mu_1)|_{B_0}\), and \(\mu_0 = \mu_1|_{B_0}\).

Proof By Definition 11.108, \((X, B, \mathcal{P} \circ \mu)\) is a measure space. By Proposition 11.12, it admits the completion \((X, B_0, \nu_0)\), where \(B_0\) satisfies (i) and (iii) and \(\nu_0(E) = \mathcal{P} \circ \mu(E), \forall E \in B\). Define a function \(\mu_0\) from \(B_0\) to \(Y\) as following. \(\forall E \in B_0\) with \(\nu_0(E) = +\infty\), \(\mu_0(E)\) is undefined. \(\forall E \in B_0\) with \(\nu_0(E) < +\infty\), there exists \(A, B, C \subseteq X\) such that \(E = A \cup B\), \(B, C \in B\), \(A \subseteq C\), and \(\mathcal{P} \circ \mu(C) = 0\). By Proposition 11.12 and its proof, \(\nu_0(E) = \mathcal{P} \circ \mu(B) < +\infty\) and \(B \in \text{dom} (\mu)\). We will set \(\mu_0(E) := \mu(B) \in Y\).

The mapping \(\mu_0\) is well defined because of the following. \(\forall E \in B_0\) with \(\nu_0(E) < +\infty\), let \(A_1, A_2, B_1, B_2, C_1, C_2 \subseteq X\) be such that

\[
E = A_1 \cup B_1 = A_2 \cup B_2
\]

\[
B_1, C_1, B_2, C_2 \in B
\]

Then, \(B \ni B_1 \setminus B_2 \subseteq E \setminus B_2 = A_2 \setminus B_2 \subseteq C_2\) and \(B \ni B_2 \setminus B_1 \subseteq E \setminus B_1 = A_1 \setminus B_1 \subseteq C_1\). This implies, by Proposition 11.4, \(0 \leq \mathcal{P} \circ \mu(B_1 \triangle B_2) = \mathcal{P} \circ \mu(B_1 \setminus B_2) + \mathcal{P} \circ \mu(B_2 \setminus B_1) \leq \mathcal{P} \circ \mu(C_2) + \mathcal{P} \circ \mu(C_1) = 0\). By Proposition 11.113, we have \(\mu(B_1) = \mu(B_2) \in Y\). This shows that \(\mu_0\) is well defined.

\(\forall E \in B \subseteq B_0, E = E \cup \emptyset\), then, \(\nu_0(E) = \mathcal{P} \circ \mu(E)\). If \(\mathcal{P} \circ \mu(E) = +\infty\), then \(\mu(E)\) is undefined and \(\mu_0(E)\) is also undefined. If \(\mathcal{P} \circ \mu(E) < +\infty\), then \(\mu_0(E) = \mu(E) \in Y\). This proves that \(\mu_0|_B = \mu\).
Next, we will show that \((X, \mathcal{B}_0, \mu_0)\) is a \(\mathcal{Y}\)-valued measure space with \(\mathcal{P} \circ \mu_0 = \nu_0\). Since \(\emptyset \in \mathcal{B}\), then \(\mu_0(\emptyset) = \mu(\emptyset) = 0\). Hence, (i) of Definition 11.108 is satisfied. (ii) of Definition 11.108 is satisfied by the definition of \(\mu_0\). Consider any \(E \in \mathcal{B}_0\) with \(\nu_0(E) < +\infty\), then \(\mu_0(E) \in \mathcal{Y}\). \(\forall\) pairwise disjoint \((E_n)_{n=1}^{\infty} \subseteq B_0\) with \(E = \bigcup_{n=1}^{\infty} E_n\), \(\forall n \in \mathbb{N}\), there exists \(A_n, B_n, C_n \subseteq X\) such that \(E_n = A_n \cup B_n, B_n, C_n \in B, A_n \subseteq C_n, \) and \(\mathcal{P} \circ \mu(C_n) = 0\). Since \(B_n \subseteq E_n, \forall n \in \mathbb{N}\), then, \((B_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) is pairwise disjoint. Let \(A := \bigcup_{n=1}^{\infty} A_n, B := \bigcup_{n=1}^{\infty} B_n \in \mathcal{B}, \) and \(C := \bigcup_{n=1}^{\infty} C_n \in \mathcal{B}\). Then, \(E = A \cup B, A \subseteq C, \) and \(\mathcal{P} \circ \mu(C) = 0\). This implies that \(\mathcal{P} \circ \mu(B) = \nu_0(E) < +\infty, \mu(B) = \sum_{n=1}^{\infty} \mu(B_n) \in \mathcal{Y}, \) and \(\sum_{n=1}^{\infty} \| \mu(B_n) \| < +\infty\). Hence, \(\mu_0(E) = \mu(B) = \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} \mu_0(E_n) \in \mathcal{Y}\). Also, \(\sum_{n=1}^{\infty} \| \mu_0(E_n) \| = \sum_{n=1}^{\infty} \| \mu(B_n) \| < +\infty\). This proves that (iii) of Definition 11.108 is satisfied and \(\mu_0\) is countably additive.

\(\forall E \in \mathcal{B}_0\) with \(\nu_0(E) < +\infty, \forall n \in \mathbb{Z}_+, \forall\) pairwise disjoint \((E_i)_{i=1}^{n} \subseteq B_0\) with \(E = \bigcup_{i=1}^{n} E_i, \forall i \in \{1, \ldots, n\}, \exists A_i, B_i, C_i \subseteq X\) such that \(A_i = A_i \cup B_i, B_i, C_i \subseteq B, A_i \subseteq C_i, \) and \(\mathcal{P} \circ \mu(C_i) = 0\). Then, \((B_i)_{i=1}^{n} \subseteq \mathcal{B}\) is pairwise disjoint. Let \(A := \bigcup_{i=1}^{n} A_i, B := \bigcup_{i=1}^{n} B_i \in \mathcal{B}, \) and \(C := \bigcup_{i=1}^{n} C_i \in \mathcal{B}\), we have \(E = A \cup B, A \subseteq C, \) and \(\mathcal{P} \circ \mu(C) = 0\). By Proposition 11.12 and its proof, \(+\infty > \nu_0(E) = \mathcal{P} \circ \mu(B) = \sum_{i=1}^{n} \mathcal{P} \circ \mu(B_i) \geq \sum_{i=1}^{n} \| \mu(B_i) \| = \sum_{i=1}^{n} \| \mu_0(E_i) \|\). Hence,

\[\nu_0(E) \geq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq B_0, \bigcup_{i=1}^{n} E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu_0(E_i) \| =: s_E\]

On the other hand,

\[+\infty > \nu_0(E) = \mathcal{P} \circ \mu(B) = \sup_{n \in \mathbb{Z}_+, (B_i)_{i=1}^{n} \subseteq B, \bigcup_{i=1}^{n} B_i = B, B_i \cap B_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(B_i) \| \leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq B_0, \bigcup_{i=1}^{n} E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu_0(E_i) \| \leq s_E\]

Hence, we have \(\nu_0(E) = s_E\) and (iv) of Definition 11.108 is satisfied. Hence, \((X, \mathcal{B}_0, \mu_0)\) is a \(\mathcal{Y}\)-valued measure space with total variation \(\mathcal{P} \circ \mu_0 = \nu_0\). Since \(\nu_0\) is complete, then \((X, \mathcal{B}_0, \mu_0)\) is a complete \(\mathcal{Y}\)-valued measure space. Since \(\nu_0|_B = \mathcal{P} \circ \mu, \) then \(\mathcal{P} \circ \mu_0|_B = \mathcal{P} \circ \mu\). Hence, (i) – (iii) are satisfied by \((X, \mathcal{B}_0, \mu_0)\). Clearly, \((X, \mathcal{B}_0, \mathcal{P} \circ \mu)\) is the completion of \((X, \mathcal{B}, \mathcal{P} \circ \mu)\).

Let \((X, \mathcal{B}_1, \mu_1)\) be any complete \(\mathcal{Y}\)-valued measure space that satisfies (i) and (ii). \((X, \mathcal{B}_1, \mathcal{P} \circ \mu)\) is a complete measure space that extends \((X, \mathcal{B}, \mathcal{P} \circ \mu)\). By Proposition 11.12 and the fact that \((X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)\) is the completion of \((X, \mathcal{B}, \mathcal{P} \circ \mu)\), then \(\mathcal{B}_0 \subseteq \mathcal{B}_1\) and \(\mathcal{P} \circ \mu_1|_{\mathcal{B}_0} = \mathcal{P} \circ \mu_0\). \(\forall E \in \mathcal{B}_0\) with \(\mathcal{P} \circ \mu_0(E) = +\infty\), we have \(\mathcal{P} \circ \mu_1(E) = +\infty\) and \(\mu_0(E)\) and \(\mu_1(E)\) are both undefined. \(\forall E \in \mathcal{B}_0\) with \(\mathcal{P} \circ \mu_0(E) < +\infty\), by (iii), there exists \(A, B, C \subseteq X\) such that \(E = A \cup B, A \subseteq C, B, C \in \mathcal{B},\) and \(\mathcal{P} \circ \mu(C) = 0\).
Then, $\mu_0(E) = \mu(B) = \mu_1(B) \in \mathcal{Y}$ and $\mathcal{P} \circ \mu_1(E) = \mathcal{P} \circ \mu_0(E) < +\infty$. Note that $\mathcal{P} \circ \mu_1(C) = \mathcal{P} \circ \mu(C) = 0$, $E = B \cup (E \setminus B)$, $B_0 \ni E \setminus B = A \setminus B \subseteq C$, and $0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0$. This leads to $\mu_1(E \setminus B) = \vartheta_y$ and, by Fact 11.110, $\mu_1(E) = \mu_1(B) + \mu_1(E \setminus B) = \mu(B) = \mu_0(E) \in \mathcal{Y}$. Therefore, $\mu_0 = \mu_1|_{B_0}$.

Therefore, $(X, B_0, \mu_0)$ is a complete $\mathcal{Y}$-valued measure space that satisfies (i), (ii), and (iii). It is the unique such space since $B_0$ is unique by (iii), and $\mu_0$ is also unique since any complete $\mathcal{Y}$-valued measure space $(X, B_0, \mu_1)$ satisfying (i) and (ii) will be such that $\mu_1 = \mu_0$ and $\mathcal{P} \circ \mu_0 = \mathcal{P} \circ \mu_1$.

This completes the proof of the proposition. □

**Proposition 11.115** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space, $A \subseteq \mathcal{B}$, and $\mathcal{B}_A := \{C \subseteq A \mid C \in \mathcal{B}\}$.

Then, the following statements hold.

(i) $A := (A, \mathcal{B}_A, \mu_A := \mu|_{\mathcal{B}_A})$ is a $\mathcal{Y}$-valued measure space, which is said to be the subspace of $\mathcal{X}$, and $(\mathcal{P} \circ \mu)|_{\mathcal{B}_A} = \mathcal{P} \circ \mu_A$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $A$ is also complete.

(iii) Let $\tilde{\mathcal{X}} := (X, \tilde{\mathcal{B}}, \tilde{\mu})$ be the completion of $\mathcal{X}$ as defined in Proposition 11.114, and $\tilde{A} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)$ be the $\mathcal{Y}$-valued measure subspace of $\tilde{\mathcal{X}}$. Then, $\tilde{A}$ is the completion of $A$.

**Proof** (i) Clearly, $\mathcal{B}_A$ is a $\sigma$-algebra on $A$ as we have shown in Proposition 11.13. By Proposition 11.13, $\nu_A := (\mathcal{P} \circ \mu)|_{\mathcal{B}_A} : \mathcal{B}_A \to [0, +\infty]$ is a measure on $(A, \mathcal{B}_A)$. Clearly, $\emptyset \in \mathcal{B}_A$ and $\mu_A(\emptyset) = \mu|_{\mathcal{B}_A}(\emptyset) = \mu(\emptyset) = \nu_y$. ∀ $E \in \mathcal{B}_A$ with $\nu_A(E) = +\infty$, we have $\mathcal{P} \circ \mu(E) = +\infty$ and $\mu(E)$ is undefined. Then, $\mu_A(E)$ is also undefined. ∀ $E \in \mathcal{B}_A$ with $\nu_A(E) < +\infty$, we have $\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty$. Then, $\mu_A(E) = \mu(E) \in \mathcal{Y}$, ∀ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}_A \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^n E_i$, we have $\sum_{i=1}^n \|\mu_A(E_i)\| = \sum_{i=1}^\infty \|\mu(E_i)\| < +\infty$, and $\mu_A(E) = \mu(E) = \sum_{i=1}^\infty \mu(E_i) \in \mathcal{Y}$, ∀ $E \in \mathcal{B}_A$ with $\mu_A(E) < +\infty$, we have $\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty$. Then, $\nu_A(E) = \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{Z}_+, \sum_{i=1}^n E_i \subseteq \mathcal{B}} \sum_{i=1}^n \|\mu_A(E_i)\| = \sup_{n \in \mathbb{Z}_+, \sum_{i=1}^n E_i \subseteq \mathcal{B}} \sum_{i=1}^n \|\mu(E_i)\|$. Hence, $A = (A, \mathcal{B}_A, \mu_A)$ is a $\mathcal{Y}$-valued measure space with $\mathcal{P} \circ \mu_A = \nu_A = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}$.

(ii) If, in addition, $\mathcal{X}$ is complete, then $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a complete measure space. Since $(A, \mathcal{B}_A, \mathcal{P} \circ \mu_A)$ is a subspace of $(X, \mathcal{B}, \mathcal{P} \circ \mu)$, then it is complete by Proposition 11.13. Hence, $A$ is complete.

(iii) Note that $\tilde{\mathcal{B}}_A = \{C \subseteq A \mid C \in \tilde{\mathcal{B}}\}$, $\tilde{\mu}_A = \tilde{\mu}|_{\tilde{\mathcal{B}}_A}$, and $\mathcal{P} \circ \tilde{\mu}_A = (\mathcal{P} \circ \tilde{\mu})|_{\tilde{\mathcal{B}}_A}$. By (ii), $\tilde{A}$ is a complete $\mathcal{Y}$-valued measure space. Since $\mathcal{B} \subseteq \mathcal{B}$, then $\tilde{\mathcal{B}} \subseteq \tilde{\mathcal{B}}_A$. Note that $\mathcal{P} \circ \tilde{\mu}_A = (\mathcal{P} \circ \tilde{\mu})|_{\tilde{\mathcal{B}}_A} = ((\mathcal{P} \circ \tilde{\mu})|_{\tilde{\mathcal{B}}_A})|_{\tilde{\mathcal{B}}_A} = ((\mathcal{P} \circ \tilde{\mu})|_{\tilde{\mathcal{B}}_A})|_{\tilde{\mathcal{B}}_A}$, where the
first equality follows from (i); the second equality follows from Proposition 11.114; the third and fourth equalities follow from the fact that $B_A \subseteq B$ and $B_A \subseteq B_A$; and the last equality follows from (i). Note also that $\mu_A = \mu|_{B_A} = (\bar{\mu}|_{B_A})|_{B_A} = (\bar{\mu}|_{B_A})|_{B_A} = \mu_A|_{B_A}$, where the first equality follows from (i), the second equality follows from Proposition 11.114; the third equality follows from the fact that $B_A \subseteq B$ and $B_A \subseteq B_A$; and the last equality follows from (i). For $E \subseteq A$, assume that $E = E_A \cup E_B$ with $E_B, E_C \in B_A$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu_A(E_C) = 0$. Then, $\mathcal{P} \circ \mu_A(E_C) = 0$ and $E_A \in B_A$ since $A$ is complete. Then, $E \in B_A$. On the other hand, for $E \subseteq A$, assume that $E \in B_A$. Then, $E \subseteq A$ and $E \in B$. By Proposition 11.114, $E = E_A \cup E_B$ with $E_B, E_C \in B$, $E_A \subseteq E_C$, and $\mathcal{P} \circ \mu(E_C) = 0$. Then, $E_B, E_C \cap A \in B_A$, $E_A \subseteq E_C \cap A$, and $0 \leq \mathcal{P} \circ \mu_A(E_C \cap A) = \mathcal{P} \circ \mu(E_C \cap A) \leq \mathcal{P} \circ \mu(E_C) = 0$, where the last inequality follows from Proposition 11.14. Hence, (i) — (iii) of Proposition 11.114 are satisfied and $\bar{A}$ is the completion of $A$.

This completes the proof of the proposition.  

\begin{proposition}
\textbf{Proposition 11.116} Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, and $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable. Define $\bar{\nu} : B \to [0, +\infty] \subseteq \mathbb{R}$ by $\bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mu$, $\forall E \in B$. Then, $\bar{\nu}$ is a measure on $(X, \mathcal{B})$. Define a function $\nu$ from $B$ to $\mathcal{Y}$ by $\nu(E)$ is undefined, $\forall E \in B$ with $\bar{\nu}(E) = +\infty$, and $\nu(E) = \int_E f \, d\mu$, $\forall E \in B$ with $\bar{\nu}(E) < +\infty$. Then, $(X, \mathcal{B}, \nu)$ is a $\mathcal{Y}$-valued measure space with $\mathcal{P} \circ \nu = \bar{\nu}$. We will call $\nu$ the $\mathcal{Y}$-valued measure with kernel $f$ over $\mathcal{X}$. Furthermore, the following statements hold.

(i) If $\mathcal{X}$ is $\sigma$-finite, then $\nu$ is also $\sigma$-finite.

(ii) $f$ is absolutely integrable over $\mathcal{X}$ if, and only if, $\nu$ is finite.

(iii) If $\mathcal{P} \circ f > 0$ a.e. in $\mathcal{X}$ and $\mathcal{X}$ is $\sigma$-finite and complete, then $(X, \mathcal{B}, \nu)$ is a $\sigma$-finite complete $\mathcal{Y}$-valued measure space.

\begin{proof}
Since $f$ is $\mathcal{B}$-measurable, by Propositions 7.21 and 11.38, $\mathcal{P} \circ f$ is $\mathcal{B}$-measurable. Then, $\bar{\nu} : B \to [0, +\infty] \subseteq \mathbb{R}$ is well defined by Definition 11.79. By Proposition 11.75, $\bar{\nu}(\emptyset) = \int_{\emptyset} (\mathcal{P} \circ f) \, d\mu = 0$. \forall pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq B$, by Proposition 11.83, we have $\bar{\nu}\left(\bigcup_{i=1}^{\infty} E_i\right) = \int_{\bigcup_{i=1}^{\infty} E_i} (\mathcal{P} \circ f) \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} (\mathcal{P} \circ f) \, d\mu = \sum_{i=1}^{\infty} \bar{\nu}(E_i)$. Hence, $\bar{\nu}$ is a measure on $(X, \mathcal{B})$. By Proposition 11.75, $\nu(\emptyset) = \int_{\emptyset} f \, d\mu = 0$. \forall $E \in B$ with $\bar{\nu}(E) = +\infty$, $\nu(E)$ is undefined. \forall $E \in B$ with $\bar{\nu}(E) < +\infty$, by Proposition 11.92, $\nu(E) = \int_E f \, d\mu = \int_E f \, d\mu_E \in \mathcal{Y}$, where $E := (E, \mathcal{B}_E, \mu_E)$ is the measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. \forall pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq B$ with $\bigcup_{i=1}^{\infty} E_i = E$, we have $\sum_{i=1}^{\infty} \nu(E_i) = \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu \leq \sum_{i=1}^{\infty} \int_{E_i} (\mathcal{P} \circ f) \, d\mu = \int_E (\mathcal{P} \circ f) \, d\mu = \bar{\nu}(E) < +\infty$, where the first inequality follows from the fact that $E_i \subseteq E$ and $0 \leq \nu(E_i) \leq \bar{\nu}(E_i) < +\infty$: the first inequality follows from Proposition 11.92; and the second equality follows from Proposition 11.83. Then, $\sum_{i=1}^{\infty} \nu(E_i) =$
11.6. BANACH SPACE VALUED MEASURES

\[ \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu_E = \int_{E} f \, d\mu_E = \int_{E} f \, d\mu = \nu(E), \]

where the second, the third, and the fourth equalities follow from Proposition 11.92.

\[ \forall E \in \mathcal{B} \text{ with } \bar{\nu}(E) < +\infty, \forall n \in \mathbb{Z}_+, \forall \text{ pairwise disjoint } (E_i)_{i=1}^{n} \subseteq \mathcal{B} \text{ with } E = \bigcup_{i=1}^{n} E_i, \int_{E} f \, d\mu \text{ is absolutely integrable over } \mathcal{E}. \sum_{i=1}^{n} \left\| \nu(E_i) \right\| = \sum_{i=1}^{n} \left\| \int_{E_i} f \, d\mu \right\| \leq \sum_{i=1}^{n} \int_{E_i} \mathcal{P} \circ f \, d\mu = \int_{E} \mathcal{P} \circ f \, d\mu = \bar{\nu}(E) < +\infty, \]

where the first equality follows from the fact that \( E_i \subseteq E \) and 0 \( \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) \leq +\infty \); the first inequality follows from Proposition 11.92; and the second equality follows from Proposition 11.83. Hence, \( s_E := \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E=\bigcup_{i=1}^{n} E_i, E \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \left\| \nu(E_i) \right\| \leq \bar{\nu}(E). \]

On the other hand, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), by Proposition 11.95, \( \exists \) a simple function \( \phi : \mathcal{E} \to \mathcal{W} \) such that \( 0 \leq \left| \int_{E} (\mathcal{P} \circ f) \, d\mu_E - \int_{E} (\mathcal{P} \circ \phi) \, d\mu_E \right| < \epsilon/2 \) and \( 0 \leq \int_{E} \mathcal{P} \circ (\phi - f) \, d\mu_E < \epsilon/2 \). Let \( \phi \) admit canonical representation \( \phi = \sum_{i=1}^{n} w_i \chi_{E_i}, \) where \( n \in \mathbb{N}, w_1, \ldots, w_n \in \mathcal{W} \) are distinct and none equals to \( \vartheta_y, E_1, \ldots, E_n \in \mathcal{B}_E \) are nonempty, pairwise disjoint, and \( \mu_E(E_i) < +\infty, \forall i \in \{1, \ldots, n\} \). Then, we have

\[ \bar{\nu}(E) = \int_{E} (\mathcal{P} \circ f) \, d\mu = \int_{E} (\mathcal{P} \circ f) \, d\mu_E < \int_{E} (\mathcal{P} \circ \phi) \, d\mu_E + \epsilon/2 \]

\[ = \sum_{i=1}^{n} \left\| \nu(E_i) \right\| + \sum_{i=1}^{n} \left\| w_i \right\| \mu_E(E_i) + \epsilon/2 - \sum_{i=1}^{n} \left\| \nu(E_i) \right\| \]

\[ = \sum_{i=1}^{n} \left\| \nu(E_i) \right\| + \epsilon/2 + \sum_{i=1}^{n} \left\| \int_{E_i} \phi \, d\mu_E \right\| - \sum_{i=1}^{n} \left\| \int_{E_i} f \, d\mu_E \right\| \]

\[ \leq s_E + \epsilon/2 + \sum_{i=1}^{n} \left\| \int_{E_i} \phi \, d\mu_E - \int_{E_i} f \, d\mu_E \right\| \]

\[ = s_E + \epsilon/2 + \sum_{i=1}^{n} \left\| \int_{E_i} (\phi - f) \, d\mu_E \right\| \]

\[ \leq s_E + \epsilon/2 + \sum_{i=1}^{n} \int_{E_i} \mathcal{P} \circ (\phi - f) \, d\mu_E \]

\[ = s_E + \epsilon/2 + \int_{\bigcup_{i=1}^{n} E_i} \mathcal{P} \circ (\phi - f) \, d\mu_E < s_E + \epsilon \]

where the second equality follows from Proposition 11.92; the third equality follows from Proposition 11.75; the fourth equality follows from Proposition 11.75 and the fact that \( 0 \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) < +\infty \); the fifth equality and the third inequality follow from Proposition 11.92.; the last equality and the last inequality follows from Proposition 11.83. Hence, by the arbitrariness of \( \epsilon \), we have \( \bar{\nu}(E) \leq s_E \). Therefore, \( \bar{\nu}(E) = s_E \).

The above shows that \((X, \mathcal{B}, \nu)\) is a \( Y \)-valued measure space with \( \mathcal{P} \circ \nu = \bar{\nu} \).
Proposition 11.117 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space, $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ be such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) = \langle x \in X \mid \|\mathbf{f}(x)\| \leq m \rangle \in \mathcal{B}$. Clearly, $X = \bigcup_{n=1}^{\infty} X_n = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} X_{m,n}$. By Bounded Convergence Theorem 11.77, we have $\nu(X_{m,n}) = \int_{X_{m,n}} (\mathcal{P} \circ f) \, dm \leq m\mu(X_{m,n}) \leq m\mu(X_n) < +\infty$. Then, $\nu$ is $\sigma$-finite. Hence, $(X, \mathcal{B}, \mu)$ is a $\sigma$-finite $\mathcal{Y}$-valued measure space.

(ii) Let $f$ be absolutely integrable over $\mathcal{X}$. Then, $\nu(X) = \int_{X} (\mathcal{P} \circ f) \, dm < +\infty$. Hence, $(X, \mathcal{B}, \nu)$ is a finite $\mathcal{Y}$-valued (pre-)measure space. On the other hand, let $\nu$ be finite. Then, $\nu(X) = \int_{X} (\mathcal{P} \circ f) \, dm < +\infty$. Hence, $f$ is absolutely integrable over $\mathcal{X}$.

(iii) By (i), $(X, \mathcal{B}, \nu)$ is a $\sigma$-finite. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) = 0$, then $\int_{E} (\mathcal{P} \circ f) \, dm = 0$. By Proposition 11.96, $\mathcal{P} \circ f = 0$ a.e. in $\mathcal{E}$. Since $\mathcal{P} \circ f > 0$ a.e. in $\mathcal{X}$. Then, $\mu(E) = 0$. $\forall A \subseteq E$, by the completeness of $\mathcal{X}$, $A \in \mathcal{B}$. Hence, $(X, \mathcal{B}, \nu)$ is a $\sigma$-finite complete $\mathcal{Y}$-valued measure space.

This completes the proof of the proposition. 

Proposition 11.117 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space, $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ be such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) = \langle x \in X \mid \|\mathbf{f}(x)\| \leq m \rangle \in \mathcal{B}$. Clearly, $X = \bigcup_{n=1}^{\infty} X_n = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} X_{m,n}$. By Bounded Convergence Theorem 11.77, we have $\nu(X_{m,n}) = \int_{X_{m,n}} (\mathcal{P} \circ f) \, dm \leq m\mu(X_{m,n}) \leq m\mu(X_n) < +\infty$. Then, $\nu$ is $\sigma$-finite. Hence, $(X, \mathcal{B}, \mu)$ is a $\sigma$-finite $\mathcal{Y}$-valued measure space.

(ii) Let $f$ be absolutely integrable over $\mathcal{X}$. Then, $\nu(X) = \int_{X} (\mathcal{P} \circ f) \, dm < +\infty$. Hence, $(X, \mathcal{B}, \nu)$ is a finite $\mathcal{Y}$-valued (pre-)measure space. On the other hand, let $\nu$ be finite. Then, $\nu(X) = \int_{X} (\mathcal{P} \circ f) \, dm < +\infty$. Hence, $f$ is absolutely integrable over $\mathcal{X}$.

(iii) By (i), $(X, \mathcal{B}, \nu)$ is a $\sigma$-finite. $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) = 0$, then $\int_{E} (\mathcal{P} \circ f) \, dm = 0$. By Proposition 11.96, $\mathcal{P} \circ f = 0$ a.e. in $\mathcal{E}$. Since $\mathcal{P} \circ f > 0$ a.e. in $\mathcal{X}$. Then, $\mu(E) = 0$. $\forall A \subseteq E$, by the completeness of $\mathcal{X}$, $A \in \mathcal{B}$. Hence, $(X, \mathcal{B}, \nu)$ is a $\sigma$-finite complete $\mathcal{Y}$-valued measure space.

This completes the proof of the proposition. 

Proposition 11.118 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a Banach space, $\forall n \in \mathcal{N}$. $\forall n, m \in \mathcal{N}$, let $(X_n \cap X_m, \mathcal{B}_{n,m}, \mu_{n,m})$ be the finite $\mathcal{Y}$-valued measure subspace of $\mathcal{X}_n$. Assume that $(\mathcal{X}_n)_{n=1}^{\infty}$ is consistent, that is $\mathcal{B}_{n,m} = \mathcal{B}_{m,n}$ and $\mu_{n,m} = \mu_{m,n}$,
Let \( X := \bigcup_{n=1}^{\infty} X_n, \mathcal{B} := \{ B \subseteq X \mid B \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N} \} \), \( \nu : \mathcal{B} \to [0, \infty] \subset \mathbb{R} \) be defined by \( \nu(E) := \sum_{n=1}^{\infty} P \circ \mu_n(E_n), \) where \( E_1 := E \cap X_1 \) and \( E_{n+1} := E \cap (X_n+1 \setminus \bigcup_{i=1}^{n} X_i), \forall n \in \mathbb{N} \), and \( \mu \) be a function from \( \mathcal{B} \) to \( \mathcal{Y} \) defined by \( \mu(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( \nu(E) = +\infty \), and \( \mu(E) := \sum_{n=1}^{\infty} \mu_n(E_n) \in \mathcal{Y}, \forall E \in \mathcal{B} \) with \( \nu(E) < +\infty \), where \( (E_n)_{n=1}^{\infty} \) is defined as above. Then, \( X := (X, \mathcal{B}, \mu) \) is a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space with \( \mathcal{P} \circ \mu = \nu \) and \( \mathcal{X}_n \) is the finite \( \mathcal{Y} \)-valued measure subspace of \( X \), \( \forall n \in \mathbb{N} \), and \( X \) is unique among any \( \mathcal{Y} \)-valued measure spaces on \( X \). The above process is said to be the generation process on \( (X_n)_{n=1}^{\infty} \) for \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space \( X \).

**Proof**

We will first show that \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \). Clearly, \( \emptyset \in \mathcal{B} \) and \( X \in \mathcal{B} \). \( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^{\infty} E_i, \forall n \in \mathbb{N}, E \cap X_n = \bigcup_{i=1}^{\infty} (E_i \cap X_n) \). Note that \( \forall X \in \mathcal{B}, E_i \in \mathcal{B} \implies E_i \cap X_n \in \mathcal{B}_n \). Since \( \mathcal{B}_n \) is a \( \sigma \)-algebra, then \( E \cap X_n \in \mathcal{B}_n \). By the arbitrariness of \( n \), we have \( E \in \mathcal{B}, \forall E \in \mathcal{B}, E \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N} \). Then, \( (X \setminus E) \cap X_n = X_n \setminus E = X_n \setminus (X_n \cap E) \in \mathcal{B}_n \). By the arbitrariness of \( n \), we have \( X \setminus E \in \mathcal{B} \). Hence, \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \). \( \forall n \in \mathbb{N}, \forall m \in \mathbb{N}, X_n \cap X_m \in \mathcal{B}_{n,m} = \mathcal{B}_{n,m} \subseteq \mathcal{B}_m \). By the arbitrariness of \( m \), we have \( X_n \in \mathcal{B} \).

Next, we will show that \( \nu \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \). Clearly, \( \nu(\emptyset) = 0 \). \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^{\infty} E_i, \forall i \in \mathbb{N}, E_i \cap X_n = \bigcup_{j=1}^{\infty} (E_i \cap X_n) \). Let \( B_i := E \cap X_n = \bigcup_{j=1}^{\infty} E_i \cap X_n \in B_i \) and \( B_{n+1} := E \cap (X_n+1 \setminus \bigcup_{j=1}^{n} E_j) = \bigcup_{i=1}^{\infty} (E_i, n+1) \in B_{n+1} \), \( \forall n \in \mathbb{N} \). Let \( B_i := E \cap X_n = \bigcup_{j=1}^{\infty} E_i \cap X_n \in B_i \) and \( B_{n+1} := E \cap (X_n+1 \setminus \bigcup_{j=1}^{n} E_j) = \bigcup_{i=1}^{\infty} (E_i, n+1) \in B_{n+1} \), \( \forall n \in \mathbb{N} \). Then, \( \nu(E) = \sum_{n=1}^{\infty} P \circ \mu_n(B_n) = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} P \circ \mu_n(E_i, n) = \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} P \circ \mu_n(E_i, n) = \sum_{n=1}^{\infty} \nu(E_i) \), where the first equality follows from the definition of \( \nu \); the second equality follows from the fact that \( P \circ \mu_n \) is a measure on \( (X_n, \mathcal{B}_n) \); the third equality follows from the fact that all summands are nonnegative; and the fourth equality follows from the definition of \( \nu \). Hence, \( \nu \) is a measure on \( (X, \mathcal{B}) \). \( \forall n \in \mathbb{N}, X_n \cap X_{n+1} := X_n \cap (X_{n+1} \setminus \bigcup_{j=1}^{n} X_j), \forall n \in \mathbb{N} \). Then, \( \nu(X_n) = \sum_{i=1}^{\infty} P \circ \mu_i(X_n,i) = \sum_{i=1}^{\infty} P \circ \mu_i((X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{n} X_j) + P \circ \mu_1(X_n \cap X_1) = \sum_{i=1}^{\infty} P \circ \mu_{n+1}(X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{n} X_j) + P \circ \mu_n(X_n \cap X_1) = \sum_{i=1}^{\infty} P \circ \mu_n((X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{n} X_j) + P \circ \mu_n(X_n \cap X_1) = \sum_{i=1}^{\infty} P \circ \mu_n(X_n \cap X_1) = \sum_{i=1}^{\infty} P \circ \mu_n(X_n \cap X_1) + P \circ \mu_n(X_n \cap X_1) = P \circ \mu_n(X_n \cap X_1) = P \circ \mu_n(X_n) = +\infty \), where the first equality follows from the definition of \( \nu \); the third equality follows from Proposition 11.115; the fourth equality follows from the consistency assumption; the fifth equality follows from Proposition 11.115; and the last equality follows from the fact that \( P \circ \mu_n \) is a measure on \( (X, \mathcal{B}_n) \). Hence, \( \nu \) is \( \sigma \)-finite.

Next, we will show that \( (X, \mathcal{B}, \mu) \) is a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space with \( P \circ \mu = \nu \). Clearly, \( \mu(\emptyset) = 0 \). \( \forall E \in \mathcal{B} \) with \( \nu(E) = +\infty \), \( \mu(E) \) is undefined. \( \forall E \in \mathcal{B} \) with \( \nu(E) < +\infty \), let \( (B_n)_{n=1}^{\infty} \) be as defined in the previous paragraph. Then, \( +\infty > \nu(E) = \sum_{n=1}^{\infty} P \circ \mu_n(B_n) \geq \sum_{n=1}^{\infty} \| \mu_n(B_n) \| \). This implies that \( \mu(E) = \sum_{n=1}^{\infty} \mu_n(B_n) \in \mathcal{Y} \) by Propo-
situation 7.27. \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq B \) with \( E = \bigcup_{i=1}^{\infty} E_i, \forall i \in \mathbb{N} \), let \( (E_{i,n})_{i=1}^{\infty} \) be as defined in the previous paragraph. Then, we have \( \sum_{i=1}^{\infty} \| \mu(E_i) \| = \sum_{i=1}^{\infty} \| \sum_{n=1}^{\infty} \mu_n(E_{i,n}) \| \leq \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \| \mu(E_{i,n}) \| \leq \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} P \circ \mu_n(E_{i,n}) = \sum_{i=1}^{\infty} \nu(E_i) = \nu(E) < +\infty \), where the first equality follows from the definition of \( \mu \); the second inequality follows from the fact that \( X_n \) is a finite \( \mathcal{Y} \)-valued measure space; the second equality follows from the definition of \( \nu \); and the third equality follows from the fact that \( \nu \) is a measure. We also have \( \sum_{i=1}^{\infty} \| \mu(E_i) \| = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \| \mu_n(E_{i,n}) \| = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \| \mu(B_{i,n,m}) \| = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu(B_{i,n,m}) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu_n(B_{i,n,m}) = \mu(B_i) \), where the first equality follows from the definition of \( \mu \); the second equality follows from the fact that the double summation is absolutely summable; the third equality follows from the fact that \( \mu_n \) is a finite \( \mathcal{Y} \)-valued measure; and the last equality follows from the definition of \( \mu \). Hence, (i) - (iii) of Definition 11.108 are satisfied.

\[
\forall E \in B \text{ with } \nu(E) < +\infty, \text{ let } (B_{n,i})_{i=1}^{\infty} \subseteq B \text{ with } E = \bigcup_{i=1}^{m} E_i, \forall i \in \{1, \ldots, m\}, \text{ be as defined in the previous paragraph. Then, } \sum_{i=1}^{m} \| \mu(E_i) \| = \sum_{i=1}^{m} \sum_{n=1}^{\infty} \| \mu_n(E_{i,n}) \| \leq \sum_{i=1}^{m} \sum_{n=1}^{\infty} P \circ \mu_n(E_{i,n}) = \sum_{i=1}^{m} \nu(E_i) = \nu(E) < +\infty. \]

Hence,

\[
s_E := \sup_{m \in \mathbb{Z}^+} \sum_{i=1}^{m} \| \mu(E_i) \| \leq \nu(E).
\]

On the other hand, \( \nu(E) = \sum_{n=1}^{\infty} P \circ \mu_n(B_n) < +\infty \), where \( B_1, B_2, \ldots \) are pairwise disjoint. Since \((X_n, \mathcal{B}_n, \mu_n)\) is a finite \( \mathcal{Y} \)-valued measure space, then, \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \exists m_n \in \mathbb{Z}^+, \exists \text{ pairwise disjoint } (B_{n,i})_{i=1}^{\infty} \subseteq \mathcal{B}_n \text{ such that } P \circ \mu_n(B_n) - 2^{-n} \epsilon < \sum_{i=1}^{m} \| \mu_n(B_{n,i}) \| \leq \sum_{i=1}^{m} \| \mu_n(B_{n,i}) \| \leq P \circ \mu_n(B_n), \forall i \in \{1, \ldots, m_n\}. \) Then, \( \nu(E) - \epsilon = \sum_{n=1}^{\infty} (P \circ \mu_n(B_n) - 2^{-n} \epsilon < \sum_{n=1}^{\infty} \sum_{i=1}^{m_n} \| \mu_n(B_{n,i}) \| = \sum_{n=1}^{\infty} \sum_{i=1}^{m_n} \mu_n(B_{n,i}) = \sum_{n=1}^{\infty} \mu_n(B_n) = \mu(B_n) \), where the first equality follows from the fact that \( \mu_n \) is a finite \( \mathcal{Y} \)-valued measure; the second equality follows from the fact that \( \mu_n \) is a finite \( \mathcal{Y} \)-valued measure; and the third equality follows from the consistency of the definition of \( \mu \).
11.6. BANACH SPACE VALUED MEASURES

assumption; the third equality follows from the fact that $\mu_{n,i} = \mu_{n}|_{B_{n,i}}$; the fourth equality follows from countable additivity of $\mathcal{P} \circ \mu$; and the first inequality follows from Proposition 11.4. Then, $Y \ni \mu(E) = \sum_{i=1}^{\infty} \mu_i(B_i) = \sum_{i=1}^{\infty} \mu_{n,i}(B_i) = \sum_{i=1}^{\infty} \mu_n(B_i) = \mu_n(E)$, where the first equality follows from the definition of $\mu$; the second equality follows from the fact that $\mu_{n,i} = \mu_n|_{B_{n,i}}$; the third equality follows from the consistency assumption; the fourth equality follows from the fact that $\mu_{n,i} = \mu_n|_{B_{n,i}}$; and the last equality follows from the fact that $\mu_n$ is a finite $Y$-valued measure. Hence, $\mu_n = \mu |_{B_n}$. This shows that $X_n$ is the subspace of $X$. By Proposition 11.117, $X$ is the unique $Y$-valued measure space on the set $\mathcal{X}$ such that $X_n$ is the finite $Y$-valued measure subspace of $X$, $\forall n \in \mathbb{N}$.

This completes the proof of the proposition. $\square$

In the application of Proposition 11.118, if the sequence $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Then, the consistency requirement is automatically satisfied as long as $X_n$ is a $Y$-valued measure space, $\forall n \in \mathbb{N}$.

Definition 11.119 Let $X := (X, \mathcal{B}, \mu)$ be a finite $Y$-valued measure space, where $Y$ is a normed linear space over $\mathbb{K}$, $\mathcal{B}$ is a normed linear space over $\mathbb{K}$, $\mathcal{W} \subseteq \mathcal{B}(Y, Z)$ be a subspace, $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, and $\mathcal{I}(\mathcal{W}) := (\mathcal{R}(\mathcal{W}), \vartriangleleft)$ be the integration system on $\mathcal{W}$. $\forall R = \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda\} \in \mathcal{R}(\mathcal{W})$, define $F_R := \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{in}}(U_\alpha)) \in Z$. This defines a net $(F_R)_{R \in \mathcal{I}(\mathcal{W})}$; $f$ is said to be integrable over $X$ if the net admits a limit in $Z$. In this case, $\lim_{R \in \mathcal{I}(\mathcal{W})} F_R \in Z$ is said to be the integral of $f$ over $X$ and denoted by $\int_X f \, d\mu$ or $\int_X f(x) \, d\mu(x)$. When $Z = \mathbb{R}$, we will denote $\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{R \in \mathcal{I}(\mathcal{W})} F_R \in \mathbb{R}$, whenever the limit exists.

Definition 11.120 Let $X := (X, \mathcal{B}, \mu)$ be a $Y$-valued measure space with $\mathcal{P} \circ \mu(X) = +\infty$, where $Y$ is a normed linear space over $\mathcal{K}$, $\mathcal{Z}$ be a normed linear space over $\mathcal{K}$, $\hat{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ be the total variation, and $f : X \to \mathcal{W} := \mathcal{B}(Y, Z)$ be $\mathcal{B}$-measurable. Define $\mathcal{M}(X)$ to be the directed system, $\mathcal{M}(\hat{X})$, of subsets of $X$ as defined in Definition 11.71. $\forall A \in \mathcal{M}(X)$, let $(\mathcal{A}, \mathcal{B}_A, \mu_A)$ be the finite $Y$-valued measure subspace of $\mathcal{X}$ and define $F_A := \int_A f \, d\mu_A \in Z$, if the integral exists. This defines a net $(F_A)_{A \in \mathcal{M}(X)}$ when all of the integrals involved exist. $f$ is said to be integrable over $X$ if the net is well-defined (i.e., all integrals involved exist) and admits a limit in $Z$. In this case, $\lim_{A \in \mathcal{M}(X)} F_A \in Z$ is said to be the integral of $f$ over $X$ and denoted by $\int_X f \, d\mu$ or $\int_X f(x) \, d\mu(x)$. When $Z = \mathbb{R}$, we will allow the net $(F_A)_{A \in \mathcal{M}(X)} \subseteq \mathbb{R}$ and denote $\int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{A \in \mathcal{M}(X)} F_A \in \mathbb{R}$, whenever the limit exists.

Proposition 11.121 Let $X := (X, \mathcal{B}, \mu)$ be a finite $Y$-valued measure space, where $Y$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathcal{K}$, $f : X \to \mathcal{W} := \mathcal{B}(Y, Z)$ be $\mathcal{B}$-measurable, and $\hat{X} := (X, \mathcal{B}, \bar{\mu})$ be a finite $Y$-valued measure space that satisfies $\mathcal{B} \subseteq \mathcal{B}$ and $\mu = \bar{\mu}$. Then, $\int_X f \, d\mu = \int_X f \, d\bar{\mu}$, whenever one of the integrals exists.
By Propositions 11.121 and 11.114, the integrals exists. This completes the proof of the proposition.

Lemma 11.122 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathfrak{y}$-valued measure space, where $\mathfrak{y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : X \to W := \mathcal{B}(\mathfrak{y}, \mathcal{Z})$ be $\mathcal{B}$-measurable, $\hat{X} \in \mathcal{B}$, and $\hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \hat{\mu})$ be the finite $\mathfrak{y}$-valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Assume that $\mathcal{P} \circ \mu(X \setminus \hat{X}) = 0$. Then, \( \int_{\hat{X}} f \, d\hat{\mu} = \int_X f|_{\hat{X}} \, d\hat{\mu} \), whenever one of the integrals exists.

Proof Let $\mathcal{J}(\mathcal{W})$ be the integration system on $\mathcal{W}$ as defined in Proposition 11.69, $(\mathcal{F}_R)_{R \in \mathbb{R}^+}$ be the net for $\int_X f \, d\mu$, and $(\hat{\mathcal{F}}_R)_{R \in \mathbb{R}^+}$ be the net for $\int_{\hat{X}} f \, d\hat{\mu}$, as defined in Definition 11.119. For $\forall R := \{(w_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in \mathcal{J}(\mathcal{W})$, we have $f_{\text{inv}}(U_\alpha) \in \mathfrak{B} \subseteq \mathcal{B}$ since $f$ is $\mathcal{B}$-measurable, and \( \mathcal{P} \circ \mu(f_{\text{inv}}(U_\alpha)) = \mathcal{P} \circ \mu(f(U_\alpha)) \in \mathfrak{Y} \), since $\mu = \hat{\mu}|_B$. Then, \( \mathcal{F}_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} w_\alpha \hat{\mu}(f_{\text{inv}}(U_\alpha)) = \hat{\mathcal{F}}_R \in \mathcal{Z} \). Hence, \( \int_X f \, d\mu = \lim_{R \to \infty} \mathcal{F}_R \). \( \mathcal{F}_R = \lim_{R \to \infty} \hat{\mathcal{F}}_R = \int_{\hat{X}} f \, d\hat{\mu}, \) whenever one of the integrals exists. This completes the proof of the proposition.

Proposition 11.123 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathfrak{y}$-valued measure space with $\mathcal{P} \circ \mu(X) = +\infty$, where $\mathfrak{y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $f : X \to W := \mathcal{B}(\mathfrak{y}, \mathcal{Z})$ be $\mathcal{B}$-measurable, and $\hat{\mathcal{X}} := (X, \hat{\mathcal{B}}, \hat{\mu})$ be the completion of $\mathcal{X}$. Then, \( \int_X f \, d\mu = \int_{\hat{X}} f \, d\hat{\mu} \), whenever one of the integrals exists.

Proof Let $\hat{\mathcal{X}} := (X, \hat{\mathcal{B}})$ and $\hat{\mathcal{X}} := (X, \hat{\mathcal{B}}, \mathcal{P} \circ \hat{\mu} := \nu)$. Then, $\hat{\mathcal{X}}$ and $\hat{\mathcal{X}}$ are measure spaces and $\hat{\mathcal{X}}$ is the completion of $\hat{\mathcal{X}}$ by Proposition 11.114. Let $\hat{\mathcal{F}}_\mathcal{A}(\hat{\mathcal{X}})$ be the net for $\int_{\hat{X}} f \, d\hat{\mu}$ and $(\hat{\mathcal{F}}_\mathcal{A})_{\hat{\mathcal{A}} \in \mathcal{M}(\hat{\mathcal{X}})}$ be the net for $\int_{\hat{X}} f \, d\hat{\mu}$ as defined in Definition 11.120. Clearly, $\mathcal{M}(\hat{\mathcal{X}}) \subseteq \mathcal{M}(\hat{\mathcal{X}})$, since $\hat{\mathcal{X}}$ is the completion of $\mathcal{X}$. We will distinguish two exhaustive cases: Case 1: $\int_{\hat{X}} f \, d\hat{\mu}$ exists; Case 2: $\int_{\hat{X}} f \, d\hat{\mu}$ exists.

Case 1: $\int_{\hat{X}} f \, d\hat{\mu}$ exists. Then, the net $(\hat{\mathcal{F}}_\mathcal{A})_{\hat{\mathcal{A}} \in \mathcal{M}(\hat{\mathcal{X}})}$ is well defined. For $\forall A \in \mathcal{M}(\mathcal{X})$, we have $A \in \mathcal{B}$ and $\nu(A) < +\infty$. Then, $A \in \mathcal{M}(\mathcal{X})$. By Propositions 11.121 and 11.114, $\hat{\mathcal{F}}_\mathcal{A} = \int_{\hat{X}} f|_A \, d\hat{\mu}_A = \int_{\hat{X}} f|_A \, d\hat{\mu}_A = F_A$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite $\mathfrak{y}$-valued measure subspace of $\mathcal{X}$ and
(A, B, A,  B) is the finite complete Y-valued measure subspace of X. Then, 

\( \mu \) and \( \nu \) are measure subspaces of \( \mathcal{M}(X) \) and \( \mathcal{M}(X) \), respectively. By Proposition 11.114 and its proof, there exist \( \tilde{A}, \tilde{B}, \tilde{C} \subset X \) with \( \tilde{A} = A \cup \tilde{B}, \tilde{B} \subset \tilde{C}, \nu(\tilde{C}) = 0 \), and \( \nu(A) = \nu(B) \). Then, \( \tilde{A} \subseteq \tilde{B} \cup \tilde{C} \subseteq A \cup B \) and \( \nu(A) \leq \nu(B) + \nu(\tilde{C}) = \nu(B) = \nu(A) < +\infty \). This shows that \( A \in \mathcal{M}(X) \) and \( A \subseteq A \). Hence, \( (F_A)_{A \in \mathcal{M}(X)} \) is a subset of \( (\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})} \). By Proposition 3.70, \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(X)} F_A = \lim_{A \in \mathcal{M}(\tilde{X})} \tilde{F}_A = \int_X f \, d\tilde{\mu} \).

Case 2: \( \int_X f \, d\mu \) exists. Then, the net \( (F_A)_{A \in \mathcal{M}(X)} \) is well defined. \( \forall A \in \mathcal{M}(\tilde{X}), \) we have \( \tilde{A} \in \mathcal{B} \) with \( \nu(\tilde{A}) < +\infty \). By Proposition 11.114 and its proof, there exist \( \tilde{A}, \tilde{B}, \tilde{C} \subset X \) with \( \tilde{A} = \tilde{A} \cup \tilde{B}, \tilde{B} \subset \tilde{C}, \nu(\tilde{C}) = 0 \), and \( \nu(\tilde{A}) = \nu(B) \). Let \( \tilde{B} = A \in \mathcal{B} \). Then, \( 0 \leq \nu(\tilde{A} \setminus \tilde{A}) = \nu(\tilde{A} \setminus \tilde{A}) = 0 \) and \( \nu(A) = \nu(\tilde{A}) < +\infty \). This shows that \( A \in \mathcal{M}(\tilde{X}) \) and \( A \subseteq A \). Let \( \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite Y-valued measure subspace of \( \mathcal{M}(\tilde{X}), \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite complete Y-valued measure subspace of \( \tilde{X} \), and \( \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite complete Y-valued measure subspace of \( \tilde{A} \). Since \( \tilde{X} \) is the completion of \( X \), then, by Proposition 11.115, \( \tilde{A} \) is the completion of \( A \). By Proposition 11.121, we have \( \int_A f_A \, d\mu_A = \int_A f_A \, d\mu_A = F_A \). By Lemma 11.122, we have \( \tilde{F}_A = \int_A f_A \, d\tilde{\mu}_A = \int_A f_A \, d\tilde{\mu}_A = F_A \). Hence, \( \tilde{F}_A = F_A \). Then, the net \( (\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})} \) is well defined.

Fix any open set \( U \subseteq \mathbb{R} \) if \( Z = \mathbb{R} \) or \( U \subseteq \mathbb{Z} \) if \( Z \neq \mathbb{R} \) with \( \int_X f \, d\mu \in U \). Since \( \lim_{A \in \mathcal{M}(X)} F_A = \int_X f \, d\mu \in U \), then \( \exists A_0 \in \mathcal{M}(X) \) such that \( \forall A \in \mathcal{M}(X) \) with \( A_0 \subseteq A \), we have \( F_A \in U \). Note that \( A_0 \in \mathcal{B} \subseteq \mathcal{B} \) and \( \nu(A_0) = \nu(A_0) < +\infty \). Then, \( A_0 \in \mathcal{M}(\tilde{X}) \). By Proposition 11.114 and its proof, \( \exists \tilde{A}, \tilde{B}, \tilde{C} \subset X \) with \( \tilde{A} = A \cup \tilde{B}, \tilde{B} \subset \tilde{C}, \nu(\tilde{C}) = 0 \), and \( \nu(\tilde{A}) = \nu(B) \). Let \( \tilde{A} := A \cup \tilde{B} \in \mathcal{B} \). Then, \( \tilde{A} \subseteq A \subseteq \tilde{A} \), \( 0 \leq \nu(\tilde{A} \setminus \tilde{A}) = \nu(\tilde{A} \setminus \tilde{A}) = 0 \) and \( \nu(\tilde{A}) = \nu(B) \). This shows that \( A \in \mathcal{M}(\tilde{X}) \). Let \( \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite Y-valued measure subspace of \( \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite complete Y-valued measure subspace of \( \tilde{X} \), and \( \mathcal{A} := (A, \mathcal{B}_A, \tilde{\mu}_A) \) be the finite complete Y-valued measure subspace of \( \tilde{A} \). Since \( \tilde{X} \) is the completion of \( X \), then, by Proposition 11.115, \( \tilde{A} \) is the completion of \( A \). By Proposition 11.121, we have \( \int_A f_A \, d\mu_A = \int_A f_A \, d\mu_A = F_A \). By Lemma 11.122, we have \( \tilde{F}_A = \int_A f_A \, d\tilde{\mu}_A = \int_A f_A \, d\tilde{\mu}_A = F_A \in U \). Therefore, we have \( \int_X f \, d\tilde{\mu} = \lim_{A \in \mathcal{M}(\tilde{X})} \tilde{F}_A = \int_X f \, d\tilde{\mu} \).

Hence, in both cases, we have \( \int_X f \, d\mu = \int_X f \, d\tilde{\mu} \). This completes the proof of the proposition. \( \square \)

**Definition 11.124** Let \( X := (X, \mathcal{B}, \mu) \) be a Y-valued measure space, where Y is a normed linear space over \( K \), and \( Z \) be a normed linear space over \( K \). \( \phi : X \to Z \) is said to be a simple function if \( \exists n \in \mathbb{Z}_+, \exists z_1, \ldots, z_n \in \mathbb{Z} \).
\( \mathbb{Z} \), and \( \exists A_1, \ldots, A_n \in B \) with \( P \circ \mu(A_i) < +\infty, i = 1, \ldots, n \), such that \( \phi(x) = \sum_{i=1}^{n} z_i \chi_{A_i}(x), \forall x \in X \). We will say that a simple function \( \phi \) is in canonical representation if \( z_1, \ldots, z_n \) are distinct and none equals to \( \vartheta \), and \( A_1, \ldots, A_n \) are nonempty and pairwise disjoint.

Clearly, every simple function admits a unique canonical representation.

**Proposition 11.125** Let \( X := (X, B, \mu) \) be a \( \mathfrak{Y} \)-valued measure space, where \( \mathfrak{Y} \) is a normed linear space over \( \mathbb{K} \), \( \mathbb{Z} \) be a normed linear space over \( \mathbb{K} \), \( \phi : X \to B(\mathfrak{Y}, \mathbb{Z}) =: W \) be a simple function in canonical representation, i.e., \( \exists z_i \in \mathbb{Z}_+, \exists w_i, w_n \in W \), which are distinct and none equals to \( \vartheta \), \( \phi(x) = \sum_{i=1}^{n} w_i \chi_{A_i}(x), \forall x \in X \). Then, \( \int_X \phi \, d\mu = \sum_{i=1}^{n} w_i \mu(A_i) =: I_1 \quad \forall \mathbb{Z} \) and \( 0 \leq \| \int_X \phi \, d\mu \| \leq \int_X P \circ \phi \, dP \circ \mu < +\infty. \)

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( P \circ \mu (X) < +\infty \); Case 2: \( P \circ \mu (X) = +\infty \).

Case 1: \( P \circ \mu (X) < +\infty \). Let \( w_{n+1} := \vartheta_W \) and \( A_{n+1} := X \setminus (\bigcup_{i=1}^{n} A_i) \in B \). Let \( \epsilon_0 := \min \{ 1, \min_{1 \leq i < j \leq n+1} \| w_i - w_j \| \} > 0 \), by the assumption. Let \( (\Phi_R)_{R \in \mathbb{B}(W)} \) be the net as defined in Definition 11.119 for \( \int_X \phi \, d\mu \).

\( \forall \varepsilon \in (0, \epsilon_0/2) \subset \mathbb{R} \), let \( U_i := B_W ( w_i, \varepsilon ) \in \mathbb{B}(W) \), \( i = 1, \ldots, n+1 \), which are clearly pairwise disjoint and nonempty. We will distinguish two exhaustive and mutually exclusive subcases: Case 1a: \( W = \bigcup_{i=1}^{n+1} U_i \); Case 1b: \( W \supset \bigcup_{i=1}^{n+1} U_i \).

Case 1a: \( W = \bigcup_{i=1}^{n+1} U_i \). Then, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( W = B_W ( \vartheta_W, \delta ) \). Then, \( \vartheta_W \) must be the trivial normed linear space, i.e., \( \mathbb{W} \) is a singleton set containing \( \vartheta_W \). There is a unique representation of \( \mathbb{W} \), which is \( R_0 := \{ (\vartheta_W, W) \} \). Then, \( \Phi_{R_0} = \vartheta_Z \). Clearly, we must have \( n = 0 \), since \( w_1, \ldots, w_n \in \mathfrak{Y} \) are distinct and none equals to \( \vartheta \). Then, \( I = \vartheta_Z \). Hence, \( \forall R \in \mathbb{J}(W) \) with \( R \preceq R_0 \), \( \| \Phi_R - I \| = \| \Phi_{R_0} - I \| = 0 \).

Case 1b: \( W \supset \bigcup_{i=1}^{n+1} U_i \). Let \( U_{n+2} = W \setminus (\bigcup_{i=1}^{n+1} U_i) \in \mathbb{B}(W) \), which is nonempty. Let \( w_{n+2} \in U_{n+2} \) be such that \( \| w_{n+2} \| < \inf_{w \in U_{n+2}} \| w \| + 1 \). Let \( R_0 := \{ (w_i, U_i) \mid i = 1, \ldots, n+2 \} \). It is easy to check that \( R_0 \in \mathbb{J}(W) \).

\( \forall \tilde{R} \in \mathbb{J}(W) \) with \( \tilde{R} \preceq \tilde{R} \), then \( \tilde{R} = \{ (\tilde{w}_\alpha, \tilde{U}_\alpha) \mid \alpha \in \Lambda \} \), where \( \Lambda \) is a finite index set, \( \{ \tilde{U}_\alpha \}_{\alpha \in \Lambda} \subseteq \mathbb{B}(W) \) are nonempty and pairwise disjoint, \( \bigcup_{\alpha \in \Lambda} \tilde{U}_\alpha = W \), \( \tilde{w}_\alpha \in \tilde{U}_\alpha \), and \( \| \tilde{w}_\alpha \| < \inf_{w \in \tilde{U}_\alpha} \| w \| + 1, \forall \alpha \in \Lambda \).

\( \forall i \in \{ 1, \ldots, n+2 \} \), by \( R_0 \preceq \tilde{R} \), \( \exists \alpha_i \in \Lambda \) such that \( w_i = \tilde{w}_{\alpha_i} \), and by \( \tilde{R} \in \mathbb{J}(W) \), \( \exists \tilde{w}_{\alpha_i} \in \tilde{U}_{\alpha_i} \) such that \( w_i \in \tilde{U}_{\alpha_i} \). Since \( \{ \tilde{U}_\alpha \}_{\alpha \in \Lambda} \subseteq \mathbb{B}(W) \) are nonempty and pairwise disjoint and \( \tilde{w}_\alpha \in \tilde{U}_\alpha \), \( \forall \alpha \in \Lambda \), then \( \alpha_{i+1} = \alpha_i \).

Again by \( R_0 \preceq \tilde{R} \), we have \( \tilde{U}_{\alpha_i} \subseteq U_i \). Hence, \( \alpha_1, \ldots, \alpha_{n+2} \) are distinct. Let \( \Lambda := \{ \alpha_1, \ldots, \alpha_{n+2} \} \). \( \forall \alpha \in \Lambda \setminus \Lambda, \phi_{\text{inv}}(U_{\alpha}) = 0 \). Then, \( \Phi_{\tilde{R}} = \sum_{\alpha \in \Lambda} \tilde{w}_\alpha \mu(\phi_{\text{inv}}(U_{\alpha})) = \sum_{i=1}^{n+2} \tilde{w}_{\alpha_i} \mu(\phi_{\text{inv}}(U_{\alpha_i})) = \sum_{i=1}^{n+2} w_i \mu(\phi_{\text{inv}}(U_{\alpha_i})). \)
Note that $\forall i \in \{1, \ldots, n+1\}$, $\phi_{\text{inv}}(\tilde{U}_\alpha) = A_i$ and $\mu(A_i) \in Y$. Furthermore, $\phi_{\text{inv}}(\tilde{U}_{\alpha+2}) = 0$. Hence, $\Phi_R = \sum_{i=1}^n w_i \mu(A_i) = I \in Z$. Then, we have $\|\Phi_R - I\| = 0$.

In both subcases, we have $\exists R_0 \in J(W), \forall R \in J(W)$ with $R_0 \leq R$, we have $\|\Phi_R - I\| < \epsilon$. Hence, $\int_X \phi \, d\mu = \lim_{R \in J(W)} \Phi_R = I$. Note that $0 \leq \|\int_X \phi \, d\mu\| = \sum_{i=1}^n \|w_i\| \|\mu(A_i)\| \leq \sum_{i=1}^n \|w_i\| \|\mathcal{P} \circ \mu(A_i)\| = \int_X \mathcal{P} \circ \phi \, d\mu < \infty$, where the second inequality follows from Proposition 7.64; the third inequality follows from Definition 11.108; the equality follows from what we have proved in this proposition; and the last inequality follows from Definition 11.124.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $\forall A \in \mathcal{M}(X)$, $\mathcal{P} \circ \mu(A) < +\infty$. $\phi|_{\mathcal{A}} : \mathcal{A} \to \mathcal{W}$ is a simple function. By Case 1, $\Phi_A = \int_A \phi|_{\mathcal{A}} \, d\mu_A \in \mathcal{Z}$ is well-defined, where $(\mathcal{A}, \mathcal{B}_A, \mu)$ is the finite $Y$-valued measure subspace of $\mathcal{X}$.

Hence, the net $(\Phi_A)_{A \in \mathcal{M}(X)}$ is well-defined. Take $A_0 := \bigcup_{i=1}^n A_i = \mathcal{B}$.

Then, $\mathcal{P} \circ \mu(A_0) = \sum_{i=1}^n \mathcal{P} \circ \mu(A_i) < +\infty$. Therefore, $A_0 \in \mathcal{M}(X)$.

Thus, $\Phi_A = \sum_{i=1}^n \|w_i\| \|\mu(A_i)\| = \int_X \mathcal{P} \circ \phi \, d\mu < +\infty$, where the second inequality follows from Proposition 7.64; the third inequality follows from Definition 11.108; the equality follows from what we have proved in this proposition; and the last inequality follows from Definition 11.124.

This completes the proof of the proposition.

Lemma 11.126 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $Y$-valued measure space, where $Y$ is a normed linear space over $\mathbb{K}$, $Z$ be a Banach space over $\mathbb{K}$, $W$ be a separable subspace of $\mathcal{B}(Y, \mathbb{Z})$, $\phi_n : X \to \mathcal{W}$ be a simple function, $\forall n \in \mathbb{N}$, $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $g_n : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, and $g : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\mathcal{W} := \text{span}(\{z \in \mathbb{Z} \mid \exists y \in Y, \exists w \in \mathcal{W} \exists \cdot z = wy\}) \subset \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$. Assume that

(i) $\lim_{n \in \mathbb{N}} \phi_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$;

(ii) $\|\phi_n(x)\| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}$;

(iii) $\int_X g_n \, d\mathcal{P} \circ \mu \in \mathbb{R}$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \int_X g_n \, d\mathcal{P} \circ \mu = \int_X g \, d\mathcal{P} \circ \mu \in \mathbb{R}$.

Then, $f$ is integrable over $\mathcal{X}$, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu \in \mathcal{W} \subseteq \mathcal{Z}$, and $0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty$.

Proof Let $\tilde{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu)$ be the finite measure space as defined in Definition 11.108. Let $E := \{x \in X \mid (\phi_n(x))_{n=1}^\infty \text{ does not converge to } f(x) \text{ or } (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x)\}$. Then, by (i), $E \in \mathcal{B}$ and...
\( \mathcal{P} \circ \mu (E) = 0 \). Let \((\mathcal{F} \cap \mathcal{B})_{R \in \mathcal{X}(\mathcal{W})}\) be the net for \( \int_{\mathcal{X}} f \, d\mu\) as defined in Definition 11.119.

\[ \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \quad \text{(i), (iii), and Lemma 11.85, } \exists \delta \in (0, \varepsilon/6) \subset \mathbb{R} \]

and \( \exists \bar{n} \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \) with \( \mathcal{P} \circ \mu (A) < \delta\), \( \forall n \in \mathbb{N} \) with \( n \geq \bar{n}\), we have \( 0 \leq \int_{A} g \, d\mathcal{P} \circ \mu < \varepsilon/6 \) and \( 0 \leq \int_{A} g_{n} \, d\mathcal{P} \circ \mu < \varepsilon/6 \). By Egoroff’s Theorem 11.55, \( \exists \bar{E}_{1} \in \mathcal{B} \) with \( \mathcal{P} \circ \mu (\bar{E}_{1}) < \delta \) such that \( \left( \phi_{n} \mid_{\mathcal{X} \setminus \bar{E}_{1}} \right)_{n=1}^{\infty} \)
converges uniformly to \( f \mid_{\mathcal{X} \setminus \bar{E}_{1}} \). Let \( E_{1} := E \cup \bar{E}_{1} \). Then, \( E_{1} \in \mathcal{B} \) with \( \mathcal{P} \circ \mu (E_{1}) = \mathcal{P} \circ \mu (\bar{E}_{1}) < \delta \) such that \( \left( \phi_{n} \mid_{\mathcal{X} \setminus E_{1}} \right)_{n=1}^{\infty} \)
converges uniformly to \( f \mid_{\mathcal{X} \setminus E_{1}} \). Then, \( \exists n_{1} \in \mathbb{N} \) with \( n_{1} \geq \bar{n} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{1} \), \( \forall x \in \mathcal{X} \setminus E_{1}, \parallel \phi_{n}(x) - f(x) \parallel < \frac{\varepsilon}{6 \sup_{x \in \phi_{n}(X) + 1} x} =: \varepsilon \in (0, \infty) \subset \mathbb{R} \). Fix any \( n \geq n_{1} \). Let \( \phi_{n} \) admit the canonical representation \( \phi_{n} = \sum_{i=1}^{\bar{n}} w_{i} \chi_{A_{i}} \), where \( \bar{n} \in \mathbb{Z}_{+}, w_{1}, ..., w_{\bar{n}} \in \mathcal{W} \) are distinct and none equals to \( \vartheta \mathcal{W} \), and \( A_{1}, ..., A_{\bar{n}} \in \mathcal{B} \) are nonempty, pairwise disjoint, and \( \mathcal{P} \circ \mu (A_{i}) < +\infty, i = 1, ..., \bar{n} \). Let \( w_{\bar{n}+1} := \vartheta \mathcal{W} \) and \( A_{\bar{n}+1} := X \setminus (\bigcup_{i=1}^{\bar{n}} A_{i}) \in \mathcal{B} \). Then, \( X = \bigcup_{i=1}^{\bar{n}+1} A_{i} \) and the sets in the union are pairwise disjoint and of finite total variation. Define \( \bar{V}_{1} := \mathcal{B}_{\mathcal{W}} (w_{1}, \varepsilon) \), \( i = 1, ..., \bar{n} + 1 \). Define \( V_{i} := \bar{V}_{1} \in \mathcal{B}_{\mathcal{W}} (\mathcal{W}) \), \( V_{i+1} := \bar{V}_{i+1} \setminus (\bigcup_{j=1}^{i} \bar{V}_{j}) \in \mathcal{B}_{\mathcal{W}} (\mathcal{W}) \), \( i = 1, ..., \bar{n} \). Let \( V_{\bar{n}+2} := \mathcal{W} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_{j}) \in \mathcal{B}_{\mathcal{W}} (\mathcal{W}) \). Clearly, we may form a representation \( \hat{R} := \{(\hat{w}_{i}, V_{i}) \mid i = 1, ..., \bar{n} + 2, \hat{V}_{i} \neq \emptyset \} \in \mathcal{J} (\mathcal{W}) \), where \( \hat{w}_{i} \in V_{i} \) are any vectors that satisfy the assumption of Proposition 11.69. \( \forall \hat{R} \in \mathcal{J} (\mathcal{W}) \) with \( \mathcal{R} \leq \hat{R} \). Let \( R := \bigcup_{i \in \Gamma} (\hat{w}_{i}, \lambda_{i}) \) such that \( \hat{U}_{\lambda} \subseteq V_{\lambda} \). Define \( \lambda \hat{R} := \{ \gamma \in \Gamma \mid \lambda_{\gamma} = \bar{n} + 2 \} \). Let \( A_{\gamma} := A_{\gamma} \cap f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap (X \setminus E_{1}) \in \mathcal{B}, i = 1, ..., \bar{n} \). Let \( \hat{A}_{\gamma} := f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap E_{1} \).

Claim 11.126.1 \( \forall \gamma \in \Gamma, f_{\mathcal{W}}(\hat{U}_{\gamma}) \subseteq E_{1}; \forall \gamma \in \Gamma_{x} := \{ \gamma \in \Gamma \mid \hat{A}_{\gamma} \neq \emptyset \}, \parallel \hat{w}_{\gamma} \mu (\hat{A}_{\gamma}) \parallel \leq (1 + \inf_{x \in \hat{A}_{\gamma}} g(x)) \mathcal{P} \circ \mu (\hat{A}_{\gamma}); \text{ and } \sum_{\gamma \in \Gamma} \hat{w}_{\gamma} \mu (f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap E_{1}) < \varepsilon/3. \)

Proof of claim: Fix any \( \gamma \in \Gamma \). \( \hat{U}_{\gamma} \subseteq V_{\bar{n}+2} = \mathcal{W} \setminus (\bigcup_{j=1}^{\bar{n}+1} \bar{V}_{j}) \). Then, \( \mathcal{B} \ni f_{\mathcal{W}}(\hat{U}_{\gamma}) \subseteq E_{1} \).

\( \forall \gamma \in \Gamma_{x}, \hat{A}_{\gamma} := f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap (E_{1} \setminus \mathcal{E}) \neq \emptyset. \forall x \in \hat{A}_{\gamma}, \) we have \( f(x) \in \hat{U}_{\gamma}, \lim_{m \in \mathbb{N}} \phi_{m}(x) = f(x), \parallel \phi_{m}(x) \parallel \leq g_{m}(x), \forall m \in \mathbb{N}, \) and \( \lim_{m \in \mathbb{N}} g_{m}(x) = g(x) \). Then, by Propositions 7.21 and 3.66, we have \( \parallel f(x) - \lim_{m \in \mathbb{N}} \phi_{m}(x) \parallel \leq \lim_{m \in \mathbb{N}} g_{m}(x) = g(x) \). Then, by \( \mathcal{R} \in \mathcal{J} (\mathcal{W}), \parallel \hat{w}_{\gamma} \parallel < \inf_{x \in \hat{U}_{\gamma}} g_{m}(x) + 1 \leq \parallel f(x) \parallel + 1 \leq g(x) + 1 \). By the arbitrariness of \( x \), we have \( \parallel \hat{w}_{\gamma} \parallel \leq \inf_{x \in \hat{A}_{\gamma}} g(x) + 1 \). Therefore, \( \parallel \hat{w}_{\gamma} \mu (\hat{A}_{\gamma}) \parallel \leq (1 + \inf_{x \in \hat{A}_{\gamma}} g(x)) \mu (\hat{A}_{\gamma}) \leq (1 + \inf_{x \in \hat{A}_{\gamma}} g(x)) \mathcal{P} \circ \mu (\hat{A}_{\gamma}) \).

Note that
\[ \sum_{\gamma \in \Gamma} \hat{w}_{\gamma} \mu (f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap E_{1}) = \sum_{\gamma \in \Gamma} \hat{w}_{\gamma} (\mu (\hat{A}_{\gamma}) + \mu (f_{\mathcal{W}}(\hat{U}_{\gamma}) \cap E_{1} \cap E_{1})) \]
The simple function \( \psi : X \to [0, \infty) \subseteq \mathbb{R} \), defined by \( \psi(x) = \sum_{\gamma \in \Gamma} (\inf_{x \in \cal{A}}(g(x)))_{\gamma}, \forall x \in X \), satisfies \( 0 \leq \psi(x) \leq g(x), \forall x \in X \) and \( \psi(x) = 0, \forall x \in X \setminus E_1 \). Then, \( \sum_{\gamma \in \Gamma} (\inf_{x \in \cal{A}}(g(x)))_{\gamma} = \int_X \psi \, d\mu = \int_{E_1} \psi \, d\mu < \epsilon/6 \), where the equalities follow from Proposition 11.75; and the first inequality follows from Proposition 11.83. Then, \( \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma (f_{\text{inv}}(U_\gamma) \cap E_1) \right\| < \mathcal{P} \circ \mu(A_1) + \epsilon/6 < \delta + \epsilon/6 \leq \epsilon/3 \).

This completes the proof of the claim. \( \square \)

Claim 11.126.2 \( \left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| < \epsilon/6. \)

Proof of claim: Since \( \| \phi_n(x) \| \leq g_n(x), \forall x \in X \), then \( \| w_i \| \leq g_n(x), \forall x \in A_i, \, i = 1, \ldots, \bar{n} + 1 \). Then, by Definition 11.79, \( \left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \leq \sum_{i=1}^{\bar{n}+1} \| w_i \| \mu(A_i \cap E_1) \leq \sum_{i=1}^{\bar{n}+1} \| w_i \| \mathcal{P} \circ \mu(A_1 \cap E_1) \leq \int_{E_1} g_n \, d\mathcal{P} \circ \mu < \epsilon/6. \)

Claim 11.126.3 \( \forall \gamma \in \Gamma \setminus \bar{\Gamma}, \, \forall i \in \{1, \ldots, \bar{n} + 1\} \), we have \( \| \hat{w}_\gamma - w_i \| \cdot \| \mu(A_{\gamma,i}) \| \leq \frac{3\epsilon}{6 \mathcal{P} \circ \mu(A_{\gamma,i})}. \)

Proof of claim: \( \forall \gamma \in \Gamma \setminus \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}, \hat{U}_\gamma \subseteq V_{i_{\gamma}}, \) where \( i_{\gamma} \in \{1, \ldots, \bar{n} + 1\} \). The result is trivial if \( A_{\gamma,i} = \emptyset \). If \( A_{\gamma,i} \neq \emptyset \), then \( \exists x \in A_{\gamma,i} = A_i \cap f_{\text{inv}}(\hat{U}_\gamma) \subseteq (X \setminus E_1) \). This implies that \( \phi_n(x) = w_i \) and \( f(x) \in B_W(w_i, \epsilon) \subseteq B_W(w_i, \epsilon) \cap V_{i_{\gamma}} \subseteq B_W(w_i, \epsilon) \cap \hat{U}_\gamma \subseteq B_W(w_i, \epsilon). \) Note that \( \hat{w}_\gamma \in \hat{U}_\gamma \subseteq B_W(w_i, \epsilon) \). Therefore, we have \( \| \hat{w}_\gamma - w_i \| \leq \| \hat{w}_\gamma - w_i \| + \| w_i - f(x) \| + \| f(x) - w_i \| < 3\epsilon = \frac{3\epsilon}{6 \mathcal{P} \circ \mu(X_{\gamma,i}) + 1} \). Hence, the result holds. \( \square \)

Note that \( F_R = \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(U_\gamma)) \) and \( \int_X \phi_n \, d\mu = \sum_{i=1}^{\bar{n}} w_i \mu(A_i) = \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i), \) by Proposition 11.125. Then,

\[
\left\| F_R - \int_X \phi_n \, d\mu \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(U_\gamma)) - \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \right\|
\]
\[ \begin{align*} &= \left\| \sum_{\gamma \in \Gamma} \tilde{w}_\gamma (\mu(\text{inv}(\tilde{U}_\gamma) \cap (X \setminus E_1)) + \mu(\text{inv}(\tilde{U}_\gamma) \cap E_1)) \\ &\quad - \sum_{i=1}^{n+1} w_i (\mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1))) \right\| \\ &\quad - \sum_{i=1}^{n+1} w_i (\mu(A_i \cap E_1)) \right\| \\ &\quad - \sum_{i=1}^{n+1} w_i (\mu(A_i \cap E_1)) \right\| \\ &< \left\| \sum_{\gamma \in \Gamma} (\tilde{w}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| + \frac{\epsilon}{3} + \frac{\epsilon}{6} \\ &= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma \setminus \Gamma} \sum_{i=1}^{n+1} (\tilde{w}_\gamma - w_i) \mu(A_{\gamma,i}) \right\| \\ &\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \Gamma} \sum_{i=1}^{n+1} \left\| \tilde{w}_\gamma - w_i \right\| \mu(A_{\gamma,i}) \\ &\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma \setminus \Gamma} \sum_{i=1}^{n+1} \frac{3\epsilon}{5P \circ \mu(X) + 1} \circ \mu(A_{\gamma,i}) \leq \frac{\epsilon}{2} + \frac{3P \circ \mu(X) \epsilon}{6P \circ \mu(X) + 1} < \epsilon \\ \end{align*} \]

where the first inequality follows from Claims 11.126.1 and 11.126.2; the seventh equality follows from Claim 11.126.1 and the fact that \( A_{\gamma,i} = \emptyset, \forall \gamma \in \Gamma, \forall i \in \{1, \ldots, n+1\} \); and the third inequality follows from Claim 11.126.3.

In summary, we have shown that \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n \geq n_1, \exists R \in \mathcal{J}(W) \) such that \( \forall R \in \mathcal{J}(W) \) with \( \bar{R} \leq R \), we have
\[ \| F_R - \int_X \phi_n \, d\mu \| < \epsilon. \] Then, the net \((F_R)_{R \in \mathbb{R}([W])} \subseteq \mathcal{Z}\) is a Cauchy net. By Proposition 4.44, the net admits a limit \(\int_X f \, d\mu \in \mathcal{Z}\). By Propositions 7.21, 7.23, 3.67, and 3.66, we have \(\| \int_X f \, d\mu - \int_X \phi_n \, d\mu \| = \lim_{R \in \mathbb{R}([W])} \| F_R - \int_X \phi_n \, d\mu \| \leq \epsilon, \forall n \geq n_1\). Hence, we have \(\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu\) and \(f\) is integrable over \(\mathcal{X}\). By Proposition 11.125, \(\int_X \phi_n \, d\mu \in W, \forall n \in \mathbb{N}\). Then, by the completeness of \(W\), we have \(\int_X f \, d\mu \in W\).

Note that \(0 \leq \| \int_X f \, d\mu \| = \lim_{n \in \mathbb{N}} \| \int_X \phi_n \, d\mu \| = \limsup_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n \, d\mathcal{P} \circ \mu \leq \lim_{n \in \mathbb{N}} \int_X g_n \, d\mathcal{P} \circ \mu = \int_X g \, d\mathcal{P} \circ \mu < +\infty\), where the first equality follows from Propositions 7.21 and 3.66; the second inequality follows from Proposition 11.125; and the third inequality follows from Proposition 11.83. This completes the proof of the lemma.

**Lemma 11.127** Let \(\mathcal{X} := (X, \mathcal{B,} \mu)\) be a finite \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space over \(\mathbb{K}\), \(\mathcal{Z}\) be a Banach space over \(\mathbb{K}\), \(\mathcal{W}\) be a separable subspace of \(\mathcal{B}(\mathcal{Y, Z})\), \(f : X \rightarrow \mathcal{W}\) be \(\mathcal{B}\)-measurable, and \(g : X \rightarrow [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that \(\| f(x) \| \leq g(x), \forall x \in X, g\) is integrable over \(\mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu)\). Then, there exists a sequence of simple functions \((\varphi_k)_{k=1}^\infty, \varphi_k : X \rightarrow \mathcal{W}, \forall k \in \mathbb{N}\) such that \(\lim_{k \in \mathbb{N}} \varphi_k = f \) a.e. in \(\mathcal{X}\), \(\| \varphi_k(x) \| \leq g(x), \forall x \in X, \forall k \in \mathbb{N}\), \(f\) is integrable over \(\mathcal{X}\) with \(\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k \, d\mu \in \mathcal{Z}\), \(\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_k \, d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \in [0, \infty) \subseteq \mathbb{R}\), and \(\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_k - f) \, d\mathcal{P} \circ \mu = 0\).

**Proof** By Proposition 11.66, there exists a sequence of simple functions \((\varphi_k)_{k=1}^\infty, \varphi_k : X \rightarrow \mathcal{W}, \forall k \in \mathbb{N}\) such that \(\| \varphi_k(x) \| \leq \| f(x) \| \leq g(x), \forall x \in X, \forall k \in \mathbb{N}\), and \(\lim_{k \in \mathbb{N}} \varphi_k = f \) a.e. in \(\mathcal{X}\). By Lemma 11.126, \(f\) is integrable over \(\mathcal{X}\) and \(\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \varphi_k \, d\mu \in \mathcal{Z}\).

Note that \(0 \leq \mathcal{P} \circ \varphi_k (x) \leq g(x), \forall x \in X, \forall k \in \mathbb{N}\), \(\mathcal{P} \circ f\) is \(\mathcal{B}\)-measurable, and \(\lim_{k \in \mathbb{N}} \mathcal{P} \circ \varphi_k = \mathcal{P} \circ f\) a.e. in \(\mathcal{X}\), by Propositions 7.21 and 11.52. Then, by Lebesgue Dominated Convergence Theorem 11.91, we have \(0 \leq \lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_k \, d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty\), where the second inequality follows from Proposition 11.83. By Lemma 11.43, \(\mathcal{P} \circ (\varphi_k - f)\) is \(\mathcal{B}\)-measurable. By Propositions 7.21, 7.23, 11.52, and 11.53, \(\lim_{k \in \mathbb{N}} \mathcal{P} \circ (\varphi_k - f) = 0\) a.e. in \(\mathcal{X}\). Note that \(0 \leq \mathcal{P} \circ (\varphi_k - f)(x) \leq 2g(x), \forall x \in X, \forall k \in \mathbb{N}\). By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have \(\lim_{k \in \mathbb{N}} \int_X \mathcal{P} \circ (\varphi_k - f) \, d\mathcal{P} \circ \mu = 0\). This completes the proof of the lemma.

**Theorem 11.128** (Lebesgue Dominated Convergence) Let \(\mathcal{X} := (X, \mathcal{B,} \mu)\) be a finite \(\mathcal{Y}\)-valued measure space, where \(\mathcal{Y}\) is a normed linear space over \(\mathbb{K}\), \(\mathcal{Z}\) be a Banach space over \(\mathbb{K}\), \(\mathcal{W}\) be a separable subspace of \(\mathcal{B}(\mathcal{Y, Z})\), \(\mathcal{W} := \text{span}\{\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \exists z = wy\}\} \subseteq \mathcal{Z}\) be the Banach subspace of \(\mathcal{Z}\), \(f_n : X \rightarrow \mathcal{W}\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), \(f : X \rightarrow \mathcal{W}\) be \(\mathcal{B}\)-measurable, \(g_n : X \rightarrow [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), and \(g : X \rightarrow [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that

1. \(\lim_{n \in \mathbb{N}} f_n = f \) a.e. in \(\mathcal{X}\), \(\lim_{n \in \mathbb{N}} g_n = g \) a.e. in \(\mathcal{X}\);
(ii) \( \| f_n(x) \| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N} \);

(iii) \( \int_X g_n \, d\mathcal{P} \circ \mu = \lim_{n \to \infty} \int_X f_n \, d\mathcal{P} \circ \mu \in \mathbb{R} \).

Then, \( f \) is integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \overline{W} \subseteq \mathbb{Z} \), and
\[
0 \leq \| \int_X f \, d\mu \| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty.
\]

Proof. \( \forall n \in \mathbb{N} \), by Lemma 11.127, there exists a sequence of simple functions (\( \tilde{\varphi}_{n,k} \)) \( \in \mathbb{R} \), \( \forall k \in \mathbb{N} \), such that \( \lim_{k \to \infty} \tilde{\varphi}_{n,k} = f_n \) a.e. in \( X \), \( \| \tilde{\varphi}_{n,k}(x) \| \leq g_n(x), \forall x \in X, \forall k \in \mathbb{N}, f_n \) is integrable over \( X \), and \( \int_X f_n \, d\mu = \lim_{k \to \infty} \int_X \tilde{\varphi}_{n,k} \, d\mu \in \mathbb{Z} \). By Lemma 11.126, \( \int_X f_n \, d\mu \in \overline{W} \), \( \forall n \in \mathbb{N} \). By Proposition 11.58, \( \lim_{k \to \infty} \tilde{\varphi}_{n,k} = f_n \) in measure in \( X \). Then, \( \exists k_n \in \mathbb{N} \) such that \( \mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{ x \in X \mid \| \tilde{\varphi}_{n,k_n} - f_n(x) \| \geq 2^{-n} \}) < 2^{-n} \) and \( \| \int_X f_n \, d\mu - \int_X \tilde{\varphi}_{n,k_n} \, d\mu \| < 2^{-n} \). Denote \( \psi_n := \tilde{\varphi}_{n,k_n} \).

By Proposition 11.58, \( \lim_{n \to \infty} f_n = f \) in measure in \( X \). \( \forall \in (0, \infty) \in \mathbb{R}, \exists n_0 \in \mathbb{N} \) such that, for all \( n \geq n_0 \), we have \( 2^{-n} < \epsilon/2 \) and \( \mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{ x \in X \mid \| f_n(x) - f(x) \| \geq \epsilon/2 \}) < \epsilon/2 \). Note that, by Lemma 11.43, \( B \supseteq F_n := \{ x \in X \mid \| \psi_n(x) - f(x) \| \geq \epsilon \} \subseteq \{ x \in X \mid \| \psi_n(x) - f_n(x) \| + \| f_n(x) - f(x) \| \geq \epsilon \} \subseteq E_n \cup E_n \). Then, we have \( \mathcal{P} \circ \mu(F_n) \leq \mathcal{P} \circ \mu(E_n) + \mathcal{P} \circ \mu(E_n) < \epsilon \). This implies that \( \lim_{n \to \infty} \psi_n = f \) in measure in \( X \). By Proposition 11.57, there exists a subsequence \( (\psi_n)_{n=1}^{\infty} \) of \( (\psi_n)_{n=1}^{\infty} \) that converges to \( f \) a.e. in \( X \). By Lemma 11.126, \( f \) is integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \psi_n \, d\mu \in \mathbb{Z} \), and \( 0 \leq \| \int_X f \, d\mu \| < +\infty \). By the fact that \( \| \int_X f_n \, d\mu - \int_X \psi_n \, d\mu \| < 2^{-n} \), \( \forall i \in \mathbb{N} \), then we have \( \lim_{n \to \infty} \int_X f_n \, d\mu = \lim_{n \to \infty} \int_X \psi_n \, d\mu = \int_X f \, d\mu \in \mathbb{Z} \).

So far, we have shown that, for the sequence \( (f_n)_{n=1}^{\infty} \), there exists a subsequence \( (f_{n_i})_{i=1}^{\infty} \) such that \( \int_X f_{n_i} \, d\mu = \int_X f \, d\mu \in \mathbb{Z} \). This then holds for any subsequence of \( (f_n)_{n=1}^{\infty} \). By Proposition 3.71, we have \( \lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \). By the completeness of \( \overline{W} \), we have \( \int_X f \, d\mu \in \overline{W} \subseteq \mathbb{Z} \).

This completes the proof of the theorem. \( \square \)

Definition 11.129 Let \( X := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( Z \) be a normed linear space over \( \mathbb{K} \), \( W := B(\mathcal{Y}, \mathcal{Z}) \), and \( f : X \to W \) be \( \mathcal{B} \)-measurable. \( f \) is said to be absolutely integrable if \( \mathcal{P} \circ f \) is integrable over \( X := (X, \mathcal{B}, \mathcal{P} \circ \mu) \), that is, \( 0 \leq \int_X \mathcal{P} \circ f \, d\mu < +\infty \).

Proposition 11.130 Let \( X := (X, \mathcal{B}, \mu) \) be a finite \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( Z \) be a Banach space over \( \mathbb{K} \), \( W \) be a separable subspace of \( B(\mathcal{Y}, \mathcal{Z}) \), \( U \) be a Banach space over \( \mathbb{K} \), \( W := \text{span}\{ \{ z \in Z \mid \exists \gamma \in \mathcal{Y}, \exists w \in W : z = \gamma \cdot w \} \} \subseteq Z \), be the Banach subspace of \( Z \), \( f_i : X \to W \) be absolutely integrable over \( X \), \( i = 1, 2 \). Then, the following statements hold.

(i) \( f_i \) is integrable over \( X \) and \( \int_X f_i \, d\mu \in \overline{W} \subseteq \mathbb{Z} \), \( i = 1, 2 \).
(ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathcal{W}$.

(iii) $\forall A \in \mathcal{B}(\mathcal{Z}, \mathcal{U}), A f_1$ is absolutely integrable over $\mathcal{X}$ and $\int_X (A f_1) \, d\mu = A \int_X f_1 \, d\mu \in \mathcal{U}$.

(iv) $\forall c \in \mathbb{K}$, $c f_1$ is absolutely integrable over $\mathcal{X}$ and $\int_X (c f_1) \, d\mu = c \int_X f_1 \, d\mu \in \mathcal{W}$.

(v) $\forall H \in \mathcal{B}$, let $\mathcal{H} := (H, \mathcal{B}_H, \mu_H)$ be the finite $\mathcal{Y}$-valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Then, $f_1 |H$ is absolutely integrable over $\mathcal{H}$ and $\int_X (f_1 |_{H,x}) \, d\mu = \int_H f_1 |_{H} \, d\mu_H \in \mathcal{W}$. We will henceforth denote $\int_H f_1 |_{H} \, d\mu_H$ by $\int_H f_1 \, d\mu$.

(vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 \, d\mu = \int_X f_2 \, d\mu \in \mathcal{W}$.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, $\sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \int_{\bigcup_{i=1}^{\infty} E_i} f_1 \, d\mu \in \mathcal{W}$.

(viii) $0 \leq \| \int_X f_1 \, d\mu \| \leq \int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty$.

**Proof**

(i) $\forall i \in \{1, 2\}$, by Lemma 11.127, there exists a sequence of simple functions $(\phi_{i,n})_{n=1}^{\infty}$, $\phi_{i,n} : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \to \infty} \phi_{i,n} = f_i$ a.e. in $\mathcal{X}$, $\| \phi_{i,n}(x) \| \leq \mathcal{P} \circ f_i(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $f_i$ is integrable over $\mathcal{X}$, and $\int_X f_i \, d\mu = \lim_{n \to \infty} \int_X \phi_{i,n} \, d\mu \in \mathcal{Z}$. By Lebesgue Dominated Convergence Theorem 11.128, $\int_X f_i \, d\mu \in \mathcal{W}$, $i = 1, 2$.

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, $f_1 + f_2$ and $\mathcal{P} \circ (f_1 + f_2)$ are $\mathcal{B}$-measurable. Note that $\mathcal{P} \circ (f_1 + f_2)(x) = \| f_1(x) + f_2(x) \| \leq \| f_1(x) \| + \| f_2(x) \| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$. By Proposition 11.83, we have $0 \leq \int_X \mathcal{P} \circ f_1 + \mathcal{P} \circ f_2 \, d\mathcal{P} \circ \mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) \, d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu + \int_X \mathcal{P} \circ f_2 \, d\mathcal{P} \circ \mu < +\infty$. Hence, $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$.

Note that $\phi_{1,n} + \phi_{2,n} : X \to \mathcal{W}$, $f_1 + f_2 : X \to \mathcal{W}$, and $\| \phi_{1,n}(x) + \phi_{2,n}(x) \| \leq \| \phi_{1,n}(x) \| + \| \phi_{2,n}(x) \| \leq (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$, $\forall n \in \mathcal{N}$. By Propositions 7.29, 11.52, and 11.53, $\lim_{n \to \infty} (\phi_{1,n} + \phi_{2,n}) = f_1 + f_2$ a.e. in $\mathcal{X}$. By Lebesgue Dominated Convergence Theorem 11.128, $f_1 + f_2$ is integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{1,n} + \phi_{2,n}) \, d\mu = \lim_{n \to \infty} (\int_X \phi_{1,n} \, d\mu + \int_X \phi_{2,n} \, d\mu) = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathcal{W}$, where the second equality follows from Proposition 11.125; and the third equality follows from Propositions 3.66, 3.67, and 7.23.

(iii) Fix any $A \in \mathcal{B}(\mathcal{Z}, \mathcal{U})$. By Propositions 11.38, 7.21, and 7.62, $A \phi_{1,n}$, $A f_1$, and $\mathcal{P} \circ (A f_1)$ are $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. Note that $\mathcal{P} \circ (A f_1)(x) = \| A f_1(x) \| \leq \| A \| \| f_1(x) \| = \| A \| \mathcal{P} \circ f_1(x)$, $\forall x \in X$. By Proposition 11.83 and the fact that $0 \leq \int_X (\mathcal{P} \circ f_1) \, d\mathcal{P} \circ \mu < +\infty$, we have $0 \leq \int_X (\mathcal{P} \circ (A f_1)) \, d\mathcal{P} \circ \mu \leq \int_X (\| A \| \mathcal{P} \circ f_1) \, d\mathcal{P} \circ \mu = \| A \| \int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty$. Hence, $A f_1$ is absolutely integrable over $\mathcal{X}$.

Note that $A f_1 : X \to \mathcal{W}$ and $A \phi_{1,n} : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$, where $\mathcal{W} := \{ F \in \mathcal{B}(\mathcal{Y}, \mathcal{U}) \mid F = A F_1, F_1 \in \mathcal{W} \}$ is a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{U})$. 

11.6. BANACH SPACE VALUED MEASURES

443
By Propositions 11.52 and 7.62, \( \lim_{n \to \infty} A \phi_{1,n} = A f_1 \) a.e. in \( \mathcal{X} \). By Proposition 11.125, we have \( \int_X (A \phi_{1,n}) \, d\mu = A \int_X \phi_{1,n} \, d\mu \in \mathcal{U} \), \( \forall n \in \mathbb{N} \). Note also that \( \|A \phi_{1,n}(x)\| \leq \|A\| \|\phi_{1,n}(x)\| \leq \|A\| P \circ f_1(x) \), \( \forall x \in X, \forall n \in \mathbb{N} \).

By Lebesgue Dominated Convergence Theorem 11.128, \( A f_1 \) is integrable over \( \mathcal{X} \) and \( \int_X (A f_1) \, d\mu = \lim_{n \to \infty} \int_X (A \phi_{1,n}) \, d\mu = \lim_{n \to \infty} A \int_X \phi_{1,n} \, d\mu = A \int_X f_1 \, d\mu \in \mathcal{U} \), where the last equality follows from Propositions 3.66 and 7.62.

(iv) This follows immediately from (iii).

(v) Fix any \( H \in \mathcal{B} \). Note that \( (\mathcal{P} \circ f_1)_{|H} = \mathcal{P} \circ (f_1)_{|H} \). Then, by Proposition 11.115, we have \( 0 \leq \int_H (P \circ f_1) \, dP \circ \mu_H = \int_H (P \circ f_1)_{|H} \, d(P \circ \mu)_{|B} = \int_H P \circ f_1 \, dP \circ \mu \leq \int_X P \circ f_1 \, dP \circ \mu < +\infty \), where the second equality and the second inequality follow from Proposition 11.83. Hence, \( f_1_{|H} \) is absolutely integrable over \( \mathcal{H} \).

\( (\phi_{1,n} H \, \mathcal{H} \, 1_{n=1}) \) is a sequence of simple functions that converges to \( f_1_{|H} \) a.e. in \( \mathcal{H} \). Note that \( \|\phi_{1,n} H (x)\| \leq (\mathcal{P} \circ f_1)_{|H} (x) = \mathcal{P} \circ (f_1)_{|H} (x) \), \( \forall x \in H, \forall n \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_H f_1_{|H} \, d\mu_H = \lim_{n \to \infty} \int_H \phi_{1,n} H \, d\mu_H \in \mathcal{W} \). Note that \( (\phi_{1,n} \chi_{H,X} H \, \mathcal{H} \, 1_{n=1}) \) is a sequence of simple functions that converges to \( f_1 \chi_{H,X} \) a.e. in \( \mathcal{X} \), and \( \|\phi_{1,n} \chi_{H,X} H (x)\| \leq \mathcal{P} \circ f_1 \), \( \forall x \in X, \forall n \in \mathbb{N} \).

Again by Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_X (f_1 \chi_{H,X}) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{1,n} \chi_{H,X}) \, d\mu = \lim_{n \to \infty} \int_H \phi_{1,n} \, d\mu_H = \int_H f_1 \, d\mu_H \in \mathcal{W} \), where the second equality follows from Proposition 11.125.

(vi) If \( f_1 = f_2 \) a.e. in \( \mathcal{X} \), then, by Proposition 11.54, \( \lim_{n \to \infty} \phi_{1,n} = f_2 \) a.e. in \( \mathcal{X} \). By Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_X f_2 \, d\mu = \lim_{n \to \infty} \int_X \phi_{1,n} \, d\mu = \int_X f_1 \, d\mu \).

(vii) \( \forall \) pairwise disjoint \( (E_i \, i=1)^\infty \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^\infty E_i \in \mathcal{B} \) and \( \tilde{E}_n := \bigcup_{i=1}^n E_i \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \lim_{n \to \infty} f_1 \chi_{\tilde{E}_n} \chi_{X} = f_1 \chi_{X} \chi_{E} \chi_{X} \), \( \forall x \in X \). Then, we have \( \int_X f_1 \chi_{X} \, d\mu = \lim_{n \to \infty} \int_X f_1 \chi_{\tilde{E}_n} \chi_{X} \, d\mu = \int_X f_1 \chi_{X} \, d\mu \subseteq \mathcal{W} \), where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.128; the last equality follows from (v); and the last inclusion follows from (v).

(viii) The result follows immediately from Lebesgue Dominated Convergence Theorem 11.128.

This completes the proof of the proposition. \( \Box \)

**Theorem 11.131 (Lebesgue Dominated Convergence)** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( \mathcal{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( B (\mathcal{Y}, \mathcal{Z}) \), \( \mathcal{W} := \text{span} \left( \{ \{ z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \mid z = w y \} \} \right) \subseteq \mathcal{Z} \) be the Banach subspace of \( \mathcal{Z} \), \( f_n : X \to \mathcal{W} \) be \( B \)-measurable, \( \forall n \in \mathbb{N} \), \( f : X \to \mathcal{W} \) be \( B \)-measurable, \( g_n : X \to [0, \infty) \subset \mathbb{R} \) be \( B \)-measurable, \( \forall n \in \mathbb{N} \), and \( g : X \to [0, \infty) \subset \mathbb{R} \) be \( B \)-measurable. Assume that

(i) \( \lim_{n \to \infty} f_n = f \) a.e. in \( \mathcal{X} \), \( \lim_{n \to \infty} g_n = g \) a.e. in \( \mathcal{X} \);
(ii) \( \|f_n(x)\| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}; \)

(iii) \( \int_X g_n \, d\mathcal{P} \circ \mu = \lim_{n \to \infty} \int_X g_n \, d\mathcal{P} \circ \mu \in \mathbb{R}. \)

Then, \( f \) is integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathcal{W} \subseteq \mathbb{Z}, \) and \( 0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty. \) Furthermore, \( f \) is absolutely integrable over \( X \).

**Proof** We will show that \( f \) is integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathcal{W}, \) and \( 0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{P} \circ \mu(X) < +\infty \); Case 2: \( \mathcal{P} \circ \mu(X) = +\infty \).

Case 1: \( \mathcal{P} \circ \mu(X) < +\infty \). Then, \( X \) is a finite \( \mathcal{Y} \)-valued measure space. By Lebesgue Dominated Convergence Theorem 11.128, \( f \) is integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathcal{W}, \) and \( 0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty \).

Case 2: \( \mathcal{P} \circ \mu(X) = +\infty \). Let \( (F_A)_{A \in \mathcal{M}(X)} \) and \( (\tilde{F}_n, A)_{A \in \mathcal{M}(X)} \) be the nets for the integrals \( \int_X f \, d\mu \) and \( \int_X f_n \, d\mu \) as defined in Definition 11.120, respectively, \( \forall n \in \mathbb{N}, \forall A \in \mathcal{M}(X), \) we have \( A \in \mathcal{B} \) and \( \mathcal{P} \circ \mu(A) < \infty \). Let \( A := (A, \mathcal{B}_A, \mu_A) \) be the finite \( \mathcal{Y} \)-valued measure subspace of \( X \) as defined in Proposition 11.115. By Proposition 11.83, \( g \mid_A \) and \( g_n \mid_A \)'s are integrable over \( A := (A, \mathcal{B}_A, \mathcal{P} \circ \mu_A), \) which is the finite measure subspace of \( X := (X, \mathcal{B}, \mathcal{P} \circ \mu) \) as shown in Proposition 11.115. By Proposition 11.83, \( \lim_{n \to \infty} g_n \mid_A = g \mid_A \) a.e. in \( A \) and \( \int_A g \mid_A \, d\mathcal{P} \circ \mu_A = \lim_{n \to \infty} \int_A g_n \mid_A \, d\mathcal{P} \circ \mu_A \in \mathbb{R}. \) By Lebesgue Dominated Convergence Theorem 11.128, we have \( f \mid_A \) and \( f_n \mid_A \)'s are integrable over \( A \), \( \int_A f \mid_A \, d\mu_A = \lim_{n \to \infty} \int_A f_n \mid_A \, d\mu_A \in \mathcal{W}, \) and \( 0 \leq \|\int_A f \mid_A \, d\mu_A\| \leq \int_A g \mid_A \, d\mathcal{P} \circ \mu_A < +\infty \). Note that \( F_A = \int_A f \mid_A \, d\mu_A \in \mathcal{W} \) and \( F_n, A = \int_A f_n \mid_A \, d\mu_A \in \mathcal{W}, \forall n \in \mathbb{N}. \) Then, the nets \( (F_A)_{A \in \mathcal{M}(X)} \) and \( (\tilde{F}_n, A)_{A \in \mathcal{M}(X)} \)'s are well defined.

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by Definition 11.79, there is a simple function \( \phi : X \to [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi(x) \leq g(x), \forall x \in X, \) and \( \int_X g \, d\mathcal{P} \circ \mu - \epsilon/5 < \int_X \phi \, d\mathcal{P} \circ \mu \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty \). Let \( A_0 := \{ x \in X : \phi(x) > 0 \}. \) Then, \( A_0 \in \mathcal{B}, \mathcal{P} \circ \mu(A_0) < +\infty, \) and \( A_0 \in \mathcal{M}(X) \), \( \forall A \in \mathcal{M}(X) \) with \( A_0 \subseteq A, \) we have \( 0 \leq \int_{A \setminus A_0} g \mid_{A \setminus A_0} \, d\mathcal{P} \circ \mu_{A \setminus A_0} = \int_A g \mid_A \, d\mathcal{P} \circ \mu_A - \int_{A_0} g \mid_{A_0} \, d\mathcal{P} \circ \mu_{A_0} \leq \int_A g \, d\mathcal{P} \circ \mu - \int_{A_0} \phi \mid_{A_0} \, d\mathcal{P} \circ \mu_{A_0} = \int_X g \, d\mathcal{P} \circ \mu - \int_X \phi \, d\mathcal{P} \circ \mu < \epsilon/5, \) where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.74. Note that \( \|F_A - F_{A_0}\| = \|\int_A f \mid_A \, d\mu_A - \int_{A_0} f \mid_{A_0} \, d\mu_{A_0}\| = \|\int_{A \setminus A_0} f \mid_{A \setminus A_0} \, d\mu_{A \setminus A_0}\| \leq \int_{A \setminus A_0} g \mid_{A \setminus A_0} \, d\mathcal{P} \circ \mu_{A \setminus A_0} < \epsilon/5, \) where the second equality and the first inequality follow form Proposition 11.130. Then, the net \( (F_A)_{A \in \mathcal{M}(X)} \) is a Cauchy net, which admits a limit, by Proposition 4.44, \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(X)} F_A \in \mathcal{W}. \) Hence, \( f \) is integrable over \( X \). By Propositions 7.21 and 3.66, \( 0 \leq \|\int_X f \, d\mu\| = \lim_{A \in \mathcal{M}(X)} \|F_A\| \leq \lim_{A \in \mathcal{M}(X)} \int_A g \mid_A \, d\mathcal{P} \circ \mu_A = \int_X g \, d\mathcal{P} \circ \mu < +\infty. \)
forall n in N, by Definition 11.79, exists a simple function φn : X → [0, ∞) ⊂ R such that 0 ≤ φn(x) ≤ g(x), ∀x ∈ X, and ∫X φn dP o μ - ε/5 < ∫X φn dP o μ < ∞. Let An := \{x ∈ X | φn(x) > 0\}. Then, An ∈ B, P o μ(An) < +∞, and An ∈ M(X). ∀A ∈ M(X) with A ⊆ X, we have 0 ≤ ∫A \( g_n\) dP o μA - ∫A φn dP o μA ≤ ∫A φn dP o μ - ∫A φn dP o μA = ∫A φn dP o μ - ∫A φn dP o μA, where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that equality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Let \( g_n := g_n \wedge g, \forall n \in N \). Then, 0 ≤ \( g_n(x) \leq g(x) \), ∀x ∈ X, and \( g_n \) is \( B \)-measurable by Proposition 11.40, \( \forall n \in N \). Then, limn∈N \( g_n \) = g a.e. in X by Proposition 11.50. By Fatou’s Lemma 11.80, ∫X g dP o μ ≤ lim infn∈N \( ∫X g_n dP o μ \). By Definition 11.79, \( 0 ≤ ∫X \( g_n \wedge g \) dP o μ ≤ ∫X g dP o μ < +∞, ∀n ∈ N \). Then, we have ∫X g dP o μ ≤ lim infn∈N ∫X g_n dP o μ ≤ lim supn∈N ∫X g_n dP o μ ≤ ∫X g dP o μ < +∞. Therefore, by Proposition 3.83, ∫X g dP o μ = limn∈N \( ∫X g_n dP o μ \). This coupled with (iii) and the fact that \( F_{a_0} = lim_{n∈N} F_{a_n} \), implies that \( \exists n_0 \in N, \forall n \in N \) with \( n_0 ≤ n \), we have 0 ≤ ∫X g dP o μ - ∫X g_n dP o μ < ε/5, \( ∫X g dP o μ - ∫X g_n dP o μ < \epsilon/5, \text{ and } \| F_{a_0} - F_{a_n} \| < \epsilon/5, \forall A ∈ M(X) \) with A ⊆ X, we have 0 ≤ ∫A \( g_n\) dP o μA ≤ ∫A φn dP o μA = ∫A \( g_n\) - \( \bar{g}_n\) dP o μA + ∫A \( \bar{g}_n\) dP o μA. Furthermore, ∫A \( g_n\) dP o μA - ∫A φn dP o μA < ε/5, where the second equality and the first inequality follow from Proposition 11.83. This implies that ∫X g dP o μ - ∫X \( g_n\) dP o μ < ε/5, where the second inequality follows from Propositions 7.21 and 3.66. Hence, ∫X g dP o μ = limn∈N ∫X \( g_n\) dP o μ ∈ W.

Then, in both cases, we have shown that \( f \) is integrable over X, ∫X f dμ = limn∈N ∫X \( f_n\) dμ ∈ W, and 0 ≤ ∫X g dP o μ < +∞.

Note that 0 ≤ P o f(x) ≤ g(x), ∀x ∈ X, \( \forall n ∈ N \), limn∈N P o f_n = P o f a.e. in X by Propositions 7.21 and 11.52, and P o f and P o f_n’s are \( B \)-measurable by Propositions 7.21 and 11.38. Then, by Lebesgue Dominated Convergence Theorem 11.91, ∫X P o f dP o μ = limn∈N ∫X P o f_n dP o μ ∈ R.

Hence, \( f \) is absolutely integrable over X.
This completes the proof of the theorem. □

**Proposition 11.132** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $\mathcal{U}$ be a Banach space over $\mathbb{K}$, $\tilde{\mathcal{W}} := \text{span} \left( \{ z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} : z = wy \} \right) \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_i : X \to \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.

(i) $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in \tilde{\mathcal{W}} \subseteq \mathcal{Z}$, $i = 1, 2$.

(ii) $f_1 + f_2$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathcal{W}$.

(iii) $\forall A \in \mathcal{B}(\mathcal{Z}, \mathcal{U})$, $A f_1$ is absolutely integrable over $\mathcal{X}$ and $\int_X (A f_1) \, d\mu = A \int_X f_1 \, d\mu \in \mathcal{U}$.

(iv) $\forall c \in \mathbb{K}$, $c f_1$ is absolutely integrable over $\mathcal{X}$ and $\int_X (c f_1) \, d\mu = c \int_X f_1 \, d\mu \in \mathcal{W}$.

(v) $\forall H \in \mathcal{B}$, let $\mathcal{H} := (H, \mathcal{B}_H, \mu_H)$ be the $\mathcal{Y}$-valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Then, $f_1|_H$ is absolutely integrable over $\mathcal{H}$ and $\int_H (f_1|_H) \, d\mu_H = \int_H f_1|_H \, d\mu_H \in \mathcal{W}$. We will henceforth denote $\int_H f_1|_H \, d\mu_H$ by $\int_H f_1 \, d\mu$.

(vi) If $f_1 = f_2$ a.e. in $\mathcal{X}$ then $\int_X f_1 \, d\mu = \int_X f_2 \, d\mu \in \tilde{\mathcal{W}}$.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, $\sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \int_{\bigcup_{i=1}^{\infty} E_i} f_1 \, d\mu \in \mathcal{W}$.

(viii) $0 \leq \left\| \int_X f_1 \, d\mu \right\| \leq \int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty$.

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. The results follow immediately from Proposition 11.130.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $f_1$, $f_2$, $\mathcal{P} \circ f_1$, and $\mathcal{P} \circ f_2$ are $\mathcal{B}$-measurable by Propositions 7.21 and 11.38. Let $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f_i \, d\mu$ as defined in Definition 11.120, $i = 1, 2$.

(i) $\forall E \in \mathcal{M}(\mathcal{X})$, by Proposition 11.130, $F_{i,E} = \int_{E} f_i|_E \, d\mu_E \in \tilde{\mathcal{W}}$, $i = 1, 2$. Then, the net $(F_{i,E})_{E \in \mathcal{M}(\mathcal{X})}$ is well-defined, $i = 1, 2$. By Lebesgue Dominated Convergence Theorem 11.131, $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in \tilde{\mathcal{W}}$, $i = 1, 2$. Then, by Definition 11.120, we have $\lim_{E \in \mathcal{M}(\mathcal{X})} F_{i,E} = \int_X f_i \, d\mu \in \tilde{\mathcal{W}}$, $i = 1, 2$.

(ii) By Propositions 11.38, 11.39, 7.21, and 7.23, $f_1 + f_2$ and $\mathcal{P} \circ (f_1 + f_2)$ are $\mathcal{B}$-measurable. Note that $\mathcal{P} \circ (f_1 + f_2)(x) = \| f_1(x) + f_2(x) \| \leq \| f_1(x) \| + \| f_2(x) \| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in \mathcal{X}$. By Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ (f_1 + f_2) \, d\mathcal{P} \circ \mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) \, d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu + \int_X \mathcal{P} \circ f_2 \, d\mathcal{P} \circ \mu < +\infty$. Hence, $f_1 + f_2$ is absolutely
integrable over $\mathcal{X}$. Let $(\tilde{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (f_1 + f_2) \, d\mu$ as defined in Definition 11.120. $\forall E \in \mathcal{M}(\mathcal{X})$, $\tilde{F}_E = \int_E (f_1 + f_2) \, d\mu_E = \int_E f_1 \, d\mu_E + \int_E f_2 \, d\mu_E = F_{1,E} + F_{2,E} \in \mathcal{W}$, where the second equality follows from Proposition 11.130. Then, $\int_X (f_1 + f_2) \, d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \tilde{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} (F_{1,E} + F_{2,E}) = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathcal{W}$, where the third equality follows from Propositions 3.66, 3.67, and 7.23. Hence, $f_1 + f_2$ is integrable over $\mathcal{X}$.

(iii) By Propositions 11.38, 7.21, and 7.62, $Af_1$ and $\mathcal{P} \circ (Af_1)$ are $\mathcal{B}$-measurable. Note that $\mathcal{P} \circ (Af_1)(x) = \|Af_1(x)\| = \|A\| \mathcal{P} \circ f_1(x)$, $\forall x \in X$. By Proposition 11.83 and the fact that $\int_X \mathcal{P} \circ f_1 \, d\mu \in \mathbb{R}$, we have $0 \leq \int_X \mathcal{P} \circ (Af_1) \, d\mu \leq \int_X (\|A\| \mathcal{P} \circ f_1) \, d\mu = \|A\| \int_X \mathcal{P} \circ f_1 \, d\mu < +\infty$. Hence, $Af_1$ is absolutely integrable over $\mathcal{X}$. Note also that $Af_1 : X \to \mathcal{W}$, where $\mathcal{W} := \{ F \in \mathcal{B}(\mathcal{Y}, \mathcal{U}) : F = AF_1, F_1 \in \mathcal{W} \}$ is a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{U})$. By (i), $\int_X (Af_1) \, d\mu \in \mathcal{U}$. Let $(\tilde{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X (Af_1) \, d\mu$ as defined in Definition 11.120. $\forall E \in \mathcal{M}(\mathcal{X})$, $\tilde{F}_E = \int_E (Af_1) \, d\mu_E = A \int_E f_1 \, d\mu_E = AF_1, E \in \mathcal{U}$, where the second equality follows from Proposition 11.130. Then, by Propositions 3.66 and 7.62, we have $\int_X (Af_1) \, d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \tilde{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} AF_1 = A \int_X f_1 \, d\mu \in \mathcal{U}$. Hence, $Af_1$ is integrable over $\mathcal{X}$.

(iv) This follows immediately from (iii).

(v) Fix any $H \in \mathcal{B}$. By Proposition 11.41, $f_1|_H$ and $\mathcal{P} \circ (f_1|_H) = \mathcal{P} \circ f_1|_H$ are $\mathcal{B}_H$-measurable. By Proposition 11.83, $0 \leq \int_H \mathcal{P} \circ (f_1|_H) \, d\mu \leq \int_X \mathcal{P} \circ f_1 \, d\mu < +\infty$. Hence, $f_1|_H$ is absolutely integrable over $\mathcal{H}$. By (i), $f_1|_H$ is integrable over $\mathcal{H}$. By Propositions 11.38, 11.39, and 7.23, $f_1\chi_{H,X}$ is $\mathcal{B}$-measurable. By Proposition 11.83, $f_1\chi_{H,X}$ is absolutely integrable over $\mathcal{X}$. Again, by (i), $f_1\chi_{H,X}$ is integrable over $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subseteq A$, we have $\left\| \int_A (f_1\chi_{H,X}) \, d\mu_A - \int_X (f_1\chi_{H,X}) \, d\mu \right\| < \epsilon/2$. We will distinguish two exhaustive and mutually exclusive subcases: Case 2a: $\mathcal{P} \circ \mu(H) = +\infty$; Case 2b: $\mathcal{P} \circ \mu(H) < +\infty$.

Case 2a: $\mathcal{P} \circ \mu(H) = +\infty$. $\exists E_0 \in \mathcal{M}(\mathcal{H})$ such that $\forall E \in \mathcal{M}(\mathcal{H})$ with $E_0 \subseteq E$, we have $\left\| \int_E f_1 \, d\mu_E - \int_H f_1 \, d\mu_H \right\| < \epsilon/2$. Let $A_1 := A_0 \cup E_0 \in \mathcal{M}(\mathcal{X})$ and $E_1 := A_1 \cap H \in \mathcal{M}(\mathcal{H})$. By Proposition 11.130, $\int_{A_1} (f_1\chi_{H,X}) |_{A_1} \, d\mu_{A_1} = \int_{H \cap A_1} f_1 |_{H \cap A_1} \, d\mu_{H \cap A_1} = \int_{E_1} f_1 |_{E_1} \, d\mu_{E_1}$. Then, we have $\left\| \int_X (f_1\chi_{H,X}) \, d\mu - \int_H f_1 \, d\mu_H \right\| \leq \left\| \int_X (f_1\chi_{H,X}) \, d\mu - \int_{A_1} (f_1\chi_{H,X}) |_{A_1} \, d\mu_{A_1} \right\| + \left\| \int_{E_1} f_1 |_{E_1} \, d\mu_{E_1} - \int_H f_1 |_H \, d\mu_H \right\| < \epsilon/2 + \epsilon/2 = \epsilon$. By the arbitrariness of $\epsilon$, we have $\int_X (f_1\chi_{H,X}) \, d\mu = \int_H f_1 \, d\mu_H \in \mathcal{W}$.

Case 2b: $\mathcal{P} \circ \mu(H) < +\infty$. Let $A_1 := A_0 \cup H \in \mathcal{M}(\mathcal{X})$. By Proposition 11.130, $\int_{A_1} (f_1\chi_{H,X}) |_{A_1} \, d\mu_{A_1} = \int_H f_1 |_H \, d\mu_H$. Then, we have $\left\| \int_X (f_1\chi_{H,X}) \, d\mu - \int_H f_1 |_H \, d\mu_H \right\| = \left\| \int_X (f_1\chi_{H,X}) \, d\mu - \int_{A_1} (f_1\chi_{H,X}) |_{A_1} \, d\mu_{A_1} \right\| < \epsilon/2$. By the arbitrariness of $\epsilon$, we have $\int_X (f_1\chi_{H,X}) \, d\mu = \int_H f_1 |_H \, d\mu_H \in \mathcal{W}$. 
11.6. BANACH SPACE VALUED MEASURES

Hence, in both subcases, \( \int_X (f_1 \chi_H, x) \, d\mu = \int_H f_1|_H \, d\mu_H \in \bar{W} \).

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

(vii) \( \forall \) pairwise disjoint \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \) and \( \bar{E}_n := \bigcup_{i=n}^{\infty} E_i \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \lim_{n \to \infty} f_1(x) \chi_{\bar{E}_n, X}(x) = f_1(x) \chi_{E, X}(x) \), \( \forall x \in X \). Then, we have \( \sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \int_{E_i} f_1 \chi_{E_i, X} \, d\mu = \lim_{n \to \infty} \int_X (f_1 \chi_{\bar{E}_n, X}) \, d\mu = \int_X f_1 \chi_{E, X} \, d\mu = \int_E f_1 \, d\mu \in \bar{W} \), where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.131; the last equality follows from (v); and the last inclusion follows from (v).

(viii) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

This completes the proof of the proposition.

\[ \square \]

Proposition 11.133 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( \mathcal{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B} \), and \( f : X \to \mathcal{W} \) be absolutely integrable over \( \mathcal{X} \). Then, there exists a sequence of simple functions \((\varphi_n)_{n=1}^{\infty}, \varphi_n : X \to \mathcal{W}, \forall n \in \mathbb{N} \), such that \( \| \varphi_n(x) \| \leq \mathcal{P} \circ f(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), \( \lim_{n \to \infty} \mathcal{P} \circ \varphi_n = f \) a.e. in \( \mathcal{X} \). Let \( \mathcal{P} \circ \varphi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \) in \( \mathcal{W} \). Note that \( \lim_{n \to \infty} \mathcal{P} \circ \varphi_n = \mathcal{P} \circ f \) a.e. in \( \mathcal{X} \), by Propositions 7.21 and 11.52. By Lebesgue Dominated Convergence Theorem 11.91, we have \( \lim_{n \to \infty} \int_X (\mathcal{P} \circ \varphi_n) \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in \mathbb{R} \). By Lemma 11.43, \( \mathcal{P} \circ (\varphi_n - f) \) is \( \mathcal{B} \)-measurable. By Propositions 7.21, 7.23, 11.53, and 11.52, we have \( \lim_{n \to \infty} \mathcal{P} \circ (\varphi_n - f) = 0 \) a.e. in \( \mathcal{X} \). Note that \( 0 \leq \mathcal{P} \circ (\varphi_n - f)(x) = \| \varphi_n(x) - f(x) \| \leq \| \varphi_n(x) \| + \| f(x) \| \leq 2 \mathcal{P} \circ f(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \). Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have \( \lim_{n \to \infty} \int_X \mathcal{P} \circ (\varphi_n - f) \, d\mu = 0 \). This completes the proof of the proposition.

\[ \square \]

Proposition 11.134 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( \mathcal{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B} \), and \( f : X \to \mathcal{W} \) be absolutely integrable over \( \mathcal{X} \). Then, there exists a sequence of simple functions \((\phi_n)_{n=1}^{\infty}, \phi_n : X \to \mathcal{W}, \forall n \in \mathbb{N} \), such that \( \| \phi_n(x) \| \leq \mathcal{P} \circ f(x), \forall x \in X \), \( \forall n \in \mathbb{N} \), \( \lim_{n \to \infty} \int_X \phi_n \, d\mu = \int_X f \, d\mu \in \mathbb{Z} \), \( \lim_{n \to \infty} \int_X \mathcal{P} \circ \phi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in \mathbb{R} \), and \( \lim_{n \to \infty} \int_X \mathcal{P} \circ (\phi_n - f) \, d\mu = 0 \).

\[ \square \]
exists a sequence of simple functions \((\phi_n)_{n=1}^{\infty}\), \(\phi_n : X \to \mathcal{W}, \forall n \in \mathbb{N}\), such that \(\|\phi_n(x)\| \leq P \circ f(x), \forall x \in X, \forall n \in \mathbb{N}\), \(\lim_{n \in \mathbb{N}} \int_X P \circ \phi_n \, dP \circ \mu = \int_X P \circ f \, dP \circ \mu \in [0, \infty) \subset \mathbb{R}\), and \(\lim_{n \in \mathbb{N}} \int_X P \circ (\phi_n - f) \, dP \circ \mu = 0\). By Propositions 11.132 and 11.125, we have \(\int_X f \, d\mu \in \mathcal{Z}\) and \(\int_X \phi_n \, d\mu \in \mathcal{Z}, \forall n \in \mathbb{N}\). By Proposition 11.132, \(\lim_{n \in \mathbb{N}} \|\int_X \phi_n \, d\mu - \int_X f \, d\mu\| = \lim_{n \in \mathbb{N}} \|\int_X (\phi_n - f) \, d\mu\| \leq \lim_{n \in \mathbb{N}} \int_X P \circ (\phi_n - f) \, dP \circ \mu = 0\). Then, \(\lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Z}\). This completes the proof of the proposition.

\[ \square \]

### 11.7 Calculation With Measures

It is easy to see that a \(\sigma\)-finite measure space \((X, \mathcal{B}, \mu)\) may be identified with a \(\sigma\)-finite \(\mathbb{R}\)-valued measure space \((X, \mathcal{B}, \bar{\mu})\) as follows: \(\bar{\mu}(E) = \mu(E) \in [0, +\infty) \subset \mathbb{R}, \forall E \in \mathcal{B}\) with \(\mu(E) < +\infty\); and \(\bar{\mu}(E)\) is undefined, \(\forall E \in \mathcal{B}\) with \(\mu(E) = +\infty\). Then, \(P \circ \bar{\mu} = \mu\). The converse of this identification is proved in the following proposition. Thus, a \(\sigma\)-finite measure on \((X, \mathcal{B})\) is a special kind of \(\sigma\)-finite \(\mathbb{R}\)-valued measure on \((X, \mathcal{B})\). We will make use of this identification to imply definitions and results for \(\sigma\)-finite measures from definitions and results for \(\sigma\)-finite \(\mathbb{R}\)-valued measures.

**Proposition 11.135** Let \(X := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite \(\mathbb{R}\)-valued measure space with \(\nu := P \circ \mu \geq 0, \forall E \in \text{dom}(\mu)\). Then, any function \(\bar{\nu} : \mathcal{B} \to \mathbb{R}\) satisfying \(\bar{\nu}(E) = \sum_{i=1}^{\infty} \bar{\nu}(E_i), \forall \text{pairwise disjoint } (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}\); and \(\bar{\nu}(E) = \mu(E), \forall E \in \text{dom}(\mu)\), must be such that \(\bar{\nu} = \nu\).

**Proof** \(\forall E \in \mathcal{B}, \) we will show that \(\bar{\nu}(E) = \nu(E)\) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \(E \in \text{dom}(\mu)\); Case 2: \(E \in \mathcal{B} \setminus \text{dom}(\mu)\).

Case 1: \(E \in \text{dom}(\mu)\). Then, \(\bar{\nu}(E) = \mu(E) \in [0, \infty) \subset \mathbb{R}\). By Definition 11.108, \(\mu(E) \leq \nu(E) < \infty\) and

\[
\nu(E) = \sup_{n \in \mathbb{N}, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \cup_{i=1}^{n} E_i = E} \sum_{i=1}^{n} |\mu(E_i)| = \mu(E) = \bar{\nu}(E)
\]

Case 2: \(E \in \mathcal{B} \setminus \text{dom}(\mu)\). Then, \(\nu(E) = \infty\). Since \(X \) is \(\sigma\)-finite, then \(\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) with \(\nu(X_n) < \infty, \forall n \in \mathbb{N}\) and \(X = \bigcup_{n=1}^{\infty} X_n\). Without loss of generality, \(, \) we may assume that \((X_n)_{n=1}^{\infty}\) is pairwise disjoint. Then, \(E = \bigcup_{n=1}^{\infty}(E \cap X_n), 0 \leq \nu(E \cap X_n) \leq \nu(X_n) < \infty, \) and \(E \cap X_n \in \text{dom}(\mu)\), \(\forall n \in \mathbb{N}\). Then, \(\bar{\nu}(E) = \sum_{n=1}^{\infty} \bar{\nu}(E \cap X_n) = \sum_{n=1}^{\infty} \nu(E \cap X_n) = \nu(E) = \infty\), where the second equality follows from Case 1.

Hence, \(\bar{\nu}(E) = \nu(E)\) in both cases. This completes the proof of the proposition.

\[ \square \]

**Proposition 11.136** Let \((X, \mathcal{B})\) be a measurable space and \(\mathcal{Y}\) and \(\mathcal{Z}\) be normed linear spaces over \(\mathbb{K}\). Then, the following statements hold.
11.7. CALCULATION WITH MEASURES

(i) Let $\mu_1$ and $\mu_2$ be measures on $(X, \mathcal{B})$, $a \in (0, \infty) \subset \mathbb{R}$. Define $\mu := \mu_1 + \mu_2 : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+$ by $\mu(E) = \mu_1(E) + \mu_2(E)$, $\forall E \in \mathcal{B}$. Then $\mu$ is a measure on $(X, \mathcal{B})$. $\mu$ is $\sigma$-finite if $\mu_1$ and $\mu_2$ are $\sigma$-finite. Define $\bar{\mu} := a\mu_1 : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+$ by $\bar{\mu}(E) = a\mu_1(E)$, $\forall E \in \mathcal{B}$, then $\bar{\mu}$ is a measure on $(X, \mathcal{B})$. $\bar{\mu}$ is $\sigma$-finite if $\mu_1$ is $\sigma$-finite.

(ii) Let $\mu$ be a finite measure on $(X, \mathcal{B})$ and $a \in [0, \infty) \subset \mathbb{R}$. Define $\bar{\mu} := a\mu : \mathcal{B} \to [0, \infty) \subset \mathbb{R}$ by $\bar{\mu}(E) = a\mu(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\mu}$ is a finite measure on $(X, \mathcal{B})$.

(iii) Let $\nu_1$ and $\nu_2$ be finite $\mathcal{Y}$-valued measures on $(X, \mathcal{B})$ and $\alpha \in \mathcal{K}$. Define $\nu := \nu_1 + \nu_2 : \mathcal{B} \to \mathcal{Y} \ni \nu(E) = \nu_1(E) + \nu_2(E)$, $\forall E \in \mathcal{B}$. Then, $\nu$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $P \circ \nu \leq P \circ \nu_1 + P \circ \nu_2$. Define $\bar{\nu} := \alpha\nu_1 : \mathcal{B} \to \mathcal{Y}$ by $\bar{\nu}(E) = \alpha\nu_1(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\nu}$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $P \circ \bar{\nu} = |\alpha| P \circ \nu_1$.

(iv) Let $\mu$ be a finite $\mathcal{K}$-valued measure on $(X, \mathcal{B})$ and $\gamma \in \mathcal{Y}$. Define $\nu := \gamma \mu : \mathcal{B} \to \mathcal{Y} \ni \nu(E) = \gamma \mu(E)$, $\forall E \in \mathcal{B}$. Then, $\nu$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $P \circ \nu = \|\gamma\| P \circ \nu$.

(v) Let $\mu$ be a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ and $A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Define $\bar{\mu} := A\mu : \mathcal{B} \to \mathcal{Z}$ by $\bar{\mu}(E) = A\mu(E)$, $\forall E \in \mathcal{B}$. Then, $\bar{\mu}$ is a finite $\mathcal{Z}$-valued measure on $(X, \mathcal{B})$ with $P \circ \bar{\mu} \leq \|A\| P \circ \mu$.

Proof

(i) Clearly, $\mu(\emptyset) = \mu_1(\emptyset) + \mu_2(\emptyset) = 0$ and $\bar{\mu}(\emptyset) = a\mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\sum_{i=1}^{\infty} (\mu_1 + \mu_2)(E_i) = \sum_{i=1}^{\infty} (\mu_1(E_i) + \mu_2(E_i)) = \sum_{i=1}^{\infty} \mu_1(E_i) + \sum_{i=1}^{\infty} \mu_2(E_i) = \mu_1(E) + \mu_2(E) = (\mu_1 + \mu_2)(E)$ and $\sum_{i=1}^{\infty} (a\mu_1)(E_i) = \sum_{i=1}^{\infty} a\mu_1(E_i) = a\sum_{i=1}^{\infty} \mu_1(E_i) = a\mu_1(E)$. Hence, $\mu_1 + \mu_2$ and $a\mu_1$ are measures on $(X, \mathcal{B})$.

When $\mu_1$ is $\sigma$-finite, there exists $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu_1(X_n) < \infty$, $\forall n \in \mathbb{N}$. Then, $(a\mu_1)(X_n) = a\mu_1(X_n) < \infty$. Hence, $a\mu_1$ is $\sigma$-finite.

When $\mu_2$ is also $\sigma$-finite, there exists $(X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu_2(X_n) < \infty$, $\forall n \in \mathbb{N}$. Then, $X = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} (X_n \cap X_m)$, $X_n \cap X_m \in \mathcal{B}$, and $\mu(X_n \cap X_m) = \mu_1(X_n \cap X_m) + \mu_2(X_n \cap X_m) \leq \mu_1(X_n) + \mu_2(X_m) < \infty$. Hence, $\mu$ is also $\sigma$-finite.

(ii) Clearly, $\bar{\mu}(\emptyset) = a\mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\sum_{i=1}^{\infty} (a\mu)(E_i) = \sum_{i=1}^{\infty} a\mu(E_i) = a\sum_{i=1}^{\infty} \mu(E_i) = (a\mu)(E)$, where the second equality follows from the fact that $\sum_{i=1}^{\infty} \mu(E_i) = \mu(E) \leq \mu(X) < +\infty$. Hence, $a\mu$ is a finite measure on $(X, \mathcal{B})$.

(iii) Clearly, $(\nu_1 + \nu_2)(\emptyset) = \nu_1(\emptyset) + \nu_2(\emptyset) = \emptyset y$ and $(\alpha\nu_1)(\emptyset) = \alpha\nu_1(\emptyset) = \emptyset y$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then,

$$\sum_{i=1}^{\infty} \|(\nu_1 + \nu_2)(E_i)\| = \sum_{i=1}^{\infty} \|\nu_1(E_i) + \nu_2(E_i)\| \leq \sum_{i=1}^{\infty} (\|\nu_1(E_i)\| + \|\nu_2(E_i)\|) = \sum_{i=1}^{\infty} \|\nu_1(E_i)\| + \sum_{i=1}^{\infty} \|\nu_2(E_i)\| < +\infty,$$

where the last inequality follows from the fact that $\nu_1$ and $\nu_2$ are finite $\mathcal{Y}$-valued measures, and $\sum_{i=1}^{\infty} \|(\alpha\nu_1)(E_i)\| = \sum_{i=1}^{\infty} |\alpha| \|\nu_1(E_i)\| = |\alpha| \sum_{i=1}^{\infty} \|\nu_1(E_i)\| < +\infty.$
Note that \( \sum_{i=1}^\infty (\nu_1 + \nu_2)(E_i) = \sum_{i=1}^\infty (\nu_1(E_i) + \nu_2(E_i)) = \sum_{i=1}^\infty \nu_1(E_i) + \sum_{i=1}^\infty \nu_2(E_i) = \nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) \in \mathcal{Y} \), where the second equality follows from Propositions 7.23, 3.66, and 3.67, and \( \sum_{i=1}^\infty (\alpha \nu_1)(E_i) = \sum_{i=1}^\infty \alpha \nu_1(E_i) = \alpha \nu_1(E) = (\alpha \nu_1)(E) \in \mathcal{Y} \), where the second equality follows from Propositions 7.23, 3.66, and 3.67. Hence, \( \nu_1 + \nu_2 \) and \( \alpha \nu_1 \) are \( \mathcal{Y} \)-valued pre-measures on \((X, \mathcal{B})\).

\( \forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall \text{ pairwise disjoint } (E_i)_{i=1}^n \subseteq \mathcal{B} \text{ with } E = \bigcup_{i=1}^n E_i, \text{ we have } \sum_{i=1}^n \| (\nu_1 + \nu_2)(E_i) \| \leq \sum_{i=1}^n \| \nu_1(E_i) \| + \sum_{i=1}^n \| \nu_2(E_i) \| \leq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E). \) Hence, \( \mathcal{P} \circ (\nu_1 + \nu_2)(E) \leq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E). \) Then, \( \mathcal{P} \circ \nu \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2. \) Since \( \nu_1 \) and \( \nu_2 \) are finite, then \( \mathcal{P} \circ \nu(X) \leq \mathcal{P} \circ \nu_1(X) + \mathcal{P} \circ \nu_2(X) < +\infty. \) Therefore, \( \nu_1 + \nu_2 \) is a finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\).

\( \forall E \in \mathcal{B}, \) we have

\[
\mathcal{P} \circ (\alpha \nu_1)(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \atop E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| (\alpha \nu_1)(E_i) \| = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \atop E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} |\alpha| \sum_{i=1}^n \| \nu_1(E_i) \| = |\alpha| \mathcal{P} \circ \nu_1(E)
\]

where the last equality follows from Proposition 3.81 and the fact that \( 0 \leq \mathcal{P} \circ \nu_1(E) \leq \mathcal{P} \circ \nu_1(X) < +\infty. \) Hence, \( \mathcal{P} \circ (\alpha \nu_1) = |\alpha| \mathcal{P} \circ \nu_1. \) Therefore, \( \alpha \nu_1 \) is a finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\).

(iv) Clearly, \( (\mu y)(\emptyset) = \mu(\emptyset) y = 0 y. \) \( \forall \text{ pairwise disjoint } (E_i)_{i=1}^\infty \subseteq \mathcal{B}, \) let \( E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}. \) Then, \( \sum_{i=1}^\infty \| (\mu y)(E_i) \| = \sum_{i=1}^\infty \| y \| [\mu(E_i)] = \| y \| \sum_{i=1}^\infty [\mu(E_i)] < +\infty. \) Note that \( \sum_{i=1}^\infty (\mu y)(E_i) = \sum_{i=1}^\infty [\mu(E)] = \mu(E) y = (\mu y)(E) \in \mathcal{Y}, \) where the second equality follows from Propositions 7.23, 3.66, and 3.67. Hence, \( \mu y \) is a \( \mathcal{Y} \)-valued pre-measure on \((X, \mathcal{B})\). \( \forall E \in \mathcal{B}, \)

\[
\mathcal{P} \circ (\mu y)(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \atop E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| (\mu y)(E_i) \| = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E \atop E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \| y \| \sum_{i=1}^n |\mu(E_i)| = \| y \| \mathcal{P} \circ \mu(E)
\]

where the last equality follows from Proposition 3.81 and the fact that \( 0 \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu(X) < +\infty. \) Hence, \( \mathcal{P} \circ (\mu y) = \| y \| \mathcal{P} \circ \mu. \) Therefore, \( \mu y \) is a finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\).

(v) Clearly, \( \bar{\mu} = A \mu = \partial_2. \) \( \forall \text{ pairwise disjoint } (E_i)_{i=1}^\infty \subseteq \mathcal{B}, \) let \( E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}. \) Then, \( \mu(E) = \sum_{i=1}^\infty \mu(E_i) \in \mathcal{Y}, \) and \( \sum_{i=1}^\infty \| \mu(E_i) \| < \infty. \) This implies that \( \bar{\mu}(E) = A \sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty A \mu(E_i) = \sum_{i=1}^\infty \bar{\mu}(E_i) \in \mathbb{Z}, \) where the second equality follows from Proposition 3.66,
and $\sum_{i=1}^{\infty} \| \tilde{\mu}(E_i) \| \leq \sum_{i=1}^{\infty} \| A \| \| \mu(E_i) \| = \| A \| \sum_{i=1}^{\infty} \| \mu(E_i) \| < \infty$. Hence, $\tilde{\mu}$ is a $\mathcal{Z}$-valued pre-measure on $(X, \mathcal{B})$. \forall E \in \mathcal{B},

$$\mathcal{P} \circ \tilde{\mu}(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| \tilde{\mu}(E_i) \|$$

$$\leq \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| A \| \| \mu(E_i) \|$$

$$= \| A \| \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| \mu(E_i) \|$$

$$= \| A \| \mathcal{P} \circ \mu(E) \leq \| A \| \mathcal{P} \circ \mu(X) < \infty$$

where the second equality follows from Proposition 3.81. Hence, $\mathcal{P} \circ \tilde{\mu} \leq \| A \| \mathcal{P} \circ \mu$ and $\tilde{\mu}$ is a finite $\mathcal{Z}$-valued pre-measure on $(X, \mathcal{B})$. Then, $\tilde{\mu}$ is a finite $\mathcal{Z}$-valued measure on $(X, \mathcal{B})$.

This completes the proof of the proposition.

**Proposition 11.137** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $\forall$ be a normed linear space over $\mathbb{K}$, and $\nu_1$ and $\nu_2$ be $\forall$-valued measures on $(X, \mathcal{B})$. Assume that, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$, we have $\nu_1(E) = \nu_2(E) \in \forall$. Then, $\nu_1 = \nu_2$ is $\sigma$-finite.

**Proof** Since $\mu$ is $\sigma$-finite, then $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mu(X_n) < \infty, \forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. By the assumption, $\mu(X_n) < \infty$ implies that $X_n \in \text{dom}(\nu_1) \cap \text{dom}(\nu_2)$ and $\mathcal{P} \circ \nu_1(X_n) < \infty, i = 1, 2$. Hence, $\nu_1$ and $\nu_2$ are $\sigma$-finite. Let $\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite measure subspace of $\mathcal{X}$, $(X_n, \mathcal{B}_n, \nu_{n,i})$ be the finite $\forall$-valued measure subspace of $(X, \mathcal{B}, \nu_i), i = 1, 2, \forall E \in \mathcal{B}_n$, we have $\mu(E) \leq \mu(X_n) < \infty$. Then, $\nu_{n,2}(E) = \nu_1(E) = \nu_2(E) = \nu_{n,2}(E)$. Thus, $\nu_{n,1} = \nu_{n,2}$. By Proposition 11.117, we have $\nu_1 = \nu_2$.

**Proposition 11.138** Let $(X, \mathcal{B})$ be a measurable space, $\forall$ be a normed linear space over $\mathbb{K}$, and $\mathcal{Z}$ be a Banach space over $\mathbb{K}$. Then, the following statements hold.

(i) Let $\mu_1$ and $\mu_2$ be $\sigma$-finite $\mathcal{Z}$-valued measures on $(X, \mathcal{B})$. Then, there exists a unique $\mathcal{Z}$-valued measure $\mu$ on $(X, \mathcal{B})$ such that $\mu(E) = \mu_1(E) + \mu_2(E) \in \mathcal{Z}$, $\forall E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Furthermore, $\mu$ is $\sigma$-finite. We will denote this $\sigma$-finite $\mathcal{Z}$-valued measure by $\mu := \mu_1 + \mu_2$. Then, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$.

(ii) Let $\mu$ be a $\sigma$-finite measure on $(X, \mathcal{B})$ and $\alpha \in [0, \infty) \subseteq \mathbb{R}$. Then, there exists a unique measure $\tilde{\mu}$ on $(X, \mathcal{B})$ such that $\tilde{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B}$ with $\mu(E) < \infty$. Furthermore, $\tilde{\mu}$ is $\sigma$-finite. We will denote this $\sigma$-finite measure by $\tilde{\mu} := \alpha \mu$. The $\sigma$-finite measure $\tilde{\mu}$ satisfies: if $\alpha > 0$, then $\tilde{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B}$; if $\alpha = 0$, then $\tilde{\mu}(E) = 0, \forall E \in \mathcal{B}$. 
(iii) Let \( \mu \) be a \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\) and \( \alpha \in \mathbb{K} \). Then, there exists a unique \( \mathcal{Y} \)-valued measure \( \bar{\mu} \) on \((X, \mathcal{B})\) such that \( \bar{\mu}(E) = \alpha \mu(E) \), \( \forall E \in \text{dom}(\mu) \). Furthermore, \( \bar{\mu} \) is \( \sigma \)-finite. We will denote this \( \sigma \)-finite \( \mathcal{Y} \)-valued measure by \( \bar{\mu} := \alpha \mu \). The \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \bar{\mu} \) satisfies: if \( \alpha \neq 0 \), then \( \bar{\mu}(E) = \alpha \mu(E) \), \( \forall E \in \text{dom}(\mu) \), and \( \bar{\mu}(E) \) is undefined, \( \forall E \in \mathcal{B} \setminus \text{dom}(\mu) \); if \( \alpha = 0 \), then \( \bar{\mu}(E) = \emptyset \), \( \forall E \in \mathcal{B} \); and \( P \circ (\alpha \mu) = |\alpha| P \circ \mu \).

(iv) Let \( \mu \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure on \((X, \mathcal{B})\) and \( y \in \mathcal{Y} \). Then, there exists a unique \( \mathcal{Y} \)-valued measure \( \bar{\mu} \) on \((X, \mathcal{B})\) such that \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \text{dom}(\mu) \). Furthermore, \( \bar{\mu} \) is \( \sigma \)-finite. We will denote this \( \sigma \)-finite \( \mathcal{Y} \)-valued measure by \( \bar{\mu} := \mu y \). The \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \bar{\mu} \) satisfies: if \( y \neq \emptyset \), then \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \text{dom}(\mu) \), and \( \bar{\mu}(E) \) is undefined, \( \forall E \in \mathcal{B} \setminus \text{dom}(\mu) \); if \( y = \emptyset \), then \( \bar{\mu}(E) = \emptyset \), \( \forall E \in \mathcal{B} \); and \( P \circ (\mu y) = \|y\| P \circ \mu \).

(v) Let \( \mu \) be a \( \sigma \)-finite \( \mathcal{Z} \)-valued measure on \((X, \mathcal{B})\) and \( A \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \).
Then, there is a unique \( \mathcal{Z} \)-valued measure \( \bar{\mu} \) on \((X, \mathcal{B})\) such that \( \bar{\mu}(E) = A\mu(E) \), \( \forall E \in \text{dom}(\mu) \). Furthermore, \( \bar{\mu} \) is \( \sigma \)-finite and \( P \circ \bar{\mu} \leq \|A\| P \circ \mu \). We will denote this \( \sigma \)-finite \( \mathcal{Z} \)-valued measure by \( \bar{\mu} := A\mu \).

**Proof**

(i) Since \( \mu_1 \) and \( \mu_2 \) are \( \sigma \)-finite, then there exists \( (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( P \circ \mu_i(X_n) < \infty \), \( \forall n \in \mathbb{N}, i = 1, 2 \). Without loss of generality, we may assume that \( (X_n)_{n=1}^{\infty} \) is pairwise disjoint. Let \( (X_n, B_n, \mu_{i,n}) \) be the finite \( \mathcal{Z} \)-valued measure subspace of \((X, \mathcal{B}, \mu_i), \forall n \in \mathbb{N}, i = 1, 2 \). By Proposition 11.136, \( X_n := (X_n, B_n, \mu_{1,n} + \mu_{2,n}) \) is a finite \( \mathcal{Z} \)-valued measure space, \( \forall n \in \mathbb{N} \). By Proposition 11.118, the generation process on \((X_n)_{n=1}^{\infty} \) yields a unique \( \sigma \)-finite \( \mathcal{Z} \)-valued measure space \( \mathcal{X} := (X, B, \mu) \) on the set \( X \).

We will show that \( P \circ \mu_1 \leq P \circ \mu_1 + P \circ \mu_2 \), \( \forall E \in \mathcal{B}, P \circ \mu(E) \in [0, \infty) \subseteq \mathbb{R}_e \).
We will show that \( P \circ \mu(E) \leq P \circ \mu_1(E) + P \circ \mu_2(E) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( P \circ \mu_1(E) + P \circ \mu_2(E) = \infty \); Case 2: \( P \circ \mu_1(E) + P \circ \mu_2(E) < \infty \). Case 1: \( P \circ \mu_1(E) + P \circ \mu_2(E) = \infty \).
Then, it is trivially true that \( P \circ \mu(E) \leq P \circ \mu_1(E) + P \circ \mu_2(E) = \infty \). Case 2: \( P \circ \mu_1(E) + P \circ \mu_2(E) < \infty \). Then, \( P \circ \mu(E) \in [0, \infty) \subseteq \mathbb{R}_e, i = 1, 2 \).
Hence, \( E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2) \). Let \( E_n := E \cap X_n \in B_n, \forall n \in \mathbb{N} \). Then, \( P \circ \mu(E) = \sum_{n=1}^{\infty} P \circ \mu(E_n) = \sum_{n=1}^{\infty} P \circ (\mu_{1,n} + \mu_{2,n})(E_n) \leq \sum_{n=1}^{\infty} (P \circ \mu_{1,n}(E_n) + P \circ \mu_{2,n}(E_n)) = \sum_{n=1}^{\infty} P \circ \mu_1(E_n) + \sum_{n=1}^{\infty} P \circ \mu_2(E_n) = P \circ \mu_1(E) + P \circ \mu_2(E) < \infty \), where the second equality follows from Propositions 11.118 and 11.115; the first inequality follows Proposition 11.136; and the third equality follows from Proposition 11.115. Hence, \( P \circ \mu \leq P \circ \mu_1 + P \circ \mu_2 \).

\( \forall E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2) \).
Then, \( P \circ \mu(E) < \infty \) and \( \mu(E) \in \mathcal{Z}, i = 1, 2 \). This implies that \( P \circ \mu(E) \leq P \circ \mu_1(E) + P \circ \mu_2(E) < \infty \), \( E \in \text{dom}(\mu) \), and \( \mu(E) \in \mathcal{Z} \). Let \( E_n := E \cap X_n \in B_n, \forall n \in \mathbb{N} \). Then, \( \mu(E) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} (\mu_{1,n} + \mu_{2,n})(E_n) = \sum_{n=1}^{\infty} (\mu_1(E_n) + \mu_2(E_n)) = \sum_{n=1}^{\infty} \mu_1(E_n) + \sum_{n=1}^{\infty} \mu_2(E_n) = \mu_1(E) + \mu_2(E) \).
11.7. CALCULATION WITH MEASURES

\[ \mu_2(E_n) = \sum_{n=1}^{\infty} \mu_1(E_n) + \sum_{n=1}^{\infty} \mu_2(E_n) = \mu_1(E) + \mu_2(E) \in \mathcal{Z}, \]
where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.118; the third equality follows from Propositions 11.115 and 11.136; the fourth equality follows from Propositions 7.23, 3.66, and 3.67; and the last equality follows from Definition 11.108. Hence, \( \mu \) is the \( \sigma \)-finite \( \mathcal{Z} \)-valued measure we seek.

(ii) Define \( \tilde{\mu} : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+ \) by \( \tilde{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B}, \alpha > 0; \) and \( \tilde{\mu}(E) = 0, \forall E \in \mathcal{B}, \alpha = 0. \) When \( \alpha \in (0, \infty) \subset \mathbb{R}, \tilde{\mu} \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \) by Proposition 11.136. When \( \alpha = 0, \) clearly \( \tilde{\mu} \) is a finite measure on \((X, \mathcal{B}).\) Hence, \( \tilde{\mu} \) is the \( \sigma \)-finite measure we seek. Let \( \tilde{\mu} \) be any measure on \((X, \mathcal{B})\) such that \( \tilde{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B} \) with \( \mu(E) < \infty. \) Since \( \mu \) is \( \sigma \)-finite, then there exists \( (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mu(X_n) < \infty, \forall n \in \mathbb{N}. \) Without loss of generality, we may assume that \( (X_n)_{n=1}^{\infty} \) is pairwise disjoint. \( \forall E \in \mathcal{B}, \) let \( E_n := E \cap X_n, \forall n \in \mathbb{N}. \) Then, \( \mu(E_n) < \infty, \forall n \in \mathbb{N}. \) This implies that \( \tilde{\mu}(E) = \sum_{n=1}^{\infty} \tilde{\mu}(E_n) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} \tilde{\mu}(E_n) = \tilde{\mu}(E) \in [0, \infty] \subset \mathbb{R}_+. \) Hence, \( \tilde{\mu} \) is unique.

(iii) Define \( \nu := |\alpha| \mathcal{P} \circ \mu \) as prescribed in (ii), which is a \( \mu \)-finite measure on \((X, \mathcal{B}).\) Define \( \tilde{\nu} \) to be a function from \( \mathcal{B} \) to \( \mathcal{Y} \) by, if \( \alpha \neq 0, \) then \( \tilde{\nu}(E) = \alpha \mu(E) \in \mathcal{Y}, \forall E \in \text{dom}(\mu), \) and \( \tilde{\nu}(E) \) is undefined, \( \forall E \in \mathcal{B} \setminus \text{dom}(\mu); \) if \( \alpha = 0, \) then \( \tilde{\nu}(E) = \emptyset, \forall E \in \mathcal{B}. \) We will show that \( \tilde{\nu} \) is a \( \mu \)-finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\) satisfying: \( \tilde{\nu}(E) = \alpha \mu(E), \forall E \in \text{dom}(\mu), \) and \( \mathcal{P} \circ \tilde{\nu} = \nu \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \alpha = 0; \) Case 2: \( \alpha \neq 0. \)

Case 1: \( \alpha = 0. \) Then, \( \tilde{\nu}(E) = \emptyset, \) and \( \nu(E) = 0, \forall E \in \mathcal{B}. \) It is easy to show that \( \tilde{\nu}(E) \) is a finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \tilde{\nu} = \nu. \) It is straightforward to check that \( \tilde{\nu}(E) = 0 \mu(E) = \emptyset, \forall E \in \text{dom}(\mu). \) Hence, this case is proved.

Case 2: \( \alpha \neq 0. \) \( \forall E \in \mathcal{B} \) with \( \nu(E) = \infty, \) then \( \mathcal{P} \circ \mu(E) = \infty. \) This leads to that \( \mu(E) \) is undefined and \( \tilde{\nu}(E) \) is also undefined. \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty, \) then \( \mathcal{P} \circ \mu(E) = \nu(E)/|\alpha| < \infty \) and \( E \in \text{dom}(\mu). \) Then, \( \tilde{\nu}(E) = \alpha \mu(E) \in \mathcal{Y}. \) \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{\infty} E_i, \) we have \( \tilde{\nu}(E) = \alpha \mu(E) = \alpha \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \alpha \mu(E_i) = \sum_{i=1}^{\infty} \tilde{\nu}(E_i) \in \mathcal{Y} \) and \( \sum_{i=1}^{\infty} \|\tilde{\nu}(E_i)\| = \sum_{i=1}^{\infty} |\alpha| \|\mu(E_i)\| \leq \sum_{i=1}^{\infty} |\alpha| \|\mathcal{P} \circ \mu(E_i)\| = |\alpha| \|\mathcal{P} \circ \mu(E)\| < \infty. \) \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty, \)

\[ \nu(E) = |\alpha| \|\mathcal{P} \circ \mu(E)\| \]
\[ = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E=\bigcup_{i=1}^{n} E_i, E_i \cap E_j=\emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} |\alpha| \|\mu(E_i)\| \]
\[ = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E=\bigcup_{i=1}^{n} E_i, E_i \cap E_j=\emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \|\tilde{\nu}(E_i)\| \]
\[ = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E=\bigcup_{i=1}^{n} E_i, E_i \cap E_j=\emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \|\tilde{\nu}(E_i)\| \]
Proposition 11.137, \( \bar{\mu} \) proved. 

Case 1: straightforward to check that \( \bar{\mu} \) is a \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) such that \( \bar{\mu}(E) = \mu(E) \), \( \forall E \in \mathcal{B} \). Hence, \( \bar{\mu} \) is unique.

Case 2: \( \forall y \in \mathcal{Y} \), then \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \mathcal{B} \). Hence, \( \bar{\mu} = \bar{\mu} \). This case is proved.

(iv) Define \( \nu := \|y\| \mathcal{P} \circ \mu \) as prescribed in (ii), which is a \( \mathcal{Y} \)-finite measure on \( (X, \mathcal{B}) \). Define \( \bar{\mu} \) to be a function from \( \mathcal{B} \) to \( \mathcal{Y} \) by, if \( y \neq \emptyset \), then \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \mathcal{B} \). Hence, \( \bar{\mu} \) is unique.

Case 1: \( y = \emptyset \); Case 2: \( y \neq \emptyset \). 

Case 1: \( y = \emptyset \). Then, \( \bar{\mu}(E) = \emptyset \) and \( \nu(E) = 0 \), \( \forall E \in \mathcal{B} \). It is easy to show that \( \bar{\mu} \) is a \( \mathcal{Y} \)-finite measure on \( (X, \mathcal{B}) \) with \( \mathcal{P} \circ \bar{\mu} = \nu \). It is straightforward to check that \( \bar{\mu}(E) = \mu(E)y = \emptyset \), \( \forall E \in \mathcal{B} \). Hence, this case is proved.

Case 2: \( y \neq \emptyset \), \( \forall E \in \mathcal{B} \) with \( \nu(E) = \infty \), then \( \mathcal{P} \circ \mu(E) = \infty \). This leads to that \( \mu(E) \) is undefined and \( \bar{\mu}(E) \) is also undefined. \( \forall E \in \mathcal{B} \) with \( \nu(E) = \infty \), then \( \mathcal{P} \circ \mu(E) = \nu(E)/\|y\| < \infty \) and \( \nu(E) = \infty \). Then, \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \mathcal{B} \). Hence, \( \bar{\mu} = \bar{\mu} \). This case is proved.

\[
\nu(E) = \|y\| \mathcal{P} \circ \mu(E)
\]

where the third equality follows from Proposition 3.81. Hence, \( \bar{\mu} \) is a \( \mathcal{Y} \)-finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) with \( \mathcal{P} \circ \bar{\mu} = \nu = \|y\| \mathcal{P} \circ \mu \). This case is proved.

Hence, \( \bar{\mu} \) is the \( \mathcal{Y} \)-finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) we seek. Let \( \bar{\mu} \) be any \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) such that \( \bar{\mu}(E) = \mu(E)y \), \( \forall E \in \mathcal{B} \). Hence, \( \bar{\mu} = \bar{\mu} \). This case is proved.

(v) Since \( \mu \) is \( \mathcal{Y} \)-finite, then \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mathcal{P} \circ \mu(X_n) < \infty \), \( \forall n \in \mathbb{N} \). Without loss of generality, we may assume that
11.7. CALCULATION WITH MEASURES

\((X_n)_{n=1}^\infty\) is pairwise disjoint. Fix any \(n \in \mathbb{N}\). Let \(X_n := (X_n, \mathcal{B}_n, \mu_n)\) be the finite \(\mathcal{Y}\)-valued measure subspace of \(X := (X, \mathcal{B}, \mu)\). By Proposition 11.136, \(X_n := (X_n, \mathcal{B}_n, \mu_n := A\mu_n)\) is a finite \(\mathcal{Z}\)-valued measure space with \(\mathcal{P} \circ \mu_n \leq \|
u\| \mathcal{P} \circ \mu_n\). By Proposition 11.118, the generation process on \((X_n)_{n=1}^\infty\) yields a unique \(\sigma\)-finite \(\mathcal{Z}\)-valued measure \(\bar{\mathcal{X}} := (X, \mathcal{B}, \bar{\mu})\) on the set \(X\).

\[
\forall E \in \mathcal{B}, \mathcal{P} \circ \bar{\mu}(E) = \sum_{n=1}^\infty \mathcal{P} \circ \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E \cap X_n) \leq \sum_{n=1}^\infty \|
u\| \mathcal{P} \circ \mu_n(E \cap X_n). \text{ If } \|
u\| = 0, \text{ then } \mathcal{P} \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty 0 = 0 = (\|
u\| \mathcal{P} \circ \mu)(E). \text{ If } \|
u\| > 0, \text{ then } \mathcal{P} \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty \|
u\| \mathcal{P} \circ \mu_n(E \cap X_n) = \|
u\| \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E \cap X_n) = \|
u\| \mathcal{P} \circ \mu(E) = (\|
u\| \mathcal{P} \circ \mu)(E). \text{ Hence, we have } \mathcal{P} \circ \bar{\mu} \leq \|
u\| \mathcal{P} \circ \mu.
\]

\[
\forall E \in \text{dom } (\bar{\mu}), \mathcal{P} \circ \bar{\mu}(E) < \infty. \text{ Then, } \mathcal{P} \circ \bar{\mu}(E) \leq (\|
u\| \mathcal{P} \circ \mu)(E) < \infty \text{ and } E \in \text{dom } (\bar{\mu}). \text{ This leads to } \bar{\mu}(E) = \sum_{n=1}^\infty \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty \mu_n(E \cap X_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E \cap X_n) = \mathcal{P} \circ \mu_n(E \cap X_n) = \mathcal{P} \circ \mu(E), \text{ where the first equality follows from Definition 11.108; the second and the fourth equalities follow from Propositions 11.115 and 11.136; the fifth equality follows from Proposition 3.66; and the last equality follows from Definition 11.108. Hence, } \bar{\mu} \text{ is the } \sigma\text{-finite } \mathcal{Z}\text{-valued measure we seek.}
\]

This completes the proof of the proposition. \(\Box\)

Note that, for \(\sigma\)-finite measures \(\mu_1\) and \(\mu_2\), by Proposition 11.135, they are identified with \(\sigma\)-finite \(\mathbb{R}\)-valued measures \(\bar{\mu}_1\) and \(\bar{\mu}_2\), respectively. Then, the \(\sigma\)-finite measure \(\mu_1 + \mu_2\) defined in Proposition 11.136 is identified with the \(\sigma\)-finite \(\mathbb{R}\)-valued measure \(\bar{\mu}_1 + \bar{\mu}_2\) defined in Proposition 11.138. Similar statement can be made for \(\alpha \mu_1\) and \(\alpha \bar{\mu}_1\), where \(\alpha \in [0, \infty) \subset \mathbb{R}\). Hence, the definitions for \(\mu_1 + \mu_2\) and \(\alpha \mu_1\) in Propositions 11.136 and 11.138 are unambiguous.

Proposition 11.139 Let \(m, n \in \mathbb{Z}_+, \mathcal{Y}_i\) be a normed linear space over \(\mathbb{K}\), \(i = 1, \ldots, m, \mathcal{Z}_j\) be a normed linear space over \(\mathbb{K}\), \(j = 1, \ldots, n\), \(\mathcal{Y} := \prod_{i=1}^m \mathcal{Y}_i, \mathcal{Z} := \prod_{j=1}^n \mathcal{Z}_j, h : \prod_{i=1}^m \prod_{j=1}^n \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \to \mathcal{B}(\mathcal{Y}, \mathcal{Z})\) be defined by,

\[
h(A_{1,1}, A_{2,1}, \ldots, A_{n,1}, \ldots, A_{1,m}, \ldots, A_{n,m}) = \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix}
\]

\[
\forall A_{j,i} \in \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j), i = 1, \ldots, m, j = 1, \ldots, n, \text{ and } p_{j,i} : \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \to \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \text{ defined by}
\]

\[
p_{j,i} \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} = A_{j,i}
\]

\[
\forall \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}), i = 1, \ldots, m, j = 1, \ldots, n. \text{ Then, } h
and $p_{j,i}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, are bound linear functions with norm less than or equal to 1.

**Proof** It is easy to show that $h$ is a linear function. Then, we have

$$
\| h \| = \sup_{\sum_{i=1}^m \sum_{j=1}^n \| A_{j,i} \|^2 \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| 
$$

$$
= \sup_{\sum_{i=1}^m \sum_{j=1}^n \| A_{j,i} \|^2 \leq 1} \sup_{\sum_{i=1}^m \| y_i \|^2 \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \right\| 
$$

$$
\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \| A_{j,i} \|^2 \leq 1} \sup_{\sum_{i=1}^m \| y_i \|^2 \leq 1} \left( \sum_{j=1}^n \left( \sum_{i=1}^m \| A_{j,i} \| \| y_i \| \right)^2 \right)^{1/2} 
$$

$$
\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \| A_{j,i} \|^2 \leq 1} \sup_{\sum_{i=1}^m \| y_i \|^2 \leq 1} \left( \sum_{j=1}^n \left( \sum_{i=1}^m \| A_{j,i} \|^2 \right) \left( \sum_{i=1}^m \| y_i \|^2 \right) \right)^{1/2} 
$$

$$
\leq \sup_{\sum_{i=1}^m \sum_{j=1}^n \| A_{j,i} \|^2 \leq 1} \left( \sum_{j=1}^n \sum_{i=1}^m \| A_{j,i} \|^2 \right)^{1/2} \leq 1 
$$

where the second inequality follows from the Cauchy-Schwarz Inequality.

Fix any $i \in \{1, \ldots, m\}$ and any $j \in \{1, \ldots, n\}$. It is easy to show that $p_{j,i}$ is a linear function. Then, we have

$$
\| p_{j,i} \| = \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \sup_{\| y_i \| \leq 1} \| A_{j,i} y_i \| 
$$

$$
= \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \begin{bmatrix} \varnothing y_1 \\ \vdots \\ \varnothing y_m \end{bmatrix} \right\| 
$$

$$
\leq \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \begin{bmatrix} \varnothing y_1 \\ \vdots \\ \varnothing y_m \end{bmatrix} \right\| 
$$
11.7. CALCULATION WITH MEASURES

Hence,

\[ j = 1 \sum_{\|y_i\| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| y_i \| \leq 1 \]

This completes the proof of the proposition. \( \square \)

**Proposition 11.140** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( K \)-valued measure space, \( \mathcal{Y}_i \) be a normed linear space over \( K \), \( \mathcal{Z}_j \) be a Banach space over \( K \), \( W_{j,i} \subseteq B(\mathcal{Y}_i, \mathcal{Z}_j) \) be a separable subspace, \( A_{j,i} : X \to W_{j,i} \), \( i = 1, 2, j = 1, 2 \), and \( A : X \to B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \) be defined by \( A(x) = \begin{bmatrix} A_{1,1}(x) & A_{1,2}(x) \\ A_{2,1}(x) & A_{2,2}(x) \end{bmatrix} \), \( \forall x \in X \). Then, \( A \) is absolutely integrable over \( \mathcal{X} \) if, and only if, \( A_{j,i}, i = 1, 2, j = 1, 2 \), are absolutely integrable over \( \mathcal{X} \). In this case, we have

\[ \int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \int_X A_{2,1} \, d\mu \int_X A_{2,2} \, d\mu \right] \subseteq B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \]

**Proof** Let \( A_{j,i}, i = 1, 2, j = 1, 2 \), be absolutely integrable over \( \mathcal{X} \). By Proposition 11.39, \( a : X \to W_{1,1} \times W_{2,1} \times W_{1,2} \times W_{2,2} \) defined by \( a(x) = (A_{1,1}(x), A_{2,1}(x), A_{1,2}(x), A_{2,2}(x)), \forall x \in X \), is \( \mathcal{B} \)-measurable. By Propositions 11.139 and 11.38, \( A = ha \) is \( \mathcal{B} \)-measurable, where \( h \) is the bounded linear function defined in Proposition 11.139 for \( m = n = 2 \). By Propositions 11.38 and 7.21, \( \mathcal{P} \circ A \) is \( \mathcal{B} \)-measurable, \( \forall x \in X \). Hence, \( A \) is absolutely integrable over \( \mathcal{X} \). By Propositions 7.22, 3.28, and 4.4, \( \prod_{i=1}^{2} \prod_{j=1}^{2} W_{j,i} \subseteq \prod_{i=1}^{2} \prod_{j=1}^{2} \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \) is a separable subspace. By Proposition 11.132, \( A_{j,i}, i = 1, 2, j = 1, 2, a, \) and \( A \) are integrable over \( \mathcal{X} \) and \( \int_X A \, d\mu = h \int_X a \, d\mu \subseteq B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \).

On the other hand, let \( A \) be absolutely integrable over \( \mathcal{X} \). Fix any \( i \in \{1, 2\} \) and any \( j \in \{1, 2\} \). Then, by Propositions 11.39, 11.38, \( A_{j,i} = p_{j,i} A \) is \( \mathcal{B} \)-measurable, where \( p_{j,i} \) is the bounded linear function as defined in Proposition 11.139. By Propositions 7.21 and 11.38, \( \mathcal{P} \circ A_{j,i} \) is \( \mathcal{B} \)-measurable, \( \forall x \in X \). Then, by Proposition 11.83, \( 0 \leq \int_X \mathcal{P} \circ A_{j,i} \, d\mu \leq \int_X \mathcal{P} \circ A(x) \, d\mu \leq \int_X \mathcal{P} \circ A \, d\mu \leq \int_X A \, d\mu \subseteq B(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \).

Hence, \( A_{j,i} \) is absolutely integrable over \( \mathcal{X} \). By Proposition 11.132, \( A_{j,i} \) is integrable over \( \mathcal{X} \) and \( \int_X A_{j,i} \, d\mu \subseteq B(\mathcal{Y}_i, \mathcal{Z}_j) \).

Hence, \( A \) is absolutely integrable over \( \mathcal{X} \) if, and only if, \( A_{j,i}, i = 1, 2, j = 1, 2 \), are absolutely integrable over \( \mathcal{X} \).
Let $A, A_{j,i}, i = 1, 2, j = 1, 2,$ be absolutely integrable over $\mathcal{X}$. We will show that $\int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \right] \in B (\gamma_1 \times \gamma_2, Z_1 \times Z_2)$ by distinguishing two exhaustive and mutually exclusives cases: Case 1: $\mathcal{P} \circ \mu (X) < \infty$; Case 2: $\mathcal{P} \circ \mu (X) = \infty$.

Case 1: $\mathcal{P} \circ \mu (X) < \infty$. By Proposition 11.133, $\forall i \in \{1, 2\}, \forall j \in \{1, 2\}$, sequence of simple functions $(\varphi_{j,i,n})_{n=1}^{\infty}$, $\varphi_{j,i,n} : X \rightarrow W_{j,i}$, $\forall n \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \varphi_{j,i,n} = A_{j,i}$ a.e. in $\mathcal{X}$, $\| \varphi_{j,i,n} (x) \| \leq \mathcal{P} \circ A_{j,i} (x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\int_X A_{j,i} \, d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_{j,i,n} \, d\mu \in B (\gamma_1, \gamma_2)$.

By Proposition 11.53, we have $\lim_{n \in \mathbb{N}} \psi_n = a$ a.e. in $\mathcal{X}$, where $\psi_n : X \rightarrow \prod_{j=1}^{2} W_{j,i}$ is defined by $\psi_n (x) = (\varphi_{1,1,n} (x), \varphi_{2,1,n} (x), \varphi_{1,2,n} (x), \varphi_{2,2,n} (x))$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Clearly, $\psi_n$ is a simple function, $\forall n \in \mathbb{N}$.

By Lebesgue Dominated Convergence Theorem 11.31, $\int_X A \, d\mu = \lim_{n \in \mathbb{N}} \int_X \psi_n \, d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_{1,1,n} \, d\mu, \int_X \varphi_{1,2,n} \, d\mu, \int_X \varphi_{2,1,n} \, d\mu, \int_X \varphi_{2,2,n} \, d\mu) = \left( \int_X A_{1,1} \, d\mu, \int_X A_{1,2} \, d\mu, \int_X A_{2,1} \, d\mu, \int_X A_{2,2} \, d\mu \right) \in \prod_{j=1}^{2} \prod_{i=1}^{2} B (\gamma_1, \gamma_2)$, where the second equality follows from Proposition 11.75; and the third equality follows from Proposition 3.67. Then, by Proposition 11.132, $\int_X A \, d\mu = h \int_X a \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \right] \in B (\gamma_1 \times \gamma_2, Z_1 \times Z_2)$.

Case 2: $\mathcal{P} \circ \mu (X) = \infty$. Let $(A_E)_{E \in \mathcal{M} (\mathcal{X})}$ and $(A_{j,i,E})_{E \in \mathcal{M} (\mathcal{X})}$ be the nets for $\int_X A \, d\mu$ and $\int_X A_{j,i} \, d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2, j = 1, 2$. $\forall E \in \mathcal{M} (\mathcal{X})$, by Case 1, $A_E = \left[ \int_X A_{1,1} \, d\mu_{E} \int_X A_{1,2} \, d\mu_{E} \right] = \left[ A_{1,1,E} A_{1,2,E} \right] \in B (\gamma_1 \times \gamma_2, Z_1 \times Z_2)$, where $E := (E, E', \mu_E)$ is the finite $\mathbb{K}$-valued measure space of $\mathcal{X}$. Since $A, A_{j,i}, i = 1, 2, j = 1, 2$ are integrable over $\mathcal{X}$, we have $\int_X A \, d\mu = \lim_{E \in \mathcal{M} (\mathcal{X})} A_E = \lim_{E \in \mathcal{M} (\mathcal{X})} \left[ A_{1,1,E} A_{1,2,E} \right] = h \lim_{E \in \mathcal{M} (\mathcal{X})} (A_{1,1,E}, A_{2,1,E}, A_{1,2,E}, A_{2,2,E}) = h (\int_X A_{1,1} \, d\mu, \int_X A_{2,1} \, d\mu, \int_X A_{1,2} \, d\mu, \int_X A_{2,2} \, d\mu) \in B (\gamma_1 \times \gamma_2, Z_1 \times Z_2)$, where the the first equality follows from Definition 11.119; the third equality follows from Propositions 11.139 and 3.66; and the fourth equality follows from Proposition 3.67 and Definition 11.119.

Hence, in both cases, we have $\int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \right] \in B (\gamma_1 \times \gamma_2, Z_1 \times Z_2)$. This completes the proof of the proposition. \hfill \Box

**Proposition 11.141** Let $(X, \mathcal{B})$ be a measurable space, $\gamma$ be a Banach space over $\mathbb{K}$, and $Z$ be the set of $\sigma$-finite $\gamma$-valued measures on $(X, \mathcal{B})$. Define vector addition and scalar multiplication on $Z$ according to Proposition 11.138 and $\partial \in Z$ by $\partial (E) = \delta_y$, $\forall E \in \mathcal{B}$. Then, $(Z, +, \cdot, \partial) =: \mathcal{M}_{\sigma} (X, \mathcal{B}, \gamma)$ is a vector space over $\mathbb{K}$. 


Define vector addition and scalar multiplication on $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ is a vector space over $\mathbb{K}$.

Proof. 

We will first show that $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ is a vector space over $\mathbb{K}$. Fix any $\mu_1, \mu_2, \mu_3 \in Z$ and fix any $\alpha, \beta \in \mathbb{K}$.

(i) $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2)(E) < \infty$, then $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Then, we have $(\mu_1 + \mu_2)(E) = \mu_1(E) + \mu_2(E) = (\mu_2 + \mu_1)(E) \in \mathcal{Y}$. By Proposition 11.137, $\mu_1 + \mu_2 = \mu_2 + \mu_1$.

(ii) $\forall E \in \mathcal{B}$ with $(\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2 + \mathcal{P} \circ \mu_3)(E) < \infty$, then $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2) \cap \text{dom}(\mu_3)$. Then, we have $((\mu_1 + \mu_2) + \mu_3)(E) = (\mu_1 + \mu_2)(E) + \mu_3(E) = (\mu_1(E) + \mu_2(E)) + \mu_3(E) = (\mu_1(E) + (\mu_2(E) + \mu_3(E))) = \mu_1(E) + (\mu_2 + \mu_3)(E) = (\mu_1 + (\mu_2 + \mu_3))(E) \in \mathcal{Y}$. By Proposition 11.137, $(\mu_1 + \mu_2) + \mu_3 = \mu_1 + (\mu_2 + \mu_3)$.

(iii) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $(\mu_1 + \theta)(E) = \mu_1(E) + \theta(E) = \mu_1(E) \in \mathcal{Y}$. By Proposition 11.137, $\mu_1 + \theta = \mu_1$.

(iv) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2)(E) < \infty$, we have $E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$ and $E \in \text{dom}(\mu_1 + \mu_2)$. Then, $(\alpha(\mu_1 + \mu_2))(E) = \alpha(\mu_1 + \mu_2) = \alpha(\mu_1 + \mu_2) = (\alpha \mu_1 + \alpha \mu_2)(E) = (\alpha \mu_1)(E) + (\alpha \mu_2)(E) = (\alpha \mu_1)(E) + (\alpha \mu_2)(E) \in \mathcal{Y}$. By Proposition 11.137, $\alpha(\mu_1 + \mu_2) = \alpha \mu_1 + \alpha \mu_2$.

(v) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $((\alpha + \beta) \mu_1)(E) = (\alpha + \beta) \mu_1(E) = (\alpha \mu_1(E) + \beta \mu_1(E)) = (\alpha \mu_1(E) + (\alpha \mu_1(E) + \beta \mu_1(E)) = (\alpha \mu_1)(E) + (\beta \mu_1)(E) \in \mathcal{Y}$. By Proposition 11.137, $(\alpha + \beta) \mu_1 = \alpha \mu_1 + \beta \mu_1$.

(vi) $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \mu_1(E) < \infty$, we have $((\alpha \beta) \mu_1)(E) = (\alpha \beta) \mu_1(E) = (\alpha \beta \mu_1(E)) = (\alpha \beta \mu_1)(E) = (\alpha \beta \mu_1)(E) \in \mathcal{Y}$. By Proposition 11.137, $\alpha \beta \mu_1 = \alpha \beta \mu_1$.

Therefore, $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) = (Z, +, \cdot, \theta)$ is a vector space over $\mathbb{K}$. This completes the proof of the proposition.

□

Proposition 11.142 Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $Z$ be the set of finite $\mathcal{Y}$-valued measures on $(X, \mathcal{B})$. Define vector addition and scalar multiplication on $Z$ according to Proposition 11.136 and $\vartheta \in Z$ by $\vartheta(E) = \vartheta_\mathcal{Y}, \forall E \in \mathcal{B}$. Then, $Z := (Z, +, \cdot, \vartheta)$ is a vector space over $\mathbb{K}$. Define $\| \cdot \| : Z \to [0, \infty) \subset \mathbb{R}$ by $\| \mu \| = P \circ \mu(X)$, $\forall \mu \in Z$. Then, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) := (Z, \mathbb{K}, \| \cdot \|)$ is a normed linear space.

Furthermore, if $\mathcal{Y}$ is a Banach space, then $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a subspace of $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ and is a Banach space over $\mathbb{K}$.

Proof. We will first show that $Z$ is a vector space over $\mathbb{K}$. Fix any $\mu_1, \mu_2, \mu_3 \in Z$ and fix any $\alpha, \beta \in \mathbb{K}$.

(i) By Proposition 11.136, $\mu_1 + \mu_2, \mu_1 + \mu_3 \in Z$ and, $\forall E \in \mathcal{B}$, $(\mu_1 + \mu_2)(E) = \mu_1(E) + \mu_2(E) = (\mu_2 + \mu_1)(E) \in \mathcal{Y}$. Hence, $\mu_1 + \mu_2 = \mu_2 + \mu_1$.

(ii) By Proposition 11.136, $(\mu_1 + \mu_2) + \mu_3, \mu_1 + (\mu_2 + \mu_3) \in Z$ and, $\forall E \in \mathcal{B}$, $((\mu_1 + \mu_2) + \mu_3)(E) = (\mu_1 + \mu_2)(E) + \mu_3(E) = (\mu_1(E) + \mu_2(E)) + \mu_3(E) = \mu_1(E) + (\mu_2 + \mu_3)(E)$.
\[ \mu_1(E) + (\mu_2(E) + \mu_3(E)) = \mu_1(E) + (\mu_2 + \mu_3)(E) = (\mu_1 + (\mu_2 + \mu_3))(E) \in \mathcal{Y}. \]

Hence, \((\mu_1 + \mu_2) + \mu_3 = \mu_1 + (\mu_2 + \mu_3)\).

(iii) By Proposition 11.136, \(\mu_1 + \vartheta \in \mathcal{Z}\) and, \(\forall E \in \mathcal{B}, (\mu_1 + \vartheta)(E) = \mu_1(E) + \vartheta(E) = \mu_1(E) \in \mathcal{Y}\). Hence, \(\mu_1 + \vartheta = \mu_1\).

(iv) By Proposition 11.136, \(\alpha (\mu_1 + \mu_2), \alpha \mu_1 + \alpha \mu_2 \in \mathcal{Z}\) and, \(\forall E \in \mathcal{B}, \alpha (\mu_1 + \mu_2)(E) = \alpha (\mu_1 + \mu_2)(E) = \alpha (\mu_1(E) + \mu_2(E)) = \alpha \mu_1(E) + \alpha \mu_2(E) = (\alpha \mu_1)(E) + (\alpha \mu_2)(E) = (\alpha \mu_1 + \alpha \mu_2)(E) \in \mathcal{Y}\). Hence, \(\alpha (\mu_1 + \mu_2) = \alpha \mu_1 + \alpha \mu_2\).

(v) By Proposition 11.136, \((\alpha + \beta)\mu_1, \alpha \mu_1 + \beta \mu_1 \in \mathcal{Z}\) and, \(\forall E \in \mathcal{B}, ((\alpha + \beta)\mu_1)(E) = (\alpha + \beta)\mu_1(E) = \alpha \mu_1(E) + \beta \mu_1(E) = (\alpha \mu_1)(E) + (\beta \mu_1)(E) = (\alpha \mu_1 + \beta \mu_1)(E) \in \mathcal{Y}\). Hence, \((\alpha + \beta)\mu_1 = \alpha \mu_1 + \beta \mu_1\).

(vi) By Proposition 11.136, \((\alpha \beta)\mu_1, \alpha (\beta \mu_1) \in \mathcal{Z}\) and, \(\forall E \in \mathcal{B}, ((\alpha \beta)\mu_1)(E) = (\alpha \beta)\mu_1(E) = \alpha (\beta \mu_1(E) = (\alpha (\beta \mu_1))(E) \in \mathcal{Y}\). Then, \((\alpha \beta)\mu_1 = \alpha (\beta \mu_1)\).

(vii) By Proposition 11.136, \(0\mu_1, 1\mu_1 \in \mathcal{Z}\) and, \(\forall E \in \mathcal{B}, (0\mu_1)(E) = 0\mu_1(E) = \vartheta(E) \in \mathcal{Y}\) and \((1\mu_1)(E) = 1\mu_1(E) = \mu_1(E) \in \mathcal{Y}\). Hence, \(0\mu_1 = \vartheta\) and \(1\mu_1 = \mu_1\).

Therefore, \(\mathcal{Z}\) is a vector space over \(\mathbb{K}\). In the following, we show that \(\|\cdot\|\) defines a norm on \(\mathcal{Z}\). Fix any \(\mu_1, \mu_2 \in \mathcal{Z}\) and fix any \(\alpha \in \mathbb{K}\).

(i) \(\|\mu_1\| = \mathcal{P} \circ \mu_1(X) = 0, \|\vartheta\| = \mathcal{P} \circ \vartheta(E) = 0\). Clearly, \(\|\vartheta\| = \mathcal{P} \circ \vartheta(X) = 0\). If \(\|\mu_1\| = 0\), we have, \(\forall E \in \mathcal{B}, 0 \leq \|\mu_1(E)\| \leq \mathcal{P} \circ \mu_1(E) \leq \mathcal{P} \circ \mu_1(X)\), which implies that \(\mu_1(E) = \vartheta y, \forall E \in \mathcal{B}\). Then, \(\mu_1 = \vartheta\). This shows that \(\|\mu_1\| = 0 \iff \mu_1 = \vartheta\).

(ii) \(\|\mu_1 + \mu_2\| = \mathcal{P} \circ (\mu_1 + \mu_2)(X) \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2(X) = \mathcal{P} \circ \mu_1(X) + \mathcal{P} \circ \mu_2(X) = \|\mu_1\| + \|\mu_2\|\), where the inequality follows from Proposition 11.136.

(iii) By Proposition 11.136, \(\|\alpha \mu_1\| = \mathcal{P} \circ (\alpha \mu_1)(X) = |\alpha| \mathcal{P} \circ \mu_1(X) = |\alpha| \|\mu_1\|\).

Hence, \(\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\) is a normed linear space.

Let \(\mathcal{Y}\) be a Banach space. Clearly, \(\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) \subseteq \mathcal{M}_e(X, \mathcal{B}, \mathcal{Y})\). Therefore, \(\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\) is a subspace of \(\mathcal{M}_e(X, \mathcal{B}, \mathcal{Y})\). Finally, we show that \(\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\) is a Banach space when \(\mathcal{Y}\) is a Banach space. Let \((\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\) be a Cauchy sequence. \(\forall \varepsilon \in (0, +\infty) \subset \mathbb{R}\), \(\exists N \in \mathbb{N}\), \(\forall m, n \in \mathbb{N}\) with \(N \leq m\) and \(N \leq n\), we have \(\|\mu_n - \mu_m\| = \mathcal{P} \circ (\mu_n - \mu_m)(X) < \varepsilon\). \(\forall E \in \mathcal{B}, \|\mu_n(E) - \mu_m(E)\| = \|\mu_n(E) - \mu_m(E)\| \leq \mathcal{P} \circ (\mu_n - \mu_m)(E) \leq \mathcal{P} \circ (\mu_n - \mu_m)(X) < \varepsilon\). Hence, \((\mu_n)_{n=1}^\infty\) is a uniform Cauchy sequence uniform in \(E \in \mathcal{B}\). By \(\mathcal{Y}\) being complete, \(\forall E \in \mathcal{B}\), there exists a unique \(\mu(E) \in \mathcal{Y}\) such that \(\lim_{n \to \infty} \mu_n(E) = \mu(E)\). This defines a function \(\mu : \mathcal{B} \to \mathcal{Y}\). Clearly, \(\mu(\emptyset) = \lim_{n \to \infty} \mu_n(\emptyset) = \emptyset\). \(\forall\) pairwise disjoint \((E_i)_{i=1}^\infty \subseteq \mathcal{B}\), let \(E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}\). Then, \(\sum_{i=1}^\infty \|\mu(E_i)\| = \sum_{i=1}^\infty \lim_{n \to \infty} \|\mu_n(E_i)\| \leq \liminf_{n \to \infty} \sum_{i=1}^\infty \|\mu_n(E_i)\| \leq \liminf_{n \to \infty} \sum_{i=1}^\infty \mathcal{P} \circ \mu_n(E_i) = \liminf_{n \to \infty} \mathcal{P} \circ \mu_n(E) \leq \liminf_{n \to \infty} \mathcal{P} \circ \mu_n(X) = \lim_{n \to \infty} \|\mu_n\| \in \mathbb{R}\), where the first equality follows from Propositions 7.21 and 3.66; the first inequality follows from Fatou’s Lemma 11.80; the second inequality follows from Definition 11.108; and the last step
follows from Propositions 4.23 and 7.21 and the fact that \((\mu_n)_{n=1}^{\infty}\) is a Cauchy sequence. This implies that 
\[ \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \lim_{n \to \infty} \mu_n(E_i) = \lim_{m \to \infty} \sum_{i=1}^{m} \mu_n(E_i) = \lim_{m \to \infty} \sum_{i=1}^{m} \mu_n(E_i) = \lim_{m \to \infty} \mu_n(\bigcup_{i=1}^{m} E_i) = \lim_{n \to \infty} \mu_n(E) = \mu(E) \in \mathcal{Y}, \]
where fourth equality follows from Iterated Limit Theorem 4.56. Hence, \(\mu\) is a \(\mathcal{Y}\)-valued pre-measure on \((X, \mathcal{B})\).

Note that
\[
\mathcal{P} \circ \mu(X) = \sup_{n \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{n} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n \right) \sum_{i=1}^{n} \| \mu(E_i) \|
\]
\[= \sup_{n \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{n} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n \right) \lim_{m \to \infty} \sum_{i=1}^{n} \| \mu_m(E_i) \|
\]
\[\leq \sup_{n \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{n} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n \right) \lim_{m \to \infty} \sum_{i=1}^{n} \mathcal{P} \circ \mu_m(E_i)
\]
\[= \lim_{m \to \infty} \mathcal{P} \circ \mu_m(X) = \lim_{m \to \infty} \| \mu_m \| \in \mathbb{R}
\]
where the second equality follows from Propositions 7.21 and 3.66; and the last step follows from Propositions 4.23 and 7.21 and the fact that \((\mu_n)_{n=1}^{\infty}\) is a Cauchy sequence. Hence, \(\mu\) is a finite \(\mathcal{Y}\)-valued measure on \((X, \mathcal{B})\) and \(\mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}).\)

Note that, \(\forall n \in \mathbb{N}\) with \(N \leq n,\)
\[\| \mu_n - \mu \| = \mathcal{P} \circ (\mu_n - \mu)(X)
\]
\[= \sup_{m \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{m} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{m} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m \right) \sum_{i=1}^{m} \| \mu_n(E_i) - \mu(E_i) \|
\]
\[= \sup_{m \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{m} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{m} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m \right) \sum_{i=1}^{m} \| \mu_n(E_i) - \mu_k(E_i) \|
\]
\[= \sup_{m \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{m} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{m} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m \right) \lim_{k \to \infty} \sum_{i=1}^{m} \| \mu_n(E_i) - \mu_k(E_i) \|
\]
\[\leq \sup_{m \in \mathbb{Z}^+} \left( (E_i)_{i=1}^{m} \subseteq \mathcal{B}, X = \bigcup_{i=1}^{m} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq m \right) \lim_{k \to \infty} \sum_{i=1}^{m} \mathcal{P} \circ (\mu_n - \mu_k)(E_i)
\]
\[= \lim_{k \to \infty} \mathcal{P} \circ (\mu_n - \mu_k)(X) = \lim_{k \to \infty} \| \mu_n - \mu_k \| \leq \epsilon
\]
where the fourth equality follows from Propositions 7.21, 7.23, 3.66, and 3.67. Hence, \(\lim_{n \to \mathbb{N}} \| \mu_n - \mu \| = 0\) and \(\lim_{n \to \mathbb{N}} \mu_n = \mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}).\)
Thus, we have shown that every Cauchy sequence in $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ converges in $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$. Hence, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is a Banach space.

This completes the proof of the proposition.

In the above proposition, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ admits a topology $\mathcal{O}_1$ that is induced by the norm. We may define a topology on the vector space $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ when $\mathcal{Y}$ is a Banach space, whose subset topology on $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ is exactly $\mathcal{O}_1$. Let $\mathcal{O}_B$ be a collection of subsets of $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ such that, $\forall B \in \mathcal{O}_B$, $\exists E \in \mathcal{B}$, let $\mathcal{B}_E := \{ A \in \mathcal{B} \mid A \subseteq E \}$, $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \nu_E \in \mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})$ such that

$$B = \{ \mu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \mid (E, \mathcal{B}_E, \mu_E) \text{ is the finite } \mathcal{Y}\text{-valued measure subspace of } (X, \mathcal{B}, \mu) \text{ and } \| \mu_E - \nu_E \|_{\mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})} < \epsilon \}$$

It is easy to show that $\mathcal{O}_B$ is a basis for a topology $\mathcal{O}$ on $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ by Proposition 3.18. We will abuse the notation and denote the topological space $(\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}), \mathcal{O})$ by $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. With this topology in place, we may define convergence for nets in $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ according to Section 3.9. For clarity, we summarize the concept of convergence for sequences in $\mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$ in the following definition.

**Definition 11.143** Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, and $\nu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$. We will say the sequence $(\mu_n)_{n=1}^\infty$ converges to $\nu$ and write $\lim_{n \to \infty} \mu_n = \nu$ if, $\forall E \in \text{dom}(\nu)$, $\lim_{n \in \mathbb{N}} \mathcal{P} \circ (\mu_n - \nu)(E) = 0$.

This definition means that if $(\mu_n)_{n=1}^\infty$ converges to $\nu$, $\forall E \in \text{dom}(\nu)$, the subspace $(E, \mathcal{B}_E, \mu_n, \nu_E)$ of $(X, \mathcal{B}, \mu_n)$ is finite for sufficiently large $n \in \mathbb{N}$ and $\lim_{n \in \mathbb{N}} \mu_n, \nu_E = \nu_E$ in $\mathcal{M}_f(E, \mathcal{B}_E, \mathcal{Y})$, where $(E, \mathcal{B}_E, \nu_E)$ is the finite $\mathcal{Y}$-valued measure subspace of $(X, \mathcal{B}, \nu)$. Furthermore, by Proposition 11.142 and its proof and Propositions 7.21 and 3.66, we have $\lim_{n \in \mathbb{N}} \mu_n(E) = \nu(E)$ and $\lim_{n \in \mathbb{N}} \mathcal{P} \circ \mu_n(E) = \mathcal{P} \circ \nu(E)$, $\forall E \in \text{dom}(\nu)$.

**Proposition 11.144** Let $(X, \mathcal{B})$ be a measurable space, $\mu_1$ and $\mu_2$ be $\sigma$-finite measures on $(X, \mathcal{B})$, $a \in [0, \infty) \subset \mathbb{R}$, $f : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, and $\mathcal{X}_i := (X, \mathcal{B}, \mu_i), i = 1, 2$. Then, the following statements hold.

(i) If $\mu_1 \leq \mu_2$ (that is $\mu_1(E) \leq \mu_2(E), \forall E \in \mathcal{B}$), then $0 \leq \int_X f \, d\mu_1 \leq \int_X f \, d\mu_2 \leq \infty$.

(ii) Let $\mu := \mu_1 + \mu_2$ (as defined in Proposition 11.136), then $0 \leq \int_X f \, d\mu = \int_X f \, d\mu_1 + \int_X f \, d\mu_2 \leq \infty$.

(iii) Let $\nu := a\mu_1$ (as defined in Proposition 11.138), then $0 \leq \int_X f \, d\nu = \int_X (af) \, d\mu_1 \leq \infty$. 
11.7. CALCULATION WITH MEASURES

Proof (i) We will distinguish two exhaustive and mutually exclusives cases: Case 1: \( \int_X f \, d\mu_2 = \infty \); Case 2: \( \int_X f \, d\mu_2 < \infty \). Case 1: \( \int_X f \, d\mu_2 = \infty \). The result holds trivially in this case. Case 2: \( \int_X f \, d\mu_2 < \infty \). By Definition 11.79, \( 0 \leq \int_X f \, d\mu_1 = \sup_{0 \leq f \leq 1} \int_X \phi \, d\mu_1 \leq 1 \), where the supremum is over all simple functions \( \phi : X_1 \to [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). Fix any such simple function \( \phi \). Let \( \phi \) admit the canonical representation \( \phi = \sum_{i=1}^n a_i \chi_{A_i}, \) where \( n \in \mathbb{Z}_+, \) \( a_1, \ldots, a_n \in (0, \infty) \subset \mathbb{R} \) are distinct, and \( A_1, \ldots, A_n \in \mathcal{B} \) are nonempty, pairwise disjoint, and \( \mu_1(A_i) < \infty, i = 1, \ldots, n \). Then, by Proposition 11.83, \( \int_X \phi \, d\mu_1 = \sum_{i=1}^n a_i \mu_1(A_i) \in [0, +\infty) \subset \mathbb{R} \). By Proposition 11.83, \( \int_X \phi \, d\mu_2 \leq \int_X f \, d\mu_2 < \infty \).

Claim 11.144.1 \( \mu_2(A_i) < \infty, i = 1, \ldots, n \).

Proof of claim: Suppose \( \exists i_0 \in \{1, \ldots, n\} \) such that \( \mu_2(A_{i_0}) = \infty \). Since \( X_2 \) is \( \sigma \)-finite, then, by Proposition 11.7, \( \exists E \in \mathcal{B} \) with \( E \subsetneq A_{i_0} \) such that \( \frac{1}{a_{i_0}} \int_X f \, d\mu_2 < \mu_2(E) < \infty \). Choose \( \phi : X_2 \to [0, \infty) \subset \mathbb{R} \) by \( \phi = a_{i_0} \chi_{E \times \mathbb{R}} \). Clearly, \( \phi \) is a simple function of \( X_2 \) and \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). By Proposition 11.75, \( \int_X \phi \, d\mu_2 = a_{i_0} \mu_2(E) > \int_X f \, d\mu_2 \), which contradicts Definition 11.79. Hence, the claim is true.

By the claim, \( \phi \) is a simple function of \( X_2 \). By Proposition 11.75, \( \int_X \phi \, d\mu_2 = \sum_{i=1}^n a_i \mu_2(A_i) \geq \sum_{i=1}^n a_i \mu_1(A_i) = \int_X \phi \, d\mu_1 \). This implies that \( \int_X f \, d\mu_2 \geq \int_X \phi \, d\mu_2 \geq \int_X \phi \, d\mu_1 \). By the arbitrariness of \( \phi \), we have \( \int_X f \, d\mu_2 \geq \int_X f \, d\mu_1 \). Hence, in both cases, we have \( 0 \leq \int_X f \, d\mu_1 \leq \int_X f \, d\mu_2 \leq \infty \).

(ii) By Proposition 11.136, \( \mu \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \). By Proposition 11.66, there exists a sequence of simple functions \( (\phi_n)_{n=1}^\infty \), \( \phi_n : X \to [0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N} \), such that \( 0 \leq \phi_n(x) \leq f(x) < \infty, \forall x \in X, \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n = f \) a.e. in \( X \), where \( X := (X, \mathcal{B}, \mu) \). Since \( \mu = \mu_1 + \mu_2 \), then \( \mu_i \leq \mu, i = 1, 2 \). This implies that \( (\phi_n)_{n=1}^\infty \) is sequence of simple functions of \( X_1 \) and \( \lim_{n \to \infty} \phi_n = f \) a.e. in \( X_1, i = 1, 2 \). By Fatou’s Lemma 11.80 and Propositions 3.83 and 11.83, we have \( 0 \leq \int_X f \, d\mu_1 \leq \liminf_{n \to \infty} \int_X \phi_n \, d\mu_1 \leq \limsup_{n \to \infty} \int_X \phi_n \, d\mu_1 \leq \int_X f \, d\mu_2 \leq \infty \). Then, \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu_1 \in [0, \infty) \subset \mathbb{R} \). By an argument that is similar to the above, we have \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu_1 \in [0, \infty) \subset \mathbb{R} \).

(iii) We will distinguish two exhaustive and mutually exclusives cases: Case 1: \( a = 0 \); Case 2: \( a > 0 \). Case 1: \( a = 0 \). Then, by Proposition 11.138, \( \nu(E) = 0, \forall E \in \mathcal{B} \). This implies that, by Definition 11.70 Proposition 11.75, \( 0 = \int_X f \, d\nu = \int_X (af) \, d\mu_1 \). Hence, the result holds.

Case 2: \( a > 0 \). By Definition 11.79, \( \int_X (af) \, d\mu_1 = \sup_{0 \leq \phi \leq af} \int_X \phi \, d\mu_1 \),
where the supremum is over all simple functions $\phi : X_1 \to [0, \infty) \subset \mathbb{R}$ such that $0 \leq \phi(x) \leq af(x) < \infty$, $\forall x \in X$. Clearly, $\phi$ is such a simple function if, and only if, $\phi/a : X \to [0, \infty) \subset \mathbb{R}$ is a simple function of $X := (X, \mathcal{B}, \nu)$ with $0 \leq \phi(x)/a \leq f(x)$, $\forall x \in X$. Then, $\int_X \phi \, d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\nu = \int_X f \, d\nu \in [0, \infty] \subset \mathbb{R}$, where the fourth equality follows from Proposition 11.75; and the fifth equality follows from Definition 11.79.

This completes the proof of the proposition.

Proposition 11.145 Let $(X, \mathcal{B}, \mu)$ be a measurable space, $Y$ be a Banach space over $\mathbb{K}$, $\mu_1, \mu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, Y)$, $Z$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(Y, Z)$, $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, and $\mu := \mu_1 + \mu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, Y)$. Assume that $f$ is absolutely integrable over $X_i := (X, \mathcal{B}, \mu_i)$, $i = 1, 2$. Then, $f$ is absolutely integrable over $X := (X, \mathcal{B}, \mu)$ and $\int_X f \, d\mu = \int_X f \, d\mu_1 + \int_X f \, d\mu_2 \in Z$.

**Proof** By Propositions 11.138 and 11.136, $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2 =: \nu$ and $\nu$ is $\sigma$-finite. By Proposition 11.144, $0 \leq \int_X \mathcal{P} \circ f \, d\mu \leq \int_X \mathcal{P} \circ f \, d\nu = \int_X \mathcal{P} \circ f \, d\mu_1 + \int_X \mathcal{P} \circ f \, d\mu_2 < \infty$. Hence, $f$ is absolutely integrable over $X$.

By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^\infty$, $\varphi_n : X \to \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\| \varphi_n(x) \| \leq \| f(x) \|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $X$, where $X := (X, \mathcal{B}, \nu)$. Since $\mathcal{P} \circ \mu \leq \nu$ and $\mathcal{P} \circ \mu_i \leq \nu$, $i = 1, 2$, then $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $X_i$, $i = 1, 2$, and $\varphi_n$ is a simple function of $X_i$; $X_1$, or $X_2$, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, we have $\int_X f \, d\mu_1 = \lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu_1 \in Z$ and $\int_X f \, d\mu_1 = \lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu_1 \in Z$, $i = 1, 2$.

$\forall n \in \mathbb{N}$, let $\varphi_n$ admit canonical representation $\varphi_n := \sum_{i=1}^{m_n} w_{i,n} \chi_{A_{i,n}, X}$. By Proposition 11.125, we have $\int_X \varphi_n \, d\mu = \sum_{i=1}^{m_n} w_{i,n} \mu(A_{i,n}) = \sum_{i=1}^{m_n} w_{i,n}(\mu(A_{i,n}) + \mu_1(A_{i,n}) + \mu_2(A_{i,n})) = \sum_{i=1}^{m_n} w_{i,n} \mu(A_{i,n}) + \sum_{i=1}^{m_n} w_{i,n} \mu_1(A_{i,n}) + \sum_{i=1}^{m_n} w_{i,n} \mu_2(A_{i,n}) = \int_X \varphi_n \, d\mu_1 + \int_X \varphi_n \, d\mu_2 \in Z$. Then, by Propositions 7.23, 3.66, and 3.67, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_n \, d\mu_1 + \int_X \varphi_n \, d\mu_2) = \lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu_1 + \lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu_2 = \int_X f \, d\mu_1 + \int_X f \, d\mu_2 \in Z$. This completes the proof of the proposition.

**Proposition 11.146** Let $X := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $Y_1$-valued measure space, where $Y_1$ is a normed linear space over $\mathbb{K}$, $Y_2$ and $Z$ be Banach spaces over $\mathbb{K}$, $A \in \mathcal{B}(Y_1, Y_2)$, $\nu := A \mu \in \mathcal{M}_\sigma(X, \mathcal{B}, Y_2)$ as defined in Proposition 11.138, $\mathcal{W}_2$ be a separable subspace of $\mathcal{B}(Y_2, Z)$, and $f : X \to \mathcal{W}_2$ be $\mathcal{B}$-measurable. Assume that $\| A \| \mathcal{P} \circ f$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Then, $f$ is absolutely integrable over $X := (X, \mathcal{B}, \nu)$, $f A$ is absolutely integrable over $X$, and $\int_X f \, d\nu = \int_X f(A) \, d\mu \in Z$.

**Proof** By Propositions 11.38, 7.64, and 7.21, $f A$ and $\mathcal{P} \circ (f A)$ are $\mathcal{B}$-measurable. By Proposition 11.83, $0 \leq \int_X \mathcal{P} \circ (f A) \, d\mathcal{P} \circ \mu \leq$
\[ f_X(||A\circ f||)d\mathcal{P}\circ \mu < \infty. \] Thus, \( fA \) is absolutely integrable over \( \mathcal{X} \). By Proposition 11.138, \( \mathcal{P}\circ \nu \leq ||A||\mathcal{P}\circ \mu \). By Proposition 11.144, 0 \leq \int_X \mathcal{P}\circ f d\mathcal{P}\circ \nu \leq \int_X \mathcal{P}\circ f d(||A||\mathcal{P}\circ \mu) = \int_X ||A||\mathcal{P}\circ f d\mathcal{P}\circ \mu < \infty. \] Hence, \( f \) is absolutely integrable over \( \mathcal{X} \).

By Proposition 11.66, there exists a sequence of simple functions \((\varphi_n)_{n=1}^{\infty}, \varphi_n : \mathcal{X} \to \mathcal{W}_2, \forall n \in \mathbb{N}\), such that \( ||\varphi_n(x)|| \leq ||f(x)||, \forall x \in \mathcal{X}, \forall n \in \mathbb{N}\), and \( \lim_{n \to \infty} \varphi_n = f \) a.e. in \( \mathcal{X} \). Since \( \mathcal{P}\circ \nu \leq ||A||\mathcal{P}\circ \mu \), then \( \lim_{n \to \infty} \varphi_n = f \) a.e. in \( \mathcal{X} \) and \( \varphi_n \) is a simple function of \( \mathcal{X}, \forall n \in \mathbb{N}\). By Lebesgue Dominated Convergence Theorem 11.131, we have \( \int_X f d\nu = \lim_{n \to \infty} \int_X \varphi_n d\nu \in \mathcal{Z} \). Clearly, \( ||\varphi_n(x)A|| \leq ||A|| ||f(x)||, \forall x \in \mathcal{X}, \forall n \in \mathbb{N}\). Note that \( fA : \mathcal{X} \to \mathcal{W}_1 \) and \( \varphi_n A : \mathcal{X} \to \mathcal{W}_1, \forall n \in \mathbb{N}\). Without \( W_1 \) := \{ \mathcal{F} \in \mathcal{B}(\mathcal{Y}_1, \mathcal{Z}) \mid \mathcal{F} = \mathcal{F}_1 A, \mathcal{F}_1 \in \mathcal{W}_2 \} \) is a separable subspace of \( \mathcal{B}(\mathcal{Y}_1, \mathcal{Z}) \). By Proposition 11.52, we have \( \lim_{n \to \infty} \varphi_n A = fA \) a.e. in \( \mathcal{X} \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X (fA) d\mu = \lim_{n \to \infty} \int_X (\varphi_n A) d\mu \in \mathcal{Z} \). By Proposition 11.125, we have \( \int_X (fA) d\mu = \int_X f d\nu \in \mathcal{Z} \). Then, \( \int_X (fA) d\mu = \int_X f d\nu \in \mathcal{Z} \).

This completes the proof of the proposition.

Proposition 11.147 Let \( n, m \in \mathbb{Z}_+ \), \( (\mathcal{X}, \mathcal{B}) \) be a measurable space, and \( \mu_{j,i} \) be a \( \sigma \)-finite \( \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \)-valued measure on \( (\mathcal{X}, \mathcal{B}) \), where \( \mathcal{Y}_i \) is a normed linear space over \( \mathbb{K} \) and \( \mathcal{Z}_j \) is a Banach space over \( \mathbb{K} \), \( i = 1, \ldots, m \), \( j = 1, \ldots, n \). Then, there exists a unique \( \sigma \)-finite \( \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \)-valued measure \( \mu \) on \( (\mathcal{X}, \mathcal{B}) \) such that \( \mu(E) = \left[ \begin{array}{c} \mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\ \vdots & \ddots & \vdots \\ \mu_{n,1}(E) & \cdots & \mu_{n,m}(E) \end{array} \right] \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \), \( \forall E \in \bigcap_{i=1}^{m} \bigcap_{j=1}^{n} \text{dom}(\mu_{j,i}) \), where \( \mathcal{Y} := \prod_{i=1}^{m} \mathcal{Y}_i \) and \( \mathcal{Z} := \prod_{j=1}^{n} \mathcal{Z}_j \). Furthermore, \( \max_{i \in \{1, \ldots, m\}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^{m} \sum_{j=1}^{n} \mathcal{P} \circ \mu_{j,i} \). \( \mu \) will be called the vector measure and denoted by \( \begin{bmatrix} \mu_{1,1} & \cdots & \mu_{1,m} \\ \vdots & \ddots & \vdots \\ \mu_{n,1} & \cdots & \mu_{n,m} \end{bmatrix} \).

Proof Consider the special case where \( \mu_{j,i} \)'s are finite. Then, \( \mu \) is uniquely defined since \( \text{dom}(\mu_{j,i}) = \mathcal{B}, i = 1, \ldots, m \), \( j = 1, \ldots, n \).

\[
\begin{bmatrix} \mu_{1,1}(\emptyset) & \cdots & \mu_{1,m}(\emptyset) \\ \vdots & \ddots & \vdots \\ \mu_{n,1}(\emptyset) & \cdots & \mu_{n,m}(\emptyset) \end{bmatrix} = \begin{bmatrix} \partial_{\mathcal{B}(\mathcal{Y}_1, \mathcal{Z}_1)} & \cdots & \partial_{\mathcal{B}(\mathcal{Y}_m, \mathcal{Z}_1)} \\ \vdots & \ddots & \vdots \\ \partial_{\mathcal{B}(\mathcal{Y}_1, \mathcal{Z}_n)} & \cdots & \partial_{\mathcal{B}(\mathcal{Y}_m, \mathcal{Z}_n)} \end{bmatrix} = \partial_{\mathcal{B}(\mathcal{Y}, \mathcal{Z})}.
\]

\( \forall \) pairwise disjoint \( \mathcal{E}_k \subset \mathcal{B} \), we have \( \sum_{k=1}^{\infty} \| \mu(\mathcal{E}_k) \| \leq \sum_{k=1}^{\infty} \| h \| \left( \sum_{i=1}^{m} \sum_{j=1}^{n} \| \mu_{j,i}(\mathcal{E}_k) \| \right)^{1/2} \leq \sum_{k=1}^{\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} \| \mu_{j,i}(\mathcal{E}_k) \| = \sum_{i=1}^{m} \sum_{j=1}^{n} \| \mu_{j,i}(\mathcal{E}_k) \| < \infty \), where the first inequality follows from Proposition 11.139 with \( h \) being the bounded linear function defined in Proposition 11.139; and the last inequality follows from the fact that \( \mu_{j,i} \) is a finite \( \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \)-valued measure on \( (\mathcal{X}, \mathcal{B}) \) and therefore is a finite \( \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \)-valued pre-measure on \( (\mathcal{X}, \mathcal{B}) \), \( i = 1, \ldots, m \), \( j = 1, \ldots, n \). We
also have \( \sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} \left[ \begin{array}{cccc} \mu_{1,1}(E_k) & \cdots & \mu_{1,m}(E_k) \\ \vdots & \ddots & \vdots \\ \mu_{n,1}(E_k) & \cdots & \mu_{n,m}(E_k) \end{array} \right] \).  

\[
\sum_{k=1}^{\infty} \mu_{1,1}(E_k) \cdots \sum_{k=1}^{\infty} \mu_{1,m}(E_k) \\
\vdots \\
\sum_{k=1}^{\infty} \mu_{n,1}(E_k) \cdots \sum_{k=1}^{\infty} \mu_{n,m}(E_k) = \mu(E) \in B(\mathcal{Y}, \mathcal{Z}), \text{ where }
\]

\( E := \bigcup_{k=1}^{\infty} E_k \in \mathcal{B} \); the second equality follows from Propositions 3.66, 3.67, and Proposition 11.139; and the third equality follows from the fact that \( \mu_{j,i} \) is a finite \( B(\mathcal{Y}, \mathcal{Z}) \)-valued pre-measure on \((X, \mathcal{B})\), \( i = 1, \ldots, m \).  

This shows that \( \mu \) is a \( \mathcal{Y} \)-valued pre-measure on \((X, \mathcal{B})\).  

Observe that, \( \forall s \in \{1, \ldots, m\}, \forall t \in \{1, \ldots, n\}, \forall E \in \mathcal{B}, \)

\[
\mathcal{P} \circ \mu_{t,s}(E) = \sup_{k \in \mathbb{Z}_+, \{E_i\}_{i=1}^{k} \subseteq \mathcal{B}, \ E=\bigcup_{i=1}^{k} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq k} \sum_{i=1}^{k} \| \mu_{t,s}(E_i) \|
\]

\[
= \sup_{k \in \mathbb{Z}_+, \{E_i\}_{i=1}^{k} \subseteq \mathcal{B}, \ E=\bigcup_{i=1}^{k} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq k} \sum_{i=1}^{k} \| \mu_{t,s}(E_i) \|
\]

\[
\leq \sup_{k \in \mathbb{Z}_+, \{E_i\}_{i=1}^{k} \subseteq \mathcal{B}, \ E=\bigcup_{i=1}^{k} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq k} \sum_{i=1}^{k} \| \mu_{t,s}(E_i) \|
\]

\[
= \mathcal{P} \circ \mu(E) \leq \sum_{s=1}^{m} \sum_{t=1}^{n} \mathcal{P} \circ \mu_{t,s}(X) < \infty
\]

where the first equality follows from Definition 11.108; the second equality and the first inequality follow from Propositions 11.139 and 7.64; and the last equality follows from Definition 11.99.  

Observe also that, \( \forall E \in \mathcal{B}, \)

\[
\mathcal{P} \circ \mu(E) = \sup_{k \in \mathbb{Z}_+, \{E_i\}_{i=1}^{k} \subseteq \mathcal{B}, \ E=\bigcup_{i=1}^{k} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq k} \sum_{i=1}^{k} \| \mu(E_i) \|
\]

\[
\leq \sup_{k \in \mathbb{Z}_+, \{E_i\}_{i=1}^{k} \subseteq \mathcal{B}, \ E=\bigcup_{i=1}^{k} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq k} \sum_{i=1}^{k} \sum_{s=1}^{m} \sum_{t=1}^{n} \| \mu_{t,s}(E_i) \|
\]

\[
\leq \sum_{s=1}^{m} \sum_{t=1}^{n} \mathcal{P} \circ \mu_{t,s}(E) \leq \sum_{s=1}^{m} \sum_{t=1}^{n} \mathcal{P} \circ \mu_{t,s}(X) < \infty
\]

where the first equality follows from Definition 11.99; the first inequality follows from Propositions 11.139, 7.64, and 7.22; the second inequality follows from Proposition 3.81; and the second equality follows from Definition 11.108.  

Then, \( \mu \) is a finite \( \mathcal{Y} \)-valued pre-measure on \((X, \mathcal{B})\) and \( \max_{i \in \{1, \ldots, m\}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^{m} \sum_{j=1}^{n} \mathcal{P} \circ \mu_{j,i} \).  

Then, \((X, \mathcal{B}, \mu)\) is a finite \( B(\mathcal{Y}, \mathcal{Z}) \)-valued measure space.  

This completes the proof for the special case.
Consider the general case where \( \mu_j \)'s are \( \sigma \)-finite. Then, \( \exists (X_k)_{k=1}^\infty \subseteq \mathcal{B} \) such that \( X = \bigcup_{k=1}^\infty X_k \) and \( \mathcal{P} \circ \mu_j(X_k) < \infty, \forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}, \forall k \in \mathbb{N} \). Without loss of generality, assume that \( (X_k)_{k=1}^\infty \) is pairwise disjoint. Let \( \mathcal{X}_{j,i,k} := (X_k, \mathcal{B}_k, \mu_{j,i,k}) \) be the finite \( \mathcal{B} (Y, \mathcal{Z}) \)-valued measure subspace of \( (X, \mathcal{B}, \mu_j), \forall k \in \mathbb{N}, \forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\} \). Then, by the special case, there exists a unique finite \( \mathcal{B} (Y, \mathcal{Z}) \)-valued measure \( \bar{\mu}_k \) on \( (X_k, \mathcal{B}_k) \) such that \( \bar{\mu}_k(E) = \begin{bmatrix}
mu_{1,1,k}(E) & \cdots & \mu_{1,m,k}(E) \\
\vdots & \ddots & \vdots \\
\mu_{n,1,k}(E) & \cdots & \mu_{n,m,k}(E)
\end{bmatrix} \in \mathcal{B} (Y, \mathcal{Z}), \forall E \in \mathcal{B}_k \). Let \( \mathcal{X}_k := (X_k, \mathcal{B}_k, \bar{\mu}_k) \) be the finite \( \mathcal{B} (Y, \mathcal{Z}) \)-valued measure space, \( \forall k \in \mathbb{N} \). By Propositions 4.31, 7.22, and 7.66, \( \mathcal{B} (Y, \mathcal{Z}) \) is a Banach space over \( \mathbb{K} \). By Proposition 11.118, there is a unique \( \sigma \)-finite \( \mathcal{B} (Y, \mathcal{Z}) \)-valued measure space \( \mathcal{X} := (X, \mathcal{B}, \mu) \) on \( X \) such that \( \bar{\mu}_k \) is the finite \( \mathcal{B} (Y, \mathcal{Z}) \)-valued measure subspace of \( \mathcal{X}, \forall k \in \mathbb{N} \). \( \forall s \in \{1, \ldots, m\}, \forall t \in \{1, \ldots, n\}, \forall E \in \mathcal{B}_k \), we have \( \mathcal{P} \circ \mu_{j,i,k}(E) = \sum_{k=1}^\infty \mathcal{P} \circ \mu_{j,i,k}(E) \cap X_k = \sum_{k=1}^\infty \mathcal{P} \circ \bar{\mu}_k(E \cap X_k) = \sum_{k=1}^\infty \mathcal{P} \circ \mu(E \cap X_k) = \mathcal{P} \circ \mu(E) \leq \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i,k}(E) = \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E), \forall E \in \mathcal{B} (Y, \mathcal{Z}) \), where the second equality follows from Proposition 11.115: the first inequality follows from the special case; the third equality follows from Proposition 11.115; and the second inequality follows from Proposition 11.115. Then, \( \max_{s \in \{1, \ldots, m\}} \mathcal{P} \circ \mu_{j,i} \leq \mathcal{P} \circ \mu = \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}, \forall E \in \bigcap_{i=1}^m \bigcap_{j=1}^n \text{dom} (\mu_{j,i}), \mathcal{P} \circ \mu(E) \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E) < \infty \) and \( \mu(E) \in \mathcal{B} (Y, \mathcal{Z}) \). This implies that

\[
\mu(E) = \sum_{k=1}^\infty \mu(E \cap X_k) = \sum_{k=1}^\infty \bar{\mu}_k(E \cap X_k)
\]

\[
= \sum_{k=1}^\infty \begin{bmatrix}
\mu_{1,1}(E \cap X_k) & \cdots & \mu_{1,m}(E \cap X_k) \\
\vdots & \ddots & \vdots \\
\mu_{n,1}(E \cap X_k) & \cdots & \mu_{n,m}(E \cap X_k)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\sum_{k=1}^\infty \mu_{1,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{1,m}(E \cap X_k) \\
\vdots & \ddots & \vdots \\
\sum_{k=1}^\infty \mu_{n,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{n,m}(E \cap X_k)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\
\vdots & \ddots & \vdots \\
\mu_{n,1}(E) & \cdots & \mu_{n,m}(E)
\end{bmatrix}
\]

where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.115; the third equality follows from the special case; the fourth equality follows from Propositions 11.139, 3.66, and 3.67; the fifth equality follows from Definition 11.108. This completes the proof.
of the proposition.

**Proposition 11.148** Let \( X_i := (X, \mathcal{B}, \mu_i) \) be a \( \sigma \)-finite \( \mathcal{Y}_i \)-valued measure space, where \( \mathcal{Y}_i \) is a Banach space over \( \mathbb{K} \), \( \mathcal{Z}_j \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W}_{j,i} \subseteq \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \) be a separable subspace, \( f_{j,i} : X \to \mathcal{W}_{j,i}, i = 1, 2, j = 1, 2, \mu := \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \) be the vector measure as defined in Proposition 11.147, \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be the \( \sigma \)-finite \( \mathcal{Y}_1 \times \mathcal{Y}_2 \)-valued measure space, and \( F : X \to \mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \) be defined by \( F(x) = \begin{bmatrix} f_{1,1}(x) & f_{1,2}(x) \\ f_{2,1}(x) & f_{2,2}(x) \end{bmatrix}, \forall x \in X \).

Assume that \( F \) is absolutely integrable over \( \mathcal{X} \). Then, \( f_{j,i}, j = 1, 2 \) are absolutely integrable over \( \mathcal{X}_i \) for some \( i \in \{1, 2\} \), \( \psi := \begin{bmatrix} \psi_{1,1} \\ \psi_{1,2} \\ \psi_{2,1} \\ \psi_{2,2} \end{bmatrix} \) be the vector measure as defined in Proposition 11.147, \( \mathcal{X}_i \) be the separable subspace. Let \( \mathcal{W}_{j,i} \subseteq \mathcal{W}_{j,i}(\mathcal{Y}_i, \mathcal{Z}_j) \) be a separable subspace. Let \( h \) be the bounded linear function as defined in Proposition 11.139 for \( n = m = 2 \). Then, \( \mathcal{W} := h(\prod_{j=1}^2 \mathcal{W}_{j,i}) \subseteq \mathcal{B}(\mathcal{Y}_1 \times \mathcal{Y}_2, \mathcal{Z}_1 \times \mathcal{Z}_2) \) is a separable subspace. By Proposition 11.66, \( \mathcal{Y}_1 \) sequence of simple functions \( \psi_n : X \to \mathcal{W}, \forall n \in \mathbb{N} \), such that \( \lim_{n \to \infty} \psi_n F = F \) a.e. in \( \mathcal{X} \), \( \| \psi_n F \| \leq \| F \|, \forall x \in X, \forall n \in \mathbb{N} \). Fix any \( i \in \{1, 2\} \) and any \( j \in \{1, 2\} \). Let \( \varphi_{j,i,n} := p_{j,i} \psi_n \). Then, \( \varphi_{j,i,n} : \mathcal{X} \to \mathcal{W}_i \subseteq \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \) is a simple function. By Proposition 11.52, \( \lim_{n \to \infty} \varphi_{j,i,n} F = F \) a.e. in \( \mathcal{X} \). Since \( \mathcal{P} \circ \mu \geq \mathcal{P} \circ \mu_i \), we have \( \varphi_{j,i,n} : \mathcal{X}_i \to \mathcal{W}_i \subseteq \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \) is a simple function, \( \forall n \in \mathbb{N} \). \( \lim_{n \to \infty} \varphi_{j,i,n} F_{j,i} \) a.e. in \( \mathcal{X}_i \), and \( F \) is absolutely integrable over \( \mathcal{X}_i \), by Proposition 11.144. Note that \( \| \varphi_{j,i,n} F \| \leq \| p_{j,i} \| \leq \| \psi_n \| \leq \| F \| , \forall x \in X, \forall n \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X F d\mu = \lim_{n \to \infty} \int_X \psi_n d\mu = \lim_{n \to \infty} \left[ \int_X \varphi_{1,1,n} F d\mu_1 + \int_X \varphi_{1,2,n} F d\mu_2 \right] \right] \in \mathcal{Z}_1 \times \mathcal{Z}_2 \).
This completes the proof of the proposition. \qed

**Theorem 11.149 (Fatou’s Lemma)** Let \((X, \mathcal{B})\) be a measurable space, \((\mu_n)_{n=1}^\infty\) be a sequence of measures on \((X, \mathcal{B})\), \(\mu\) be a measure on \((X, \mathcal{B})\), \(f_n : X \to [0, \infty] \subset \mathbb{R}_+\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), \(f : X \to [0, \infty] \subset \mathbb{R}_+\) be \(\mathcal{B}\)-measurable. Assume that \(\lim_{n \to \infty} f_n = f\) a.e. in \(X\), where \(X := (X, \mathcal{B}, \mu)\), and \(\lim_{n \to \infty} \mu_n = \mu\) \((\text{weak})\), \(\forall E \in \mathcal{B}\) with \(\mu(E) < \infty\). Then, 
\[
0 \leq \int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu_n \leq \infty.
\]

**Proof** By Definition 11.79, \(\int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu\), where the supremum is over all simple functions \(\phi : X \to [0, \infty] \subset \mathbb{R}\) that satisfies \(0 \leq \phi(x) \leq f(x)\), \(\forall x \in X\). Fix any such simple function \(\phi\). Then, \(\exists M \in [0, \infty) \subset \mathbb{R}\) such that \(\phi(x) \leq M\), \(\forall x \in X\) and \(\mu(E) := \mu(\{x \in X \mid \phi(x) > 0\}) < \infty\). By \(\lim_{n \to \infty} \mu_n(E) = \mu(E) < \infty\), \(\exists n_0 \in \mathbb{N}\) such that \(\mu_n(E) < \infty\), \(\forall n \in \mathbb{N}\) with \(n_0 \leq n\). Thus, \(\forall n \in \mathbb{N}\) \(\exists \phi_n\) that satisfies \(0 \leq \phi_n(x) \leq f_n(x)\), \(\forall x \in X\) and \(\mu_n(E) \leq \mu_n(E) \leq \mu_n(E) < \infty\), \(\forall n \in \mathbb{N}\). By Proposition 11.75, \(\int_X \phi \, d\mu = : c_1 \in [0, \infty) \subset \mathbb{R}\). Let \(E := \{x \in X \mid f_n(x) \to f(x)\} \not\to\). By the assumption \(\hat{E} \in \mathcal{B}\) and \(\mu(\hat{E}) = 0\).

Fix any \(\epsilon \in (0, 1) \subset \mathbb{R}\). Define \(E_n := \left\{x \in X \mid f_n(x) \geq (1 - \frac{\epsilon}{2 + 2c_1})\phi(x)\right\},\) \(\forall n \in \mathbb{N}\). By Propositions 7.23, 11.38, and 11.39, \(E_n \in \mathcal{B}\), \(\forall n \in \mathbb{N}\). Clearly, \(\forall x \in X \setminus (\hat{E} \cup \hat{E})\), \(\forall E_n \in \mathcal{B}\), \(\forall n \in \mathbb{N}\). Then, \(X \setminus E_n \subseteq \hat{E} \cup \hat{E},\) \(\forall n \in \mathbb{N}\). By Proposition 11.5, we have \(\lim_{n \to \infty} \mu(X \setminus E_n) = \mu(\hat{E}) = 0\). Then, \(\exists n_1 \in \mathbb{N}\) with \(n_0 \leq n_1\) such that \(\mu(X \setminus E_{n_1}) < \frac{\epsilon}{4 + 4\epsilon^2}\). By the assumption \(\lim_{n \to \infty} \mu(X \setminus E_{n_1}) = \mu(X \setminus E_{n_1}) < \frac{\epsilon}{4 + 4\epsilon^2}\). Then, \(\exists n_2 \in \mathbb{N}\) with \(n_1 \leq n_2\), \(\forall n \in \mathbb{N}\) with \(n_2 \leq n\), we have \(\mu_n(X \setminus E_{n_2}) < \frac{\epsilon}{4 + 4\epsilon^2}\). Fix any \(n \in \mathbb{N}\) with \(n_2 \leq n\), \(X \setminus E_{n_1} \supseteq X \setminus E_n\), which implies that \(0 \leq \mu(X \setminus E_{n_1}) \leq \mu(X \setminus E_{n_1}) < \frac{\epsilon}{4 + 4\epsilon^2}\). Thus, \(\forall n \in \mathbb{N}\) \(\exists \phi_n\) that satisfies \(0 \leq \phi_n(x) \leq f_n(x)\), \(\forall x \in X\) and \(\mu_n(E) \leq \mu_n(X \setminus E_{n_1}) \leq \mu(X \setminus E_{n_1}) < \frac{\epsilon}{4 + 4\epsilon^2}\).

This completes the proof of the lemma. \qed

**Theorem 11.150 (Lebesgue Dominated Convergence)** Let \((X, \mathcal{B})\) be a measurable space, \((\mu_n)_{n=1}^\infty\) be a sequence of measures on \((X, \mathcal{B})\), \(\mu\) be a measure on \((X, \mathcal{B})\), \(f_n : X \to \mathbb{R}\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), \(f : X \to \mathbb{R}\)
be $\mathcal{B}$-measurable, $g_n : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, $g : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable. Assume that

(i) $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\mathcal{X}$, where $\mathcal{X} := (X, \mathcal{B}, \mu)$;

(ii) $|f_n(x)| \leq g_n(x)$, $\forall x \in X$, $\forall n \in \mathbb{N};$

(iii) $\int_X g_n \, d\mu_n \in \mathbb{R}$, $\forall n \in \mathbb{N}$, and $\int_X g \, d\mu = \lim_{n \in \mathbb{N}} \int_X g_n \, d\mu_n \in \mathbb{R};$

(iv) $\lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E)$, $\forall E \in \mathcal{B}$ with $\mu(E) < \infty$.

Then, $f$ is absolutely integrable over $\mathcal{X}$, $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \in \mathbb{R}$, and $0 \leq |\int_X f \, d\mu| \leq \int_X g \, d\mu < \infty$.

Proof By Proposition 11.92, $f_n$ is absolutely integrable over $\mathcal{X}_n := (X, \mathcal{B}, \mu_n)$ and $\int_X f_n \, d\mu_n \in \mathbb{R}$. Clearly, $|f(x)| \leq g(x)$ a.e. in $\mathcal{X}$. By Proposition 11.93, $f$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, $\int_X f \, d\mu \in \mathbb{R}$ and $0 \leq |\int_X f \, d\mu| \leq \int_X g \, d\mu \leq \int_X f \, d\mu < \infty$.

By Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (g_n - f_n) = g - f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} (g_n + f_n) = g + f$ a.e. in $\mathcal{X}$. Clearly, $(g_n - f_n)(x) \geq 0$ and $(g_n + f_n)(x) \geq 0$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Propositions 7.23, 11.52, and 11.53, $\lim_{n \in \mathbb{N}} (g_n - f_n) = \lim_{n \in \mathbb{N}} (g_n + f_n) = 0 = (g - f) \lor 0 = (g + f) \lor 0 = g + f$ a.e. in $\mathcal{X}$. By Fatou’s Lemma 11.149 and Propositions 11.92 and 3.83, we have

$$0 \leq \liminf_{n \in \mathbb{N}} \int_X (g_n - f_n) \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X g_n \, d\mu_n - \int_X f_n \, d\mu_n = g \, d\mu - \int_X f \, d\mu \leq \int_X g \, d\mu = \limsup_{n \in \mathbb{N}} \int_X g_n \, d\mu_n + \int_X f_n \, d\mu_n \leq \int_X g \, d\mu + \int_X f \, d\mu.$$

Then, by Proposition 3.83, we have

$$\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \leq \int_X f \, d\mu \leq \limsup_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \leq \int_X f \, d\mu \in \mathbb{R}.$$ 

This further implies that $\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \in \mathbb{R}$. This completes the proof of the theorem. \qed

In the preceding two results, we work with measure spaces which may not be $\sigma$-finite, and we assume a mode of convergence for the measures to be set-wise convergence rather than the stronger “norm” convergence alluded to in Definition 11.143. The next result shows that Lebesgue Dominated Convergence Theorem holds in general under the “norm” convergence assumption on the Banach space valued measures.

**Theorem 11.151 (Lebesgue Dominated Convergence)** Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $(\mu_n)_{n=1}^{\infty} \subseteq \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\mu \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $\mathcal{W} := \text{span}\{\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \cdot z = wy\}\} \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_n : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, $f : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $g_n : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, $g : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable. Assume that
11.7. Calculation with Measures

(i) \( \lim_{n \to \infty} f_n = f \) a.e. in \( \mathcal{X} \) and \( \lim_{n \to \infty} g_n = g \) a.e. in \( \mathcal{X} \), where \( \mathcal{X} := (X, \mathcal{B}, \mu) \);

(ii) \( \|
fn(x)\| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}; \)

(iii) \( I_n g_n \mathcal{P} \circ \mu_n \in \mathbb{R}, \forall n \in \mathbb{N}, \) and \( \int_X g \mathcal{P} \circ \mu = \lim_{n \to \infty} \int_X g_n \mathcal{P} \circ \mu_n \in \mathbb{R}; \)

(iv) \( \lim_{n \to \infty} \mu_n = \mu \) as defined in Definition 11.143.

Then, \( f \) is absolutely integrable over \( \mathcal{X} \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu_n \in \mathcal{W} \subseteq \mathcal{Z} \), and \( 0 \leq \|
\int_X f \, d\mu \| \leq \int_X g \mathcal{P} \circ \mu < \infty. \)

**Proof**

By Proposition 11.132, \( f_n \) is absolutely integrable over \( \mathcal{X}_n := (X, \mathcal{B}, \mu_n) \) and \( \int_X f_n \, d\mu_n \in \mathcal{W} \). Clearly, \( \|f(x)\| \leq g(x) \) a.e. in \( \mathcal{X} \). By Proposition 11.83, \( f \) is absolutely integrable over \( \mathcal{X} \). By Proposition 11.132, \( \int_X f \, d\mu \in \mathcal{W} \) and \( 0 \leq \|
\int_X f \, d\mu \| \leq \int_X g \mathcal{P} \circ \mu \leq \int_X g \mathcal{P} \circ \mu < \infty. \)

We will distinguish two exhaustive and mutually exclusives cases: Case 1: \( \mathcal{P} \circ \mu(X) < \infty \); Case 2: \( \mathcal{P} \circ \mu(X) = \infty \).

Case 1: \( \mathcal{P} \circ \mu(X) < \infty \). By \( \lim_{n \to \infty} \mu_n = \mu \), we have \( \lim_{n \to \infty} \mathcal{P} \circ \mu_n(X) = \mathcal{P} \circ \mu(X) < \infty \). Without loss of generality, we may assume that \( \mathcal{P} \circ \mu_n(X) < \infty \), \( \forall n \in \mathbb{N} \). Fix any \( \varepsilon \in (0, \infty) \subseteq \mathbb{R} \). By Proposition 11.84, \( \exists \delta \in (0, \varepsilon/6) \subseteq \mathbb{R} \) such that \( \forall A \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(A) < \delta \), we have \( \int_A g \mathcal{P} \circ \mu < \varepsilon/6. \)

By Egoroff’s Theorem 11.55, \( \exists \tilde{E} \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(\tilde{E}) < \delta/2 \) such that \( \left( f_n|_{X \setminus \tilde{E}} \right)_{n=1}^\infty \) converges uniformly to \( f|_{X \setminus \tilde{E}} \). Then, \( \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( \|f_n(x) - f(x)\| < \frac{\varepsilon}{4\mathcal{P} \circ \mu(X) + 1}, \forall x \in X \setminus \tilde{E}. \) By Proposition 11.5, there exists \( M \in \mathbb{N} \) such that \( E_1 := \{ x \in X \mid g(x) \geq M \} \in \mathcal{B} \) and \( \mathcal{P} \circ \mu(E_1) < \delta/2. \) Let \( E := \tilde{E} \cup E_1 \in \mathcal{B}. \) Then, \( \mathcal{P} \circ \mu(E) < \delta. \) Hence, \( \int_E g \mathcal{P} \circ \mu < \varepsilon/6. \)

By Lebesgue Dominated Convergence Theorem 11.150, we have \( \lim_{n \to \infty} \int_X g_n \mathcal{P} \circ \mu_n = \lim_{n \to \infty} \int_X g_n \mathcal{X}, \mathcal{E}, \cdot, \mu_n \mathcal{P} \circ \mathcal{E}(X) \mathcal{P} \circ \mu = \int_X g \mathcal{P} \circ \mu < \varepsilon/6. \)

Then, \( \exists n_1 \in \mathbb{N} \) with \( n_0 \leq n_1, \forall n \in \mathbb{N} \) with \( n_1 \leq n \), we have \( 0 \leq \int_{E_1} g \mathcal{P} \circ \mu_n < \varepsilon/3. \) By the assumption (iv), \( \lim_{n \to \infty} \mathcal{P} \circ (\mu_n - \mu)(X) = 0. \) This implies that \( \exists n_2 \in \mathbb{N} \) with \( n_1 \leq n_2, \forall n \in \mathbb{N} \) with \( n_2 \leq n \), we have \( \mathcal{P} \circ (\mu_n - \mu)(X) < \frac{\varepsilon}{3M^2} \) and \( \mathcal{P} \circ \mu_n(X) < \mathcal{P} \circ \mu(X) + 1/4. \)

Fix any \( n \in \mathbb{N} \) with \( n_2 \leq n \). Clearly, \( \int_{X \setminus E} f \, d\mu_n \in \mathcal{Z} \), since \( \|f(x)\| \leq M \) a.e. in \( \mathcal{E}_n \), where \( \mathcal{E}_n := (X \setminus E, \mathcal{B}_{X \setminus E}, \mu_n, X \setminus E) \) is the finite \( \mathfrak{L} \)-valued measure subspace of \( \mathcal{X}_n. \) We have

\[
\|\int_X f_n \, d\mu_n - \int_X f \, d\mu\|
\[
= \|\int_E f_n \, d\mu_n + \int_{X \setminus E} f_n \, d\mu_n - \int_X f \, d\mu_n + \int_X f \, d\mu_n
\[
- \int_{X \setminus E} f \, d\mu - \int_E f \, d\mu\|.
\]
Let $\nu\in\mathcal{M}(X)$ be absolutely integrable over $X$. Hence, we have $\lim_{n\to\infty}\int_X f_n\,d\nu_n = \int_X f\,d\nu \in \mathcal{W}$. This case is proved.

Case 2: $P\circ\mu(X) = \infty$. Fix any $\epsilon \in (0, \infty) \cap \mathbb{R}$. By Definition 11.79, there exists a simple function $\phi : X \to [0, \infty) \cap \mathbb{R}$ such that $0 \leq \phi(x) \leq g(x)$, $\forall x \in X$ and $\int_X g\,d\mu = \int_X \phi\,d\mu < \int_X g\,d\mu < \infty$. Let $A_0 := \{ x \in X \mid \phi(x) > 0 \}$. Then, $A_0 \in \mathcal{B}$ and $P(\mu(A_0)) < \infty$. Note that, by Propositions 11.83 and 11.75, $0 \leq \int_{X\setminus A_0} g\,d\mu = \int_X g\,d\mu - \int_{A_0} g\,d\mu \leq \int_X g\,d\mu - \int_{A_0} \phi\,d\mu = \int_X g\,d\mu - \int_X \phi\,d\mu < \epsilon/4$. This implies that $0 \leq \int_X f\,d\mu - \int_{A_0} f\,d\mu = \int_{X\setminus A_0} f\,d\mu = \int_{X\setminus A_0} (P\circ\mu - \phi)\,d\mu$. This follows from Proposition 11.132; and the third inequality follows from Proposition 11.83. Hence, we have $\lim_{n\to\infty}\int_X f_n\,d\mu_n = \int_X f\,d\nu \in \mathcal{W}$. This case is proved.

This completes the proof of the theorem.

**Proposition 11.152** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : X \to \mathcal{W}$ be absolutely integrable over $X$. Define $\nu : \mathcal{B} \to \mathcal{Z}$ by $\nu(E) = \int_E f\,d\mu$, $\forall E \in \mathcal{B}$. Then, $\tilde{\mathcal{X}} := (X, \mathcal{B}, \nu)$ is a finite $\mathcal{Z}$-valued measure space and $0 \leq P \circ \nu(E) \leq \int_E P \circ f\,d\mu \leq 4\int_X f\,d\mu$.
\[ f \in \mathcal{P} \circ f \mathcal{P} \circ \mu < \infty, \forall E \in \mathcal{B}. \]

**Proof**

By Proposition 11.125, \( \nu(\emptyset) = \int_\emptyset f d\mu = 0 \). For every \( E \in \mathcal{B} \), by Proposition 11.132, \( \nu(E) = \int_E f d\mu \in \mathcal{Z} \) and \( 0 \leq \| \nu(E) \| \leq \int_E f d\mathcal{P} \circ \mu < \infty \). Since Proposition 11.125, \( \nu \) is a \( \mathcal{Z} \)-valued measure. Then, there exists a unique \( \mathcal{Z} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \). Hence, \( \nu \) is a finite \( \mathcal{Z} \)-valued measure on \( (X, \mathcal{B}) \). For every \( E \in \mathcal{B} \), \( \mathcal{P} \circ \nu \) is a finite \( \mathcal{Z} \)-valued measure. Thus, \( \nu \) is a finite \( \mathcal{Z} \)-valued measure on \( (X, \mathcal{B}) \). This completes the proof of the proposition. \( \Box \)

**Proposition 11.153**

Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathbb{K} \), \( \mathcal{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \), and \( f : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable. Then, there exists a unique \( \mathcal{Z} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) such that \( \nu(E) = \int_E f d\mu \in \mathcal{Z}, \forall E \in \mathcal{B} \) with \( \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu < \infty \). Furthermore, \( \nu \) is \( \sigma \)-finite and \( 0 \leq \mu \circ \nu \leq \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu, \forall E \in \mathcal{B} \).

**Proof**

Since \( \mathcal{X} \) is \( \sigma \)-finite, there exists \( \exists(X_n)_{n=1}^\infty \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^\infty X_n \) and \( \mathcal{P} \circ \mu(X_n) < \infty \), \( \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( (X_n)_{n=1}^\infty \) is pairwise disjoint. For every \( m \in \mathbb{N} \), let \( \bar{X}_m := \{ x \in X \mid m - 1 < \| f(x) \| < m \} \subseteq \mathcal{B} \). Clearly, \( (\bar{X}_m)_{m=1}^\infty \subseteq \mathcal{B} \) is pairwise disjoint and \( X = \bigcup_{m=1}^\infty \bar{X}_m \). Clearly, \( \mu_{m,n} := (X_m \cap X_n, \mathcal{B}_{m,n}, \nu_{m,n}) \) is the finite \( \mathcal{Z} \)-valued measure subspace of \( \mathcal{X} \). Then, \( f|_{\bar{X}_m \cap X_n} \) is absolutely integrable over \( \mathcal{X}_{m,n} \). By Propositions 11.152 and 11.112, we may define a finite \( \mathcal{Z} \)-valued measure space \( \mathcal{X}_{m,n} := (\bar{X}_m \cap X_n, \mathcal{B}_{m,n}, \nu_{m,n}) \) by \( \nu_{m,n}(E) = \int_E f d\mu \in \mathcal{Z}, \forall E \in \mathcal{B}_{m,n} \). Clearly, the sets \( \bar{X}_m \cap X_n, m \in \mathbb{N}, n \in \mathbb{N} \), are pairwise disjoint. Then, by Proposition 11.118, the generation process on \( (\mathcal{X}_{m,n})_{m \in \mathbb{N}, n \in \mathbb{N}} \) yields the unique \( \sigma \)-finite \( \mathcal{Z} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) such that \( \mathcal{X}_{m,n} \) is a finite \( \mathcal{Z} \)-valued measure subspace of \( \mathcal{X} := (X, \mathcal{B}, \nu), \forall m, n \in \mathbb{N} \).

Fix any \( E \in \mathcal{B} \) with \( \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu < \infty \) for every \( m, n \in \mathbb{N} \), let \( E_{m,n} := E \cap \bar{X}_m \cap X_n \in \mathcal{B} \). Then, \( \mathcal{P} \circ \nu(E) = \sum_{m=1}^\infty \sum_{n=1}^\infty \mathcal{P} \circ \nu_{m,n}(E_{m,n}) \leq \sum_{m=1}^\infty \sum_{n=1}^\infty \int_{E_{m,n}} \mathcal{P} \circ f d\mathcal{P} \circ \mu = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu < \infty \), where the first equality follows from Proposition 11.118; the first inequality follows from Proposition 11.152; and the second equality follows from Proposition 11.83. This implies that \( E \in \text{dom} \nu \) and \( \nu(E) = \sum_{m=1}^\infty \sum_{n=1}^\infty \nu_{m,n}(E_{m,n}) = \sum_{m=1}^\infty \sum_{n=1}^\infty \int_{E_{m,n}} f d\mu = \int_E f d\mu \in \mathcal{Z}, \) where the first equality follows from Proposition 11.118; the second equality follows from Proposition 11.152;
and the third equality follows from Proposition 11.132. Hence, \( \nu \) is the \( \sigma \)-finite \( \mathcal{Z} \)-valued measure we seek.

This completes the proof of the proposition.

**Proposition 11.154** Let \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( \mathcal{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B}(\mathbb{K}, \mathcal{Z}) = \mathcal{Z} \), and \( f : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable. Then, there exists a unique \( \mathcal{Z} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) such that \( \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \in [0, \infty] \subseteq \mathbb{R}_+, \forall E \in \mathcal{B}, \) and \( \nu(E) = \int_E f \, d\mu \in \mathcal{Z}, \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \nu(E) < \infty \). Furthermore, \( \nu \) is \( \sigma \)-finite.

**Proof** By Proposition 11.153, there exists a unique \( \mathcal{Z} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) such that \( \nu(E) = \int_E f \, d\mu \in \mathcal{Z}, \forall E \in \mathcal{B} \) with \( \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty \). Furthermore, \( \nu \) is \( \sigma \)-finite and \( 0 \leq \mathcal{P} \circ \nu(E) \leq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu, \forall E \in \mathcal{B} \). All we need to show is that \( \mathcal{P} \circ \nu(E) \geq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu, \forall E \in \mathcal{B} \).

Define \( \tilde{\nu} : \mathcal{B} \to [0, \infty] \subseteq \mathbb{R}_+ \) by \( \tilde{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu, \forall E \in \mathcal{B} \). By Proposition 11.116, \( \tilde{\nu} \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \). Fix any \( E \in \mathcal{B} \).

We will distinguish two exhaustive and mutually exclusives cases: Case 1: \( E \in \mathcal{B} \setminus \text{dom}(\nu) \); Case 2: \( E \in \text{dom}(\nu) \). Clearly, we have \( \mathcal{P} \circ \nu(E) = \infty \geq \tilde{\nu}(E) \). This case is proved. Case 2: \( E \in \text{dom}(\nu) \).

Since \( \tilde{\nu} \) is \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^\infty \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^\infty X_n \) and \( \tilde{\nu}(X_n) < \infty, \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( (X_n)_{n=1}^\infty \) is pairwise disjoint. Let \( E_n := E \cap X_n, \forall n \in \mathbb{N} \). Then, \( \forall n \in \mathbb{N} \), we have \( \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \tilde{\nu}(E_n) < \infty, \forall n \in \mathbb{N} \), \( \forall F \in \mathcal{B} \) with \( F \subseteq E_n \), we have \( 0 \leq \mathcal{P} \circ \nu(F) \leq \int_F \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \tilde{\nu}(F) \leq \tilde{\nu}(E_n) < \infty \) and \( \nu(F) = \int_F f \, d\mu \in \mathcal{Z} \). Fix any \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \). Fix any \( n \in \mathbb{N} \). Let \( E_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n}) \) be the \( \sigma \)-finite \( \mathbb{K} \)-valued measure subspace of \( X \). By Proposition 11.94, \( \exists \) simple function \( \varphi_n : E_n \to \mathcal{W} \) such that \( \| \varphi_n(x) \| \leq \mathcal{P} \circ f(x), \forall x \in E_n \), \( \int_{E_n} \mathcal{P} \circ \varphi_n \, d\mathcal{P} \circ \mu - \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < 2^{-n-2}\epsilon \), and \( \int_{E_n} \mathcal{P} \circ (f|_{E_n} - \varphi_n) \, d\mathcal{P} \circ \mu < 2^{-n-1}\epsilon \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m_n} w_{i,n} \chi_{E_{i,n}}, \) where \( m_n \in \mathbb{Z}_+, w_{1,n}, \ldots, w_{m_n}, n \in \mathcal{W} \) are distinct and none equals to \( \varphi \), and \( E_{i,1}, \ldots, E_{i,m_n}, n \in \mathcal{B} \) are subsets of \( E_n \), nonempty, pairwise disjoint, and \( \mathcal{P} \circ \mu(E_{i,n}) < \infty, i = 1, \ldots, m_n \), \( \forall i \in \{1, \ldots, m_n\}, \exists m_{i,n} \in \mathbb{Z}_+, \exists \text{ pairwise disjoint } (E_{j,i,n})_{j=1}^{m_{i,n}} \subseteq \mathcal{B} \) with \( E_{i,n} = \bigcup_{j=1}^{m_{i,n}} E_{j,i,n} \) such that \( \mathcal{P} \circ \mu(E_{i,n}) < \sum_{j=1}^{m_{i,n}} |\mu(E_{j,i,n})| + \frac{2^{-n-2}\epsilon}{(1+m_n)\|w_{i,n}\|} \).

This leads to

\[
\tilde{\nu}(E_n) = \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \int_{E_n} \mathcal{P} \circ \varphi_n \, d\mathcal{P} \circ \mu + 2^{-n-2}\epsilon
\]

\[
= \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\varphi(E_{j,i,n})\| + \sum_{i=1}^{m_n} \|w_{i,n}\| \mathcal{P} \circ \mu(E_{i,n}) + 2^{-n-2}\epsilon
\]

\[
- \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \|\nu(E_{j,i,n})\|
\]
\[ \begin{align*}
\leq & \ \mathcal{P} \circ \nu(E_n) + \sum_{i=1}^{m_n} \| w_{i,n} \| \left( \sum_{j=1}^{m_{i,n}} |\mu(E_{j,i,n})| + \frac{2^{-n-2}\epsilon}{(1 + m_n) \| w_{i,n} \|} \right) \\
+ & 2^{-n-2}\epsilon - \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \| \nu(E_{j,i,n}) \| \\
< & \ \mathcal{P} \circ \nu(E_n) + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \| w_{i,n} \mu(E_{j,i,n}) \| + 2^{-n-1}\epsilon - \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \| \nu(E_{j,i,n}) \| \\
= & \ \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \left( \| \int_{E_{j,i,n}} \varphi_n \ d\mu \| - \| \int_{E_{j,i,n}} f \ d\mu \| \right) \\
\leq & \ \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \left\| \int_{E_{j,i,n}} (\varphi_n - f|_{E_n}) \ d\mu \right\| \\
\leq & \ \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \sum_{i=1}^{m_n} \int_{E_{j,i,n}} \mathcal{P} \circ (\varphi_n - f|_{E_n}) \ d\mathcal{P} \circ \mu \\
= & \ \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \int_{\bigcup_{i=1}^{m_n} \bigcup_{j=1}^{m_{i,n}} E_{j,i,n}} \mathcal{P} \circ (\varphi_n - f|_{E_n}) \ d\mathcal{P} \circ \mu \\
\leq & \ \mathcal{P} \circ \nu(E_n) + 2^{-n-1}\epsilon + \int_{E_n} \mathcal{P} \circ (\varphi_n - f|_{E_n}) \ d\mathcal{P} \circ \mu \\
< & \ \mathcal{P} \circ \nu(E_n) + 2^{-n}\epsilon
\end{align*} \]

where the second equality follows from Proposition 11.75; the third equality follows from Proposition 11.125; the fourth and the fifth inequalities follow from Proposition 11.132; and the fourth equality and the sixth inequality follow from Proposition 11.83. Then, \( \hat{\nu}(E) = \sum_{n=1}^{\infty} \nu(E_n) < \sum_{n=1}^{\infty} (\mathcal{P} \circ \nu(E_n) + 2^{-n}\epsilon) = \mathcal{P} \circ \nu(E) + \epsilon \). By the arbitrariness of \( \epsilon \), we have \( \hat{\nu}(E) \leq \mathcal{P} \circ \nu(E) \). This case is proved.

Hence, \( \mathcal{P} \circ \nu(E) = \hat{\nu}(E) = \int_E \mathcal{P} \circ f \ d\mathcal{P} \circ \mu \), \( \forall E \in \mathcal{B} \). This completes the proof of the proposition. \( \square \)

**Definition 11.155** Let \((X, \mathcal{B})\) be a measurable space, \(\mu_1\) and \(\mu_2\) be measures on \((X, \mathcal{B})\). We will say that \(\mu_1\) is absolutely continuous with respect to \(\mu_2\) if \(\forall A \in \mathcal{B} \) with \(\mu_2(A) = 0\), we have \(\mu_1(A) = 0\). In this case, we will denote \(\mu_1 \ll \mu_2\). \(\mu_1\) and \(\mu_2\) are said to be mutually singular if \(\exists A, B \in \mathcal{B} \) with \(A \cap B = \emptyset\) and \(X = A \cup B\), such that \(\mu_1(A) = \mu_2(B) = 0\). In this case, we will denote \(\mu_1 \perp \mu_2\).

**Proposition 11.156** Let \(\mu\), \(\lambda_1\), and \(\lambda_2\) be measures on a measurable space \((X, \mathcal{B})\), \(\mathcal{Y}\) be a normed linear space over \(\mathbb{K}\), \(\nu_1\) and \(\nu_2\) be \(\sigma\)-finite \(\mathcal{Y}\)-valued measures on \((X, \mathcal{B})\), \(a, b \in (0, \infty) \subset \mathbb{R}\), and \(\alpha, \beta \in \mathbb{K}\). Then, the following statements hold.

(i) \(\lambda_1 \ll \mu\) and \(\lambda_2 \ll \mu\) if, and only if, \(a\lambda_1 + b\lambda_2 \ll \mu\).
(ii) \( \lambda_1 \perp \mu \) and \( \lambda_2 \perp \mu \) if, and only if, \( a \lambda_1 + b \lambda_2 \perp \mu \).

(iii) If \( \lambda_1 \ll \mu \) and \( \lambda_2 \perp \mu \), then \( \lambda_1 \perp \lambda_2 \).

(iv) If \( \lambda_1 \perp \lambda_1 \), then \( \lambda_1(E) = 0 \), \( \forall E \in \mathcal{B} \).

(v) If \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2 \), then \( \mathcal{P} \circ (\alpha \nu_1) \perp \mathcal{P} \circ (\beta \nu_2) \).

(vi) If \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_1 \), then \( \nu_1(E) = \vartheta_y \), \( \forall E \in \mathcal{B} \).

If, in addition, \( \mathcal{Y} \) is complete, then the following statements hold.

(vii) If \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2 \), then \( \mathcal{P} \circ (\nu_1 + \nu_2) = \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \).

(viii) If \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2 \), then \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) = |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2 \).

(ix) If \( \mathcal{P} \circ \nu_1 \ll \mu \) and \( \mathcal{P} \circ \nu_2 \ll \mu \), then \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) \ll \mu \).

(x) If \( \mathcal{P} \circ \nu_1 \perp \mu \) and \( \mathcal{P} \circ \nu_2 \perp \mu \), then \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) \perp \mu \).

**Proof**

(i) By Proposition 11.136, \( a \lambda_1 + b \lambda_2 \) is a measure on \( (X, \mathcal{B}) \).

“Sufficiency” \( \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \), by \( a \lambda_1 + b \lambda_2 \ll \mu \), we have \( a \lambda_1(E) + b \lambda_2(E) = 0 \). Then, \( \lambda_1(E) = \lambda_2(E) = 0 \). Hence, \( \lambda_1 \ll \mu \) and \( \lambda_2 \ll \mu \).

“Necessity” \( \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \), by \( \lambda_1 \ll \mu \) and \( \lambda_2 \ll \mu \), we have \( \lambda_1(E) = \lambda_2(E) = 0 \). Then, \( a \lambda_1(E) + b \lambda_2(E) = 0 \). Hence, \( a \lambda_1 + b \lambda_2 \ll \mu \).

(ii) By Proposition 11.136, \( a \lambda_1 + b \lambda_2 \) is a measure on \( (X, \mathcal{B}) \). “Sufficiency” Let \( a \lambda_1 + b \lambda_2 \perp \mu \). Then, \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( (a \lambda_1 + b \lambda_2)(A) = 0 = \mu(B) \). Then, \( \lambda_1(A) = \lambda_2(A) = \mu(B) = 0 \). Hence, \( \lambda_1 \perp \mu \) and \( \lambda_2 \perp \mu \). “Necessity” Let \( \lambda_1 \perp \mu \) and \( \lambda_2 \perp \mu \). Then, \( \exists A_i, B_i \in \mathcal{B} \) with \( A_i = X \setminus B_i \) such that \( \lambda_1(A_i) = \mu(B_i) = 0 \), \( i = 1, 2 \). Let \( A := A_1 \cap A_2 \in \mathcal{B} \) and \( B := B_1 \cup B_2 \in \mathcal{B} \). Then, \( A = X \setminus B \) and \( 0 \leq a \lambda_1(A) + b \lambda_2(A) = (a \lambda_1 + b \lambda_2)(A) \leq a \lambda_1(A_1) + b \lambda_2(A_2) = 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0 \). Hence, \( a \lambda_1 + b \lambda_2 \perp \mu \).

(iii) By \( \lambda_2 \perp \mu \), \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( \lambda_2(A) = 0 = \mu(B) \). By \( \lambda_1 \ll \mu \), we have \( \lambda_1(B) = 0 \). Hence, \( \lambda_1 \perp \lambda_2 \).

(iv) By \( \lambda_1 \perp \lambda_1 \), \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( \lambda_1(A) = \lambda_1(B) = 0 \). Then, \( \lambda_1(X) = \lambda_1(A) + \lambda_1(B) = 0 \). Hence, the result holds.

(v) By Proposition 11.138, \( \alpha \nu_1 \) and \( \beta \nu_2 \) are \( \sigma \)-finite \( \mathcal{Y} \)-valued measures on \( (X, \mathcal{B}) \) with \( \mathcal{P} \circ (\alpha \nu_1) = |\alpha| \mathcal{P} \circ \nu_1 \) and \( \mathcal{P} \circ (\beta \nu_2) = |\beta| \mathcal{P} \circ \nu_2 \). By \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2 \), \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( \mathcal{P} \circ \nu_1(A) = 0 = \mathcal{P} \circ \nu_2(B) \). Then, by Proposition 11.138, \( \mathcal{P} \circ (\alpha \nu_1)(A) = |\alpha| \mathcal{P} \circ \nu_1(A) = 0 = |\beta| \mathcal{P} \circ \nu_2(B) = \mathcal{P} \circ (\beta \nu_2)(B) \). Hence, \( \mathcal{P} \circ (\alpha \nu_1) \perp \mathcal{P} \circ (\beta \nu_2) \).

(vi) By (iv), \( \mathcal{P} \circ \nu_1(E) = 0 \), \( \forall E \in \mathcal{B} \). Then, \( \nu_1(E) = \vartheta_y \), \( \forall E \in \mathcal{B} \).

Let \( \mathcal{Y} \) be a Banach space over \( \mathbb{K} \), then \( \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \) is well-defined by Proposition 11.142.

(vii) By Proposition 11.138, \( \nu_1 + \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \) and \( \mathcal{P} \circ (\nu_1 + \nu_2) \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \). By \( \mathcal{P} \circ \nu_1 \perp \mathcal{P} \circ \nu_2 \), \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( \mathcal{P} \circ \nu_1(A) = 0 = \mathcal{P} \circ \nu_2(B) \). Fix any \( E \in \mathcal{B} \). We will show that
11.7. Calculation with Measures

\( \mathcal{P} \circ (\nu_1 + \nu_2)(E) \geq \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) < \infty \); Case 2: \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \infty \).

Case 1: \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) < \infty \). Then, \( \mathcal{P} \circ \nu_1(E) < \infty \) and \( E \in \text{dom}(\nu_i), \ i = 1, 2 \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_1 \in \mathbb{Z}^+, \exists \text{pairwise disjoint } (E_i)_{i=1}^{n_1} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{n_1} E_i \) such that \( \mathcal{P} \circ \nu_1(E) < \sum_{i=1}^{n_1} \| \nu_1(E_i) \| + \epsilon \). Without loss of generality, assume \( \exists n_2 \in \mathbb{Z}^+, \exists \text{pairwise disjoint } (F_i)_{i=1}^{n_2} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{n_2} F_i \) such that \( \mathcal{P} \circ \nu_2(E) < \sum_{i=1}^{n_2} \| \nu_2(F_i) \| + \epsilon/2 < +\infty \). Then,

\[
\mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) < \sum_{i=1}^{n_1} \| \nu_1(E_i) \| + \sum_{i=1}^{n_2} \| \nu_2(F_i) \| + \epsilon
\]

where the second equality follows from the fact that \( \forall A \in \mathcal{B} \) with \( A \subseteq A \), we have \( \mathcal{P} \circ \nu_1(A) = 0 \) and \( \nu_1(A) = \emptyset \), and the fact that \( \forall B \in \mathcal{B} \) with \( B \subseteq B \), we have \( \mathcal{P} \circ \nu_2(B) = 0 \) and \( \nu_2(B) = \emptyset \); the third equality follows from Proposition 11.138; and the second inequality and the last two equalities follow from Definition 11.108. By the arbitrariness of \( \epsilon \), we have \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E) \). This case is proved.

Case 2: \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \infty \). Without loss of generality, assume \( \mathcal{P} \circ \nu_1(E) = \infty \). Since \( \nu_1 \) and \( \nu_2 \) are \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mathcal{P} \circ \nu_1(X_n) < \infty, \forall n \in \mathbb{N}, i = 1, 2 \). Without loss of generality, we may assume that \( X_n \subseteq X_{n+1}, \forall n \in \mathbb{N} \). Let \( E_n := E \cap X_n, \forall n \in \mathbb{N} \). Then, \( E = \bigcup_{n=1}^{\infty} E_n \) and \( E_n \subseteq E_{n+1}, \forall n \in \mathbb{N} \). By Proposition 11.7, we have \( \lim_{n} \mathcal{P} \circ \nu_1(E_n) = \mathcal{P} \circ \nu_1(E) = \infty \). \( \forall M \in (0, \infty) \subset \mathbb{R}, \exists \mu_0 \in \mathbb{N} \) such that \( \mathcal{P} \circ \nu_1(E_{\mu_0}) > M + 1 \). Clearly, \( \mathcal{P} \circ \nu_1(E_{\mu_0}) \leq \mathcal{P} \circ \nu_1(X_{\mu_0}) < \infty, i = 1, 2 \). Then, \( \exists n_1 \in \mathbb{Z}^+, \exists \text{pairwise disjoint } (E_i)_{i=1}^{n_1} \subseteq \mathcal{B} \) with \( E_{\mu_0} = \bigcup_{i=1}^{n_1} E_i \) such that \( \mathcal{P} \circ \nu_1(E_{\mu_0}) < \sum_{i=1}^{n_1} \| \nu_1(E_i) \| + 1 \). Then,

\[
M < \sum_{i=1}^{n_1} \| \nu_1(E_i) \| = \sum_{i=1}^{n_1} \| \nu_1(E_i \cap A) + \nu_1(E_i \cap B) \|
\]
where the first equality follows from Definition 11.108; the second equality follows from the fact that \( \forall A \in \mathcal{B} \) with \( A \subseteq A \), we have \( \mathcal{P} \circ \nu_1(A) = 0 \) and \( \nu_1(A) = \nu_y \), and the fact that \( \forall B \in \mathcal{B} \) with \( B \subseteq B \), we have \( \mathcal{P} \circ \nu_2(B) = 0 \) and \( \nu_2(B) = \nu_y \); the third equality follows from Proposition 11.138; and the second inequality follows from Definition 11.108. By the arbitrariness of \( M \), we have \( \mathcal{P} \circ (\nu_1 + \nu_2)(E) = \infty \). Then, \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) = \mathcal{P} \circ (\nu_1 + \nu_2)(E) = \infty \). This case is proved.

In both cases, we have \( \mathcal{P} \circ \nu_1(E) + \mathcal{P} \circ \nu_2(E) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E) \). By the arbitrariness of \( E \), we have \( \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \leq \mathcal{P} \circ (\nu_1 + \nu_2) \). Hence, \( \mathcal{P} \circ (\nu_1 + \nu_2) = \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \).

(viii) By Proposition 11.138, \( \alpha \nu_1, \beta \nu_2, \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, y) \). By (v), we have \( \mathcal{P} \circ (\alpha \nu_1) \perp \mathcal{P} \circ (\beta \nu_2) \). By (vii) and Proposition 11.138, we have \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) = \mathcal{P} \circ (\alpha \nu_1) + \mathcal{P} \circ (\beta \nu_2) = |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2 \).

(ix) By Proposition 11.138, \( \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, y) \). \( \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \), by \( \mathcal{P} \circ \nu_i \ll \mu \), we have \( \mathcal{P} \circ \nu_i(E) = 0 \), \( i = 1, 2 \). Then, \( 0 \leq \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2)(E) \leq |\alpha| \mathcal{P} \circ \nu_1(E) + |\beta| \mathcal{P} \circ \nu_2(E) = 0 \), where the second inequality follows from Proposition 11.138. Hence, \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) \ll \mu \).

(x) By Proposition 11.138, \( \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, y) \). By \( \mathcal{P} \circ \nu_i \perp \mu \), \( \exists A_i, B_i, A \subseteq \mathcal{B} \) such that \( \mathcal{P} \circ \nu_i(A_i) = \mu(B_i) = 0 \), \( i = 1, 2 \). Let \( A := A_1 \cap A_2 \in \mathcal{B} \) and \( B := B_1 \cup B_2 \in \mathcal{B} \). Then, \( A = X \setminus \mathcal{B} \) and \( 0 \leq \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2)(A) \leq |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2(A) \leq |\alpha| \mathcal{P} \circ \nu_1(A_1) + |\beta| \mathcal{P} \circ \nu_2(A_2) \leq 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0 \), where the second inequality follows from Proposition 11.138. Hence, \( \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2) \perp \mu \).

This completes the proof of the proposition. \( \square \)

**Proposition 11.157** Let \((X, \mathcal{B}, \mu)\) be a \( y \)-valued measure space, where \( y \) is a normed linear space, \( A \in \mathcal{B} \), and \( B := X \setminus A \). Define \( \mu_A \) to be a function from \( \mathcal{B} \) to \( y \) by: \( \forall E \in \mathcal{B} \) with \( E \cap A \in \text{dom}(\mu) \), \( \mu_A(E) := \mu(E \cap A) \in y \); \( \forall E \in \mathcal{B} \) with \( E \cap A \in \mathcal{B} \) \( \setminus \text{dom}(\mu) \), \( \mu_A(E) \) is undefined. Define \( \mu_B \) to be a function from \( \mathcal{B} \) to \( y \) similarly as \( \mu_A \) with \( A \) replaced with \( B \). Then, the following statements hold.

(i) \( \mu_A \) and \( \mu_B \) are \( y \)-valued measures on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \mu_A(E) = \mathcal{P} \circ \mu(E \cap A) \) and \( \mathcal{P} \circ \mu_B(E) = \mathcal{P} \circ \mu(E \cap B) \), \( \forall E \in \mathcal{B} \).

(ii) \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \) and \( \mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B \).

(iii) If \( \mu \) is \( \sigma \)-finite and \( y \) is a Banach space, then \( \mu_A \) and \( \mu_B \) are \( \sigma \)-finite and \( \mu = \mu_A + \mu_B \).
11.8. THE RADON-NIKODYM THEOREM

Proof  
(i) Define $\nu_A : \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\nu_A(E) = \mathcal{P} \circ \mu(E \cap A)$, $\forall E \in \mathcal{B}$. It is straightforward to check that $\nu_A$ is a measure on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$ with $\nu_A(E) = \infty$, then $\mathcal{P} \circ \mu(E \cap A) = \infty$, $E \cap A \in \mathcal{B} \setminus \text{dom}(\mu)$, and $\mu_A(E)$ is undefined. $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, then $\mathcal{P} \circ \mu(E \cap A) < \infty$, $E \cap A \in \text{dom}(\mu)$, and $\mu_A(E) = \mu(E \cap A) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, $\forall \mathcal{P} \circ \mu(E \cap A) < \infty$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$, $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$.

$$
\nu_A(E) = \mathcal{P} \circ \mu(E \cap A) \\
= \sup_{n \in \mathbb{Z}^+} (E_E)_{i=1}^n, E \cap A = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n, \sum_{i=1}^n \mu(E_i) \\
= \sup_{n \in \mathbb{Z}^+} (E_E)_{i=1}^n, E \cap A = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n, \sum_{i=1}^n \mu_A(E_i) \\
= \sup_{n \in \mathbb{Z}^+} (E_E)_{i=1}^n, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n, \sum_{i=1}^n \mu_A(E_i)
$$

Hence, $\mu_A$ is a $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_A = \nu_A$. By symmetry, $\mu_A$ is a $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_A = \nu_A$.

(ii) Clearly, $\mathcal{P} \circ \mu_A(B) = 0 = \mathcal{P} \circ \mu_B(A)$. Then, $\mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B$. $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu_A(E) + \mathcal{P} \circ \mu_B(E) = \mathcal{P} \circ \mu(E \cap A) + \mathcal{P} \circ \mu(E \cap B) = \mathcal{P} \circ \mu(E)$. Hence, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B$.

(iii) If $\mu$ is $\sigma$-finite and $\mathcal{Y}$ is a Banach space, then $\mathcal{P} \circ \mu$ is $\sigma$-finite and $\mathcal{P} \circ \mu_A$ and $\mathcal{P} \circ \mu_B$ are $\sigma$-finite. Therefore, $\mu_A$ and $\mu_B$ are $\sigma$-finite. Then, $\mu_A + \mu_B$ is well-defined as in Proposition 11.138. $\forall E \in \text{dom}(\mu_A) \cap \text{dom}(\mu_B)$, we have $E \cap A \in \text{dom}(\mu)$ and $E \cap B \in \text{dom}(\mu)$. Then, $\mathcal{P} \circ \mu(E) = \mathcal{P} \circ \mu(E \cap A) + \mathcal{P} \circ \mu(E \cap B) < \infty$ and $E \in \text{dom}(\mu)$. Hence, $\mu(E) = \mu(E \cap A) + \mu(E \cap B) = \mu_A(E) + \mu_B(E)$. By Proposition 11.138, $\mu = \mu_A + \mu_B$.

This completes the proof of the proposition.

11.8 The Radon-Nikodym Theorem

Definition 11.158 Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$-valued pre-measure space (or a $\sigma$-finite $\mathbb{R}$-valued measure space). $A \in \mathcal{B}$ is said to be a positive set with respect to $\mu$ if $\forall E \in \text{dom}(\mu)$ with $E \subseteq A$, we have $\mu(E) \in [0, \infty) \subset \mathbb{R}$. $B \in \mathcal{B}$ is said to be a negative set with respect to $\mu$ if $\forall E \in \text{dom}(\mu)$ with $E \subseteq B$, we have $\mu(E) \in (-\infty, 0) \subset \mathbb{R}$.

Lemma 11.159 Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$-valued pre-measure space (or a $\sigma$-finite $\mathbb{R}$-valued measure space). Then, every measurable subset of a positive
set is a positive set; the union of countably many positive sets is a positive set. Similar statements hold for negative sets.

**Proof** The first statement follows directly from Definition 11.158. For the second statement, let \((A_n)_{n=1}^\infty \subseteq \mathcal{B}\) be a sequence of positive sets and 
\[ A := \bigcup_{n=1}^\infty A_n, \quad \forall E \in \text{dom}(\mu) \text{ with } E \subseteq A, \quad \text{let } E_n := E \cap A_n \setminus \bigcup_{i=n+1}^\infty A_i \in \mathcal{B}, \quad \forall n \in \mathbb{N}. \]
Then, 
\[ E = \bigcup_{n=1}^\infty E_n, \quad (E_n)_{n=1}^\infty \] is pairwise disjoint, 
\[ E_n \in \text{dom}(\mu), \quad E_n \subseteq A_n, \quad \mu(E_n) \in [0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N}. \]
This leads to 
\[ \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \in [0, \infty) \subseteq \mathbb{R}. \]
Hence, \(A\) is a positive set. For the case of finitely many positive sets, we may form a sequence of positive sets by appending \(0\), which is clearly a positive set, to the collection. This completes the proof of the lemma. \(\square\)

**Lemma 11.160** Let \((X, \mathcal{B}, \mu)\) be an \(\mathbb{R}\)-valued pre-measure space (or a \(\sigma\)-finite \(\mathbb{R}\)-valued measure space) and \(E \in \mathcal{B}\). Assume that \(E \in \text{dom}(\mu)\) and 
\(\mu(E) \in (0, +\infty) \subseteq \mathbb{R}\). Then, there exists a positive set \(A \subseteq E\) with 
\(\mu(A) \in (0, +\infty) \subseteq \mathbb{R}\). Similar statement holds for negative sets.

**Proof** Note that, \(\forall E \in \mathcal{B}\) with \(E \subseteq E\), we have \(\overline{E} \in \text{dom}(\mu)\) and 
\(\mu(\overline{E}) \in \mathbb{R}\). Either \(E\) is a positive set, in which case the result holds, or \(E\) contains a subset of negative measure. In the latter case, let \(n_1 \in \mathbb{N}\) be the smallest natural number such that there is a measurable set \(E_1 \subseteq E\) with 
\(\mu(E_1) < -1/n_1\). Then, \(\mu(E \setminus E_1) = \mu(E) - \mu(E_1) \in (0, +\infty) \subseteq \mathbb{R}\).
Inductively, assume that we have obtained \(n_j \in \mathbb{N}\), \(E_j \in \mathcal{B}\), \(\forall j \in \{1, \ldots, k\} \subseteq \mathbb{N}\), such that 
\(E_j \subseteq E \setminus \bigcup_{i=1}^{j-1} E_i\), \(\mu(E_j) < -1/n_j\), and 
\(\mu(E \setminus \bigcup_{i=1}^{j-1} E_i) \in (0, +\infty) \subseteq \mathbb{R}\), \(\forall j \in \{1, \ldots, k\}\). Either \(E \setminus \bigcup_{i=1}^{k} E_i\) is a positive set, in which case the result holds, or it contains a subset of negative measure. In the latter case, let \(n_{k+1} \in \mathbb{N}\) be the smallest natural number such that there is a measurable set \(E_{k+1} \subseteq E \setminus \bigcup_{j=1}^{k} E_i\) with 
\(\mu(E_{k+1}) < -1/n_{k+1}\). Then, \(\mu(E \setminus \bigcup_{j=1}^{k+1} E_i) = \mu(E) - \mu(E_{k+1}) \in (0, +\infty) \subseteq \mathbb{R}\).
Thus, either the result holds or we have \((n_k)_{k=1}^{\infty} \subseteq \mathbb{N}\) and 
\((E_k)_{k=1}^{\infty} \subseteq \mathcal{B}\).
Let \(A := E \setminus \bigcup_{k=1}^{\infty} E_k \in \mathcal{B}\). Then, \(E = A \cup \bigcup_{k=1}^{\infty} E_k\), where the sets in the union are pairwise disjoint. Then, 
\(\mu(E) = \mu(A) + \sum_{k=1}^{\infty} \mu(E_k)\) and 
\(\mu(A) \mid \sum_{k=1}^{\infty} \mu(E_k) < +\infty\). Then, 
\(\sum_{k=1}^{\infty} 1/n_k < +\infty, \lim_{k \to \infty} n_k = \infty, \) and 
\(\mu(A) \in (0, +\infty) \subseteq \mathbb{R}\), \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists k_0 \in \mathbb{N}\) such that 
\((n_{k_0} - 1)^{-1} < \epsilon\). Since \(A \subseteq E \setminus \bigcup_{j=1}^{k_0} E_j\), then \(A\) can not contain any measurable subset with measure less than \(-1/(n_{k_0} - 1) > -\epsilon\). Thus, \(A\) contains no measurable subset with measure less than \(-\epsilon\). By the arbitrariness of \(\epsilon\), \(A\) is a positive set. This completes the proof of the lemma. \(\square\)

**Theorem 11.161** (Hahn Decomposition Theorem) Let \((X, \mathcal{B}, \mu) =: \mathcal{X}\) be an \(\mathbb{R}\)-valued pre-measure space (or a \(\sigma\)-finite \(\mathbb{R}\)-valued measure space). Then, there is a positive set \(A \in \mathcal{B}\) and a negative set \(B \in \mathcal{B}\) such that 
\(X = A \cup B\) and \(A \cap B = \emptyset\).
11.8. THE RADON-NIKODYM THEOREM

Proof  We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X}$ is an $\mathbb{R}$-valued pre-measure space; Case 2: $\mathcal{X}$ is a $\sigma$-finite $\mathbb{R}$-valued measure space.

Case 1: $\mathcal{X}$ is an $\mathbb{R}$-valued pre-measure space. Let $\lambda := \sup \{\mu(A) \in \mathbb{R} \mid A \text{ is a positive set} \} \in \mathbb{R}$. Since $\emptyset$ is a positive set, then $\lambda \geq 0$. Let $(A_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ be a sequence of positive sets such that $\lim_{i \in \mathbb{N}} \mu(A_i) = \lambda$. Let $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$. By Lemma 11.159, $A$ is a positive set. This implies that $\mu(A) \leq \lambda$. Note that $\mu(A) = \mu(A_i) + \mu(A \setminus A_i) \geq \mu(A_i), \forall i \in \mathbb{N}$. Then, $\lambda = \mu(A) \in \mathbb{R}$. Let $B := X \setminus A$. Suppose that there exists $E \subseteq B$ such that $E$ is a positive set with $\mu(E) > 0$. Then $A \cup E$ is a positive set by Lemma 11.159 and $\mu(A \cup E) = \mu(A) + \mu(E) > \lambda$. This contradicts with the definition of $\lambda$. Hence, $B$ does not contain any subset that is a positive set with positive measure. Then, by Lemma 11.160, $B$ does not contain any subset with positive measure. Hence, $B$ is a negative set. This case is proved.

Case 2: $\mathcal{X}$ is a $\sigma$-finite $\mathbb{R}$-valued pre-measure space. Then, $\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\mathcal{P} \circ \mu(X_n) < \infty, \forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Let $X_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathbb{R}$-valued measure subspace of $\mathcal{X}, \forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. $X_n$ is a finite $\mathbb{R}$-valued pre-measure space. By Case 1, there exists $A_n, B_n \in \mathcal{B}_n \subseteq \mathcal{B}$ such that $A_n = X_n \setminus B_n$ and $A_n$ is a positive set with respect to $\mu_n$ and $B_n$ is a negative set with respect to $\mu_n$. It is straightforward to see that $A_n$ is a positive set with respect to $\mu$ and $B_n$ is a negative set with respect to $\mu$. Let $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{B}$ and $B := \bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$. Then, by Lemma 11.159, $A$ is a positive set with respect to $\mu$ and $B$ is a negative set with respect to $\mu$. $A \cup B = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} (A_n \cup B_n) = \bigcup_{n=1}^{\infty} X_n = X$ and $A \cap B = (\bigcup_{n=1}^{\infty} A_n) \cap (\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} (A_n \cap B_m) = \emptyset$. This case is proved.

This completes the proof of the theorem.

□

Theorem 11.162 (Jordan Decomposition Theorem) Let $(X, \mathcal{B}, \mu)$ be an $\mathbb{R}$-valued pre-measure space. Then, there exists a unique pair of mutually singular finite measures $\mu_+$ and $\mu_-$ on $(X, \mathcal{B})$ such that $\mu = \mu_+ - \mu_-$. Furthermore, $\mathcal{P} \circ \mu = \mu_+ + \mu_-$ and $(X, \mathcal{B}, \mu)$ is a finite $\mathbb{R}$-valued measure space.

Proof  By Hahn Decomposition Theorem 11.161, $\exists A, B \in \mathcal{B}$ with $A \cap B = \emptyset$ and $X = A \cup B$ such that $A$ is a positive set with respect to $\mu$ and $B$ is a negative set with respect to $\mu$. Define $\mu_+ : \mathcal{B} \to [0, \infty) \subset \mathbb{R}$ by $\mu_+(E) = \mu(A \cap E)$, $\forall E \in \mathcal{B}$ and $\mu_- : \mathcal{B} \to [0, \infty) \subset \mathbb{R}$ by $\mu_-(E) = -\mu(B \cap E)$, $\forall E \in \mathcal{B}$. It is straightforward to show that $\mu_+$ and $\mu_-$ are finite measures on $(X, \mathcal{B})$. Clearly, $\mu_+(B) = \mu_-(A) = 0$. Then, $\mu_+ \perp \mu_-$. It is straightforward to show that $\mu = \mu_+ - \mu_-$. Then, by Proposition 11.134, $\mu$ is a finite $\mathbb{R}$-valued measure on $(X, \mathcal{B})$. By Proposition 11.156, $\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_+ + \mathcal{P} \circ \mu_- = \mu_+ + \mu_-$. 
Let \( \mu_+ \) and \( \mu_- \) be any pair of mutually singular finite measures on \((X, \mathcal{B})\) such that \( \mu = \mu_+ - \mu_- \). Then, \( \exists A, B \in \mathcal{B} \) with \( A \cap B = \emptyset \) and \( A \cup B = X \) such that \( \mu_+(B) = \mu_-(\bar{A}) = 0 \). Note that \( \mu_+(A \cap \bar{A}) = \mu_+(A) - \mu_+(\bar{A}) \leq \mu_+(A) = \mu_+(A) - \mu_-(A) = \mu_+(A \cap \bar{A}) + \mu_+(A \setminus \bar{A}) - \mu_-(A \cap \bar{A}) - \mu_-(A \setminus \bar{A}) \leq \mu_+(A) - \mu_-(\bar{A} \setminus A) \leq \mu_+(A) = \mu_+(A) - \mu_-(\bar{A} \setminus A). \)

By the Hahn Decomposition Theorem 11.161, there exists \( P \circ \mu \) and \( \sigma \) such that \( \mu = \mu_+ + \mu_- \). Furthermore, \( P \circ \mu = \mu_+ + \mu_- \).

**Proof** By Hahn Decomposition Theorem 11.161, there exists \( A, B \in \mathcal{B} \) such that \( A = X \setminus B \), \( A \) is a positive set with respect to \( \mu \), and \( B \) is a negative set with respect to \( \mu \). Define \( \mu_A \) to be a function from \( \mathcal{B} \) to \( \mathbb{R} \) by: \( \mu_A(E) = \mu(E \cap A) \in [0, \infty) \cap \mathbb{R} \) if \( E \in \mathcal{B} \) and \( E \cap A \in \text{dom}(\mu) \). Define \( \mu_B \) to be a function from \( \mathcal{B} \) to \( \mathbb{R} \) by: \( \mu_B(E) = \mu(E \cap B) \in (\infty, 0] \cap \mathbb{R} \) if \( E \in \mathcal{B} \) and \( E \cap B \in \text{dom}(\mu) \). Then, by Proposition 11.157, \( \mu_A \) and \( \mu_B \) are \( \sigma \)-finite \( \mathbb{R} \)-valued measures on \((X, \mathcal{B})\) such that \( \mu = \mu_A + \mu_B \), \( P \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B \), and \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \). Since \( \mu_A(E) \in [0, \infty) \cap \mathbb{R} \), \( \forall E \in \text{dom}(\mu_A) \), then \( \mu_A \) is identified with a \( \sigma \)-finite measure \( \mu_+ \) on \((X, \mathcal{B})\) according to Proposition 11.135. By Proposition 11.138, \( -\mu_B \) is a \( \sigma \)-finite \( \mathbb{R} \)-valued measure on \((X, \mathcal{B})\). Since \( -\mu_B(E) \in [0, \infty) \cap \mathbb{R} \), \( \forall E \in \text{dom}(\mu_B) \), then \( -\mu_B \) is identified with a \( \sigma \)-finite measure \( \mu_- \) on \((X, \mathcal{B})\) according to Proposition 11.135. Then, \( \mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ (-\mu_B) = \mu_+ + \mu_- \). Since \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \), then, \( \mu_+ \perp \mu_- \). \( \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(E) = \mu_+(E) + \mu_-(E) < \infty \), we have \( \mu(E) = (\mu_+(A) + \mu_B(E)) = \mu_+(E) + \mu_B(E) = \mu_+(E) - \mu_-(E) = (\mu_+ - \mu_-)(E) \in \mathbb{R} \).

By Propositions 11.137 and 11.138, we have \( \mu = \mu_+ - \mu_- \). Hence, the pair \( \mu_+ \) and \( \mu_- \) is the pair of mutually singular \( \sigma \)-finite measures on \((X, \mathcal{B})\) that we seek.

Next, we will show that the pair is unique. Let \( \hat{\mu}_+ \) and \( \hat{\mu}_- \) be another pair of mutually singular \( \sigma \)-finite measures on \((X, \mathcal{B})\) such that \( \mu = \hat{\mu}_+ - \hat{\mu}_- \) and \( \mathcal{P} \circ \mu = \hat{\mu}_+ + \hat{\mu}_- \). Since the four measures \( \mu_+, \mu_-, \hat{\mu}_+, \text{ and } \hat{\mu}_- \) are \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mu_+(X_n) \in \mathbb{R}, \mu_-(X_n) \in \mathbb{R}, \text{ and } \hat{\mu}_+(X_n) \in \mathbb{R}, \forall n \in \mathbb{N} \). Then, \( \mathcal{P} \circ \mu_+ = \mathcal{P} \circ \mu_- = \hat{\mu}_+ \).
\(\mu(X_n) < \infty, \forall n \in \mathbb{N}\). Fix any \(n \in \mathbb{N}\). Let \(\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)\) be the finite \(\mathbb{R}\)-valued measure subspace of \(\mathcal{X}\), \((X, \mathcal{B}_n, \mu_{n,+})\) be the finite measure subspace of \((X, \mathcal{B}, \mu_+)\), \((X, \mathcal{B}_n, \mu_{n,-})\) be the finite measure subspace of \((X, \mathcal{B}, \mu_-)\), \((X, \mathcal{B}_n, \hat{\mu}_{n,+})\) be the finite measure subspace of \((X, \mathcal{B}, \hat{\mu}_+)\), and \((X, \mathcal{B}_n, \hat{\mu}_{n,-})\) be the finite measure subspace of \((X, \mathcal{B}, \hat{\mu}_-)\). Then, it is easy to see that \(\mu_n = \mu_{n,+} - \mu_{n,-} = \hat{\mu}_{n,+} - \hat{\mu}_{n,-}, \mu_{n,+} \perp \mu_{n,-}, \mu_{n,+} \perp \mu_{n,-}\). By Jordan Decomposition Theorem 11.162, we have \(\mu_{n,+} = \hat{\mu}_{n,+}\) and \(\mu_{n,-} = \hat{\mu}_{n,-}\). By Proposition 11.117, we have \(\mu_+ = \hat{\mu}_+\) and \(\mu_- = \hat{\mu}_-\). Hence, the pair is unique.

This completes the proof of the theorem. \(\square\)

**Proposition 11.164** Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a \(\mathcal{C}\)-valued pre-measure space. Then, \(\mu = \mu_+ + \mu_-, \) where \(\mu_r : \mathcal{B} \to \mathbb{R}\) and \(\mu_i : \mathcal{B} \to \mathbb{R}\) are finite \(\mathbb{R}\)-valued measures on \((X, \mathcal{B})\), and \(\max(\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)) \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E), \forall E \in \mathcal{B}\). Then, \(\mathcal{X}\) is a finite \(\mathcal{C}\)-valued measure space.

**Proof** Define \(\mu_r : \mathcal{B} \to \mathbb{R}\) and \(\mu_i : \mathcal{B} \to \mathbb{R}\) by \(\mu_r(E) = \text{Re}(\mu(E))\) and \(\mu_i(E) = \text{Im}(\mu(E))\), \(\forall E \in \mathcal{B}\). Clearly, \(\mu = \mu_r + i\mu_i\). \(\mu_r(\emptyset) = 0\), and \(\mu_i(\emptyset) = 0\). \(\forall\) pairwise disjoint sequence \((E_n)_{n=1}^\infty \subseteq \mathcal{B}\), \(\sum_{n=1}^\infty |\mu_r(E_n)| \leq \sum_{n=1}^\infty |\mu_i(E_n)| < \infty\), where the last inequality follows from the fact that \(\mu\) is a \(\mathcal{C}\)-value pre-measure. We also have \(\sum_{n=1}^\infty \mu_r(E_n) = \sum_{n=1}^\infty \text{Re}(\mu(E_n)) = \text{Re}(\sum_{n=1}^\infty \mu(E_n)) = \text{Re}(\mu(\cup_{n=1}^\infty E_n)) = \mu_r(\cup_{n=1}^\infty E_n)\). Hence, \(\mu_r\) is an \(\mathbb{R}\)-valued pre-measure on \((X, \mathcal{B})\). Note that \(\sum_{n=1}^\infty \mu_i(E_n) \leq \sum_{n=1}^\infty |\mu_i(E_n)| < \infty\) and \(\sum_{n=1}^\infty \mu_i(E_n) = \sum_{n=1}^\infty \text{Im}(\mu(E_n)) = \text{Im}(\sum_{n=1}^\infty \mu(E_n)) = \text{Im}(\mu(\cup_{n=1}^\infty E_n)) = \mu_i(\cup_{n=1}^\infty E_n)\). Hence, \(\mu_i\) is an \(\mathbb{R}\)-valued pre-measure on \((X, \mathcal{B})\). By Jordan Decomposition Theorem 11.162, \(\mu_r\) and \(\mu_i\) are finite \(\mathbb{R}\)-valued measures on \((X, \mathcal{B})\).

\(\forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall\) pairwise disjoint \((E_k)_{k=1}^n \subseteq \mathcal{B}\) with \(E = \bigcup_{k=1}^n E_k\), \(\sum_{k=1}^n |\mu_r(E_k)| \leq \sum_{k=1}^n |\mu_i(E_k)| \leq \mathcal{P} \circ \mu(E)\). Hence, \(\mathcal{P} \circ \mu_r(E) \leq \mathcal{P} \circ \mu(E)\). \(\forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall\) pairwise disjoint \((E_k)_{k=1}^n \subseteq \mathcal{B}\) with \(E = \bigcup_{k=1}^n E_k\), \(\sum_{k=1}^n |\mu_i(E_k)| \leq \sum_{k=1}^n |\mu_i(E_k)| \leq \mathcal{P} \circ \mu(E)\). Hence, \(\mathcal{P} \circ \mu_i(E) \leq \mathcal{P} \circ \mu(E)\). \(\forall E \in \mathcal{B}, \forall n \in \mathbb{Z}_+, \forall\) pairwise disjoint \((E_k)_{k=1}^n \subseteq \mathcal{B}\) with \(E = \bigcup_{k=1}^n E_k\), \(\sum_{k=1}^n |\mu_i(E_k)| = \sum_{k=1}^n \sqrt{(\mu_r(E_k))^2 + (\mu_i(E_k))^2} \leq \sum_{k=1}^n |\mu_r(E_k)| + |\mu_i(E_k)| \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)\). Hence, \(\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E)\). Then, \(\mathcal{P} \circ \mu(X) < \infty\). Hence, \(\mathcal{X}\) is a finite \(\mathcal{C}\)-valued measure space.

This completes the proof of the proposition. \(\square\)

**Proposition 11.165** Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite \(\mathcal{C}\)-valued measure space. Then, there exists a pair of \(\sigma\)-finite \(\mathbb{R}\)-valued measures \(\mu_r\) and \(\mu_i\) on \((X, \mathcal{B})\) such that \(\mu = \mu_r + i\mu_i\) and \(0 \leq \max(\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)) \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E), \forall E \in \mathcal{B}\).

**Proof** Since \(\mathcal{X}\) is \(\sigma\)-finite, then \(\exists \left(X_n\right)_{n=1}^\infty \subseteq \mathcal{B}\) such that \(X = \bigcup_{n=1}^\infty X_n \) and \(\mathcal{P} \circ \mu(X_n) < \infty, \forall n \in \mathbb{N}\). Without loss of generality,
we may assume that \((X_n)_{n=1}^\infty\) is pairwise disjoint. Fix any \(n \in \mathbb{N}\). Let \(\mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n)\) be the finite \(\mathbb{C}\)-valued measure subspace of \(\mathcal{X}\). Then, \(\mathcal{X}_n\) is a finite \(\mathbb{C}\)-valued pre-measure space. By Proposition 11.164, there exists finite \(\mathbb{R}\)-valued measures \(\mu_{n,r}\) and \(\mu_{n,i}\) on \((X_n, \mathcal{B}_n)\) such that \(\mu_n = \mu_{n,r} + i \mu_{n,i}\) and \(0 \leq \min\{\mathcal{P} \circ \mu_{n,r}(E), \mathcal{P} \circ \mu_{n,i}(E)\} \leq \mathcal{P} \circ \mu_n(E) \leq \mathcal{P} \circ \mu_{n,r}(E) + \mathcal{P} \circ \mu_{n,i}(E) < \infty, \forall E \in \mathcal{B}_n\). By Proposition 11.118, the generation process on \((X_n, \mathcal{B}_n, \mu_{n,r})\) yields a unique \(\sigma\)-finite \(\mathbb{R}\)-valued measure space \((X, \mathcal{B}, \mu_r)\) on \(X\) and the generation process on \((X_n, \mathcal{B}_n, \mu_{n,i})\) yields a unique \(\sigma\)-finite \(\mathbb{R}\)-valued measure space \((X, \mathcal{B}, \mu_i)\) on \(X\).

\[\forall E \in \mathcal{B}, \text{ let } E_n := E \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N}. \text{ Then, } \mathcal{P} \circ \mu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_n(E_n), \text{ where the first equality follows from the fact that } \mathcal{P} \circ \mu \text{ is a measure; and the second equality follows from Propositions 11.118 and 11.115. By Propositions 11.118 and 11.115, we have } \mathcal{P} \circ \mu(E) > \sum_{n=1}^\infty \mathcal{P} \circ \mu_{n,r}(E_n) = \sum_{n=1}^\infty \mathcal{P} \circ \mu_{n,r}(E_n) = \mathcal{P} \circ \mu_r(E). \text{ Similarly, we have } \mathcal{P} \circ \mu(E) < \sum_{n=1}^\infty \mathcal{P} \circ \mu_{n,i}(E_n) = \mathcal{P} \circ \mu_i(E). \text{ Hence, we have } 0 \leq \min\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_r(E) + \mathcal{P} \circ \mu_i(E) \leq \infty, \forall E \in \mathcal{B}. \]

By Proposition 11.138, \(\mu_r + i \mu_i\) is a \(\sigma\)-finite \(\mathbb{C}\)-valued measure on \((X, \mathcal{B})\), \(\forall n \in \mathbb{N}, \forall E \in \mathcal{B}_n, \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu(X_n) < \infty \) and \(E \in \text{dom}(\mu)\). This implies that \(\max\{\mathcal{P} \circ \mu_r(E), \mathcal{P} \circ \mu_i(E)\} \leq \mathcal{P} \circ \mu(E) < \infty \) and \(E \in \text{dom}(\mu_r) \cap \text{dom}(\mu_i)\). Then, \((\mu_r + i \mu_i)(E) = \mu_r(E) + i \mu_i(E) = \mu_{n,r}(E) + i \mu_{n,i}(E) = (\mu_{n,r} + i \mu_{n,i})(E) = \mu_n(E) = \mu(E)\), where the first equality follows from Proposition 11.138; the second equality follows from Proposition 11.115; the third equality follows from Proposition 11.138; and the last equality follows from Proposition 11.115. Hence, \(\mathcal{X}_n\) is the \(\sigma\)-finite \(\mathbb{C}\)-valued measure subspace of \((X, \mathcal{B}, \mu_r + i \mu_i), \forall n \in \mathbb{N}\). By Proposition 11.117, we have \(\mu = \mu_r + i \mu_i\).

This completes the proof of the proposition. \(\square\)

**Definition 11.166** Let \((X, \mathcal{B})\) be a measurable space, \(\lambda_1\) and \(\lambda_2\) be measures on \((X, \mathcal{B})\), and \(f : X \to [0, \infty) \subset \mathbb{R}\) be \(\mathcal{B}\)-measurable. We will say that \(f\) is a Radon-Nikodym derivative of \(\lambda_1\) with respect to \(\lambda_2\) if \(\lambda_1(E) = \int_E f d\lambda_2, \forall E \in \mathcal{B}\). Let \(\mu\) be a \(\mathbb{K}\)-valued measure on \((X, \mathcal{B})\), \(\mathcal{Y}\) be a normed linear space over \(\mathbb{K}\), \(\nu\) be a \(\mathcal{Y}\)-valued measure on \((X, \mathcal{B})\), and \(f : X \to \mathcal{Y}\) be \(\mathcal{B}\)-measurable. We will say that \(f\) is a Radon-Nikodym derivative of \(\nu\) with respect to \(\mu\) if \(\mathcal{P} \circ f\) is a Radon-Nikodym derivative of \(\mathcal{P} \circ \nu\) with respect to \(\mathcal{P} \circ \mu\) (that is, \(\mathcal{P} \circ f d\mathcal{P} \circ \nu = \int_E \mathcal{P} \circ f d\mathcal{P} \circ \mu \in [0, \infty] \subset \mathbb{R}_{\text{c}}, \forall E \in \mathcal{B}\) and \(\nu(E) = \int_E f d\mu \in \mathcal{Y}, \forall E \in \text{dom}(\nu)\)).

It is easy to see that the definition above is compatible with the identification of \(\sigma\)-finite \(\mathbb{R}\)-valued measures and \(\sigma\)-finite measures eluded to in Section 11.7. It is straightforward to show that if \(f : X \to \mathcal{Y}\) is a Radon-Nikodym derivative of \(\nu\) with respect to \(\mu\), then \(\mathcal{P} \circ \nu \ll \mathcal{P} \circ \mu\). When \(\mu\) is \(\sigma\)-finite, \(\mathcal{Y}\) is a Banach space, and \(\mathcal{W}\) is a separable subspace of \(\mathcal{Y}\), then Proposition 11.154 guarantees that any measurable function \(f : X \to \mathcal{W}\)
is a Radon-Nikodym derivative of some \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu \) with respect to \( \mu \). The next result gives the sufficient condition under which the Radon-Nikodym derivative is unique. The Radon-Nikodym Theorems then give sufficient conditions on \( \nu \) and \( \mu \) under which the Radon-Nikodym derivative exists.

**Proposition 11.167** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( \mathcal{Y} \) be a separable Banach space over \( \mathbb{K} \), \( \nu \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \), and \( f : X \to \mathcal{Y} \) be a Radon-Nikodym derivative of the \( \nu \) with respect to \( \mu \), and \( g : X \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Then, \( g \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \) if, and only if, \( f = g \) a.e. in \( \mathcal{X} \).

**Proof** “Sufficiency” Let \( g = f \) a.e. in \( \mathcal{X} \). Then, by Propositions 7.21, 7.23, 11.38, and 11.39, \( \mathcal{P} \circ g = \mathcal{P} \circ f \) a.e. in \( \mathcal{X} \). \( \forall E \in \mathcal{B}, \mathcal{P} \circ \nu(E) = \int_{E} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_{E} \mathcal{P} \circ g \, d\mathcal{P} \circ \mu \), where the second equality follows from Proposition 11.83. \( \forall E \in \text{dom}(\nu), \infty > \mathcal{P} \circ \nu(E) = \int_{E} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_{E} \mathcal{P} \circ g \, d\mathcal{P} \circ \mu \). Then, \( f|_{E} \) and \( g|_{E} \) are absolutely integrable over \( \mathcal{E} := (E, \mathcal{B}_{E}, \mu_{E}) \), which is the \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \). This implies that, by Proposition 11.132, \( \nu(E) = \int_{E} f \, d\mu = \int_{E} g \, d\mu \in \mathcal{Y} \). Hence, \( g \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \).

“Necessity” Let \( g : X \to \mathcal{Y} \) be another Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \). Define \( \lambda := \nu - \nu \), which is a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \), by Proposition 11.138. Then, by Proposition 11.137, \( \lambda(E) = \theta_{\mathcal{Y}} \) and \( \mathcal{P} \circ \lambda(E) = 0 \), \( \forall E \in \mathcal{B} \). By Proposition 11.154, we may define a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \hat{\lambda} \) on \( (X, \mathcal{B}) \) such that \( \mathcal{P} \circ \hat{\lambda}(E) = \int_{E} \mathcal{P} \circ (f - g) \, d\mathcal{P} \circ \mu \), \( \forall E \in \mathcal{B} \), and \( \hat{\lambda}(E) = \int_{E} (f - g) \, d\mu \), \( \forall E \in \mathcal{B} \) with \( \int_{E} \mathcal{P} \circ (f - g) \, d\mathcal{P} \circ \mu < \infty \). \( \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \nu(E) < \infty \), then \( f|_{E} \) and \( g|_{E} \) are absolutely integrable over \( \mathcal{E} \). By Proposition 11.132, \( (f - g)|_{E} \) is absolutely integrable over \( \mathcal{E} \). Then, we have \( \theta_{\mathcal{Y}} = \lambda(E) = \nu(E) - \nu(E) = \int_{E} f \, d\mu - \int_{E} g \, d\mu = \int_{E} (f - g) \, d\mu = \hat{\lambda}(E) \). By Proposition 11.137, we have \( \lambda = \hat{\lambda} \). Then, \( \forall E \in \mathcal{B}, \mathcal{E} \in \text{dom}(\lambda), 0 = \mathcal{P} \circ \lambda(E) = \mathcal{P} \circ \hat{\lambda}(E) = \int_{E} \mathcal{P} \circ (f - g) \, d\mathcal{P} \circ \mu \). By Proposition 11.96, \( \mathcal{P} \circ (f - g) = 0 \) a.e. in \( \mathcal{X} \) and \( f = g \) a.e. in \( \mathcal{X} \).

This completes the proof of the proposition. \( \square \)

**Proposition 11.168** Let \( (X, \mathcal{B}) \) be a measurable space, \( \lambda \) and \( \lambda_{1} \) be \( \sigma \)-finite measures on \( (X, \mathcal{B}); \hat{X} := (X, \mathcal{B}, \lambda), \hat{X}_{1} := (X, \mathcal{B}, \lambda_{1}), \mu, \mu_{1} \in \mathcal{M}_{\sigma}(X, \mathcal{B}, \mathbb{K}), \mathcal{X} := (X, \mathcal{B}, \mu), \mathcal{X}_{1} := (X, \mathcal{B}, \mu_{1}), \mathcal{Y} \) be a separable Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B}(\mathbb{Y}, \mathcal{Z}), \nu_{1}, \nu_{2} \in \mathcal{M}_{\sigma}(X, \mathcal{B}, \mathcal{Y}), \hat{X}_{1} := (X, \mathcal{B}, \nu_{1}), \frac{d\lambda}{d\nu_{i}} := \hat{f}_{i} : X \to \mathcal{Y} \) a.e. in \( \mathcal{X}, \, i = 1, 2, \frac{d\lambda_{1}}{d\nu_{i}} := \hat{f} : X \to [0, \infty) \subset \mathbb{R} \) a.e. in \( \hat{X}_{1}, \mu \in \mathcal{B}_{\mathcal{Y}}, \alpha_{0} \in \mathbb{K}, A \in \mathcal{B}(\mathbb{Y}, \mathcal{Z}), \) and \( y_{0} \in \mathcal{Y} \). Then, the following statements hold.

(i) \( \int_{X} f \, d\lambda = \int_{X} (\hat{f} \hat{f}) \, d\lambda_{1} \in [0, \infty) \subset \mathbb{R}_{e} \).
(ii) $g$ is absolutely integrable over $X_1$ if, and only if, $(\mathcal{P} \circ g)(\mathcal{P} \circ \tilde{f}_1)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. In which case, $\int_X g \, d\nu_1 = \int_X (g \tilde{f}_1) \, d\mu \in \mathbb{R}$.

(iii) $\frac{d(\nu_1 + \nu_2)}{d\nu} = \tilde{f}_1 + \tilde{f}_2 = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}$ a.e. in $X$.

(iv) $A\tilde{f}_1$ is a Radon-Nikodym derivative of $A\nu_1$ with respect to $\mu$.

(v) $\frac{d(\alpha \nu_1)}{d\nu} = \alpha \tilde{f}_1 = \alpha_0 \frac{d\nu_1}{d\mu}$ a.e. in $X$.

(vi) $\frac{d(\phi \nu_1)}{d\mu} = f y_0 = \left(\frac{d\nu_1}{d\mu}\right) y_0$ a.e. in $X_1$.

(viii) $\frac{d\nu_1}{d\mu} = \hat{f}$ a.e. in $X$ if, and only if, $\hat{f} f = 1$ a.e. in $X_1$.

**Proof**

(i) By Definition 11.166, $\hat{f}$ is $\mathcal{B}$-measurable and $\lambda(E) = \int_E \hat{f} \, d\lambda_1 \in \mathbb{R}, \forall E \in \mathcal{B}$. By Propositions 7.23, 11.38, and 11.39, $\hat{f} f$ is $\mathcal{B}$-measurable. By Proposition 11.66, there exists a sequence of simple functions $(\tilde{\phi}_n)_{n=1}^{\infty}, \tilde{\phi}_n : X \rightarrow [0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}$, such that $0 \leq \tilde{\phi}_n(x) \leq \hat{f}(x), \forall x \in X, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \tilde{\phi}_n = \hat{f}$ a.e. in $X_1$, where $X_1 := (X, \mathcal{B}, \lambda + \lambda_1)$ is a $\sigma$-finite measure space by Proposition 11.136. $\forall n \in \mathbb{N}$, define $\phi_n := V_n \cup \tilde{\phi}_1$. Then, $\phi_n$ is a simple function of $X$ to $[0, \infty) \subset \mathbb{R}$, $0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq \hat{f}(x), \forall x \in X, \forall n \in \mathbb{N}$. By Proposition 11.50, $\lim_{n \in \mathbb{N}} \phi_n = \hat{f}$ a.e. in $X_1$. Since $\lambda \leq \lambda + \lambda_1$, then $\phi_n$ is a simple function of $X, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = \hat{f}$ a.e. in $X_1$. By Monotone Convergence Theorem 11.81, $\int_X f \, d\lambda = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\lambda \in \mathbb{R}$.

(ii) Since $\frac{d\nu_1}{d\mu} = \tilde{f}_1$ a.e. in $X$, then $\frac{d(\mathcal{P} \circ \nu_1)}{d\mathcal{P} \circ \mu} = \mathcal{P} \circ \tilde{f}_1$ a.e. in $X$. By (i), $\int_X \mathcal{P} \circ g \, d\mathcal{P} \circ \nu_1 = \int_X (\mathcal{P} \circ g)(\mathcal{P} \circ \tilde{f}_1) \, d\mathcal{P} \circ \mu \in \mathbb{R}$. Hence, $g$ is absolutely integrable over $X_1$ if, and only if, $(\mathcal{P} \circ g)(\mathcal{P} \circ \tilde{f}_1)$ is integrable over $(X, \mathcal{B}, \mathcal{P} \circ \mu)$. Let $g$ be absolutely integrable over $X_1$, then $\int_X \mathcal{P} \circ g \, d\mathcal{P} \circ \nu_1 = \int_X (\mathcal{P} \circ g)(\mathcal{P} \circ \tilde{f}_1) \, d\mathcal{P} \circ \mu$ a.e. in $X_1$.

By Proposition 11.66, there exists a sequence of simple functions $(\phi_n)_{n=1}^{\infty}, \phi_n : X_1 \rightarrow \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\Vert \phi_n(x) \Vert \leq ||g(x)||, \forall x \in X, \forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} \phi_n = g$ a.e. in $X_1$, where $X_1 := (X, \mathcal{B}, \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1)$ is a $\sigma$-finite measure space by Proposition 11.136. Since $\mathcal{P} \circ \nu_1 \leq \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1$, then $\phi_n$ is a simple function.
of \( \mathcal{X}_1 \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n = g \) a.e. in \( \mathcal{X}_1 \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X g \, d\nu_1 = \lim_{n \to \infty} \int_X \phi_n \, d\nu_1 \in \mathcal{Z} \). Fix any \( n \in \mathbb{N} \). Let \( \phi_n \) admit the canonical representation \( \phi_n = \sum_{i=1}^{n} w_i \chi_{A_i} \). Note that \( \mathcal{W} := \text{span} \{ \{ z \in \mathcal{Z} \mid z = wy, w \in \mathcal{W}, y \in \mathcal{Y} \} \} \) is a separable subspace of \( \mathcal{Z} \). Then, by Propositions 11.125 and 11.132, we have \( \int_X \phi_n \, d\nu_1 = \sum_{i=1}^{n} w_i \nu_1(A_i) = \sum_{i=1}^{n} w_i \int_{A_i} f_1 \, d\mu = \sum_{i=1}^{n} \int_X (\chi_{A_i} \cdot w_i f_1) \, d\mu = \int_X (\sum_{i=1}^{n} \chi_{A_i} \cdot w_i f_1) \, d\mu = \int_X (\phi_n f_1) \, d\mu \). Note that \( \mathcal{P} \circ (\phi_n f_1)(x) = \| \phi_n(x) f_1(x) \| \leq \| \phi_n(x) \| \| f_1(x) \| \leq (\mathcal{P} \circ g(x)) (\mathcal{P} \circ f_1(x)) \), \( \forall x \in \mathcal{X} \), \( \forall n \in \mathbb{N} \). Note also that \( \phi_n f_1 : \mathcal{X} \to \mathcal{W} \subseteq \mathcal{Z} \) and \( g f_1 : \mathcal{X} \to \mathcal{W} \), \( \forall n \in \mathbb{N} \). By Propositions 7.65, 11.38, and 11.39, \( \phi_n f_1 \) and \( g f_1 \) are \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). By Proposition 11.50, \( \lim_{n \to \infty} \phi_n f_1 = g f_1 \) a.e. in \( \mathcal{X}_1 \). Since \( \mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu + \mathcal{P} \circ \nu_1 \), then \( \lim_{n \to \infty} \phi_n f_1 = g f_1 \) a.e. in \( \mathcal{X} \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X (g f_1) \, d\mu = \lim_{n \to \infty} \int_X (\phi_n f_1) \, d\mu = \lim_{n \to \infty} \int_X \phi_n d\nu_1 = \int_X g \, d\nu_1 \in \mathcal{Z} \).

(iii) By Proposition 11.154, \( f_1 + f_2 \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu \) on \( (\mathcal{X}, \mathcal{B}) \) with respect to \( \mu \). \( \forall E \in \mathcal{B} \) with \( (\mathcal{P} \circ \mu \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \nu)(E) < \infty \), then \( \mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < \infty \), \( i \equiv 1, 2 \). This implies that \( f_i = f_i \mid_E \), \( i = 1, 2 \), are absolutely integrable over \( \mathcal{E} := (E, \mathcal{B}_E, \mu_E) \), which is the \( \mathcal{K} \)-valued measure subspace of \( \mathcal{X} \). Then, \( \nu(E) = \int_E f_1 \, d\mu = \int_E f_1 \, d\mu + \int_E f_2 \, d\mu = \nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) \in \mathcal{Y} \), where the second equality follows from Proposition 11.132; and the last equality follows from Proposition 11.138. Note that \( \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \( (\mathcal{X}, \mathcal{B}) \) by Proposition 11.136. By Proposition 11.137, we have \( \nu = \nu_1 + \nu_2 \). Hence, by Proposition 11.167, \( \frac{d(\nu_1 + \nu_2)}{d\mu} = f_1 + f_2 \) a.e. in \( \mathcal{X} \).

(iv) Note that \( A f_1 : \mathcal{X} \to \mathcal{R}(A) \subseteq \mathcal{Z} \), where \( \mathcal{R}(A) \) is a separable subspace of \( \mathcal{Z} \) since \( \mathcal{Y} \) is separable. By Proposition 11.154, \( A f_1 \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathcal{Z} \)-valued measure \( \nu \) on \( (\mathcal{X}, \mathcal{B}) \) with respect to \( \mu \). \( \forall E \in \mathcal{B} \) with \( (\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \nu)(E) < \infty \), then \( \mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < \infty \), which is the \( \mathcal{K} \)-valued measure subspace of \( \mathcal{X} \). Then, \( \nu_1(E) = \int_E f_1 \, d\mu = \int_E f_1 \, d\mu = \nu_1(E) = (\mathcal{P} \circ \nu_1)(E) \in \mathcal{Z} \), where the second equality follows from Proposition 11.132; and the last equality follows from Proposition 11.138. Note that \( \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 + \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \( (\mathcal{X}, \mathcal{B}) \) by Proposition 11.136. By Proposition 11.137, we have \( \nu = \nu_1 \). Hence, \( A f_1 \) is a Radon-Nikodym derivative of \( A \nu_1 \) with respect to \( \mu \).

(v) and (vi) follows immediately from (iv) and Proposition 11.167.

(vii) By Proposition 11.154, \( f_1 f \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu \) on \( (\mathcal{X}, \mathcal{B}) \) with respect to \( \mu_1 \). \( \forall E \in \mathcal{B} \) with \( (\mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu)(E) < \infty \), then \( \mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < \infty \), \( \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ (f_1 f) \, d\mathcal{P} \circ \mu_1 = \int_E (\mathcal{P} \circ f_1)(\mathcal{P} \circ f) \, d\mathcal{P} \circ \mu_1 < \infty \), and \( \nu(E) = \int_E (f_1 f) \, d\mu_1 = \int_E f_1 \, d\mu = \nu_1(E) \in \mathcal{Y} \), where the second equality follows from (ii). Note that \( \mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \( (\mathcal{X}, \mathcal{B}) \) by Proposition 11.136. By Proposition 11.137, we have
\( \tilde{\nu} = \nu_1 \). Hence, by Proposition 11.167, \( \frac{d \mu}{d \nu} = \tilde{f}_1 \) a.e. in \( \mathcal{X} \).

(viii) “Necessity” Let \( \frac{d \mu}{d \nu} = \tilde{f} \) a.e. in \( \mathcal{X} \). By (vii), \( \frac{d \mu}{d \nu} = \tilde{f} f \) a.e. in \( \mathcal{X} \).

Clearly, the constant function 1 of \( X \) is a Radon-Nikodym derivative of \( \mu_1 \) with respect to \( \mu_1 \). By Proposition 11.167, we have \( \tilde{f} f = 1 \) a.e. in \( \mathcal{X} \).

“Sufficiency” Let \( \tilde{f} \) be such that \( \tilde{f} f = 1 \) a.e. in \( \mathcal{X} \). By Proposition 11.154, \( \tilde{f} \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathbb{K} \)-valued measure \( \tilde{\mu} \) on \((X, \mathcal{B})\) with respect to \( \mu \). \( \forall E \in \mathcal{B} \) with \((\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \tilde{\mu})(E) < \infty \), \( \tilde{\mu}(E) = \int_E \tilde{f} d\mu = \int_E (\tilde{f} f) d\mu_1 = \int_E 1 d\mu_1 = \mu_1(E) \in \mathbb{K} \), where the second equality follows from (ii); the third equality follows from Proposition 11.132; and the last equality follows from Proposition 11.125. By Proposition 11.137, we have \( \tilde{\mu} = \mu_1 \). Then, by Proposition 11.167, \( \frac{d \mu}{d \nu} = \tilde{f} \) a.e. in \( \mathcal{X} \).

This completes the proof of the proposition. \( \square \)

**Theorem 11.169 (Radon-Nikodym)** Let \( \mathcal{X} := (X, \mathcal{B}, \nu) \) be a \( \sigma \)-finite measure space, \( \nu \) be a \( \sigma \)-finite \( \mathbb{R} \)-valued measure on \((X, \mathcal{B})\). Assume that \( \mathcal{P} \circ \nu \ll \mu \). Then, there exists \( f : X \to \mathbb{R} \) such that \( f \) is the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \). The function \( f \) is unique in the sense that \( g : X \to \mathbb{R} \) is another function with this property if and only if, \( g \) is \( \mathcal{B} \)-measurable and \( f = g \) a.e. in \( \mathcal{X} \). Hence, \( \frac{d \nu}{d \mu} = f \) a.e. in \( \mathcal{X} \).

Furthermore, when \( \nu \) is a \( \sigma \)-finite measure on \((X, \mathcal{B})\), then we may take \( f : X \to [0, \infty) \subset \mathbb{R} \).

**Proof** We will show the existence of \( f \) in three steps. First, consider the special case that \( \mathcal{X} \) is a finite measure space and \( \nu \) is a finite measure on \((X, \mathcal{B})\). Define \( \mathcal{F} := \{ f : X \to [0, \infty) \subset \mathbb{R} \mid f \) is \( \mathcal{B} \)-measurable, and \( \forall A \in \mathcal{B}, \int_A f d\mu \leq \nu(A) \} \). Clearly, the identically 0 function belongs to \( \mathcal{F} \). Let \( \lambda := \sup_{f \in \mathcal{F}} \int_X f d\mu \in \mathbb{R}_+ \). Then, \( \lambda \geq 0 \). Clearly, \( \lambda \leq \nu(X) + \infty \), by the definition of \( \mathcal{F} \). \( \forall f_1, f_2 \in \mathcal{F} \), by Propositions 11.38, 11.39, 7.23, \( f_1 \vee f_2 \) is \( \mathcal{B} \)-measurable. Then, \( B := \{ x \in X \mid f_1(x) - f_2(x) > 0 \} \in \mathcal{B} \) and \( C := \{ x \in X \mid f_1(x) - f_2(x) \leq 0 \} \in \mathcal{B} \). Clearly, \( B = X \setminus C \). By Proposition 11.40, \( f_1 \vee f_2 \) is \( \mathcal{B} \)-measurable. \( \forall A \in \mathcal{B} \), by Proposition 11.83, \( \int_A (f_1 \vee f_2) d\mu = \int_{A \cap B} (f_1 \vee f_2) d\mu + \int_{A \cap C} (f_1 \vee f_2) d\mu = \int_{A \cap B} f_1 d\mu + \int_{A \cap C} f_2 d\mu \leq \nu(A \cap B) + \nu(A \setminus C) = \nu(A) \). Therefore, \( f_1 \vee f_2 \in \mathcal{F} \). By the definition of \( \lambda \), there exists \( (f_n)_{n = 1}^{\infty} \subseteq \mathcal{F} \) such that \( \lambda = \lim_{n \to \infty} \int_X f_n d\mu \). \( \forall n \in \mathbb{N} \), let \( f_n := \bigvee_{i=1}^n f_i \in \mathcal{F} \). Then, \( f_n(x) \leq f_{n+1}(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \). By Proposition 11.83 and the definition of \( \lambda \), we have \( \int_X f_n d\mu \leq \int_X f_n d\mu \leq \lambda \). Then, \( \lambda = \lim_{n \to \infty} \int_X f_n d\mu \in [0, \nu(X)] \subset \mathbb{R} \). By Proposition 11.82 and Monotone Convergence Theorem 11.81, \( \exists f : X \to [0, \infty) \subset \mathbb{R} \), which is \( \mathcal{B} \)-measurable, such that \( \lim_{n \to \infty} f_n = f \) a.e. in \( \mathcal{X} \) and \( \int_X f d\mu = \lim_{n \to \infty} \int_X f_n d\mu = \lambda \). \( \forall A \in \mathcal{B} \), by Monotone Convergence Theorem 11.81, we have \( \int_A f d\mu = \int_X (f \chi_A) d\mu = \lim_{n \to \infty} \int_X (f_n \chi_A) d\mu = \lim_{n \to \infty} \int_A f_n d\mu \leq \nu(A) \). Then, \( f \in \mathcal{F} \). Clearly, \( f \) is integrable over \( \mathcal{X} \).
Claim 11.169.1 \( \nu(E) = \int_E f \, d\mu, \forall E \in \mathcal{B} \).

Proof of claim: Suppose that \( \exists E_0 \in \mathcal{B} \) such that \( \nu(E_0) > \int_{E_0} f \, d\mu \geq 0 \). Then, \( \exists \epsilon_0 > 0 \) such that \( \nu(E_0) > \epsilon_0 \mu(E_0) + \int_{E_0} f \, d\mu \). Define \( \tilde{\nu} : \mathbb{B} \to \mathbb{R} \) by \( \tilde{\nu}(E) = \nu(E) - \epsilon_0 \mu(E) - \int_E f \, d\mu \). By Propositions 11.116 and 11.136, \( \tilde{\nu} \) is a finite \( \mathbb{R} \)-valued measure with \( \tilde{\nu}(E_0) \in (0, \infty) \subset \mathbb{R} \). By Lemma 11.160, \( \exists A_0 \in \mathcal{B} \) with \( A_0 \subseteq E_0 \) such that \( A_0 \) is a positive set for \( \tilde{\nu} \). Therefore, the claim holds. □

Next step, we consider the case when \( X \) is a \( \sigma \)-finite measure space and \( \nu \) is a \( \sigma \)-finite measure on \((X, \mathcal{B})\). Then, \( X = \bigcup_{n=1}^{\infty} X_n \) where \( (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \), \( \nu(X_n) < +\infty \), and \( \mu(X_n) < +\infty \), \( \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( X_n \)'s are pairwise disjoint. \( \forall n \in \mathbb{N} \), let \( (X_n, \mathcal{B}_n, \nu_n) \) be the finite measure subspace of \( X \) and \( (X, \mathcal{B}, \nu) \) be the finite measure subspace of \((X, \mathcal{B}, \nu)\) as defined in Proposition 11.13. By \( \nu \ll \mu \), we have \( \nu_n \ll \mu_n \). Then, by the first step, \( \exists f_{n} : X_n \to [0, \infty) \subset \mathbb{R} \) such that \( f_{n} \) is \( \mathcal{B}_n \)-measurable and \( \nu_n(E_n) = \int_{E_n} f_{n} \, d\mu_n, \forall E_n \in \mathcal{B}_n \). Define \( f : X \to [0, \infty) \subset \mathbb{R} \) by \( f(x) = f_{n}(x), \forall x \in X_n, \forall n \in \mathbb{N} \). By Proposition 11.41, \( f \) is \( \mathcal{B} \)-measurable. \( \forall E \in \mathcal{B} \), \( \nu(E) = \sum_{n=1}^{\infty} \nu(E \cap X_n) = \sum_{n=1}^{\infty} \int_{E \cap X_n} f_{n} \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f_{n} \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f \, d\mu = \int_{E} f \, d\mu \), where the fifth and sixth equalities follow from Proposition 11.83. Hence, \( f \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \).

The third step, we consider the case when \( X \) is a \( \sigma \)-finite measure space and \( \nu \) is an \( \sigma \)-finite \( \mathbb{R} \)-valued measure on \((X, \mathcal{B})\). By Jordan Decomposition Theorem 11.163, there exists a pair of mutually singular \( \sigma \)-finite measures \( \nu_+ \) and \( \nu_- \) on \((X, \mathcal{B})\) such that \( \nu = \nu_+ + \nu_- \) and \( \nu \circ \mathcal{P} = \nu_+ + \nu_- \). By Proposition 11.75, \( \nu \ll \mu \), we have \( \nu_+ \ll \mu \) and \( \nu_- \ll \mu \). By the second step, \( \exists f_{+} : X \to [0, \infty) \subset \mathbb{R} \) and \( \exists f_{-} : X \to [0, \infty) \subset \mathbb{R} \) such that \( f_{+} \) and \( f_{-} \) are Radon-Nikodym derivatives of \( \nu_+ \) and \( \nu_- \) with respect to \( \mu \), respectively. Let \( \mathcal{P} = \nu \circ \mathcal{P} \). For \( \exists E \in \text{dom} (\nu) \), we have \( \mathcal{P} \circ \nu(E) = \nu_+(E) + \nu_-(E) = \int_{E} f_{+} \, d\mu + \int_{E} f_{-} \, d\mu < \infty \). This leads to \( \int_{E} f_{+} \, d\mu < \infty \) and \( \int_{E} f_{-} \, d\mu < \infty \). Then, \( f_{+} \) and \( f_{-} \) are (absolutely) integrable over \( E := (E, \mathcal{B}_E, \mu_E) \), which is the \( \sigma \)-finite measure subspace of \( X \). Then, \( \nu(E) = \nu_+(E) - \nu_-(E) = \int_{E} f_{+} \, d\mu - \int_{E} f_{-} \, d\mu = \int_{E} f \, d\mu \in \mathbb{R} \), where the third equality follows from
Proposition 11.92. By Proposition 11.116, \( f \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathbb{R} \)-valued measure \( \nu \) on \((X, \mathcal{B})\) with respect to \( \mu \). The above shows that \( \nu(E) = \nu(E) \in \mathbb{R} \), \( \forall E \in \text{dom}(\nu) \). By Proposition 11.137, we have \( \tilde{\nu} = \nu \). Hence, \( f \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \).

This completes the three steps to show the existence of \( f \). The uniqueness of \( f \) follows directly from Proposition 11.167. This completes the proof of the theorem.

\[ \square \]

Proposition 11.170 Let \( X := (X, \mathcal{B}) \) be a measurable space, \( Y \) be a separable normed linear space over \( K \), and \( f : X \to Y \). Assume that \( Y^* \) is separable, \( D_* \subseteq Y^* \) is a countable dense set, and, \( \forall y_* \in D_* \), \( f_{y_*} : X \to K \) defined by \( f_{y_*}(x) = \langle \langle y_*, f(x) \rangle \rangle \), \( \forall x \in X \), is \( \mathcal{B} \)-measurable. Then, \( f \) is \( \mathcal{B} \)-measurable.

**Proof** Let \( D \subseteq Y \) be a countable dense set. By Proposition 4.4, \( Y \) is second countable with a countable basis \( E := \{ B_y(y_0, r) \mid y_0 \in D, r \in (\mathbb{Q}, r > 0) \} \). By Proposition 11.34, all we need to show is that \( \forall \mathcal{B}_y(y_0, r) \in E \), \( V := f_{\text{inv}}(\mathcal{B}_y(y_0, r)) \in \mathcal{B} \). \( \exists N \in \mathbb{N} \) such that \( r > 1/N \). We will show that \( V = U := \bigcup_{n=N}^{\infty} \bigcap_{y_* \in D_*} f_{y_*} \text{inv}(\mathcal{B}_K(\langle \langle y_*, y_0 \rangle \rangle), \| y_* \| (r - 1/n)) \in \mathcal{B} \). \( \forall n \in \mathbb{N} \) with \( n \geq N \), \( \forall y_* \in D_* \) with \( y_* \neq \partial y_* \), \( f_{y_*} \) is \( \mathcal{B} \)-measurable implies that \( f_{y_*} \text{inv}(\mathcal{B}_K(\langle \langle y_*, y_0 \rangle \rangle), \| y_* \| (r - 1/n)) \in \mathcal{B} \). Hence, \( U \in \mathcal{B} \).

\( \forall x \in V \), we have \( \| f(x) - y_0 \| < r \). Then, \( \exists n \in \mathbb{N} \) with \( n \geq N \) such that \( \| f(x) - y_0 \| < r - 1/n \). \( \forall y_* \in D_* \) with \( y_* \neq \partial y_* \), we have \( \| f_{y_*}(x) - \langle \langle y_*, y_0 \rangle \rangle \| = \| \langle \langle y_*, f(x) - y_0 \rangle \rangle \| \leq \| y_* \| \| f(x) - y_0 \| < \| y_* \| (r - 1/n) \). Hence, \( x \in f_{y_*} \text{inv}(\mathcal{B}_K(\langle \langle y_*, y_0 \rangle \rangle), \| y_* \| (r - 1/n)) \). Then, \( x \in U \). By the arbitrariness of \( x \), we have \( V \subseteq U \).

On the other hand, \( \forall x \in U \), \( \exists n \in \mathbb{N} \) with \( n \geq N \), \( \forall y_* \in D_* \) with \( y_* \neq \partial y_* \), we have \( \| f_{y_*}(x) - \langle \langle y_*, y_0 \rangle \rangle \| < \| y_* \| (r - 1/n) \). Then, \( \| \langle \langle y_*, f(x) - y_0 \rangle \rangle \| < \| y_* \| (r - 1/n) \). If \( f(x) = y_0 \), then \( x \in V \). If \( f(x) \neq y_0 \), by Proposition 7.85, \( \exists y_{*0} \in Y^* \) with \( \| y_{*0} \| = 1 \) such that \( \| f(x) - y_0 \| = \| f_{y_{*0}}(x) - y_{*0} \| \). Since \( D_* \) is dense in \( Y^* \), then, by Proposition 4.13, \( \exists (y_{*i})_{i=1}^{\infty} \subseteq D_* \) that converges to \( y_{*0} \). Without loss of generality, assume \( y_{*i} \neq \partial y_* \), \( \forall i \in \mathbb{N} \). By Propositions 3.66, 3.67, 7.21, and 7.22, we have \( \| f(x) - y_0 \| = \lim_{i \in \mathbb{N}} \| \langle \langle y_{*i}, f(x) - y_0 \rangle \rangle \| \leq \lim_{i \in \mathbb{N}} \| y_{*i} \| (r - 1/n) = r - 1/n < r \). Hence, \( x \in V \). This implies that \( U \subseteq V \).

Hence, \( V = U \in \mathcal{B} \). This completes the proof of the proposition.

\[ \square \]

Theorem 11.171 (Radon-Nikodym) Let \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( K \)-valued measure space, \( Y \) be a separable reflexive Banach space over \( K \), \( \nu \) be a \( \sigma \)-finite \( Y \)-valued measure on \((X, \mathcal{B})\). Assume that \( \mathcal{P} \circ \nu \ll \mathcal{P} \circ \mu \) and \( Y^* \) is separable. Then, there exists \( f : X \to Y \) such that \( f \) is the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \). The function \( f \) is unique in the sense that \( g : X \to Y \) is another function with this property if, and only if, \( g \) is \( \mathcal{B} \)-measurable and \( f = g \) a.e. in \( X \). Hence, \( \frac{d\nu}{d\mu} = f \) a.e. in \( X \).
11.8. The Radon-Nikodym Theorem

**Proof** First, we consider the special case when \( \mathcal{X} \) is a \( \sigma \)-finite measure space and \( \mathcal{Y} = \mathbb{C} \). By Proposition 11.165, \( \nu = \nu_r + iv \), where \( \nu_r \) and \( \nu_i \) are \( \sigma \)-finite \( \mathbb{R} \)-valued measures on \((X, \mathcal{B})\), and \( \max\{P \circ \nu_r, P \circ \nu_i\} \leq P \circ \nu \), \( \forall E \in \mathcal{B} \). Then, \( \mathcal{P} \circ \nu \ll \mu \) implies that \( \mathcal{P} \circ \nu_r \ll \mu \) and \( \mathcal{P} \circ \nu_i \ll \mu \). By Radon-Nikodym Theorem 11.169, there exist \( f_r : X \to \mathbb{R} \) and \( f_i : X \to \mathbb{R} \) such that \( \frac{d\nu_r}{d\mu} = f_r \) a.e. in \( \mathcal{X} \) and \( \frac{d\nu_i}{d\mu} = f_i \) a.e. in \( \mathcal{X} \). By Proposition 11.168, we have \( \frac{d\nu_r}{d\mu} = \frac{d(\nu_r + iv)}{d\mu} = \frac{d\nu_r}{d\mu} + i\frac{d\nu_i}{d\mu} = f_r + if_i =: f : X \to \mathbb{C} \) a.e. in \( \mathcal{X} \). This completes the proof of the special case.

Next, we consider the special case when \( \mathcal{X} \) is a \( \sigma \)-finite measure space.

\( \forall y \in \mathcal{Y} \), define \( \nu_y \) to be the \( \sigma \)-finite \( \mathbb{K} \)-valued measure on \((X, \mathcal{B})\) such that \( \nu_y(E) = \langle \langle y, \nu(E) \rangle \rangle, \forall E \in \text{dom}(\nu), \) as shown in Proposition 11.138. Then, by Proposition 11.138, \( \mathcal{P} \circ \nu_y \ll \|y\| \mathcal{P} \circ \nu \). Then, \( \mathcal{P} \circ \nu_y \ll \mu \) since \( \mathcal{P} \circ \nu \ll \mu \). By Radon-Nikodym Theorem 11.169 and the special case, \( \exists! f_y : X \to \mathbb{K} \) such that \( \frac{d\nu_y}{d\mu} = f_y \) a.e. in \( \mathcal{X} \). \( \forall y_1, y_2 \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}, \forall E \in \text{dom}(\nu), \) we have \( E \in \text{dom}(\nu_{y_1}) \cap \text{dom}(\nu_{y_2}) \cap \text{dom}(\nu_{y_1 + y_2}) \).

Then, \( \mathbb{K} \ni \nu_{y_1 + y_2}(E) = \int_E f_{y_1 + y_2} d\mu = \langle \langle y_1 + y_2, \nu(E) \rangle \rangle = \alpha \langle \langle y_1, \nu(E) \rangle \rangle + \beta \langle \langle y_2, \nu(E) \rangle \rangle = \alpha f_{y_1}(E) + \beta f_{y_2}(E) = \alpha \int_E f_{y_1} d\mu + \beta \int_E f_{y_2} d\mu = \int_E \mathcal{P} \circ \nu_{y_1} d\mu + \beta \int_E f_{y_2} d\mu < \infty, \int_E \mathcal{P} \circ \nu_{y_1} d\mu < \infty, \text{ and } \int_E \mathcal{P} \circ f_{y_2} d\mu < \infty. \)

Then, \( f_{y_1} \) and \( f_{y_2} \) are absolutely integrable over \( \mathcal{E} := (E, \mathcal{B}_E, \mu_E) \), which is the \( \sigma \)-finite measure subspace of \( \mathcal{X} \). By Proposition 11.92, we have \( \nu_{y_1 + y_2}(E) = \int_E f_{y_1 + y_2} d\mu = \int_E (\alpha f_{y_1} + \beta f_{y_2}) d\mu \in \mathbb{K}, \forall E \in \mathcal{B} \) since \( \mathcal{P} \circ \nu(E) < \infty \). By Propositions 11.116 and 11.137, \( \alpha f_{y_1} + \beta f_{y_2} \) is a Radon-Nikodym derivative of \( \nu_{y_1 + y_2} \) with respect to \( \mu \). Then, by Proposition 11.167,

\[ f_{y_1 + y_2} = \alpha f_{y_1} + \beta f_{y_2} \text{ a.e. in } \mathcal{X}; \quad \forall y_1, y_2 \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K} \]

Note that \( \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \((X, \mathcal{B})\) and \( \mathcal{P} \circ \nu \ll \mu \). Again, by Radon-Nikodym Theorem 11.169, \( \exists ! \tilde{f} : X \to [0, \infty) \subset \mathbb{R} \) such that \( \frac{d\mathcal{P} \circ \nu}{d\mu} = \tilde{f} \) a.e. in \( \mathcal{X} \). This yields \( \mathcal{P} \circ \nu(E) = \int_E \tilde{f} d\mu, \forall E \in \mathcal{B}. \) \( \forall y \in \mathcal{Y}^* \), by Definition 11.166, \( \mathcal{P} \circ \nu_y(E) = \int_E (\mathcal{P} \circ f_y) d\mu \leq \langle \|y\| \mathcal{P} \circ \nu(E) \rangle, \forall E \in \mathcal{B}. \) When \( \|y\| > 0, \mathcal{P} \circ \nu_y(E) \leq \langle \|y\| \mathcal{P} \circ \nu(E) \rangle = \|y\| \int_E \tilde{f} d\mu = \int_E (\|y\| \tilde{f}) d\mu, \) where the first equality follows from Proposition 11.136; and the third equality follows from Proposition 11.83. When \( \|y\| = 0, \mathcal{P} \circ \nu_y(E) \leq 0 = \int_E (\|y\| \|f\|) d\mu, \) where the inequality follows from Proposition 11.138; and the equality follows from Proposition 11.75. Hence, \( \mathcal{P} \circ \nu_y(E) = \int_E (\mathcal{P} \circ f_y) d\mu \leq \int_E (\|y\| \tilde{f}) d\mu, \forall E \in \mathcal{B}, \forall y \in \mathcal{Y}^* \).

Then, by Propositions 11.96,

\[ \mathcal{P} \circ f_y \leq \|y\| \|f\| \text{ a.e. in } \mathcal{X}; \quad \forall y \in \mathcal{Y}^* \]
Let $D \subseteq \mathcal{Y}^*$ be a countable dense set and $\hat{D} := \{\sum_{i=1}^{n} \alpha_i y_i \in \mathcal{Y}^* \mid n \in \mathbb{Z}_+, \alpha_1, \ldots, \alpha_n \in \mathbb{K}_Q, y_1, \ldots, y_n \in D\}$. Then, $\hat{D} \subseteq \mathcal{Y}^*$ is also a countable dense set. Note that, $\forall y_1, y_2 \in \hat{D}$, $\forall \alpha, \beta \in \mathbb{K}_Q$, we have $\alpha y_1 + \beta y_2 \in \hat{D}$.

Define

$$E_0 := \{x \in X \mid \exists y_1, y_2 \in \hat{D}, \exists \alpha, \beta \in \mathbb{K}_Q \ni \alpha f_{y_1} + \beta f_{y_2} \neq f(y_1 + \beta y_2, x) \neq f(y_1, x) \} \cup \{ \bar{f}(x) \}$$

Then, $E_0 \in \mathcal{B}$ and $\mu(E_0) = 0$.

We have the following results.

$$f_{\alpha y_1 + \beta y_2}(x) = \alpha f_{y_1}(x) + \beta f_{y_2}(x); \quad \forall x \in X \setminus E_0, \forall y_1, y_2 \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}_Q$$

and

$$|f_{y_1}(x)| \leq \|y_1\| \bar{f}(x); \quad \forall x \in X \setminus E_0, \forall y_1 \in \hat{D}$$

$\forall x \in X \setminus E_0, \forall y_1 \in \mathcal{Y}^*$, by Proposition 4.13, $\exists (y_1)_{i=1}^{\infty} \subseteq \hat{D}$ that converges to $y_1$. Then, this sequence is a Cauchy sequence. By (11.4) and (11.5), $(f_{y_1}(x))_{i=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence and admits limit $F(x, y_1) \in \mathbb{K}$. It should be clear that $F(x, y_1)$ is independent of the choice of the sequence $(y_1)_{i=1}^{\infty}$, since let $(\tilde{y}_1)_{i=1}^{\infty} \subseteq \hat{D}$ be any sequence converging to $y_1$, the combined sequence $(y_1, \tilde{y}_1, y_2, \tilde{y}_2, \ldots) \subseteq \hat{D}$ converges to $y_1$. Therefore, the sequence $(f_{y_1}(x), f_{\tilde{y}_1}(x), f_{y_2}(x), f_{\tilde{y}_2}(x), \ldots) \subseteq \mathbb{K}$ is a Cauchy sequence and admits a limit, which has to equal to $F(x, y_1)$. Then, $\forall x \in X \setminus E_0$, $\forall y_1 \in \mathcal{Y}^*, F(x, y_1) \in \mathbb{K}$. $\forall x \in X \setminus E_0, \forall y_1 \in \hat{D}$, $F(x, y_1) = \bar{f}(x)$.

$\forall x \in X \setminus E_0, \forall y_1, y_2 \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}$, $\exists (\tilde{y}_1)_{i=1}^{\infty} \subseteq \hat{D}$ that converges to $y_1$, $\exists (\tilde{y}_2)_{i=1}^{\infty} \subseteq \hat{D}$ that converges to $y_2$, $\exists (\alpha_i)_{i=1}^{\infty} \subseteq \mathbb{K}_Q$ that converges to $\alpha$, and $\exists (\beta_i)_{i=1}^{\infty} \subseteq \mathbb{K}_Q$ that converges to $\beta$. By Propositions 7.23, 3.66, and 3.67, we have $\lim_{i \in \mathbb{N}} (\alpha_i \tilde{y}_1, \beta_i \tilde{y}_2) = (\alpha y_1 + \beta y_2)$, where the sequence on the left-hand-side is contained in $\hat{D}$. This implies that $F(\alpha y_1 + \beta y_2) = \lim_{i \in \mathbb{N}} f_{\alpha_i \tilde{y}_1, \beta_i \tilde{y}_2}(x) = \lim_{i \in \mathbb{N}} (\alpha_i f_{\tilde{y}_1}(x) + \beta_i f_{\tilde{y}_2}(x)) = \alpha F(x, y_1) + \beta F(x, y_2)$, where the second equality follows from (11.4); and the last equality follows from Propositions 3.66 and 7.21. Hence, $\forall x \in X \setminus E_0$, $F(x, \cdot) : \mathcal{Y}^* \to \mathbb{K}$ is a bounded linear functional on $\mathcal{Y}^*$. Since $\mathcal{Y}$ is reflexive, then $\exists f(x) \in \mathcal{Y}$ such that $F(x, y_1) = \langle y_1, f(x) \rangle$, $\forall y_1 \in \mathcal{Y}^*$ and $\|f(x)\| \leq \bar{f}(x)$. Therefore, we may define $f : X \to \mathcal{Y}$ by assigning $f(x) = \bar{y}_x$, $\forall x \in E_0$.

$\forall y_1 \in \mathcal{Y}^*$, define $F_{y_1} : X \to \mathbb{K}$ by $F_{y_1}(x) = \langle y_1, f(x) \rangle$, $\forall x \in X$. Fix any $y_1 \in \hat{D}$. Then, $\forall x \in X \setminus E_0$, $F_{\tilde{y}_1}(x) = \langle \tilde{y}_1, f(x) \rangle = F(x, y_1) = \tilde{y}_1(x)$ and $\forall x \in E_0$, $F_{\tilde{y}_1}(x) = \langle \tilde{y}_1, f(x) \rangle = 0$. By Proposition 11.41, $F_{\tilde{y}_1}$ is $\mathcal{B}$-measurable. By the arbitrariness of $\tilde{y}_1$ and Proposition 11.170, $f$ is $\mathcal{B}$-measurable.
Claim 11.171.1 \( \forall y_0 \in \mathbb{Y}^* \), \( f_{y_0}(x) = \langle \langle y_0, f(x) \rangle \rangle \) a.e. \( x \in \mathcal{X} \).

Proof of claim: Let \( \hat{D} := D \cup \{ y_0 \} \) and \( \hat{D} := \{ \sum_{i=1}^{\infty} \alpha_i y_i \in \mathbb{Y}^* \mid n \in \mathbb{Z}^+, \alpha_1, \ldots, \alpha_n \in \mathbb{K}_0, y_1, \ldots, y_n \in \hat{D} \} \). Then, \( \hat{D} \subseteq \mathbb{Y}^* \) is also a countable dense set and \( \hat{D} \subseteq \hat{D} \). Define

\[
E_1 := \{ x \in \mathcal{X} \mid \exists y_1, y_2 \in \hat{D}, \exists \alpha, \beta \in \mathbb{K}_0 \ni f_{y_1} + \beta f_{y_2}(x) \neq \alpha f_{y_1}(x) + \beta f_{y_2}(x) \text{ or } \| f_{y_1}(x) \| > || y_1 || \| \hat{f}(x) \| \}
\]

Then, \( E_1 \in \mathcal{B}, \mu(E_1) = 0 \) and \( E_0 \subseteq E_1 \).

\( \forall x \in X \setminus E_1, \forall y_1, y_2 \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}_0 \), we have \( f_{y_1} + \beta f_{y_2}(x) = \alpha f_{y_1}(x) + \beta f_{y_2}(x) \) and \( | f_{y_1}(x) | \leq || y_1 || \| \hat{f}(x) \| \). By Proposition 4.13, \( \exists (y_i)_{i=1}^{\infty} \subseteq \hat{D} \) that converges to \( y_0 \). Then, this sequence is a Cauchy sequence. By the above, \( (f_{y_i}(x))_{i=1}^{\infty} \subseteq \mathbb{K} \) is a Cauchy sequence and admits \( \hat{F}(x, y_0) \in \mathbb{K} \). It should be clear that \( \hat{F}(x, y_0) \) is independent of the choice of the sequence \( (y_i)_{i=1}^{\infty} \). Two particular sequences \( (y_i)_{i=1}^{\infty} \subseteq \hat{D} \) converging to \( y_0 \) and \( (y_i)_{i=1}^{\infty} \subseteq \hat{D} \) leads to \( f_{y_0}(x) = \hat{F}(x, y_0) = \hat{F}(x, y_0) \). Therefore, \( f_{y_0}(x) = F(x, y_0) = \langle \langle y_0, f(x) \rangle \rangle = F_{y_0}(x), \forall x \in X \setminus E_1 \). Note that \( f_{y_0} \) is \( \mathcal{B} \)-measurable and, by Propositions 7.72 and 11.38, \( F_{y_0} \) is \( \mathcal{B} \)-measurable. Then, \( E_1 \supseteq \{ x \in X \mid f_{y_0}(x) \neq F_{y_0}(x) \} = \{ x \in X \mid | f_{y_0}(x) - F_{y_0}(x) | > 0 \} \in \mathcal{B} \). Hence, \( f_{y_0} = F_{y_0} \) a.e. in \( \mathcal{X} \).

Therefore, the claim holds. \( \square \)

Note that \( \mathcal{P} \circ f(x) \leq \hat{f}(x), \forall x \in \mathcal{X} \), and \( \mathcal{P} \circ f \) is \( \mathcal{B} \)-measurable, by Propositions 7.21 and 11.38. \( \forall E \in \text{dom}(\nu), \mathcal{P} \circ \nu(E) = \int_E \hat{f} \, d\mu < \infty \). This implies that \( \int_E \mathcal{P} \circ f \, d\mu < \infty \). By Proposition 11.92, we have \( \int_E f \, d\mu \in \mathbb{Y}, \forall y_0 \in \mathbb{Y}^*, \mathcal{P} \circ \nu_y(E) \leq (\| y_0 \| \mathcal{P} \circ \nu)(E) \leq \infty, E \in \text{dom}(\nu_y), \mathcal{P} \circ \nu_y(E) = \int_E f \, d\mu = \int_E \langle \langle y_0, f \rangle \rangle \, d\mu(x) = \| y_0 \| \mathcal{P} \circ \nu \, d\mu \). where the last three equalities follow from Proposition 11.92. Hence, by Proposition 7.85, \( \nu(E) = \int_E f \, d\mu \in \mathbb{Y}, \forall E \in \text{dom}(\nu) \). Hence, by Propositions 11.116 and 11.137, \( f \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \). By Proposition 11.167, \( f \) is unique as desired and \( \frac{d\nu}{d\mu} = f \) a.e. in \( \mathcal{X} \).

Finally, we consider the general case. Let \( \mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu), \) which is a \( \sigma \)-finite measure space. By the second special case, \( \exists f_1 : X \to \mathbb{Y} \) such that \( \frac{d\nu}{d\mu} = f_1 \) a.e. in \( \mathcal{X} \). Again by the second special case, \( \exists f_2 : X \to \mathcal{K} \) such that \( \frac{d\nu}{d\mu} = f_2 \) a.e. in \( \mathcal{X} \). By Definition 11.166, \( \frac{d\nu}{d\mu} = \mathcal{P} \circ f_2 \) a.e. in \( \mathcal{X} \). Clearly, the constant function 1 is the Radon-Nikodym derivative of \( \mathcal{P} \circ \mu \) with respect to \( \mathcal{P} \circ \mu \). Then, \( \mathcal{P} \circ f_2 = 1 \) a.e. in \( \mathcal{X} \) and \( \exists f_3 : X \to \mathcal{K} \) such that \( f_3 f_2 = 1 \) a.e. in \( \mathcal{X} \). By Proposition 11.168, \( \frac{d\nu}{d\mu} = f_3 \) a.e. in \( \mathcal{X} \) and \( \mathcal{P} \circ f_2 = f_3 \) a.e. in \( \mathcal{X} \).

This completes the proof of the theorem. \( \square \)
Theorem 11.172 (Lebesgue Decomposition) Let $X := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $Y$ a Banach space, and $\nu$ be a $\sigma$-finite $Y$-valued measure on $(X, \mathcal{B})$. Then, there exists a unique pair of $\sigma$-finite $Y$-valued measures $\nu_0$ and $\nu_1$ on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \nu_0 \perp \mu$, $\mathcal{P} \circ \nu_1 \ll \mu$, $\nu = \nu_0 + \nu_1$, and $\mathcal{P} \circ \nu = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1$. Furthermore, the following statements hold.

(i) if $\nu$ is a finite $Y$-valued measure on $(X, \mathcal{B})$, then $\nu_0$ and $\nu_1$ are finite $Y$-valued measures on $(X, \mathcal{B})$.

(ii) if $\nu$ is a $\sigma$-finite measure on $(X, \mathcal{B})$, then $\nu_0$ and $\nu_1$ are $\sigma$-finite measures on $(X, \mathcal{B})$.

(iii) if $\nu$ is a finite measure on $(X, \mathcal{B})$, then $\nu_0$ and $\nu_1$ are finite measures on $(X, \mathcal{B})$.

Proof. Let $\lambda := \mu + \mathcal{P} \circ \nu$. By Proposition 11.136, $\lambda$ is a $\sigma$-finite measure on $(X, \mathcal{B})$. Let $\mathcal{X} := (X, \mathcal{B}, \lambda)$. Clearly, we have $\mu \ll \lambda$ and $\mathcal{P} \circ \nu \ll \lambda$. By Radon-Nikodym Theorem 11.169, $\exists f : X \to [0, \infty) \subset \mathbb{R}$ such that $f$ is the Radon-Nikodym derivative of $\mu$ with respect to $\lambda$. Then, $\mu(E) = \int_E f \, d\lambda$, $\forall E \in \mathcal{B}$. Let $A := \{ x \in X \mid f(x) > 0 \} \in \mathcal{B}$ and $B := \{ x \in X \mid f(x) = 0 \} \in \mathcal{B}$. Then, $A \cap B = \emptyset$ and $A \cup B = X$. By Proposition 11.75, $\mu(B) = \int_B f \, d\lambda = 0$. Define $\nu_0$ to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by: $\nu_0(E) = \nu_0(\nu) \notin \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap B \in \text{dom}(\nu)$; and $\nu_0(E) = \nu(\nu \cap E) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap B \in \text{dom}(\nu)$. Define $\nu_1$ to be a function from $\mathcal{B}$ to $\mathcal{Y}$ by: $\nu_1(E) = \nu_1(\nu) \notin \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap A \in \text{dom}(\nu)$; and $\nu_1(E) = \nu(\nu \cap A) \in \mathcal{Y}$, $\forall E \in \mathcal{B}$ with $E \cap A \in \text{dom}(\nu)$. By Proposition 11.157, $\nu_0$ and $\nu_1$ are $\sigma$-finite $Y$-valued measures on $(X, \mathcal{B})$ such that $\nu = \nu_0 + \nu_1$, $\mathcal{P} \circ \nu = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1$, $\mathcal{P} \circ \nu_0(E) = \mathcal{P} \circ \nu(\nu \cap B)$, and $\mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \nu(\nu \cap A)$, $\forall E \in \mathcal{B}$. This implies that $\mathcal{P} \circ \nu_0(A) = \mathcal{P} \circ \nu(\nu \cap B) = 0$. Then, $\mathcal{P} \circ \nu_0 \perp \mu$.

$\forall E \in \mathcal{B}$ with $\mu(E) = 0$, we have $\int_E f \, d\lambda = 0$. Let $\tilde{E} := (E, \mathcal{B}_E, \lambda_E)$ be the $\sigma$-finite measure subspace of $\mathcal{X}$ as defined in Proposition 11.13. Then, by Proposition 11.96, we have $f = 0$ a.e. in $\tilde{E}$. Then, $\lambda(\nu \cap A) = 0$ and $\mathcal{P} \circ \nu(\nu \cap A) = 0$, since $\mathcal{P} \circ \nu \ll \lambda$. Hence, $\mathcal{P} \circ \nu_1 \ll \mu$. This shows that $\nu_0$ and $\nu_1$ is the pair of $\sigma$-finite $Y$-valued measures on $(X, \mathcal{B})$ that we seek.

Next, we show the uniqueness of the pair $\nu_0$ and $\nu_1$. Let $\tilde{\nu}_0$ and $\tilde{\nu}_1$ be another pair of $\sigma$-finite $Y$-valued measures on $(X, \mathcal{B})$ such that $\mathcal{P} \circ \tilde{\nu}_0 \perp \mu$, $\mathcal{P} \circ \tilde{\nu}_1 \ll \mu$, $\nu = \tilde{\nu}_0 + \tilde{\nu}_1$, and $\mathcal{P} \circ \nu = \mathcal{P} \circ \tilde{\nu}_0 + \mathcal{P} \circ \tilde{\nu}_1$. Then, $\exists \tilde{A}, \tilde{B} \in \mathcal{B}$ with $\tilde{A} = X \setminus \tilde{B}$ such that $\mathcal{P} \circ \tilde{\nu}_0(\tilde{A}) = \mu(\tilde{A}) = 0$. $\forall E \in \mathcal{B}$ with $\lambda(E) < \infty$, we have $\mu(E) < \infty$, $\mathcal{P} \circ \nu(E) = \mathcal{P} \circ \nu_0(E) + \mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \nu_0(E) + \mathcal{P} \circ \nu_1(E) < \infty$. Then, $\tilde{\nu}_0(E) = \tilde{\nu}_0(\nu \cap A) + \tilde{\nu}_0(\nu \cap B) = \tilde{\nu}_0(\nu \cap A) + \tilde{\nu}_0(\nu \cap B) \ll \nu(\nu \cap A) + \nu(\nu \cap B)$, $\forall E \in \mathcal{B}$ with $\lambda(E) < \infty$. The second equality follows from the fact that $\mathcal{P} \circ \tilde{\nu}_0(\tilde{A}) = 0$; the third equality follows from $\mu(\tilde{B}) = 0$ and $\mathcal{P} \circ \tilde{\nu}_1 \ll \mu$; the fifth equality follows from the fact that $\mathcal{P} \circ \tilde{\nu}_0(\tilde{A}) = 0$, $\mu(\tilde{B}) = 0$ and $\mathcal{P} \circ \tilde{\nu}_1 \ll \mu$; and the sixth equality...
follows from the fact that $\mu(\dot{B}) = 0$ and $\mathcal{P} \circ \nu_1 \ll \mu$. We also have
\[ \hat{\nu}_1(E) = \hat{\nu}_1(E \cap \dot{A}) + \hat{\nu}_1(E \cap \dot{B}) = \nu_1(E \cap \dot{A}) + \nu_0(E \cap \dot{A}) = \nu_1(E \cap \dot{A}) = \nu_1(E) = \nu_1(E \cap \dot{A}) + \nu_1(E \cap \dot{A}) \]
and the second equality follows from the fact that $\mu(\dot{B}) = 0$ and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; the third equality follows from $\mathcal{P} \circ \hat{\nu}_0(\dot{A}) = 0$; the eighth equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\dot{A}) = 0$, $\mu(B) = 0$, and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; and the ninth equality follows from the fact that $\mu(\dot{B}) = 0$ and $\mathcal{P} \circ \nu_1 \ll \mu$. By Proposition 11.137, we have $\hat{\nu}_0 = \nu_0$ and $\hat{\nu}_1 = \nu_1$. Hence, the pair $\nu_0$ and $\nu_1$ is unique.

(i) If $\nu$ is finite, then $\mathcal{P} \circ \nu(X) = \mathcal{P} \circ \nu_0(X) + \mathcal{P} \circ \nu_1(X) < \infty$. Then, $\nu_0$ and $\nu_1$ are finite.

(ii) If $\nu$ is a $\sigma$-finite measure on $(X, \mathcal{B})$, by Proposition 11.135, it is identified with a $\sigma$-finite $\mathbb{R}$-valued measure $\tilde{\nu}$ on $(X, \mathcal{B})$ such that $\tilde{\nu}(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\tilde{\nu})$. By the general case, there exists a pair of $\sigma$-finite $\mathbb{R}$-valued measures $\nu_0$ and $\nu_1$ on $(X, \mathcal{B})$ that satisfies the desired properties. By the proof of the general case, $\nu_0(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\nu_0)$, and $\nu_1(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\nu_1)$. By Proposition 11.135, $\nu_0$ and $\nu_1$ are identified with $\sigma$-finite measures $\nu_0$ and $\nu_1$, respectively. Then, $\nu_0 = \mathcal{P} \circ \nu_0 \ll \mu$, $\nu_1 = \mathcal{P} \circ \nu_1 \ll \mu$, and $\nu = \mathcal{P} \circ \tilde{\nu} = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1 = \nu_0 + \nu_1$. Then, $\nu_0$ and $\nu_1$ are $\sigma$-finite measures on $(X, \mathcal{B})$.

(iii) If $\nu$ is a finite measure on $(X, \mathcal{B})$, then, by (i) and (ii), $\nu_0$ and $\nu_1$ are finite measures on $(X, \mathcal{B})$.

This completes the proof of the theorem. 

\section{11.9 L_p Spaces}

\textbf{Example 11.173} Let $p \in [1, \infty) \subset \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathcal{K})$ be the vector space of functions of $\mathcal{X}$ to $\mathcal{Y}$ as defined in Example 6.20 with the usual vector addition, scalar multiplication, and the null vector $0$. Define $l : [0, \infty) \subset \mathbb{R} \to [0, \infty) \subset \mathbb{R}$ by $l(t) = t^p$, $\forall t \in [0, \infty) \subset \mathbb{R}$. We will introduce the notation $\mathcal{P}_p \circ f$ to denote $l \circ \mathcal{P} \circ f$. Let $Z_p := \{ f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \ | \ f \text{ is } \mathcal{B}\text{-measurable and } \mathcal{P}_p \circ f \text{ is integrable over } \mathcal{X} \}$. Define $\| \cdot \|_p : Z_p \to [0, \infty) \subset \mathbb{R}$ by $\| f \|_p = (\int_X (\mathcal{P}_p \circ f) \, d\mu)^{1/p}$, $\forall f \in Z_p$. We will next show that $Z_p$ is a subspace of $(\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathcal{K})$ and $\| \cdot \|_p$ defines a pseudo-norm on $Z_p$.

Clearly, $\vartheta \in Z_p \neq 0$, $\forall f, g \in Z_p$, $\forall \alpha \in \mathbb{K}$, by Propositions 7.23, 11.38, and 11.39, $f + g$ and $\alpha f$ are $\mathcal{B}$-measurable. By Minkowski's Inequality 11.174, $f + g \in Z_p$ and $\| f + g \|_p \leq \| f \|_p + \| g \|_p$. $\| \alpha f \|_p = (\int_X (\mathcal{P}_p \circ (\alpha f)) \, d\mu)^{1/p} = (|\alpha|^p \int_X (\mathcal{P}_p \circ f) \, d\mu)^{1/p} = |\alpha| \| f \|_p$, $\forall f \in Z_p$, where the second equality follows from Proposition 11.92. Hence, $\alpha f \in Z_p$. 

\section*{References, Notes, and Further Reading}
The above shows that \( Z_p \) is a subspace of \((\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K})\). Hence, \((Z_p, \mathbb{K})\) is a vector space. Clearly, \( \|\cdot\|_p = 0 \). Then, combined with the above, we have \( \|\cdot\|_p \) defines a pseudo-norm on \((Z_p, \mathbb{K})\). By Proposition 7.47, the quotient space of \((Z_p, \mathbb{K})\) modulo \( \|\cdot\|_p \) is a normed linear space, to be denoted \( L_p(\mathcal{X}, \mathcal{Y}) \). We will denote the vector space \((Z_p, \mathbb{K})\) with the pseudo-norm \( \|\cdot\|_{\mathbb{K}} \) by \( L_p(\mathcal{X}, \mathcal{Y}) \).

\[ \forall f \in Z_p \text{ with } \|f\|_p = 0, \text{ we have } \int_X P \circ f \, d\mu = 0. \]

By Proposition 11.96, we have \( P \circ f = 0 \) a.e. in \( \mathcal{X} \) and \( f = \delta_y \) a.e. in \( \mathcal{X} \). On the other hand, \( P \circ f = 0 \) a.e. in \( \mathcal{X} \) and \( f \) being \( \mathcal{B} \)-measurable, then, by Proposition 11.83, \( \|f\|_p = 0 \) and \( f \in Z_p \). Hence, \( \forall f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}), \ f \in Z_p \) and \( \|f\|_p = 0 \) if, and only if, \( f = \delta_y \) a.e. in \( \mathcal{X} \) and \( f \) is \( \mathcal{B} \)-measurable. We will denote the norm in \( L_p(\mathcal{X}, \mathcal{Y}) \) by \( \|\cdot\|_p \) and elements in \( L_p(\mathcal{X}, \mathcal{Y}) \) by \( [f] \), where \( f \in L_p(\mathcal{X}, \mathcal{Y}) \).

\[ \left( 1 - \lambda \right) \frac{\|g(x)\|_p}{\|g\|_p} \geq \lambda \left( \frac{\|f(x)\|_p}{\|f\|_p} \right)^p \]

This implies \( \|f + g\|_p \leq \|f\|_p + \|g\|_p \) and \( \|f\|_p \leq \|f\|_{\mathbb{K}} \).

**Theorem 11.174 (Minkowski’s Inequality)** Let \( p \in [1, \infty) \subset \mathbb{R}, \mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathcal{Y} \) be a separable normed linear space over \( \mathbb{K} \), and \( L_p(\mathcal{X}, \mathcal{Y}) \) be the vector space over \( \mathbb{K} \) with the pseudo-norm \( \|\cdot\|_p \) as defined in Example 11.173. \( \forall f, g \in L_p(\mathcal{X}, \mathcal{Y}), \) then, \( f + g \in L_p(\mathcal{X}, \mathcal{Y}) \) and

\[ \left( 1 - \lambda \right) \frac{\|g(x)\|_p}{\|g\|_p} \geq \lambda \left( \frac{\|f(x)\|_p}{\|f\|_p} \right)^p \]

When \( 1 < p < \infty \), equality holds if, and only if, \( \exists \alpha, \beta \in [0, \infty) \subset \mathbb{R} \), which are not both zeros, such that \( \alpha P \circ f = \beta P \circ g \) a.e. in \( \mathcal{X} \) and \( P \circ (f + g) = P \circ f + P \circ g \) a.e. in \( \mathcal{X} \).

**Proof** Clearly, \( f \) and \( g \) are \( \mathcal{B} \)-measurable. By Propositions 7.23, 11.38, and 11.39, \( f + g \) is \( \mathcal{B} \)-measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \|f\|_p \|g\|_p = 0 \); Case 2: \( \|f\|_p \|g\|_p > 0 \).

Case 1: \( \|f\|_p \|g\|_p = 0 \). Without loss of generality, assume that \( \|g\|_p = 0 \).

Then, \( \int_X (P \circ f + g) \, d\mu = 0 \). By Proposition 11.96, we have \( P \circ g = 0 \) a.e. in \( \mathcal{X} \). Thus, \( g = \delta_y \) a.e. in \( \mathcal{X} \) and \( f = g \) a.e. in \( \mathcal{X} \). By Proposition 11.83, the integral \( \int_X (P \circ f + g) \, d\mu = 0 \). Hence, \( \|f + g\|_p = \|f\|_p + \|g\|_p \in \mathbb{R} \) and \( f + g \in L_p(\mathcal{X}, \mathcal{Y}) \) by \( P \circ f = P \circ g \) a.e. in \( \mathcal{X} \).

Case 2: \( \|f\|_p \|g\|_p > 0 \). Let \( \lambda := \frac{\|f\|_p}{\|f\|_p + \|g\|_p} \in (0, 1) \subset \mathbb{R} \). \( \forall x \in \mathcal{X} \), we have

\[ \left( 1 - \lambda \right) \frac{\|g(x)\|_p}{\|g\|_p} \geq \lambda \left( \frac{\|f(x)\|_p}{\|f\|_p} \right)^p \]

where the second inequality follows from the convexity of the function \( t^p \) on \( [0, \infty) \subset \mathbb{R} \). In the above, when \( p > 1 \), equality holds if, and only if, \( \|f(x) + g(x)\| = \|f(x)\| + \|g(x)\| \) and \( \frac{\|f(x)\|_p}{\|f\|_p} = \frac{\|g(x)\|_p}{\|g\|_p} \).
By Proposition 11.83, we have $\int_X \left( \frac{(P_\alpha \circ f + g)}{|f|_p + |g|_p} \right) \, d\mu \leq \int_X \left( \frac{\lambda}{|f|_p} \right) \, d\mu = \frac{\lambda}{|f|_p} \int_X (P_\alpha \circ f) \, d\mu + \frac{1-\lambda}{|g|_p} \int_X (P_\alpha \circ g) \, d\mu = 1$. This implies that $\int_X (P_\alpha \circ (f + g)) \, d\mu \leq (\|f\|_p + \|g\|_p)^p < +\infty$ and $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. Therefore, $f + g \in L_p(\mathcal{X}, y)$.

Fix any $p \in (1, +\infty) \subset \mathbb{R}$. By Proposition 11.97, equality holds if, and only if, $P \circ (f + g) = P \circ f + P \circ g$ a.e. in $\mathcal{X}$ and $\frac{P_\alpha}{|f|_p} = \frac{P_\beta}{|g|_p}$ a.e. in $\mathcal{X}$. Equality implies that $\alpha = 1/\|f\|_p$ and $\beta = 1/\|g\|_p$ such that $\alpha P \circ f = \beta P \circ g$ a.e. in $\mathcal{X}$ and $P \circ (f + g) = P \circ f + P \circ g$ a.e. in $\mathcal{X}$. On the other hand, if $\exists \alpha, \beta \in (0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha P \circ f = \beta P \circ g$ a.e. in $\mathcal{X}$ and $P \circ (f + g) = P \circ f + P \circ g$ a.e. in $\mathcal{X}$. Without loss of generality, assume that $\beta \neq 0$. Then, $P \circ g = \alpha_1 P \circ f$ a.e. in $\mathcal{X}$ with $\alpha_1 := \alpha / \beta$. It is easy to show that $\|g\|_p = \alpha_1 \|f\|_p$. This implies that $\alpha_1 = \|g\|_p / \|f\|_p$. Then, equality holds. Hence, equality holds, if, and only if, $\exists \alpha, \beta \in (0, \infty) \subset \mathbb{R}$, which are not both zeros, such that $\alpha P \circ f = \beta P \circ g$ a.e. in $\mathcal{X}$ and $P \circ (f + g) = P \circ f + P \circ g$ a.e. in $\mathcal{X}$.

Hence, the result holds in both cases. This completes the proof of the theorem.

**Definition 11.175** Let $\mathcal{X} := (X, B, \mu)$ be a measure space and $f : X \to \mathbb{R}$ be $B$-measurable. The essential supremum of $f$ is

$$\text{ess sup } f(x) := \inf \{ M \in \mathbb{R} \mid \mu(\{x \in X \mid f(x) > M\}) = 0\} \subset \mathbb{R}_e$$

**Proposition 11.176** Let $\mathcal{X} := (X, B, \mu)$ be a measure space and $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be $B$-measurable. Then,

1. $\text{ess sup}_{x \in \mathcal{X}} (f(x) + g(x)) \leq \text{ess sup}_{x \in \mathcal{X}} f(x) + \text{ess sup}_{x \in \mathcal{X}} g(x)$;
2. If $f \leq g$ a.e. in $\mathcal{X}$, then $\text{ess sup}_{x \in \mathcal{X}} f(x) \leq \text{ess sup}_{x \in \mathcal{X}} g(x)$;
3. $\forall \alpha \in (0, \infty) \subset \mathbb{R}$, $\text{ess sup}_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \text{ess sup}_{x \in \mathcal{X}} f(x)$; and $\forall \alpha \in \{0, \infty\} \subset \mathbb{R}$, $\text{ess sup}_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \text{ess sup}_{x \in \mathcal{X}} f(x)$ if $\text{ess sup}_{x \in \mathcal{X}} f(x) \in \mathbb{R}$.
4. $\lambda := \text{ess sup}_{x \in \mathcal{X}} f(x) \in \mathbb{R}_e$. Then, $f(x) \leq \lambda$ a.e. $x \in \mathcal{X}$.

**Proof**

(i) We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mu(X) = 0$; Case 2: $\mu(X) > 0$.

Case 1: $\mu(X) = 0$. Then, $\forall M \in \mathbb{R}$, $0 \leq \mu(\{x \in X \mid f(x) > M\}) \leq \mu(X) = 0$. Then, $\text{ess sup}_{x \in \mathcal{X}} f(x) = -\infty$. Similarly, $\text{ess sup}_{x \in \mathcal{X}} g(x) = -\infty$. Then, the (i) holds.

Case 2: $\mu(X) > 0$. Note that $\bigcup_{n=1}^{\infty} A_n := \bigcup_{n=1}^{\infty} \{x \in X \mid f(x) > -n\} = X$. By Proposition 11.97, $\exists n \in \mathbb{N}$ such that $\mu(A_n) > 0$. Then, $\text{ess sup}_{x \in \mathcal{X}} f(x) \geq -n > -\infty$. Similarly, $\text{ess sup}_{x \in \mathcal{X}} g(x) > -\infty$ and $\text{ess sup}_{x \in \mathcal{X}} (f(x) + g(x)) > -\infty$. Then, $\lambda := \text{ess sup}_{x \in \mathcal{X}} f(x) + \text{ess sup}_{x \in \mathcal{X}} g(x)$.
ess sup_{x \in X} g(x) \in (-\infty, +\infty] \subset \mathbb{R}. If \( \lambda = +\infty \), then (i) holds. On the other hand, if \( \lambda \in \mathbb{R} \), \( \forall M \in \mathbb{R} \) with \( \lambda < M \), \( \exists M_1, M_2 \in \mathbb{R} \) with
\[
\text{ess sup}_{x \in X} f(x) < M_1 \quad \text{and} \quad \text{ess sup}_{x \in X} g(x) < M_2 \quad \text{and} \quad M_1 + M_2 = M.
\]
Then, \( \mu(\{ x \in X \mid f(x) > M_1 \}) = 0 = \mu(\{ x \in X \mid g(x) > M_2 \}) \).
By Propositions 7.23, 11.38, and 11.39, \( f + g \) is \( \mathcal{B} \)-measurable. Hence,
\[
\mu(\{ x \in X \mid f(x) + g(x) > M \}) = 0 \quad \text{and} \quad \text{ess sup}_{x \in X} (f(x) + g(x)) \leq M. \] By the arbitrariness of \( M \), (i) holds.

(ii) Let \( \lambda := \text{ess sup}_{x \in X} g(x) \in \mathbb{R} \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \lambda = +\infty \); Case 2: \( \lambda < +\infty \). Case 1: \( \lambda = +\infty \). The result holds. Case 2: \( \lambda < +\infty \). \( \forall M \in \mathbb{R} \) with \( \lambda < M \), \( \mu(\{ x \in X \mid g(x) > M \}) = 0 \). Then, \( \{ x \in X \mid f(x) > M \} \subseteq \{ x \in X \mid g(x) > M \} \cup \{ x \in X \mid f(x) - g(x) > 0 \} \). By Propositions 7.23, 11.38, and 11.39, \( f - g \) is \( \mathcal{B} \)-measurable. Then, \( 0 \leq \mu(\{ x \in X \mid f(x) > M \}) \leq \mu(\{ x \in X \mid g(x) > M \}) + \mu(\{ x \in X \mid f(x) - g(x) > 0 \}) = 0 \). Hence, \( \text{ess sup}_{x \in X} f(x) \leq M \). By the arbitrariness of \( M \), the result holds.

(iii) Let \( \alpha \in (0, +\infty) \subset \mathbb{R} \). Then, \( \text{ess sup}_{x \in X} (\alpha f(x)) = \inf \{ M \in \mathbb{R} \mid \mu(\{ x \in X \mid \alpha f(x) > M \}) = 0 \} = \inf \{ \alpha M \in \mathbb{R} \mid \mu(\{ x \in X \mid f(x) > M \}) = 0 \} = \alpha \text{ess sup}_{x \in X} f(x) \), where the third equality follows from Proposition 3.81.

Let \( \alpha = 0 \) and \( \text{ess sup}_{x \in X} f(x) \in \mathbb{R} \). Then, \( \mu(X) > 0 \). Then, \( \text{ess sup}_{x \in X} (\alpha f(x)) = \text{ess sup}_{x \in X} 0 = 0 = \alpha \text{ess sup}_{x \in X} f(x) \).

(iv) We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( \mu(X) = 0 \); Case 2: \( \mu(X) > 0 \) and \( \lambda < +\infty \); Case 3: \( \mu(X) > 0 \) and \( \lambda = +\infty \). Case 1: \( \mu(X) = 0 \). Then, \( \lambda = -\infty \). Clearly, the result holds. Case 2: \( \mu(X) > 0 \) and \( \lambda < +\infty \). Then, \( \lambda \in \mathbb{R} \). \( \forall n \in \mathbb{N} \), \( \mu(E_n) := \mu(\{ x \in X \mid f(x) > \lambda + 1/n \}) = 0 \). By Proposition 11.17, \( \mu(E) := \mu(\{ x \in X \mid f(x) > \lambda \}) = \mu(\bigcup_{n=1}^{\infty} E_n) = \lim_{n \to \infty} \mu(E_n) = 0 \). Hence, \( f(x) \leq \lambda \) a.e. \( x \in X \). Case 3: \( \mu(X) > 0 \) and \( \lambda = +\infty \). Clearly, the result holds. Hence, (iv) holds in all three cases.

This completes the proof of the proposition.

\( \square \)

**Example 11.177** Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a separable normed linear space over \( K \), and \( (\mathcal{M}(X, Y), \mathcal{K}) \) be the vector space of functions of \( X \) to \( Y \) as defined in Example 6.20 with the usual vector addition, scalar multiplication, and the null vector \( \vartheta \). Let \( Z_{\infty} := \{ f \in \mathcal{M}(X, Y) \mid f \text{ is } \mathcal{B}\text{-measurable and } \text{ess sup}_{x \in X} \| f(x) \| < +\infty \} \). Define \( \| \cdot \|_{\infty} : Z_{\infty} \to [0, \infty) \subset \mathbb{R} \) by \( \| f \|_{\infty} = \max \{ \text{ess sup}_{x \in X} \| f(x) \| , 0 \} \), \( \forall f \in Z_{\infty} \). We will next show that \( Z_{\infty} \) is a subspace of \( (\mathcal{M}(X, Y), \mathcal{K}) \) and \( \| \cdot \|_{\infty} \) defines a pseudo-norm on \( Z_{\infty} \). Clearly, \( \vartheta \in Z_{\infty} \neq 0 \). \( \forall f, g \in Z_{\infty}, \forall a \in K \), by Propositions 7.23, 11.38, and 11.39, \( f + g \) and \( af \) are \( \mathcal{B} \)-measurable. By Proposition 11.176,
\[
\| f + g \|_{\infty} = \max \{ \text{ess sup}_{x \in X} \| f(x) + g(x) \| , 0 \} \leq \max \{ \text{ess sup}_{x \in X} (\| f(x) \| + \| g(x) \| ) , 0 \}
\]
Theorem 11.178 (Hölder’s Inequality) is no confusion in which norm convergence occurs.)

\[
\| f \|_\infty + \| g \|_\infty < +\infty
\]

and

\[
\| \alpha f \|_\infty = \max \{ \text{ess sup}_{x \in X} \| \alpha f(x) \|, 0 \} = \max \{ \text{ess sup}_{x \in X} \| f(x) \|, 0 \}
\]

\[
= \begin{cases}
\text{ess sup}_{x \in X} \| f(x) \| & \text{if } \mu(X) > 0 \\
0 & \text{if } \mu(X) = 0
\end{cases}
\]

\[
= 0 = \| f \|_\infty < +\infty
\]

Then, \( f + g, \alpha f \in Z_\infty \). Hence, \( Z_\infty \) is a subspace of \((\mathcal{M}(X,\mathcal{Y}),\mathcal{K})\). This implies that \((Z_\infty, \mathcal{K})\) is a vector space. Clearly, \( \| \cdot \|_\infty \) defines a pseudo-norm on \((Z_\infty, \mathcal{K})\). By Proposition 7.47, the quotient space \((Z_\infty, \mathcal{K})\) modulo \( \| \cdot \|_\infty \) is a normed linear space, to be denoted \( L_\infty(X,\mathcal{Y}) \).

We will denote the vector space \((Z_\infty, \mathcal{K})\) with the pseudo-norm \( \| \cdot \|_\infty \) by \( L_\infty(X,\mathcal{Y}) \).

\[ \forall f \in Z_\infty \text{ with } \| f \|_\infty = 0, \text{ we have } \text{ess sup}_{x \in X} \| f(x) \| \leq 0. \text{ Then, by Proposition 11.176, } \mathcal{P} \circ f = 0 \text{ a.e. in } X. \text{ Hence, } f = \vartheta_y \text{ a.e. in } X. \text{ On the other hand, } \forall f \in \mathcal{M}(X,\mathcal{Y}) \text{ with } f = \vartheta_y \text{ a.e. in } X \text{ and } f \text{ being } \mathcal{B}\text{-measurable, we have } \| f \|_\infty = 0 \text{ and } f \in Z_\infty. \text{ Hence, } \forall f \in \mathcal{M}(X,\mathcal{Y}), \]

\[ f \in Z_\infty \text{ and } \| f \|_\infty = 0 \text{ if, and only if, } f = \vartheta_y \text{ a.e. in } X \text{ and } f \text{ is } \mathcal{B}\text{-measurable. We will denote the norm in } L_\infty(X,\mathcal{Y}) \text{ by } \| \cdot \|_\infty \text{ and elements in } L_\infty(X,\mathcal{Y}) \text{ by } [f], \text{ where } f \in L_\infty(X,\mathcal{Y}). \]

In the following, we will write \( \lim_{n \to \infty} z_n = z \) in \( L_p(X,\mathcal{Y}) \), when the sequence \( (z_n)_{n=1}^\infty \subseteq L_p(X,\mathcal{Y}) \) converges to \( z \in L_p(X,\mathcal{Y}) \) in \( L_p(X,\mathcal{Y}) \) pseudo-norm. We will simply write \( \lim_{n \to \infty} z_n = z \) if there is no confusion in which pseudo-norm convergence occurs. For \( z \in L_p(X,\mathcal{Y}) \), we will denote the corresponding equivalence class in \( L_p(X,\mathcal{Y}) \) by \([z]\).

Then, for \( (z_n)_{n=1}^\infty \subseteq L_p(X,\mathcal{Y}) \), \( \lim_{n \to \infty} z_n = z \in L_p(X,\mathcal{Y}) \) if, and only if, \( \lim_{n \to \infty} [z_n] = [z] \) in \( L_p(X,\mathcal{Y}) \) (or simply \( \lim_{n \to \infty} [z_n] = [z] \) when there is no confusion in which norm convergence occurs.)

Theorem 11.178 (Hölder’s Inequality) Let \( p \in [1, +\infty) \subseteq \mathbb{R} \) and \( q \in (1, +\infty) \subseteq \mathbb{R}_c \) with \( 1/p + 1/q = 1 \), \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma\)-finite measure space, and \( \mathcal{Y} \) be a separable normed linear space over \( \mathbb{K} \) with \( \mathcal{Y}^* \) being separable. Then, \( \forall f \in L_p(X,\mathcal{Y}), \forall g = L_q(X,\mathcal{Y}^*), \text{ the function } r : X \to \mathbb{K}, \text{ defined by } r(x) = \langle \langle g(x), f(x) \rangle \rangle, \forall x \in X, \text{ is absolutely integrable over } X \text{ and } \]

\[
\int_X \| \langle g(x), f(x) \rangle \| \, d\mu(x) \leq \| f \|_p \| g \|_q
\]

When \( q < \infty \), equality holds if, and only if, \( |\langle g(x), f(x) \rangle| = \| f(x) \| \cdot \| g(x) \| \text{ a.e. } x \in X \text{ and } \exists \alpha, \beta \in \mathbb{R}, \text{ which are not both zeros, such that } \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \text{ a.e. in } X. \)
Proof By Propositions 7.65, 11.38, 11.39, and 7.21, \( r \) and \( P \circ r \) are \( \mathcal{B} \)-measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( q = \infty \); Case 2: \( 1 < q < +\infty \). Case 1: \( q = \infty \). Then, \( p = 1 \). By Proposition 11.176, \( \|g(x)\| \leq \|g\|_\infty \) a.e. \( x \in \mathcal{X} \). Note that 
\[
\|\langle\langle g(x), f(x)\rangle\rangle\| \leq \|g(x)\| \|f(x)\| \leq \|g\|_\infty \|f(x)\| \text{ a.e. } x \in \mathcal{X} .
\]
Then, by Propositions 11.83 and 11.92,
\[
\int_{\mathcal{X}} |\langle\langle g(x), f(x)\rangle\rangle| \, d\mu(x) \leq \|g\|_\infty \int_{\mathcal{X}} P \circ f \, d\mu = \|f\|_p \|g\|_\infty
\]

Case 2: \( 1 < q < +\infty \). Then, \( 1 < p < +\infty \). We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: \( \|f\|_p \|g\|_q > 0 \); Case 2b: \( \|f\|_p \|g\|_q = 0 \). Case 2a: \( \|f\|_p \|g\|_q = 0 \). Without loss of generality, assume \( \|g\|_q = 0 \). Then, \( g = \delta_p \) a.e. in \( \mathcal{X} \). This implies that \( |\langle\langle g(x), f(x)\rangle\rangle| = 0 \) a.e. \( x \in \mathcal{X} \) and, by Propositions 11.83 and 11.75,
\[
\int_{\mathcal{X}} |\langle\langle g(x), f(x)\rangle\rangle| \, d\mu(x) = 0 = \|f\|_p \|g\|_q
\]

Equality holds \( \Rightarrow \alpha = 0, \beta = 1, \alpha \mathcal{P}_p \circ f = 0 = \beta \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \), and \( |\langle\langle g(x), f(x)\rangle\rangle| = 0 = \|f(x)\| \|g(x)\| \text{ a.e. } x \in \mathcal{X} \). This subcase is proved.

Case 2b: \( \|f\|_p \|g\|_q > 0 \). Then, \( \|f\|_p > 0 \) and \( \|g\|_q > 0 \). \( \forall x \in \mathcal{X} \), by Lemma 7.7 with \( a = \left( \frac{\|f(x)\|}{\|f\|_p} \right)^p, b = \left( \frac{\|g(x)\|}{\|g\|_q} \right)^q, \) and \( \lambda = 1/p \), we have
\[
\frac{|\langle\langle g(x), f(x)\rangle\rangle|}{\|f\|_p \|g\|_q} \leq \frac{\|g(x)\| \|f(x)\|}{\|g\|_q \|f\|_p} \leq \frac{1}{p} \frac{\|f(x)\|^p}{\|f\|_p^p} + \frac{1}{q} \frac{\|g(x)\|^q}{\|g\|_q^q}
\]
with equality holding if, and only if, \( |\langle\langle g(x), f(x)\rangle\rangle| = \|g(x)\| \|f(x)\| \) and \( \frac{\|f(x)\|^p}{\|f\|_p^p} = \frac{\|g(x)\|^q}{\|g\|_q^q} \). Integrating the above inequality over \( \mathcal{X} \), we have, by Propositions 11.83 and 11.97,
\[
\int_{\mathcal{X}} |\langle\langle g(x), f(x)\rangle\rangle| \, d\mu(x) \leq \frac{1}{p} + \frac{1}{q} = 1
\]

with equality holding if, and only if, \( |\langle\langle g(x), f(x)\rangle\rangle| = \|g(x)\| \|f(x)\| \text{ a.e. } x \in \mathcal{X} \) and \( \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \). Equality \( \Rightarrow \alpha = 1/\|f\|_p^p, \beta = 1/\|g\|_q^q, \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \) and \( |\langle\langle g(x), f(x)\rangle\rangle| = \|g(x)\| \|f(x)\| \text{ a.e. } x \in \mathcal{X} \). On the other hand, if \( |\langle\langle g(x), f(x)\rangle\rangle| = \|g(x)\| \|f(x)\| \text{ a.e. } x \in \mathcal{X} \) and \( \exists \alpha, \beta \in \mathbb{R}, \) which are not both zeros, such that \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \), then, without loss of generality, assume \( \beta \neq 0 \). Let \( \alpha_1 = \alpha/\beta \). Then, \( \alpha_1 \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \). Hence, \( \alpha_1 \|f\|_p = \|g\|_q^q \|f\|_p^p \), which further implies that \( \alpha_1 = \|g\|_q^q / \|f\|_p^p \). Hence, \( \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \text{ a.e. in } \mathcal{X} \). This implies equality. Therefore, equality holds.
if, and only if, \[ |\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\| \text{ a.e. } x \in X \] and \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \text{ a.e. in } X \). This sub-case is proved.

This completes the proof of the theorem. \( \square \)

When \( p = 2 = q \), the Hölder’s inequality becomes the well-known Cauchy-Schwarz Inequality:

\[
\left( \int_X |\langle g(x), f(x) \rangle| \, d\mu \right) \leq \left( \int_X \mathcal{P}_2 \circ f \, d\mu \right)^{1/2} \left( \int_X \mathcal{P}_2 \circ g \, d\mu \right)^{1/2}
\]

**Example 11.179** Let \( p \in [1, \infty) \subset \mathbb{R} \), \( X:=(X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( Y \) be a separable Banach space over \( \mathbb{K} \), and \( L_p(X,Y) \) the normed linear space over \( \mathbb{K} \) as defined in Example 11.173. We will show that \( L_p(X,Y) \) is a Banach space by Proposition 7.27. Define \( l: [0, \infty) \subset \mathbb{R} \to [0, \infty) \subset \mathbb{R} \) by \( l(t)=t^p \), \( \forall t \in [0, \infty) \subset \mathbb{R} \). Fix any \( ([f_n])_{n=1}^\infty \subseteq L_p(X,Y) \) with \( f_n \in L_p(X,Y), \forall n \in \mathbb{N} \), and \( \sum_{n=1}^\infty \|f_n\|_p := M \in [0, \infty) \subset \mathbb{R} \). \( \forall n \in \mathbb{N} \), define \( g_n : X \to [0, \infty) \subset \mathbb{R} \) by \( g_n(x) = \sum_{i=1}^n f_i(x) \), \( \forall x \in X \). By Propositions 7.23, 7.21, 11.38, and 11.39, \( g_n \) is \( \mathcal{B} \)-measurable. Note that \( \mathcal{P} \circ f_n \in L_p(X, \mathbb{R}) \). Then, \( g_n \in L_p(X, \mathbb{R}) \) and \( \|g_n\|_p \leq \sum_{i=1}^n \|\mathcal{P} \circ f_i\|_p = \sum_{i=1}^n \|f_i\|_p = \sum_{i=1}^n \|[f_i]\|_p \leq M \).

This implies that \( \int_X (l \circ g_n) \, d\mu = M^p \). Clearly, \( (g_n(x))^p \leq (g_{n+1}(x))^p \), \( \forall x \in X, \forall n \in \mathbb{N} \). By Proposition 11.82 and Monotone Convergence Theorem 11.81, \( \exists \tilde{g} : X \to [0, \infty) \subset \mathbb{R} \), which is \( \mathcal{B} \)-measurable, such that \( \lim_{n \in \mathbb{N}} \int_X (l \circ g_n) \, d\mu \leq \int_X (l \circ \tilde{g}) \, d\mu \).

Let \( E := \{ x \in X \mid (g_n(x))_{n=1}^\infty \text{ does not converge to } g(x) \} \). Then, \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). \( \forall x \in X \setminus E \), \( \lim_{n \in \mathbb{N}} g_n(x) = g(x) \in \mathbb{R} \) and \( \lim_{n \in \mathbb{N}} \sum_{i=1}^n \|f_i(x)\| = \|g(x)\| \in \mathbb{R} \). By Proposition 7.27 and the completeness of \( Y \), we have \( \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_i(x) := \lim_{n \in \mathbb{N}} s_n(x) := (s(x)) \in Y \). Define \( s(x) = s_n(x) = \tilde{g}, \forall x \in E \) and \( s(x) = \tilde{g}, \forall x \in X \setminus E \). By Proposition 11.41, \( s_n \) is \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). Clearly, \( \lim_{n \in \mathbb{N}} s_n(x) = s(x), \forall x \in X \). By Proposition 11.48, \( s \) is \( \mathcal{B} \)-measurable. By Lemma 11.43, \( s_n = \sum_{i=1}^n f_i \) a.e. in \( X \). This yields that \( [s]_n = [\sum_{i=1}^n f_i] = [\sum_{i=1}^n f_i] \) and \( s_n \in L_p(X, \mathbb{R}) \). Then, by Proposition 11.50, \( \lim_{n \in \mathbb{N}} \sum_{i=1}^n f_i = s \) a.e. in \( X \). Note that \( \lim_{n \in \mathbb{N}} \mathcal{P}_p \circ (s_n(x) - s(x)) = 0, \forall x \in X \). Note also \( \|s_n(x) - s(x)\|^p \leq \|s_n(x)\|^p + \|s(x)\|^p \leq 2\|g_n(x) + g(x)\|^p \leq 2^p(g(x))^p \), \( \forall x \in X \setminus E, \forall n \in \mathbb{N} \), and \( \|s_n(x) - s(x)\|^p = 0 \leq 2^p(g(x))^p \), \( \forall x \in E, \forall n \in \mathbb{N} \). Hence, \( \|s_n(x) - s(x)\|^p \leq 2^p(g(x))^p \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.91, \( \lim_{n \in \mathbb{N}} \int_X (\mathcal{P}_p \circ (s_n - s)) \, d\mu = 0 \). This implies that \( \lim_{n \in \mathbb{N}} \|s_n - s\|_p = 0 \). Hence, \( s \in L_p(X, \mathbb{K}) \) and \( \lim_{n \in \mathbb{N}} f_i = \lim_{n \in \mathbb{N}} s_n \). Hence, \( ([f_n])_{n=1}^\infty \) is summable in \( L_p(X, \mathbb{R}) \). By Proposition 7.27 and the arbitrariness of \( ([f_n])_{n=1}^\infty \), we have \( L_p(X,Y) \) is complete. This shows that \( L_p(X,Y) \) is a Banach space when \( X \) is a \( \sigma \)-finite measure space and \( Y \) is a separable Banach space.
over $\mathbb{K}$ as defined in Example 11.177. We will show that $L_\infty(\mathcal{X}, \mathcal{Y})$ is a Banach space. Fix any Cauchy sequence $(\{f_n\})_{n=1}^\infty \subseteq L_\infty(\mathcal{X}, \mathcal{Y})$ with $f_n \in L_\infty(\mathcal{X}, \mathcal{Y})$, $\forall n \in \mathbb{N}$. Let $A := \bigcup_{i=1}^\infty \bigcup_{n=N_i}^\infty \bigcup_{m=0}^\infty A_{n,m,k} \in \mathcal{B}$. Clearly, $\mu(A) = 0$. \phantom{a} 

Proposition 11.181 Let $p \in [1, \infty) \subseteq \mathbb{R}$, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $L_p(\mathcal{X}, \mathcal{Y})$ be the normed linear space over $\mathbb{K}$ as defined in Example 11.173, and $f \in L_p(\mathcal{X}, \mathcal{Y})$. Then, there exists a sequence of simple functions $(\phi_i)_{i=1}^\infty$, $\phi_i : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall i \in \mathbb{N}$, such that $\lim_{n \in \mathbb{N}} \phi_i = f$ a.e. in $\mathcal{X}$, $\|\phi_i(x)\| \leq \|f(x)\|$, $\forall x \in \mathcal{X}$, $\forall i \in \mathbb{N}$, $\lim_{i \in \mathbb{N}} \int_X P_p \circ (\phi_i - f) \, d\mu = 0$, and $\lim_{i \in \mathbb{N}} \phi_i = f$ in $L_p(\mathcal{X}, \mathcal{Y})$.

Proof Since $f \in L_p(\mathcal{X}, \mathcal{Y})$, then $f$ is $\mathcal{B}$-measurable and $\mathcal{P}_p \circ f$ is integrable over $\mathcal{X}$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_i)_{i=1}^\infty$, $\phi_i : \mathcal{X} \rightarrow \mathcal{Y}$, $\forall i \in \mathbb{N}$, such that $\|\phi_i(x)\| \leq \|f(x)\|$, $\forall x \in \mathcal{X}$, $\forall i \in \mathbb{N}$, and $\lim_{i \in \mathbb{N}} \phi_i = f$ a.e. in $\mathcal{X}$. By Propositions 7.23, 7.21, 11.38, and 11.39, $\mathcal{P}_p \circ (\phi_i - f)$ is $\mathcal{B}$-measurable, $\forall i \in \mathbb{N}$. Note that, by Propositions 7.23, 7.21, 11.52, and 11.53, $\lim_{i \in \mathbb{N}} \mathcal{P}_p \circ (\phi_i - f) = 0$ a.e. in $\mathcal{X}$ and $\mathcal{P}_p \circ (\phi_i - f)(x) \leq 2^p \mathcal{P}_p \circ f(x)$, $\forall x \in \mathcal{X}$, $\forall i \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91, we have $\lim_{i \in \mathbb{N}} \int_X \mathcal{P}_p \circ (\phi_i - f) \, d\mu = 0$. Hence, $\lim_{i \in \mathbb{N}} \phi_i = f$. This completes the proof of the proposition. \hfill $\Box$

Proposition 11.182 Let $p \in [1, \infty) \subseteq \mathbb{R}$, $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a $\sigma$-finite normal topological measure space, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, and $f \in L_p(\mathcal{X}, \mathcal{Y})$. Then, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists$ a continuous function $g : \mathcal{X} \rightarrow \mathcal{Y}$ such that $g \in L_p(\mathcal{X}, \mathcal{Y})$ and $\|g - f\|_p < \epsilon$.

Proof Fix $\epsilon \in (0, \infty) \subseteq \mathbb{R}$. By Proposition 11.181, there exists a simple function $\phi : \mathcal{X} \rightarrow \mathcal{Y}$ such that $\|\phi(x)\| \leq \|f(x)\|$, $\forall x \in \mathcal{X}$, and $\|\phi - f\|_p < \epsilon/2$. Let $\phi$ admit the canonical representation $\phi = \sum_{i=1}^n y_i \chi_{A_i}$, where $n \in \mathbb{Z}_+$, $y_1, \ldots, y_n \in \mathcal{Y}$ are distinct and none
equals to \( \vartheta_y, A_1, \ldots, A_n \in \mathcal{B} \) are pairwise disjoint, nonempty, and of finite measure. \( \forall i \in \{1, \ldots, n\} \), by \( X \) being a topological measure space and Proposition 11.27, \( \exists U_i, X \setminus F_i \in \mathcal{O}_X \) such that \( F_i \subseteq A_i \subseteq U_i \) and \( \mu(U_i \setminus F_i) = \mu(U_i \setminus A_i) + \mu(A_i \setminus F_i) < \frac{\epsilon}{2(n+1)\|y\|} \). By \( X \) being a normal topological space and Urysohn’s Lemma 3.55, there exists a continuous function \( g_i : X \to [0, 1] \subset \mathbb{R} \) such that \( g_i(x) = 1, \forall x \in F_i \) and \( g_i(x) = 0, \forall x \in X \setminus U_i \). By Proposition 11.37, \( g_i \) is \( \mathcal{B} \)-measurable. By Definition 11.79, \( \|g_i - \chi_{A_i,x}\|_p \leq \frac{\epsilon}{2(n+1)\|y\|} \).

Define \( g : X \to \mathbb{Y} \) by \( g(x) = \sum_{i=1}^{n} y_i g_i(x), \forall x \in X \). By Propositions 7.23, 3.12, and 3.32, \( g \) is continuous. By Proposition 11.37, \( g \) is \( \mathcal{B} \)-measurable. Then, \( \|g - f\|_p \leq \|g - \phi\|_p + \|\phi - f\|_p \leq \sum_{i=1}^{n} \|y_i(g_i - \chi_{A_i,x})\|_p + \epsilon/2 = \sum_{i=1}^{n} \|y_i\| \|g_i - \chi_{A_i,x}\|_p + \epsilon/2 \leq \sum_{i=1}^{n} \frac{\|y_i\|\epsilon}{2(n+1)\|y\|} + \epsilon/2 < \epsilon \). Hence, \( g \in \mathcal{L}_p(X, \mathbb{Y}) \).

This completes the proof of the proposition. \( \square \)

**Proposition 11.183** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathbb{Y} \) be a separable Banach space, \( f : X \to \mathbb{Y} \) be \( \mathcal{B} \)-measurable, \( \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \), \( f|_E \) be absolutely integrable over \( E := (E, \mathcal{B}_E, \mu_E), \) which is the finite measure subspace of \( \mathcal{X} \), and \( M \in [0, \infty) \subset \mathbb{R} \). Assume that, \( \forall E \in \mathcal{B} \) with \( 0 < \mu(E) < \infty \), we have \( \|f|_E \|_{L_p(E)} \leq M \). Then, \( \mathcal{P} \circ f \leq M \) a.e. in \( \mathcal{X} \).

**Proof** Consider the open set \( O := \{y \in \mathbb{Y} \mid \|y\| > M\} \subseteq \mathbb{Y} \).

Since \( \mathbb{Y} \) is separable, by Propositions 4.38 and 4.4, \( O \) is second countable and separable. Let \( D \subseteq O \) be a countable dense set in \( O \) (the relative closure of \( D \) with respect to \( O \) equals to \( O \)). It is easy to show that \( \mathcal{M} := \{B_y(y, r) \subseteq O \mid y \in D, r \in \mathbb{Q}, r > 0\} \) is a countable basis for \( O \). Let \( E := f_{\text{inv}}(O) = f_{\text{inv}}(\bigcup_{B_y(y, r) \in \mathcal{M}} B_y(y, r)) = \bigcup_{B_y(y, r) \in \mathcal{M}} f_{\text{inv}}(B_y(y, r)) \).

We will show that \( \mu(f_{\text{inv}}(B_y(y, r))) = 0, \forall B_y(y, r) \in \mathcal{M} \), by an argument of contradiction. Suppose that \( \exists y \in D, \exists r \in \mathbb{Q} \) with \( r > 0 \) such that \( B_y(y, r) \subseteq O \) and \( \mu(E) := \mu(f_{\text{inv}}(B_y(y, r))) > 0 \). Since \( \mathcal{X} \) is \( \sigma \)-finite, then \( \exists E \in \mathcal{B} \) with \( E \subseteq E \) such that \( 0 < \mu(E) < +\infty \). \( \forall x \in E \), we have \( x \in \bar{E} \) and \( \|f(x) - y\| < r \). Then, \( \|f(x) - y\| = \frac{1}{\mu(E)} \int_E \frac{f(x) - y}{\mu(E)} d\mu \leq \frac{1}{\mu(E)} \int_E \frac{0}{\mu(E)} d\mu = 0 \).

Then, \( 0 \leq \mu(E) \leq \sum_{B_y(y, r) \in \mathcal{M}} \mu(f_{\text{inv}}(B_y(y, r))) = 0. \) Hence, \( \mathcal{P} \circ f \leq M \) a.e. in \( \mathcal{X} \). \( \square \)
Lemma 11.184 Let \( p, q \in (1, \infty) \subset \mathbb{R} \) with \( 1/p + 1/q = 1 \), \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( Y \) be a separable normed linear space over \( \mathbb{K} \) with \( Y^* \) being separable, \( Z := L_p(X, Y) \) be the normed linear space over \( \mathbb{K} \) as defined in Example 11.173, and \( g : X \to Y^* \) be \( \mathcal{B} \)-measurable. Assume that

(i) \( \forall E \in \mathcal{B} \) with \( \mu(E) < +\infty \), \( g \) is absolutely integrable over \( \mathcal{E} := (E, \mathcal{B}_E, \mu_E) \), which is the finite measure subspace of \( X \);

(ii) \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \forall \) simple function \( \phi : X \to Y \) (\( \phi \in L_p(X, Y) \)), the function \( \langle \langle \langle g(\cdot), \phi(\cdot) \rangle \rangle \rangle : X \to \mathbb{K} \) is absolutely integrable over \( X \) and \( \int_X \langle \langle \langle (g(x), \phi(x)) \rangle \rangle \rangle \, d\mu(x) \leq M \| \phi \|_p \).

Then, \( g \in \bar{L}_q(X, Y^*) \) and \( \| g \|_q \leq M \).

Proof By Proposition 11.116, define a \( \sigma \)-finite \( Y^* \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) by: \( \nu(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( \int_E \mathcal{P} \circ g \, d\mu = \infty \); \( \nu(E) = \int_E g \, d\mu \in Y^*, \forall E \in \mathcal{B} \) with \( \int_E \mathcal{P} \circ g \, d\mu < \infty \). Then, \( \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ g \, d\mu, \forall E \in \mathcal{B} \). By (i), \( \mathcal{P} \circ \nu(E) < \infty, \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \). By Proposition 11.92, \( \nu(E) = 0 \), \( \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \). Fix any \( E \in \mathcal{B} \) with \( \mu(E) < \infty \). Define \( \mathcal{P}_q \circ \nu(E) \in [0, \infty] \subset \mathbb{R} \) by

\[
\mathcal{P}_q \circ \nu(E) := \sup_{n \in \mathbb{Z}^+, \{E_i\}_{i=1}^n \subseteq \mathcal{B}, \bigcup_{i=1}^n E_i = E} \sum_{i=1}^n \frac{\|\nu(E_i)\|^q}{(\mu(E_i))^q} \mu(E_i)
\]

\( \forall n \in \mathbb{Z}^+, \forall \) pairwise disjoint \( \{E_i\}_{i=1}^n \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^n E_i, \forall \epsilon \in (0, 1) \subset \mathbb{R}, \forall i \in \{1, \ldots, n\} \) with \( \mu(E_i) > 0 \), by Lemma 7.75, \( \exists \gamma_i \in Y \) such that \( \|\gamma_i\| = \|\nu(E_i)\|^{q-1}/(\mu(E_i))^{q-1} \) and \( \langle \langle \nu(E_i), y_i \rangle \rangle \geq (1 - \epsilon) \|\nu(E_i)\| \|y_i\| = (1 - \epsilon) \|\nu(E_i)\|^{q-1}/(\mu(E_i))^{q-1} \). Define a simple function \( \phi = \sum_{i=1}^n \|\nu(E_i)\| y_i \chi_{E_i} \). Then,

\[
\| \phi \|_p = \left( \sum_{i=1}^n \frac{\|\nu(E_i)\|^{(q-1)p}}{(\mu(E_i))^{(1-q)p}} \mu(E_i) \right)^{1/p} = \left( \sum_{i=1}^n \frac{\|\nu(E_i)\|^{q}}{(\mu(E_i))^{q}} \mu(E_i) \right)^{1/p}
\]

By (ii), we have

\[
\sum_{i=1}^n (1 - \epsilon) \|\nu(E_i)\|^{q}/(\mu(E_i))^{q-1} \leq \sum_{i=1}^n \langle \langle \nu(E_i), y_i \rangle \rangle \leq \sum_{i=1}^n \langle \langle \int_{E_i} g \, d\mu, y_i \rangle \rangle = \sum_{i=1}^n \left| \langle \langle \int_{E_i} g \, d\mu \rangle \rangle \right|
\]

\[
= \sum_{i=1}^n \left| \langle \langle \int_{E_i} g \, d\mu, y_i \rangle \rangle \right| = \sum_{i=1}^n \left| \langle \langle \int_{E_i} \langle \langle g(x), y_i \rangle \rangle \rangle \, d\mu(x) \rangle \rangle \right|
\]
\[\int_X \left| \langle g(x), \phi(x) \rangle \right| d\mu(x) \leq M \left( \sum_{i=1}^{\bar{n}} \| \nu(E_i) \| q (\mu(E_i))^{1-q} \right)^{1/p}\]

where the second and third equalities follow from Proposition 11.92. Hence, we have

\[\sum_{i=1}^{\bar{n}} \mu(E_i) > 0 \| \nu(E_i) \| q (\mu(E_i))^{-q} \mu(E_i) \leq M^q/(1 - \epsilon)^q.\]

By the arbitrariness of \(\epsilon\), we have

\[\int_X \left| \langle g(x), \phi(x) \rangle \right| d\mu(x) \leq M^q.\]

Then, \(P_q \circ \nu(E) \leq M^q.\)

Since \(X\) is \(\sigma\)-finite, then \(\exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B}\) such that \(X = \bigcup_{n=1}^{\infty} X_n\) and \(\mu(X_n) < \infty, \forall n \in \mathbb{N}\). Without loss of generality, we may assume that \(X_n \subseteq X_{n+1}, \forall n \in \mathbb{N}\). By Proposition 11.66, there exists a sequence of simple functions \((\psi_i)_{i=1}^{\infty}, \psi_i : X \to [0, \infty), \forall i \in \mathbb{N}\), such that \(\| \psi_i(x) \| \leq \| g(x) \|, \forall x \in X, \forall i \in \mathbb{N}\), and \(\lim_{i \in \mathbb{N}} \psi_i = g\) a.e. in \(X\). Fix any \(n \in \mathbb{N}\), let \(E_n := \{ x \in X_n \ | \ \| g(x) \| \leq n \} \in \mathcal{B}\) and \(\mathcal{E}_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n})\) be the finite measure subspace of \(X\). Then, \((P \circ g)|_{E_n}\) and \((P_q \circ g)|_{E_n}\) are integrable over \(\mathcal{E}_n\). By Propositions 7.21, 7.23, 11.52, and 11.53, \(\lim_{i \in \mathbb{N}} (P \circ (\psi_i - g))|_{E_n} = 0\) a.e. in \(\mathcal{E}_n\). Note that \((P \circ (\psi_i - g))|_{E_n}(x) \leq 2 (P \circ g)|_{E_n}(x), \forall x \in E_n, \forall i \in \mathbb{N}\). By Lebesgue Dominated Convergence Theorem 11.91, we have \(\lim_{i \in \mathbb{N}} \int_{E_n} P \circ (\psi_i - g) d\mu = 0\). Note also that \((P_q \circ \psi_i)|_{E_n}(x) \leq (P_q \circ g)|_{E_n}(x), \forall x \in E_n, \forall i \in \mathbb{N}\). By Propositions 7.21 and 11.52, \(\lim_{i \in \mathbb{N}} (P_q \circ \psi_i)|_{E_n} = (P_q \circ g)|_{E_n}\) a.e. in \(\mathcal{E}_n\).

Again, by Lebesgue Dominated Convergence Theorem 11.91, we have \(\lim_{i \in \mathbb{N}} \int_{E_n} P_q \circ \psi_i d\mu = \int_{E_n} P_q \circ g d\mu\). \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), \(\exists \epsilon_0 \in \mathbb{N}\) such that

\[0 \leq \int_{E_n} P \circ (\psi_{i_0} - g) d\mu < \frac{\epsilon}{2\mu_{E_n}}\]

and

\[\left| \int_{E_n} P_q \circ \psi_{i_0} d\mu - \int_{E_n} P_q \circ g d\mu \right| < \epsilon/2.\]

Let \(\psi_{i_0}\) admit the canonical representation \(\psi_{i_0} = \sum_{j=1}^{\bar{n}} y_{s,j} \chi_{A_{s,j}}.\) Then,

\[\int_{E_n} P_q \circ g d\mu < \int_{E_n} P_q \circ \psi_{i_0} d\mu + \epsilon/2\]

\[= \sum_{\mu(A_j \cap E_n) > 0} \| \nu(A_j \cap E_n) \| q (\mu(A_j \cap E_n))^{1-q} + \epsilon/2\]

\[= \sum_{\mu(A_j \cap E_n) > 0} \| \nu(A_j \cap E_n) \| q (\mu(A_j \cap E_n))^{1-q} + \sum_{j=1}^{\bar{n}} y_{s,j} \| q (\mu(A_j \cap E_n))^{1-q} + \sum_{\mu(A_j \cap E_n) > 0} \| \nu(A_j \cap E_n) \| q (\mu(A_j \cap E_n))^{1-q} + \epsilon/2\]

\[\text{Let } A_j = A_j \cap E_n, j = 1, \ldots, \bar{n}. \text{ Then,}\]

\[\int_{E_n} P_q \circ g d\mu < \sum_{\mu(A_j) > 0} \| \nu(A_j) \| q (\mu(A_j))^{1-q} + \epsilon/2\]

\[= \sum_{\mu(A_j) > 0} \| \nu(A_j) \| q (\mu(A_j))^{1-q} + \sum_{\mu(A_j) > 0} \| y_{s,j} \| q (\mu(A_j))^{1-q} + \epsilon/2.\]
\[ \begin{align*}
&\leq \mathcal{P}_q \circ \nu(E_n) - \sum_{\mu(A_j) > 0}^{\bar{n}} \| \nu(A_j) \|^q (\mu(A_j))^{1-q} \\
&+ \sum_{\mu(A_j) > 0}^{\bar{n}} \left\| \int_{A_j} \psi_{t_j} \, d\mu \right\|^q (\mu(A_j))^{1-q} + \epsilon/2 \\
&\leq M^q + \sum_{\mu(A_j) > 0}^{\bar{n}} (\mu(A_j))^{1-q} \left( \left\| \int_{A_j} \psi_{t_j} \, d\mu \right\|^q - \left\| \int_{A_j} g \, d\mu \right\|^q \right) + \epsilon/2 \\
&= M^q + \sum_{\mu(A_j) > 0}^{\bar{n}} (\mu(A_j))^{1-q} \left( \left\| \int_{A_j} \psi_{t_j} \, d\mu \right\| - \left\| \int_{A_j} g \, d\mu \right\| \right) q \\
&\quad \cdot \left( t_j \left( \left\| \int_{A_j} \psi_{t_j} \, d\mu \right\| + (1 - t_j) \left\| \int_{A_j} g \, d\mu \right\| \right)^{q-1} + \epsilon/2 \\
&\leq M^q + \sum_{\mu(A_j) > 0}^{\bar{n}} qn^{q-1} \left\| \int_{A_j} (\psi_{t_j} - g) \, d\mu \right\| + \epsilon/2 \\
&\leq M^q + \sum_{\mu(A_j) > 0}^{\bar{n}} qn^{q-1} \int_{A_j} \mathcal{P} \circ (\psi_{t_j} - g) \, d\mu + \epsilon/2 \\
&\leq M^q + qn^{q-1} \int_{E_n} \mathcal{P} \circ (\psi_{t_j} - g) \, d\mu + \epsilon/2 < M^q + \epsilon
\end{align*} \]

where the first equality follows from the Mean Value Theorem 9.20 and \( t_j \in (0, 1) \subset \mathbb{R}, \ j = 1, \ldots, \bar{n} \); the fourth inequality follows from Proposition 11.92 and the fact that \( \| \psi_{t_j}(x) \| \leq \| g(x) \| \leq n, \ \forall x \in E_n \); the second equality and the fifth inequality follow from Proposition 11.92; and the sixth inequality follows from Proposition 11.83. By arbitrariness of \( \epsilon \), we have \( \int_{E_n} \mathcal{P}_q \circ g \, d\mu \leq M^q \). Clearly, we have \( E_n \subseteq E_{n+1} \), \( \forall n \in \mathbb{N} \), and \( \bigcup_{n=1}^{\infty} E_n = X \). Then, by Monotone Convergence Theorem 11.81, we have \( \int_X \mathcal{P}_q \circ g \, d\mu = \lim_{n \in \mathbb{N}} \int_{E_n} (\mathcal{P}_q \circ g) \chi_{E_n,x} \, d\mu = \lim_{n \in \mathbb{N}} \int_{E_n} \mathcal{P}_q \circ g \, d\mu \leq M^q \). Hence, \( \| g \|_q \leq M \) and \( g \in \mathcal{L}_q(X, \mathcal{Y}) \).

This completes the proof of the lemma. \( \square \)

**Lemma 11.185** Let \( p \in [1, \infty) \subset \mathbb{R}, \ \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathcal{Y} \) be a separable reflexive Banach space over \( \mathbb{K} \) with \( \mathcal{Y}^* \) being separable, and \( \mathcal{Z} := \mathcal{L}_p(X, \mathcal{Y}) \) be the Banach space over \( \mathbb{K} \) as defined in
Example 11.179. Then, \( \forall f \in \mathbb{Z}^* \), \( \exists g : X \rightarrow \mathbb{Y}^* \), which is \( B \)-measurable, such that

(i) \( \forall E \in B \) with \( \mu(E) < +\infty \), \( g \) is absolutely integrable over \( E := (E, B_E, \mu_E) \), which is the finite measure subspace of \( X \);

(ii) \( \forall \) simple function \( \phi : X \rightarrow \mathbb{Y} \) (\( \phi \in \tilde{Z} := \mathbb{L}_p(X, \mathbb{Y}) \)), the function \( \langle \langle g(\cdot), \phi(\cdot) \rangle \rangle : X \rightarrow \mathbb{K} \) is absolutely integrable over \( X \), \( f(\langle \langle \phi \rangle \rangle) = \int_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \) and \( \left| \int_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \right| \leq \| f \|_p \| \phi \|_p \).

Furthermore, \( g \) is unique in the sense that \( \tilde{g} : X \rightarrow \mathbb{Y}^* \) is another function with the above properties if, and only if, \( \tilde{g} \) is \( B \)-measurable and \( g = \tilde{g} \) a.e. in \( X \).

**Proof** Fix any \( f \in \mathbb{Z}^* \). \( \forall E \in B \) with \( \mu(E) < \infty \), \( \forall y \in \mathbb{Y} \), let \( z_{E,y} := y_{X,E} \in \tilde{Z} \). Define \( f_E : \mathbb{Y} \rightarrow \mathbb{K} \) by \( f_E(y) = f([z_{E,y}]) \), \( \forall y \in \mathbb{Y} \). Since \( f \in \mathbb{Z}^* \), then \( f_E \) is linear and continuous and \( f_E = \chi_{E,Y} \). Define \( \bar{E}_X \) with \( \bar{E}_X = \bigcup_{i=1}^{\infty} E_i \). Then, \( \sum_{i=1}^{\infty} \| y_{E_i} \| leq \| \mu(\bar{E}) \|^{1/p} < \infty \).

**Proof of claim:** \( \sum_{i=1}^{\infty} \| y_{E_i} \| = \sum_{i=1}^{\infty} \sup_{y \in \mathbb{Y}, \| y \| leq 1} | f_{E_i}(y) | \). \( \forall i \in \mathbb{N} \), by Propositions 7.85 and 7.90, \( \exists y_i \in \mathbb{Y} \) with \( \| y_i \| leq 1 \) such that \( \| y_{E_i} \| = f_{E_i}(y_i) = f([z_{E_i,y_i}]) \). Then, \( \forall n \in \mathbb{N} \), \( \sum_{i=1}^{n} \| y_{E_i} \| = \sum_{i=1}^{n} f([z_{E_i,y_i}]) \leq \| f \| \| \sum_{i=1}^{n} z_{E_i,y_i} \|_p = \| f \| \| \sum_{i=1}^{n} z_{E_i,y_i} \|_p = \| f \| \cdot (\mu(E))^{|1/p|} < \infty \), by the linearity of \( f \).

**Claim 11.185.1** \( \forall E \in B \) with \( \mu(E) < \infty \), \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq B \) with \( E = \bigcup_{i=1}^{\infty} E_i \). Then, \( \sum_{i=1}^{\infty} \| y_{E_i} \| = \| \sum_{i=1}^{\infty} \| y_{E_i} \| \leq \| f \| \cdot (\mu(E))^{|1/p|} \). This completes the proof of the claim.

**Claim 11.185.2** \( \forall E \in B \) with \( \mu(E) < \infty \), \( \forall \) pairwise disjoint \( (E_n)_{n=1}^\infty \subseteq B \) with \( E = \bigcup_{n=1}^{\infty} E_n \), we have \( \sum_{n=1}^{\infty} y_{E,n} = \sum_{n=1}^{\infty} y_{E_n} \in \mathbb{Y}^* \).

**Proof of claim:** By Claim 11.185.1 and Propositions 7.27 and 7.72, \( \sum_{n=1}^{\infty} y_{E,n} \in \mathbb{Y}^* \). \( \forall y \in \mathbb{Y} \),

\[
\langle \sum_{n=1}^{\infty} y_{E,n}, y \rangle = \lim_{n \to \infty} \langle \sum_{i=1}^{n} y_{E_i}, y \rangle = \sum_{i=1}^{n} \langle y_{E_i}, y \rangle = \lim_{n \to \infty} \sum_{i=1}^{n} f([z_{E_i,y}]) = \lim_{n \to \infty} f(\sum_{i=1}^{n} z_{E_i,y})
\]

where the first equality follows from Propositions 7.72 and 3.66; and the last equality follows from the linearity of \( f \). Note that \( \lim_{n \to \infty} \sum_{i=1}^{n} z_{E_i,y}(x) = z_{E,y}(x) \), \( \forall x \in X \), and \( \| \sum_{i=1}^{n} z_{E_i,y}(x) - z_{E,y}(x) \|_p \leq \| y \|_p \chi_{E,X}(x) \), \( \forall x \in X \).
∀n ∈ N. Then, by Lebesgue Dominated Convergence Theorem 11.91, 
\lim_{n\to\infty} \int \mathcal{P}_p \left( (\sum_{i=1}^n z_{E_i}, y - z_{E_i}, y) \right) \, d\mu = 0 \quad \text{and} \quad \lim_{n\to\infty} \sum_{i=1}^n z_{E_i}, y = z_{E_i}, y \text{ in } \mathcal{Z}.

By Propositions 7.72 and 3.66, \lim_{n\to\infty} f(\sum_{i=1}^n z_{E_i}, y) = f(\langle y_{E_i}, y \rangle).\ Hence, \langle \sum_{i=1}^n y_{E_i}, y \rangle, \forall y \in Y.\n
This implies that, by Proposition 7.85, y_{E_i} = \sum_{i=1}^\infty \chi_{E_i}.\n
Since X is σ-finite, then \exists (X_n)_{n=1}^\infty \subseteq B \text{ such that } X = \bigcup_{n=1}^\infty X_n \text{ and } \mu(X_n) < \infty, \forall n \in N.\ Without loss of generality, we may assume that (X_n)_{n=1}^\infty is pairwise disjoint. Fix any n \in N, let \mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n) be the finite measure subspace of \mathcal{X}. We may define a function \nu_n : B_n \to Y by \nu_n(E) = f_E = y_{E}, \forall E \in \mathcal{B}_n.\ Clearly, \nu_n(\emptyset) = y_{\emptyset} = \emptyset_{y}.\n
∀ pairwise disjoint \{E_i\}_{i=1}^\infty \subseteq \mathcal{B}_n, let E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}_n. By Claim 11.185.1, \sum_{i=1}^\infty ||\nu_n(E_i)|| \leq ||f|| (\mu(E))^{1/p} \leq ||f|| (\mu(X_n))^{1/p} < \infty.\n
By Claim 11.185.2, \nu_n(E) = \sum_{i=1}^\infty \nu_n(E_i) \in Y. This shows that \nu_n is a Y*-valued pre-measure on (X_n, \mathcal{B}_n). By Claim 11.185.1, \mathcal{P} \circ \nu_n(X_n) \leq ||f|| (\mu(X_n))^{1/p} < +\infty. Then, \nu_n is finite. Hence, (X_n, \mathcal{B}_n, \nu_n) is a finite Y*-valued measure space. By Proposition 11.118, the generation process on ((X_n, \mathcal{B}_n, \nu_n))_{n=1}^\infty yields a unique σ-finite Y*-valued measure space \((X, \mathcal{B}, \nu)\) on X.

Next, we will show that \mathcal{P} \circ \nu(E) \leq ||f|| (\mu(E))^{1/p} < \infty and \nu(E) = y_{E}, \forall E \in \mathcal{B} \text{ with } \mu(E) < \infty. Fix any E \in \mathcal{B} \text{ with } \mu(E) < \infty. Let E_n := X_n \cap E \in \mathcal{B}_n, \forall n \in N. By Proposition 11.118, \mathcal{P} \circ \nu(E) = \sum_{n=1}^\infty \mathcal{P} \circ \nu_n(E_n), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}, \forall n \in N, \exists m_n \in \mathbb{Z}_+, \exists pairwise disjoint \{E_{n,i}\}_{i=1}^{m_n} \subseteq \mathcal{B}_n,\text{ with } E_n = \bigcup_{i=1}^{m_n} E_{n,i}, \text{ such that } \mathcal{P} \circ \nu_n(E_n) \leq \sum_{i=1}^{m_n} ||\nu_n(E_{n,i})|| + 2^{-n} \epsilon < \infty. Then,

\mathcal{P} \circ \nu(E) \leq \sum_{n=1}^{\infty} \sum_{i=1}^{m_n} ||\nu_n(E_{n,i})|| + 2^{-n} \epsilon \leq \sum_{n=1}^{\infty} ||y_{E_{n,i}}|| + \epsilon \leq ||f|| (\mu(E))^{1/p} + \epsilon < \infty

where the third inequality follows from Claim 11.185.1. By the arbitrariness of \epsilon, we have \mathcal{P} \circ \nu(E) \leq ||f|| (\mu(E))^{1/p} < \infty. Then, E \in \text{dom}(\nu) and \nu(E) = \sum_{n=1}^\infty \nu_n(E_n) \in Y. By Claim 11.185.2, we have \nu(E) = \sum_{n=1}^\infty y_{E_n} = y_{E}. Thus, by Proposition 11.137, \nu is uniquely defined independent of the choice of \((X_n)_{n=1}^\infty\).

∀E \in \mathcal{B} \text{ with } \mu(E) = 0, \mathcal{P} \circ \nu(E) \leq 0. Thus, \mathcal{P} \circ \nu \ll \mu. By Radon-Nikodym Theorem 11.171 and Proposition 7.90, \exists \gamma : X \to Y \text{ such that } g is the Radon-Nikodym derivative of \nu with respect to \mu. \forall E \in \mathcal{B} \text{ with } \mu(E) < \infty, we have \mathcal{P} \circ \gamma(E) = \int_E \gamma \, d\mu < \infty. Thus, (i) holds. \forall simple function \phi : X \to \mathbb{Y}, \exists M \in [0, +\infty) \subseteq \mathbb{R} \text{ such that } ||\phi(x)|| \leq M, \forall x \in X, \text{ and } \mu(E) := \mu\{x \in X \mid \phi(x) \neq \emptyset_y\} < \infty. Then, by Proposition 7.72, ||\langle \gamma(x), \phi(x) \rangle|| \leq M ||\gamma(x)|| \chi_{E,x}(x) = M \mathcal{P} \circ \gamma(x) \chi_{E,x}(x), \forall x \in X. By Propositions 7.72, 11.38, and 11.39, the function \langle \gamma(\cdot), \phi(\cdot) \rangle : X \to \mathbb{Y} \text{ is } \mathcal{B} \text{-measurable. Hence, the function } \langle \gamma(\cdot), \phi(\cdot) \rangle : X \to \mathbb{K} \text{ is absolutely integrable over } X. \text{ Let } \phi \text{ admit the canonical representation } \phi = \sum_{i=1}^n \chi_{E_i}.\ni
11.9. L_P SPACES

Then,
\[ f(\phi) = f\left(\sum_{i=1}^{n} y_i xE_i, X\right) = \sum_{i=1}^{n} f(y_i xE_i, X) = \sum_{i=1}^{n} f(\chi_{xE_i}, X) \]

\[ = n \sum_{i=1}^{n} \langle \langle yE_i, y_i \rangle \rangle = n \sum_{i=1}^{n} \langle \langle \mu(E_i), y_i \rangle \rangle = n \sum_{i=1}^{n} \left( \int_{E_i} g \, d\mu, y_i \right) \]

\[ = \sum_{i=1}^{n} \int_{E_i} \langle \langle g(x), y_i \rangle \rangle \, d\mu(x) = \int_{X} \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \in K \]

where the last two equalities follow from Proposition 11.92. Then, \[ |f_X(\langle \langle g(x), \phi(x) \rangle \rangle) | \, d\mu(x) = \| f \| \cdot \| \phi \| = \| f \| \| \phi \|^p. \] Thus (ii) holds. Hence, \( g \) is the function we seek. Since \( \nu \) is uniquely defined and \( g \) is also uniquely defined, then \( g \) is unique as desired.

This completes the proof of the lemma. \( \square \)

**Theorem 11.186 (Riesz Representation Theorem)** Let \( p \in [1, \infty) \subset \mathbb{R} \), \( q \in (1, \infty) \subset \mathbb{R} \) with \( 1/p + 1/q = 1 \), \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( Y \) be a separable reflexive Banach space over \( K \) with \( Y^* \) being separable, and \( Z := L_p(X, Y) \) be the Banach space over \( K \) as defined in Example 11.179. Then, \( Z^* = L_q(X, Y^*) \) isometrically isomorphically. In particular, \( \forall f \in Z^*, \exists [g] \in L_q(X, Y^*) \) with \( g \in \bar{L}_q(X, Y^*) \) such that \( f([z]) = \langle \langle [g], [z] \rangle \rangle := f_X(\langle \langle g(x), z(x) \rangle \rangle) \, d\mu(x) \) and the function \( \langle \langle g(\cdot), z(\cdot) \rangle \rangle : X \to K \) is absolutely integrable over \( X \), \( \forall z \in Z := \bar{L}_p(X, Y) \), and \( \| f \| = \| g \|_q \).

**Proof** \( \forall f \in Z^* \), by Lemma 11.185, \( \exists g : X \to Y^* \), which is \( \mathcal{B} \)-measurable, such that

(i) \( \forall E \in \mathcal{B} \) with \( \mu(E) < +\infty \), \( g \) is absolutely integrable over \( E := (E, \mathcal{B}_E, \mu_E) \), which is the finite measure subspace of \( X \);

(ii) \( \forall \) simple function \( \phi : X \to Y \) \( (\phi \in \bar{L}_p(X, Y)) \), the function \( \langle \langle g(\cdot), \phi(\cdot) \rangle \rangle : X \to K \) is absolutely integrable over \( X \), \( f(\phi) = f_X(\langle \langle g(x), \phi(x) \rangle \rangle) \, d\mu(x) \) and \( \| f(\phi) \|_p \leq \| f \| \| \phi \|_p \).

Furthermore, \( g \) is unique in the sense that \( \hat{g} : X \to Y^* \) is another function with the above properties if, and only if, \( \hat{g} \) is \( \mathcal{B} \)-measurable and \( g = \hat{g} \) a.e. in \( X \).

We will show that \( g \in \bar{L}_q(X, Y^*) \) and \( \| g \|_q \leq \| f \| < +\infty \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( 1 < p < +\infty \); Case 2: \( p = 1 \). Case 1: \( 1 < p < +\infty \). By Lemma 11.184, we have \( g \in \bar{L}_q(X, Y^*) \) and \( \| g \|_q \leq \| f \| < +\infty \). Case 2: \( p = 1 \). \( \forall E \in \mathcal{B} \) with \( 0 < \mu(E) < \infty \), by (i) and Proposition 11.92, \( y_{E,E} := \int_{E} g \, d\mu, y_{E,E} \in Y^* \) and, \( \forall y \in Y \), \( \| \langle \langle y_{E,E}, y \rangle \rangle \| = \| \int_{E} \langle \langle g(x), y \rangle \rangle \, d\mu(x) \| = \| \int_{X} \langle \langle g(x), y_{E,E}(x) \rangle \rangle \, d\mu(x) \| \leq \| f \| \| y_{E,E} \|_1 = \| f \| \| y \| \mu(E) \), where
the first two equalities follow from Proposition 11.92; and the inequality follows from (ii). Then, \( \| y_E \| / \mu(E) \leq \| f \| \). By Proposition 11.183, we have \( \mathcal{P} \circ g \leq \| f \| \) a.e. in \( \mathcal{X} \). Hence \( \| g \|_\infty \leq \| f \| \) and \( g \in L_\infty(\mathcal{X}, \mathcal{Y}^*) \).

Hence, in both cases, we have \( g \in L_q(\mathcal{X}, \mathcal{Y}^*) \) and \( \| g \|_q \leq \| f \| < +\infty \).

\( \forall z \in \tilde{Z} \), by Proposition 11.181, there exists a sequence of simple functions \( (\varphi_i)_{i=1}^\infty, \varphi_i : \mathcal{X} \to \mathcal{Y}, \forall i \in \mathbb{N} \), such that \( \lim_{i \to \infty} \varphi_i = z \) a.e. in \( \mathcal{X} \). \( \| \varphi_i(x) \| \leq \| z(x) \|, \forall x \in \mathcal{X}, \forall i \in \mathbb{N} \) and \( \lim_{i \to \infty} \varphi_i = z \) in \( \tilde{Z} \). By Propositions 7.72 and 3.66, \( f([z]) = \lim_{i \to \infty} f([\varphi_i]) = \lim_{i \to \infty} \int_X \langle \langle g(x), \varphi_i(x) \rangle \rangle \, d\mu(x) \). Note that, by Propositions 7.72, 11.52, and 11.53, \( \lim_{i \to \infty} \langle \langle g(x), \varphi_i(x) \rangle \rangle = \langle \langle g(x), z(x) \rangle \rangle \) a.e. \( x \in \mathcal{X} \) and \( \| \langle \langle g(x), \varphi_i(x) \rangle \rangle \| \leq \| g(x) \| \| \varphi_i(x) \| \leq \| g(x) \| \| z(x) \|, \forall x \in \mathcal{X}, \forall i \in \mathbb{N} \). By Hölder’s Inequality 11.178, \( \int_X \langle \langle g(x), \varphi_i(x) \rangle \rangle \, d\mu(x) \leq \| g \|_q \| z \|_p < \infty \). By Propositions 7.72, 11.38, and 11.39, \( \langle \langle g(\cdot), \varphi_i(\cdot) \rangle \rangle : \mathcal{X} \to \mathcal{Y} \) and \( \langle \langle g(\cdot), z(\cdot) \rangle \rangle : \mathcal{X} \to \mathcal{Y} \) are \( \mathcal{B} \)-measurable, \( \forall i \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.91, \( f([z]) = \lim_{i \to \infty} \int_X \langle \langle g(x), \varphi_i(x) \rangle \rangle \, d\mu(x) = \int_X \langle \langle g(x), z(x) \rangle \rangle \, d\mu(x) \) and the function \( \langle \langle g(\cdot), z(\cdot) \rangle \rangle : \mathcal{X} \to \mathcal{Y} \) is absolutely integrable over \( \mathcal{X} \). Note that

\[
\| f \| = \sup_{[z] \in \tilde{Z}, \| z \|_p \leq 1} \| f([z]) \| = \sup_{z \in \tilde{Z}, \| z \|_p \leq 1} \left| \int_X \langle \langle g(x), z(x) \rangle \rangle \, d\mu(x) \right|
\]

\[
\leq \sup_{z \in \tilde{Z}, \| z \|_p \leq 1} \int_X \| \langle \langle g(x), z(x) \rangle \rangle \| \cdot d\mu(x)
\]

\[
\leq \sup_{z \in \tilde{Z}, \| z \|_p \leq 1} \| g \|_q \| z \|_p \leq \| g \|_q
\]

where the first inequality follows from Proposition 11.92; and the second inequality follows from Hölder’s Inequality 11.178. Hence, we have \( \| f \| = \| g \|_q \).

Thus, we may define a function \( \Phi : \tilde{Z}^* \to L_q(\mathcal{X}, \mathcal{Y}^*) \) by \( \Phi(f) = [g], \forall f \in \mathcal{Z} \). Clearly, \( \Phi \) is well defined and \( \| \Phi(f) \|_q = \| f \|, \forall f \in \tilde{Z}^* \). \( \forall f_1, f_2 \in \mathcal{Z}^*, \forall \alpha, \beta \in \mathbb{K} \), let \( [g_1] = \Phi(f_1) \) and \( [g_2] = \Phi(f_2) \). Then, \( \Phi(\alpha f_1 + \beta f_2) = \alpha \Phi(f_1) + \beta \Phi(f_2) \).

Hence, \( \Phi \) is linear. \( \forall [g] \in L_q(\mathcal{X}, \mathcal{Y}^*) \), we may define \( f : \mathcal{Z} \to \mathcal{X} \) by \( f([z]) = \int_X \langle \langle g(x), z(x) \rangle \rangle \, d\mu(x) \), \( \forall z \in \tilde{Z} \). By Hölder’s Inequality 11.178 and Proposition 11.92, we have \( f \in \mathcal{Z}^* \). Then, \( \Phi(f) = [g] \) and \( \Phi \) is surjective. \( \forall f_1, f_2 \in \mathcal{Z}^* \), \( \Phi(f_1) = \Phi(f_2) \), we have \( 0 = \| \Phi(f_1) - \Phi(f_2) \|_q = \| \Phi(f_1 - f_2) \|_q = \| f_1 - f_2 \|_q \). Hence, \( \Phi \) is injective. Therefore, \( \Phi : \tilde{Z}^* \to L_q(\mathcal{X}, \mathcal{Y}^*) \) is a isometrical isomorphism. This completes the proof of the theorem. \( \square \)
11.10 Dual of $\mathcal{C}(X, \mathcal{Y})$ and $C_c(X, \mathcal{Y})$

**Proposition 11.187** Let $X := (X, \rho)$ be a metric space and $X := (X, \mathcal{B}(X), \mu)$ be a metric measure space. Then, the following statements hold.

(i) $\forall$ compact subset $K \subseteq X$, we have $K \in \mathcal{B}(X)$.

(ii) If $X$ is locally finite and $\sigma$-compact, then it is $\sigma$-finite.

(iii) If $X$ is finite, then it is locally finite.

**Proof**

(i) Fix any compact subset $K \subseteq X$. We will show that $K = \bigcap_{n=1}^{\infty} \bigcup_{x \in X_n} \mathcal{B}_X(x, 1/n) =: E$, where $X_n$'s are finite subsets of $K$. Clearly, we have $E \in \mathcal{B}(X)$.

$\forall n \in \mathbb{N}, K \subseteq \bigcup_{x \in K} \mathcal{B}_X(x, 1/n)$. Since $K$ is compact, there exists a finite set $X_n \subseteq K$ such that $K \subseteq \bigcup_{x \in X_n} \mathcal{B}_X(x, 1/n)$. Then, $K \subseteq E$. On the other hand, $\forall \bar{x} \in \bar{K}$, by Proposition 5.5, $K$ is closed in $X$. By Proposition 4.10, dist($\bar{x}, K$) > 0. Then, $\exists n_0 \in \mathbb{N}$ such that dist($\bar{x}, K$) > 1/n_0. Then, $\bar{x} \not\in \mathcal{B}_X(x, 1/n_0)$, $\forall x \in K$, which implies that $\bar{x} \not\in \bigcup_{x \in X_n} \mathcal{B}_X(x, 1/n_0)$, which further implies that $\bar{x} \in \bar{E}$. Hence, $\bar{E} \subseteq K$.

Thus, we have $K = E \in \mathcal{B}(X)$.

(ii) By $X$ being $\sigma$-compact, there exists $(K_n)_{n=1}^{\infty} \subseteq X$ such that $K_n$ is compact, $\forall n \in \mathbb{N}$. By $X$ being locally finite, we have $K_n \in \mathcal{B}(X)$ and $\mu(K_n) < \infty$. Then, $X$ is $\sigma$-finite.

(iii) $\forall$ compact subset $K \subseteq X$, by (i), $K \in \mathcal{B}(X)$. Since $X$ is finite, then $\mu(K) \leq \mu(X) < \infty$. Then, $X$ is locally finite.

This completes the proof of the proposition.

**Proposition 11.188** Let $X := (X, \mathcal{O})$ be a topological space, $(X_n)_{n=1}^{\infty} \subseteq \mathcal{O}$, $X_n := (X_n, \mathcal{O}_n)$ be a topological subspace of $X$, and $(X_n, \mathcal{B}_n(X_n), \nu_n) =: \mathcal{X}_n$ be a finite topological measure space, $\forall n \in \mathbb{N}$. Assume that $(\mathcal{X}_n)_{n=1}^{\infty}$ satisfies the assumptions of Proposition 11.118. By Proposition 11.118, the generation process on $(\mathcal{X}_n)_{n=1}^{\infty}$ yields a unique $\sigma$-finite measure space $\bar{X} := (X, \mathcal{B}, \mu)$ such that $\mathcal{X}_n$ is the finite measure subspace of $\bar{X}$, $\forall n \in \mathbb{N}$. Then, $\mathcal{X} := (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$ is the unique $\sigma$-finite topological measure space on $\mathcal{X}$ such that $\mathcal{X}_n$ is the finite topological measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$.

**Proof**

$\forall E \in \mathcal{B}(\mathcal{X})$, by Proposition 11.25, we have $E \cap X_n \in \mathcal{B}_n(X_n)$, $\forall n \in \mathbb{N}$. Define $E_{(1)} := E \cap X_1$ and $E_{(n+1)} := (E \cap X_{n+1}) \setminus (\bigcup_{i=1}^{n} E_{(i)}), \forall n \in \mathbb{N}$. Then, $(E_{(n)})_{n=1}^{\infty}$ is pairwise disjoint, $\bigcup_{n=1}^{\infty} E_{(n)} = E$, $E_{(n)} \in \mathcal{B}_n(X_n) \subseteq \mathcal{B}_n(X)$, $\forall n \in \mathbb{N}$. Define $\bar{\mu} : \mathcal{B}_n(X) \rightarrow [0, \infty) \subset \mathbb{R}_+$ by $\bar{\mu}(E) = \sum_{n=1}^{\infty} \nu_n(E_{(n)}) \in [0, \infty] \subset \mathbb{R}_+$, $\forall E \in \mathcal{B}_n(X)$.

We will show that $(X, \mathcal{B}(\mathcal{X}), \bar{\mu})$ is a $\sigma$-finite measure space and $\mathcal{X}_n$ is the finite measure subspace of it, $\forall n \in \mathbb{N}$. It is easy to show that $\bar{\mu}(\emptyset) = 0$.

$\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}(\mathcal{X})$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}(\mathcal{X})$. Note
that \( E_{(1)} = E \cap X_1 = \bigcup_{i=1}^{\infty} (E_i \cap X_1) = \bigcup_{i=1}^{\infty} E_{i,(1)} \). Inductively, \( E_{(n+1)} = (E \cap X_{n+1}) \setminus \bigcup_{i=1}^{n} E_{i,(i)} = (E \cap X_{n+1}) \setminus \bigcup_{i=1}^{n} (E_i \cap X_{n+1}) \). Then, \( \mu(E) = \sum_{n=1}^{\infty} \mu_n(E_{(n)}) = \sum_{n=1}^{\infty} \mu_n(\bigcup_{j=1}^{\infty} E_{j,(n)}) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \mu_n(E_{j,(n)}) \)

where the fourth equality follows from the fact that all summands are non-negative real numbers. Hence, \((X, B_{(\mathcal{X})}, \mu)\) is a measure space. It is \(\sigma\)-finite since \(\bar{\mu}(X_n) = \sum_{i=1}^{\infty} \mu_i(X_{n,(i)}) = \sum_{i=1}^{\infty} \mu_n(X_{n,(i)}) = \mu_n(X_n) < \infty, \forall n \in \mathbb{N}\), where the second equality follows from the consistency assumption of Proposition 11.118 on \((X_n)_{n=1}^{\infty}\).

Fix any \( n \in \mathbb{N} \), \( \forall E \in B_{(\mathcal{X})} \), \( \bar{\mu}(E) = \sum_{i=1}^{\infty} \mu_i(E_{(i)}) = \sum_{i=1}^{\infty} \mu_n(E_{n,(i)}) = \mu_n(\bigcup_{j=1}^{\infty} E_{j,(n)}) = \mu_n(E) \), where the second equality follows from the consistency assumption of Proposition 11.118 on \((X_n)_{n=1}^{\infty}\).

By Propositions 11.25 and 11.13, we have \( X_n \) is the finite measure subspace of \((X, B_{(\mathcal{X})}, \bar{\mu})\), \( \forall n \in \mathbb{N} \). By Proposition 11.118, we have \( B = B_{(\mathcal{X})} \) and \( \mu = \bar{\mu} \). Then, \( X = (\mathcal{X}, B_{(\mathcal{X})}, \bar{\mu}) \).

Finally, we will show that \( X \) is a topological measure space. Fix any \( E \in B_{(\mathcal{X})} \). Then, \( E = \bigcup_{n=1}^{\infty} E_{(n)} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N} \), since \( X_n \) is a topological measure space and \( E_{(n)} \in B_{(\mathcal{X})} \), \( \exists O_n \in O_n \) with \( E_{(n)} \subseteq O_n \), such that \( \mu_n(O_n \setminus E_{(n)}) < 2^{-n} \epsilon \). Since \( X_n \) is a topological subspace of \( \mathcal{X} \) and \( X_n \in O \), then \( O_n \in O_n \subseteq O \subseteq B_{(\mathcal{X})} \). Let \( O := \bigcup_{n=1}^{\infty} O_n \in O \). Then, \( E = \bigcup_{n=1}^{\infty} E_{(n)} \subseteq \bigcup_{n=1}^{\infty} O_n = O \) and

\[
\mu(O \setminus E) = \bar{\mu}(O \setminus E) = \bar{\mu}\left(\bigcup_{n=1}^{\infty} O_n \setminus \left( \bigcup_{n=1}^{\infty} E_{(n)} \right) \right) \leq \bar{\mu}\left(\bigcup_{n=1}^{\infty} (O_n \cap \tilde{E}_{(n)}) \right) \leq \bar{\mu}\left(\bigcup_{n=1}^{\infty} (O_n \setminus E_{(n)}) \right) < \epsilon
\]

Hence, \( X \) is a topological measure space. By Proposition 11.29, \( X_n \) is the finite topological measure subspace of \( X \), \( \forall n \in \mathbb{N} \). This completes the proof of the proposition.

**Proposition 11.189** Let \( \mathcal{X} := (X, \rho) \) be a metric space with the natural topology \( O \), \( (X_n)_{n=1}^{\infty} \subseteq O \), \( X_n := (X_n, \rho) \) be a metric subspace of \( \mathcal{X} \),
and $\mathcal{X}_n := (X_n, \mathcal{B}(X_n), \mu_n)$ be a finite metric measure space, $\forall n \in \mathbb{N}$. Assume that $(\mathcal{X}_n)_{n=1}^{\infty}$ satisfies the assumptions of Proposition 11.118. By Proposition 11.188, the generation process on $(\mathcal{X}_n)_{n=1}^{\infty}$ yields a unique $\sigma$-finite metric measure space $\mathcal{X} := (X, \mathcal{B}(X), \mu)$ such that $\mathcal{X}_n$ is the finite metric measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$. Then, $\mathcal{X}$ is locally finite.

**Proof** Fix any compact subset $K \subseteq X$. By Proposition 11.187, $K \in \mathcal{B}(X)$. Clearly, $K \subseteq \bigcup_{n=1}^{\infty} X_n = X$. By the compactness of $K$, $\exists N \in \mathbb{N}$ such that $K \subseteq \bigcup_{n=1}^{N} X_n$. Then, $\mu(K) \leq \sum_{n=1}^{N} \mu(X_n) = \sum_{n=1}^{N} \mu_n(X_n) < \infty$, where the first inequality follows from Proposition 11.6; and the equality follows from Proposition 11.13. Hence, $\mathcal{X}$ is locally finite.

**Proposition 11.190** Let $p \in [1, \infty) \subset \mathbb{R}$, $X_1$ be a finite-dimensional Banach space, $X_2$ be a $\sigma$-compact metric space, $Y_1$ be a finite-dimensional Banach space over $X$, $Y_2$ be a separable normed linear space over $K$, $X \subseteq X_1$ be a subset with subset topology $O$, $X := (X, \mathcal{O})$, and $\mathcal{X} := (X, \mathcal{B}(X), \mu)$ be a metric measure space. Then, the following statements hold.

(i) $Y_1$ is separable.

(ii) $L_p(Y_2) =: Z_2$ is a separable normed linear space.

(iii) If $X$ is compact, then $\mathcal{C}(X, Y_1)$ is a separable Banach space.

(iv) If $X$ is a locally finite $\sigma$-compact metric measure space, then $L_p(\mathcal{X}, Y_1)$ is a separable Banach space.

(v) $X_2$ is separable.

**Proof** Let $Y_1$ be $m$-dimensional, where $m \in \mathbb{Z}_+$, and $\{e_1, \ldots, e_m\} \subseteq Y_1$ be a set of basis vectors. Then, $Y_1$ is isomorphic to $K^m$. The norm $\| \cdot \|_{Y_1}$ induces a norm $\| \cdot \|_1$ on $K^m$ as defined by, $\forall a := (\alpha_1, \ldots, \alpha_m) \in K^m$, $\| a \|_1 := \| \sum_{i=1}^{m} \alpha_i e_i \|_{Y_1} \in \mathbb{R}$. Then, $Y_1$ is isometrically isomorphic to the normed linear space $(K^m, K^m, \| \cdot \|_1)$, which is a Banach space by Theorem 7.36. By Theorem 7.38, $\| \cdot \|_1$ is equivalent to the Euclidean norm $| \cdot |$ on $K^m$.

(i) Let $K_Q := Q$ if $K = \mathbb{R}$; and $K_Q := \{ a + ib \in \mathbb{C} \mid a, b \in Q \}$, if $K = \mathbb{C}$. Clearly, $K_Q$ is a countable dense subset of $K$. Then, $K_Q^m$ is a countable dense subset of $(K^m, K^m, \| \cdot \|_1)$. Since $\| \cdot \|_1$ is equivalent to $| \cdot |$, then $K_Q^m$ is a countable dense subset of $(K^m, K^m, \| \cdot \|_1)$. Hence, $Y_1$ is separable, since $(K^m, K^m, \| \cdot \|_1)$ is separable.

(ii) Let $D \subseteq Y_2$ be a countable dense subset and, $\forall i \in \mathbb{N}$, $D_i := \{ (y_1, \ldots, y_i, \vartheta y_2, \vartheta y_2, \ldots) \in \mathbb{Z}_2 \mid y_j \in D, 1 \leq j \leq i \}$. Then, $D_i$ is countable, $\forall i \in \mathbb{N}$. Let $D := \bigcup_{i=1}^{\infty} D_i \subseteq \mathbb{Z}_2$, which is also countable. We will show that $D$ is dense in $\mathbb{Z}_2 = L_p(Y_2)$. $\forall z := (y_1, y_2, \ldots) \in \mathbb{Z}_2$, $\forall e \in (0, +\infty) \subset \mathbb{R}$, we have $\sum_{i=1}^{\infty} \| y_i \|_{Y_2} < \infty$. Then, $\exists N \in \mathbb{N}$ such that $\sum_{i=N+1}^{\infty} \| y_i \|_{Y_2} < \epsilon^p/2$. $\forall i \in \{1, \ldots, N\}$, $\exists \tilde{y}_i \in D$ such that $\| y_i - \tilde{y}_i \|_{Y_2} < \epsilon^p/2$. Therefore, $D$ is dense in $\mathbb{Z}_2 = L_p(Y_2)$.
Let \( \tilde{z} := (\tilde{y}_1, \ldots, \tilde{y}_N, \vartheta_{y_2}, \vartheta_{y_2}, \ldots) \in \tilde{D} \). Then, \( \|z - \tilde{z}\|_p = \left( \sum_{i=1}^{\infty} \|y_i - \tilde{y}_i\|_p + \sum_{i=N+1}^{\infty} \|y_i\|_p \right)^{1/p} < \left( \sum_{i=1}^{\infty} \frac{\|y_i\|_p}{\epsilon} + \frac{\epsilon^p}{2} \right)^{1/p} = \epsilon \).

Hence, \( \tilde{D} \) is dense in \( Z_2 \). Therefore, \( Z_2 \) is separable.

(iii) This result is standard, which follows directly from Corollary 7.57 of Stone-Weierstrass Theorem, and therefore the proof is omitted.

(iv) First, consider the special case where \( X \) is a finite compact metric measure space. By (iii), let \( D \subseteq C(X, Y_1) \) be a countable dense subset.

By Proposition 4.11, \( X' \) is a normal topological space. Then, by Proposition 11.182 and (i), \( \forall f \in \hat{L}_p(X, Y_1), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}, \exists y \in \hat{L}_p(X, Y_1) \) such that \( g \) is continuous and \( \|g - f\|_p < \epsilon/2 \). Clearly, \( g \in C(X, Y_1) \). Then, \( \exists h \in D \) such that \( \|h - g\|_{C(X, Y_1)} < \frac{\epsilon}{2\mu(X, Y_1)} \). Since \( X \) is compact and \( h \in C(X, Y_1) \), then \( \exists M \subseteq [0, \infty) \subseteq \mathbb{R} \) such that \( \mathcal{P} \circ h \leq M < \infty \), \( \forall x \in X \). By Proposition 11.37, \( h \) is \( B \)-measurable. Hence, \( h \in \hat{L}_p(X, Y_1) \) is dense.

Then, \( \|f - h\|_p \leq \|f - g\|_p + \|g - h\|_p < \epsilon/2 + \left( \int_X \mathcal{P} \circ (g - h) \, d\mu \right)^{1/p} \leq \epsilon/2 + \left( \|g - h\|_{C(X, Y_1)} \mu(X) \right)^{1/p} < \epsilon \). Hence, \( D \subseteq \hat{L}_p(X, Y_1) \) is dense. Therefore, \( L_p(X, Y_1) \) is separable. By Example 11.179, \( L_p(X, Y_1) \) is a Banach space. Hence, \( L_p(X, Y_1) \) is a separable Banach space.

Now, consider the general case where \( X \) is a locally finite \( \sigma \)-compact metric measure space. Then, \( \exists (X_n)_{n=1}^\infty \subseteq X \) such that \( X = \bigcup_{n=1}^\infty X_n \) and \( X_n \) is compact, \( \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( X_n \subseteq X_{n+1} \), \( \forall n \in \mathbb{N} \). Since \( X \) is locally finite, then \( X_n \in \mathcal{B}_B(X) \) and \( \mu(X_n) < \infty, \forall n \in \mathbb{N} \). Fix any \( n \in \mathbb{N} \). By Proposition 4.37, let \( X_n \) be the metric subspace of \( X \). Let \( X_n := (X_n, B_n, \mu_n) \) be the finite measure space of \( X \). By Proposition 11.29, \( B_n = \mathcal{B}_B(X_n) \) and \( X_n \) is a finite compact metric measure space. By the special case, \( L_p(X_n, Y_1) \) is a separable Banach space. Let \( D_n \subseteq \hat{L}_p(X_n, Y_1) \) be a countable dense subset.

\[ \forall f \in \hat{L}_p(X_n, Y_1), \text{ define } \tilde{f} \in \hat{L}_p(X, Y_1) \text{ by } \tilde{f}(x) = \begin{cases} f(x) & \forall x \in X_n, \\ \vartheta_{y_1} & \forall x \in X \setminus X_n \end{cases}, \forall x \in X \text{. Let } D := \{ [\tilde{f}] \in L_p(X, Y_1) \mid \exists n \in \mathbb{N}, [f] \in D_n \}. \text{ Clearly, } D \subseteq L_p(X, Y_1) \text{ is countable.} \]

We will show that \( D \) is dense in \( L_p(X, Y_1) \). Fix any \( f \in \hat{L}_p(X, Y_1) \). We have \( \int_X \mathcal{P} \circ f \, d\mu < \infty \). By Monotone Convergence Theorem 11.81, \( \int_X \mathcal{P} \circ f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (f \chi_{X_n,x}) \, d\mu \). Let \( f_n := f \chi_{X_n,x} \in \hat{L}_p(X, Y_1), \forall n \in \mathbb{N} \). Then,
11.10. DUAL OF $\mathcal{C}(X, Y)$ AND $\mathcal{C}_c(X, Y)$

where the fourth equality follows from Proposition 11.83. Hence, $\lim_{n \to \infty} f_n = f$ in $L_p(X, \mathcal{Y}, \mu)$. \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}$ such that $\|f_n - f\|_p < \epsilon/2$.

Let $\hat{f}_{n_0} := f_{n_0}|_{X_{n_0}}$. Then, $\hat{f}_{n_0} \in L_p(X_{n_0}, \mathcal{Y}_1)$. This implies that $\exists \hat{g} \in D_{n_0}$ with $g \in L_p(X_{n_0}, \mathcal{Y}_1)$ such that $\|\hat{f}_{n_0} - g\|_p < \epsilon/2$. Let $\hat{g} \in L_p(X, \mathcal{Y}_1)$ be defined as in the second paragraph from last. Then, $[\hat{g}] \in D$ and

$$
\|f - \hat{g}\|_p \leq \|f - f_{n_0}\|_p + \|f_{n_0} - \hat{g}\|_p < \epsilon/2 + \bigg(\int_X \mathcal{P}_p \circ (f_{n_0} - \hat{g}) \, d\mu\bigg)^{1/p}
$$

$$
= \epsilon/2 + \bigg(\int_{X_{n_0}} \mathcal{P}_n \circ (f_{n_0} - \hat{g}) \chi_{X_{n_0}} \, d\mu\bigg)^{1/p}
$$

$$
= \epsilon/2 + \bigg(\int_{X_{n_0}} \mathcal{P}_n \circ (\hat{f}_{n_0} - \hat{g}) \, d\mu_{n_0}\bigg)^{1/p}
$$

$$
= \epsilon/2 + \|\hat{f}_{n_0} - \hat{g}\|_p < \epsilon
$$

where the second equality follows from Proposition 11.83. Hence, $D$ is dense in $L_p(X, \mathcal{Y}_1)$. Then, by Example 11.179 and Proposition 11.187, $L_p(X, \mathcal{Y}_1)$ is a separable Banach space.

(v) Since $X_2$ is $\sigma$-compact, then $\exists$ compact sets $(K_n)_{n=1}^\infty$ such that $X_2 = \bigcup_{n=1}^\infty K_n, \forall n \in \mathbb{N}, \forall i \in \mathbb{N}, K_n \subseteq \bigcup_{x \in K_n} \mathcal{B}_{X_2}(x, 1/i)$. By the compactness of $K_n$, $\exists$ finite subset $D_{n,i} \subseteq K_n$ such that $K_n \subseteq \bigcup_{x \in D_{n,i}} \mathcal{B}_{X_2}(x, 1/i)$. Then, $D := \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty D_{n,i} \subseteq X_2$ is a countable dense set. Hence, $X_2$ is separable.

This completes the proof of the proposition. \(\square\)

**Definition 11.191** Let $X := (X, O)$ be a topological space, $Y$ be a normed linear space, and $(X, B, \mu)$ be a $Y$-valued measure space on the same set $X$. The triple $X := (X, B, \mu)$ is said to be a $Y$-valued topological measure space if $B = \mathcal{B}_B(X)$ and $\mathcal{X} := (X, B_B(X), \mathcal{P} \circ \mu)$ is a topological measure space. We will say that $X$ is finite, $\sigma$-finite, or locally finite if $\mathcal{X}$ is so. We will say that $X$ is Tychonoff, Hausdorff, regular, completely regular, normal, first countable, second countable, separable, second category everywhere, connected, locally connected, compact, countably compact, sequentially compact, locally compact, $\sigma$-compact, or paracompact if $X$ is so.

Let $X := (X, \rho)$ be a metric space with the natural topology $O$, $Y$ be a normed linear space, and $(X, \mathcal{B}_B(X), \mu)$ be a $Y$-valued measure space on the same set $X$. The triple $X := (X, \mathcal{B}_B(X), \mu)$ is said to be a $Y$-valued metric measure space if $((X, O), \mathcal{B}_B(X), \mu)$ is a $Y$-valued topological measure space. $X$ is said to be complete or totally bounded if $X$ is so.

Let $X := (X, K, \|\cdot\|)$ be a normed linear space over the field $K$, $O$ be the natural topology on $X$ generated by the norm $\|\cdot\|$, $Y$ be a normed linear space, and $(X, \mathcal{B}_B(X), \mu)$ be a $Y$-valued measure space on the same set $X$. The triple $X := (X, \mathcal{B}_B(X), \mu)$ is said to be a $Y$-valued normed linear measure
space if \( (X, \mathcal{O}, \mathcal{B}_B(X), \mu) \) is a \( \gamma \)-valued topological measure space. When \( X \) is a Banach space, then \( X \) is said to be a \( \gamma \)-valued Banach measure space.

Depending on whether \( K = \mathbb{R} \) or \( K = \mathbb{C} \), we will say that \( X \) is a \( \gamma \)-valued real or complex Banach measure space.

**Proposition 11.192** Let \( X := (X, \mathcal{O}) \) be a topological space, \( Y \) be a normed linear space over \( K \). Define \( \tilde{Z} := \{ \mu \in \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \mid (X, \mathcal{B}_B(X), \mu) \) is a finite \( \gamma \)-valued topological measure space \}. Then, \( \tilde{Z} \) is a closed subspace of \( \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \) and \( (\tilde{Z}, K, || \cdot ||_{\mathcal{M}_f(X, \mathcal{B}_B(X), Y)}) =: \mathcal{M}_{f1}(X, Y) \) is a normed linear space.

If, in addition, \( Y \) is a Banach space. Then, \( \mathcal{M}_{f1}(X, Y) \) is a Banach space. Furthermore, define \( \tilde{Z} := \{ \mu \in \mathcal{M}_\sigma(X, \mathcal{B}_B(X), Y) \mid (X, \mathcal{B}_B(X), \mu) \) is a \( \sigma \)-finite \( \gamma \)-valued topological measure space \( \} \subseteq \mathcal{M}_\sigma(X, \mathcal{B}_B(X), Y) \). Let \( \tilde{Z} \) admit the subset topology \( \mathcal{O}_{\tilde{Z}} \). Then, \( \tilde{Z} := \mathcal{M}_{\sigma1}(X, Y) \) is a subspace of \( \mathcal{M}_\sigma(X, \mathcal{B}_B(X), Y) \). We will abuse the notation and denote the topological space \( (\tilde{Z}, \mathcal{O}_{\tilde{Z}}) \) by \( \mathcal{M}_{\sigma1}(X, Y) \).

**Proof** Let \( Y \) be a normed linear space. We will show that \( \tilde{Z} \) is a subspace of \( \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \). Let \( \forall \alpha_1, \alpha_2 \in K, \forall \mu_1, \mu_2 \in Z \), by Proposition 11.36, \( \mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \in \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \). \( \forall E \in \mathcal{B}_B(X), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R} \), \( \forall i \in O \) with \( E \subseteq O_i \) such that \( \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|) \). Let \( O := O_1 \cap O_2 \in O \). Clearly, \( E \subseteq O \) and \( \mathcal{P} \circ \mu(O \setminus E) \leq (\sum_{i=1}^{\infty} |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) = \sum_{i=1}^{\infty} |\alpha_i| \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon, \) where the first inequality and the first equality follow from Proposition 11.136. This shows that \( (X, \mathcal{B}_B(X), \mathcal{P} \circ \mu) \) is a topological measure space. Then, \( (X, \mathcal{B}_B(X), \mu) \) is a finite \( \gamma \)-valued topological measure space and \( \mu \in \tilde{Z} \). Clearly, \( \forall \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \), \( \tilde{Z} \neq \emptyset \). Hence, \( \tilde{Z} \) is a subspace of \( \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \). Then, \( \mathcal{M}_{f1}(X, Y) \) is a normed linear space since \( \tilde{Z} = \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \) is a normed linear space.

Next, we will show that \( \tilde{Z} \) is closed. \( \forall \mu \in \tilde{Z} \subseteq \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \), by Proposition 4.13, \( \exists (\mu_n)_{n=1}^{\infty} \subseteq \tilde{Z} \) such that \( \lim_{n \in N} \mu_n = \mu \). Then, \( \lim_{n \in N} \| \mu_n - \mu \|_{\mathcal{M}_f(X, \mathcal{B}_B(X), Y)} = \lim_{n \in N} \mathcal{P} \circ (\mu_n - \mu)(X) = 0 \). \( \forall E \in \mathcal{B}_B(X), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R} \), \( \exists n \in N \) such that \( \mathcal{P} \circ (\mu_n - \mu)(X) < \epsilon/2 \). By \( \mu_n \in \tilde{Z} \), \( \exists O \in O \) with \( E \subseteq O \) such that \( \mathcal{P} \circ \mu_n(O \setminus E) < \epsilon/2 \). Then, \( \mathcal{P} \circ \mu(O \setminus E) = \mathcal{P} \circ (\mu_n - \mu_n + \mu)(O \setminus E) \leq \mathcal{P} \circ (\mu_n + \mathcal{P} \circ (\mu_n - \mu))(O \setminus E) = \mathcal{P} \circ \mu(O \setminus E) + \mathcal{P} \circ (\mu_n - \mu)(O \setminus E) < \epsilon/2 + \mathcal{P} \circ (\mu_n - \mu)(X) < \epsilon, \) where the first equality and the first inequality follow from Proposition 11.136. Hence, \( \mu \in \tilde{Z} \). By the arbitrariness of \( \mu \), we have \( \overline{\tilde{Z}} = \tilde{Z} \) and \( \tilde{Z} \) is closed.

Let \( Y \) be a Banach space. By Proposition 11.142, \( \mathcal{M}_f(X, \mathcal{B}_B(X), Y) \) is a Banach space. By Proposition 4.39, \( \mathcal{M}_{f1}(X, Y) \) is a Banach space.

Finally, we will show that \( \tilde{Z} \) is a subspace of \( \mathcal{M}_\sigma(X, \mathcal{B}_B(X), Y) \). \( \forall \alpha_1, \alpha_2 \in K, \forall \mu_1, \mu_2 \in \tilde{Z} \), by Proposition 11.138, \( \mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \subseteq \mathcal{M}_\sigma(X, \mathcal{B}_B(X), Y) \). \( \forall E \in \mathcal{B}_B(X), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R} \), \( \forall i = 1, 2 \), by \( \mu_i \in \tilde{Z} \), \( \exists O_i \in O \) with \( E \subseteq O_i \) such that \( \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|) \). Let \( O := O_1 \cap O_2 \in O \). Clearly, \( E \subseteq O \) and \( \mathcal{P} \circ \mu(O \setminus E) \leq (\sum_{i=1}^{\infty} |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) \leq (\sum_{i=1}^{\infty} |\alpha_i| \mathcal{P} \circ \mu_i(O_i \setminus E) \leq \sum_{i=1}^{\infty} |\alpha_i| \mathcal{P} \circ \mu_i(O_i \setminus E) < \epsilon, \) where the first equality and the first inequality follow from Proposition 11.136.
Proposition 11.193

Let \( X : (X, O) \) be a topological space and \( \mu_o : O \to [0, \infty) \subset \mathbb{R} \). Assume that

1. \( \mu_o \) is an outer measure. It induces a finite complete measure space \((X, \mathcal{B}, \tilde{\mu})\), where \( \mathcal{B} := \{ E \subseteq X \mid E \) is measurable with respect to \( \tilde{\mu}_o \}\)

   and \( \tilde{\mu} := \tilde{\mu}_o |_{\mathcal{B}} \).

2. \( \mathcal{B}_X \subseteq \mathcal{B} \) and the triple \( \tilde{X} := (X, \mathcal{B}_X, \tilde{\mu}) \) is a finite complete measure space with \( \mu(O) = \mu_o(O) \), \( \forall O \subseteq X \).

3. The measure \( \mu \) is unique in the sense that if \( \tilde{\mu} \) be another measure on \((X, \mathcal{B}_X)\) satisfying \( \tilde{\mu}(O) = \mu_o(O) = \mu(O) \), \( \forall O \subseteq \mathcal{O} \), then \( \tilde{\mu} = \mu \).

Proof

1. By (i), we have \( \tilde{\mu}_o(\emptyset) = 0 \). \forall A \subseteq B \subseteq X, we have \( 0 \leq \tilde{\mu}_o(A) = \inf_{O \subseteq A, A \subseteq O} \mu_o(O) \leq \inf_{O \subseteq B, B \subseteq O} \mu_o(O) = \tilde{\mu}_o(B) \leq \mu_o(X) < \infty \). \forall E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq X, \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \forall i \in \mathbb{N}, \exists O_i \subseteq O \in \mathcal{O} \) with \( E_i \subseteq O_i \) such that \( \mu_o(O_i) < \tilde{\mu}_o(E_i) + 2^{-i} \epsilon \). Then, \( \tilde{\mu}_o(E) = \inf_{O \subseteq E, E \subseteq O} \mu_o(O) \leq \mu_o(\bigcup_{i=1}^{\infty} O_i) \leq \sum_{i=1}^{\infty} \mu_o(O_i) \leq \sum_{i=1}^{\infty} \tilde{\mu}_o(E_i) + \epsilon \), where the second inequality follows from (iii). Hence, \( \tilde{\mu}_o : X \to [0, \infty) \subset \mathbb{R} \) is an outer measure. It is easy to see that \( \tilde{\mu}_o(O) = \mu_o(O), \forall O \subseteq \mathcal{O} \). By Theorem 11.17, \((X, \bar{\mathcal{B}}, \tilde{\mu})\) is a finite complete measure space.
2. \( \forall O \in \mathcal{O}, \forall E \subseteq X, \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}, \exists O_1 \in \mathcal{O} \) with \( E \subseteq O_1 \) such that \( \mu_o(O_1) < \bar{\mu}_o(E) + \epsilon/2 \). By (v), \( \exists U \in \mathcal{O} \) with \( U \subseteq U \subseteq O_1 \cap O \) such that \( \mu_o(O \cap O_1) < \mu_o(U) + \epsilon/2 \). Then, \( \bar{\mu}_o(U) > \mu_o(U) - \epsilon/2 \geq \mu_o((O_1 \setminus U) - \epsilon/2 = \mu_o(O_1 \setminus U) + \mu_o(U) - \epsilon/2 = \bar{\mu}_o(O_1 \setminus U) + \mu_o(U) - \epsilon/2 > \bar{\mu}_o(O_1 \setminus (O_1 \cap O)) + \mu_o(O \cap O_1) - \epsilon = \bar{\mu}_o(O_1 \setminus O) + \bar{\mu}_o(O \cap O_1) - \epsilon \geq \bar{\mu}_o(E \setminus O) + \bar{\mu}_o(E \cap O) - \epsilon \geq \bar{\mu}_o(E) - \epsilon \), where the second inequality follows from (ii); the first equality follows from (iv); and the third, fourth, and the last inequalities follow from the fact that \( \bar{\mu}_o \) is an outer measure. By the arbitrariness of \( \epsilon \), we have \( \bar{\mu}_o(E \setminus O) = \bar{\mu}_o(E \cap O) \). By the arbitrariness of \( O \), we have \( \mathcal{O} \subseteq \mathcal{B} \). Since \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \), then \( \mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{B} \). By Proposition 11.13, \( (X, \mathcal{B}_B(\mathcal{X}), \mu) \) is a measure space. Clearly, \( \mu(O) = \bar{\mu}(O) = \bar{\mu}_o(O), \forall O \in \mathcal{O} \). This coupled with the definition of \( \bar{\mu}_o \) leads to the conclusion that \( \mathcal{X} \) is a topological measure space. Clearly, \( \mu(X) = \mu_o(X) < \infty \). Hence, \( \mathcal{X} \) is a finite topological measure space.

3. Let \( \bar{\mu} \) be another measure on \( (X, \mathcal{B}_B(\mathcal{X})) \) satisfying \( \bar{\mu}(O) = \mu_o(O) = \mu(O), \forall O \in \mathcal{O} \). Then, \( \bar{\mu} \) is finite. Suppose \( \mu \neq \bar{\mu} \). Then, \( \exists E \in \mathcal{B}_B(\mathcal{X}) \) such that \( \bar{\mu}(E) \neq \mu(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu(O) \). Since \( \bar{\mu}(E) \leq \bar{\mu}(O), \forall O \in \mathcal{O} \) with \( E \subseteq O \). Then, we must have \( \bar{\mu}(E) < \mu(E) \). Since \( \mathcal{X} \) is a topological measure space, then, by Proposition 11.27, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \mu(F \setminus E) < \mu(E \setminus F)/2 \). This implies that \( \mu(F) = \mu(X) - \mu(F \setminus E) > \mu(E)/2 + \bar{\mu}(E)/2 > \bar{\mu}(E) \). This completes the proof of the proposition.

Proposition 11.194 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a compact Hausdorff topological space, \( \mathcal{Y} \) be a Banach space, \( \bar{\mu} \) be a function that assigns a vector \( \bar{\mu}(F) \in \mathcal{Y} \) for each closed subset \( F \subseteq X \), \( \mathcal{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \nu) \) be a finite topological measure space. Assume that

\begin{enumerate}
\item \( \bar{\mu}(F_1 \cup F_2) = \mu(F_1) + \mu(F_2), \forall X \setminus F_1, X \setminus F_2 \in \mathcal{O} \) with \( F_1 \cap F_2 = \emptyset \);
\item \( \| \bar{\mu}(F_1) - \mu(F_2) \| \leq \nu(O), \forall X \setminus F_1, X \setminus F_2, O \in \mathcal{O} \) with \( F_1 \triangle F_2 \subseteq O \).
\end{enumerate}

Then, there exists a unique \( \mu \in \mathcal{M}_f(X, \mathcal{Y}) \) such that \( \mu(F) = \bar{\mu}(F), \forall X \setminus F \in \mathcal{O} \). Furthermore, \( \mathcal{P} \circ \mu \leq \nu \).

Proof Fix any \( E \in \mathcal{B}_B(\mathcal{X}) \). Let \( \mathcal{A}_E := \{ F \subseteq E \mid X \setminus F \in \mathcal{O} \} \) and \( \mathcal{A}_E := (\mathcal{A}_E, \subseteq) \). Clearly, \( \mathcal{A}_E \) is a directed system.

Claim 11.194.1 The net \( (\bar{\mu}(F))_{F \in \mathcal{A}_E} \) is Cauchy.
Proof of claim: \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R} \), by \( \mathcal{X} \) being a topological measure space, \( \forall V \in \mathcal{O} \) with \( E \subseteq V \) such that \( \nu(V \setminus E) < \epsilon/2 \). By Proposition 11.27, \( \exists \mathcal{X} \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \nu(E \setminus F) < \epsilon/2 \). \( \forall E_1 \in \mathcal{A}_E \) with \( E \subseteq V \), \( \exists F_1 \in \mathcal{A}_E \) such that \( \nu(F \setminus F_1) \leq \nu(V \setminus F) = \nu(V \setminus E) + \nu(E \setminus F) < \epsilon \). Hence, the net is Cauchy. This completes the proof of the claim.

By Proposition 4.44, we may define \( \mu(E) = \lim_{E \in \mathcal{A}_E} \bar{\mu}(F) \in \mathcal{Y} \). Thus, we have defined a function \( \mu : \mathcal{B}_E(\mathcal{X}) \to \mathcal{Y} \). We will show that \( \mu \) is the \( \mathcal{Y} \)-valued measure we seek. Clearly, \( \mu(F) = \bar{\mu}(E) \), \( \forall X \setminus F \in \mathcal{O} \). By (i), \( \bar{\mu}(\emptyset) + \bar{\mu}(\emptyset) = \bar{\mu}(\emptyset) \) and then \( \mu(\emptyset) = \bar{\mu}(\emptyset) = \emptyset \).

\( \forall E \in \mathcal{B}_E(\mathcal{X}), \forall \epsilon \in (0, +\infty) \subset \mathbb{R} \), \( \exists V \in \mathcal{O} \) with \( E \subseteq V \) such that \( \nu(V \setminus E) < \epsilon/2 \). By \( \mu(E) = \lim_{E \in \mathcal{A}_E} \bar{\mu}(F) \), \( \exists F \in \mathcal{A}_E \) such that \( \| \mu(E) - \bar{\mu}(F) \| < \epsilon/2 \). Clearly, \( F \subseteq E \subseteq V \). Then, \( \| \mu(E) \| < \| \bar{\mu}(F) \| + \epsilon/2 < \nu(V \setminus E) + \epsilon/2 < \nu(V \setminus E) + \epsilon \), where the second inequality follows from (ii). By the arbitrariness of \( \epsilon \), we have \( \| \mu(E) \| \leq \nu(V \setminus E) \).

Fix any pairwise disjoint \( (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_E(\mathcal{X}) \), let \( E := \bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_E(\mathcal{X}) \). \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \exists V \in \mathcal{O} \) with \( E_n \subseteq V \) such that \( \nu(V \setminus E_n) < 2^{-n-1}\epsilon/5 \). By Proposition 11.27, \( \exists \mathcal{X} \setminus F \in \mathcal{O} \) with \( F \subseteq E_n \) such that \( \nu(E_n \setminus F) < 2^{-n-1}\epsilon/5 \). By \( \mu(E_n) = \lim_{F \in \mathcal{A}_E} \bar{\mu}(F) \), \( \exists F \in \mathcal{A}_E \) such that \( \nu(E_n \setminus F) < 2^{-n-1}\epsilon/5 \) and \( \| \mu(E_n) - \bar{\mu}(F) \| < 2^{-n-1}\epsilon/5 \). Clearly, \( F_n \subseteq E_n \subseteq V_n \), \( \exists \mathcal{X} \setminus F \subseteq \mathcal{X} \in \mathcal{O} \) with \( E \subseteq V \) such that \( \nu(V \setminus E) < \epsilon/5 \). By Proposition 11.27, \( \exists \mathcal{X} \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \nu(E \setminus F) < \epsilon/5 \). By \( \mu(E) = \lim_{E \in \mathcal{A}_E} \bar{\mu}(F) \), \( \exists \mathcal{X} \setminus F \in \mathcal{A}_E \) with \( F \subseteq E \) such that \( \| \mu(E) - \bar{\mu}(F) \| < \epsilon/5 \). Clearly, \( F \subseteq \mathcal{X} \setminus F \subseteq E \subseteq V \). By Proposition 5.5, \( \mathcal{X} \setminus F \) is compact. Since \( \mathcal{X} \setminus F \subseteq E \subseteq V \), \( \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( \| \mu(E) - \sum_{i=1}^{n} \mu(E_i) \| < \nu(V \setminus E) + \nu(E \setminus F) + \epsilon/5 \), where the second inequality follows from (i). Clearly, \( \mathcal{X} \setminus F \) and \( (\bigcup_{i=1}^{n} F_i) \) are closed sets. \( \mathcal{X} \setminus F \subseteq \mathcal{X} \setminus (\bigcup_{i=1}^{n} F_i) = \mathcal{X} \setminus (\bigcup_{i=1}^{n} F_i) \), \( \forall \epsilon > 0 \), \( \| \mu(E) - \sum_{i=1}^{n} \mu(E_i) \| < 2\epsilon/5 + \nu((\bigcup_{i=1}^{n} F_i) \setminus V) + \nu(E \setminus F) < \epsilon \). Then, \( \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \in \mathcal{Y} \).

\( \forall \text{pairwise disjoint } (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_E(\mathcal{X}), \| \mu(E_n) \| < \nu(E_n) \subseteq \nu(E) \). This shows that \( \mu \) is a \( \mathcal{Y} \)-valued pre-measure on \( (E, \mathcal{B}_E(\mathcal{X})) \).

\( \forall E \in \mathcal{B}_E(\mathcal{X}), \forall n \in \mathbb{Z}_+ \), \( \forall \text{pairwise disjoint } (E_i)_{i=1}^{n} \subseteq \mathcal{B}_E(\mathcal{X}) \) with \( E = \bigcup_{i=1}^{n} E_i \), \( \sum_{i=1}^{n} \| \mu(E_i) \| < \nu(E) \). Hence, \( \mu(E) \leq \nu(E) \). By the arbitrariness of \( E \), we have \( \mathcal{P} \circ \mu \leq \nu \). Hence, \( \mu \) is a finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}_E(\mathcal{X})) \). It is easy to show that \( (X, \mathcal{B}_E(\mathcal{X}), \mu) \) is a \( \mathcal{Y} \)-valued topological measure space. Then, \( \mu \in \mathcal{M}_f(\mathcal{X}, \mathcal{Y}) \).
Finally, we need to show that $\mu \in M_{fr}(\mathcal{X}, \mathcal{Y})$ is unique. Let $\tilde{\mu} \in M_{fr}(\mathcal{X}, \mathcal{Y})$ be such that: $\tilde{\mu}(F) = \tilde{\mu}(F) = \mu(F), \forall F \in \mathcal{O}$. By Proposition 11.27, $\exists X \setminus F \in \mathcal{O}$ with $F \subseteq E$ such that $\mathcal{P} \circ (E \setminus F) < \epsilon/2$. Again by Proposition 11.27, $\exists F_1 \in \mathcal{O}$ with $F_1 \subseteq E$ such that $\mathcal{P} \circ (E \setminus F_1) < \epsilon/2$. Let $\tilde{F} := F \cup F_1 \subseteq E$, which is clearly closed. Then, $\| \mu(E) - \tilde{\mu}(E) \| \leq \| \mu(E) - \mu(F) \| + \| \mu(F) - \tilde{\mu}(F) \| + \| \tilde{\mu}(F) - \tilde{\mu}(E) \| = \| \mu(E \setminus \tilde{F}) \| + 0 + \| \tilde{\mu}(E \setminus \tilde{F}) \| \leq \mathcal{P} \circ \mu(E \setminus \tilde{F}) + \mathcal{P} \circ \tilde{\mu}(E \setminus \tilde{F}) < \epsilon$. By the arbitrariness of $\epsilon$, we have $\mu(E) = \tilde{\mu}(E)$. Hence, $\mu = \tilde{\mu}$.

This completes the proof of the proposition.

**Definition 11.195** Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space. $E \subseteq X$ is said to be a $G_\delta$ if $\exists (O_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $E = \bigcap_{i=1}^\infty O_i$. $\bar{E} \subseteq X$ is said to be an $F_\sigma$ if exists $(\bar{F}_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $\bar{E} = \bigcup_{i=1}^\infty F_i$.

**Proposition 11.196** Let $\mathcal{X} := (X, \rho)$ be a locally compact separable metric space with the natural topology $\mathcal{O}$, $\mathcal{E} := \{ E \subseteq X \mid E$ is a compact $G_\delta \}$, and $B_\sigma(\mathcal{X})$ be the $\sigma$-algebra generated by $\mathcal{E}$. Then, $B_\sigma(\mathcal{X}) = B_B(\mathcal{X})$.

**Proof** $\forall E \in \mathcal{E}$, by Proposition 5.5, $E$ is closed. Then, $E \in B_B(\mathcal{X})$. Hence, $\mathcal{E} \subseteq B_B(\mathcal{X})$. Since $B_B(\mathcal{X})$ is a $\sigma$-algebra, then $B_\sigma(\mathcal{X}) \subseteq B_B(\mathcal{X})$.

$\forall x \in X$, by Definition 4.49, $\exists O_x \in \mathcal{O}$ such that $x \in O_x$ and $O_x$ is compact. Then, $X = \bigcup_{x \in X} O_x$. By Propositions 4.44 and 3.24, $\exists$ a countable set $D \subseteq X$ such that $X = \bigcup_{x \in D} O_x = \bigcup_{x \in D} \bar{O}_x$.

$\forall \tilde{F} \in \mathcal{O}$, $\forall x \in D$, $F_x := \tilde{F} \cap \bar{O}_x$ is compact and closed by Propositions 5.5 and 3.5. Then, $F_x = \bigcap_{i=1}^\infty \bigcup_{\tilde{E} \in F_x} B_{\tilde{E}}(\tilde{x}, 1/n)$, by Proposition 4.10. This implies that $F_x$ is a $G_\delta$ in addition to being compact. Then, $F_x \in \mathcal{E} \subseteq B_\sigma(\mathcal{X})$. This implies that $F = \bigcup_{x \in D} F_x \in B_\sigma(\mathcal{X})$. Hence, $F \in B_\sigma(\mathcal{X})$. By the arbitrariness of $\tilde{F}$, we have $\mathcal{O} \subseteq B_\sigma(\mathcal{X})$. By $B_\sigma(\mathcal{X})$ being a $\sigma$-algebra, we have $B_B(\mathcal{X}) \subseteq B_\sigma(\mathcal{X})$.

Therefore, $B_\sigma(\mathcal{X}) = B_B(\mathcal{X})$. This completes the proof of the proposition.

**Lemma 11.197** Let $\mathcal{X} := (X, \mathcal{O})$ be a normal topological space. Then, $\forall \mathcal{E} \in \mathcal{O}$ with $E$ being a $G_\delta$, there exists a continuous function $\phi : X \to [0, 1] \subseteq \mathbb{R}$ such that $E = \{ x \in X \mid \phi(x) = 1 \}$.

**Proof** Since $E$ is a $G_\delta$, then $\exists (O_i)_{i=1}^\infty \subseteq \mathcal{O}$ such that $E = \bigcap_{i=1}^\infty O_i$. $\forall i \in \mathbb{N}$, by Urysohn’s Lemma 3.55, there exists a continuous function $\phi_i : X \to [0, 1] \subseteq \mathbb{R}$ such that $\phi_i|_E = 1$ and $\phi_i|_{\bar{O}_i} = 0$. Let $\phi := \sum_{i=1}^\infty 2^{-i}\phi_i$. By Proposition 4.26, $\phi : X \to [0, 1] \subseteq \mathbb{R}$ is continuous. Clearly $\phi|_E = 1$ and $\phi(x) < 1, \forall x \in \bigcup_{i=1}^\infty \bar{O}_i = \tilde{E}$. Hence, the result holds. This completes the proof of the lemma.
Theorem 11.198 Let $\mathcal{X} := (X, \rho)$ be a locally compact separable metric space with the natural topology $\mathcal{O}$ and $\mu$ be a finite measure on $(X, B_0(\mathcal{X}))$. Then, $\mathcal{X} := (\mathcal{X}, B_0(\mathcal{X}), \mu)$ is a finite measure space. Thus, $\mathcal{M}_f(X, B_0(\mathcal{X})) = \mathcal{M}_{ft}(\mathcal{X})$. As a consequence, $\mathcal{M}_f(X, B_0(\mathcal{X}), \gamma) = \mathcal{M}_{ft}(\mathcal{X}, \gamma)$, where $\gamma$ is any normed linear space.

Proof By Propositions 4.4, 3.24, and 5.72, $\mathcal{X}$ is $\sigma$-compact. Let $\mathcal{R} \subseteq B_0(\mathcal{X})$ be such that $E \in \mathcal{R}$ if $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, we have (i) $\exists \Omega \in \mathcal{O}$ with $O$ being $\sigma$-compact and $E \subseteq O$ such that $\mu(O \setminus E) < \epsilon$; (ii) $\exists \bar{F} \in \mathcal{O}$ with $F$ being a compact $G_\delta$ and $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$. We will show that $\mathcal{R} = B_0(\mathcal{X})$.

Claim 11.198.1 $\forall \bar{F} \in \mathcal{O}$, $F = \bigcap_{n=1}^{\infty} \bigcup_{x \in F} B_X(x, 1/n) =: \bar{F}$ is a $G_\delta$.

Proof of claim: Clearly, $F \subseteq \bar{F}$. $\forall x_0 \in \bar{F}$, by Proposition 4.10, $e_0 := \text{dist}(x_0, F) > 0$. Then, $\forall n \in \mathbb{N}$ with $1/n < e_0$, $x_0 \notin \bigcup_{x \in F} B_X(x, 1/n)$ and $x_0 \notin \bar{F}$. Hence, $\bar{F} \subseteq F$ and $F = \bar{F}$. Clearly, $\bar{F}$ is a $G_\delta$. This completes the proof of the claim.

Claim 11.198.2 $\forall O \in \mathcal{O}$, $O$ is $\sigma$-compact.

Proof of claim: By Claim 11.198.1, $\bar{O}$ is a $G_\delta$. Then, $\exists (O_i)_{i=1}^{\infty} \subseteq \mathcal{O}$ such that $\bar{O} = \bigcap_{i=1}^{\infty} O_i$ and $O = \bigcup_{i=1}^{\infty} \bar{O}_i$. By $\mathcal{X}$ being $\sigma$-compact, $\exists$ compact sets $(K_i)_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} K_i$. Then, $O = O \cap X = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} (\bar{O}_i \cap K_j)$. $\forall i, j \in \mathbb{N}$, by Propositions 3.5 and 5.5, $\bar{O}_i \cap K_j$ is compact. Hence, $O$ is $\sigma$-compact.

$\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{R}$, let $E := \bigcup_{i=1}^{\infty} E_i \in B_0(\mathcal{X})$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists O_i \in \mathcal{O}$ with $O_i$ being $\sigma$-compact and $E_i \subseteq O_i$ such that $\mu(O_i \setminus E_i) < 2^{-i-1}\epsilon$. $\exists \bar{F}_i \in \mathcal{O}$ with $F_i$ being a compact $G_\delta$ and $F_i \subseteq E_i$ such that $\mu(E_i \setminus F_i) < 2^{-i-1}\epsilon$. Let $O := \bigcup_{i=1}^{\infty} O_i$, which is $\sigma$-compact by Claim 11.198.2. Then, $E \subseteq O$ and $\mu(O \setminus E) = \mu((\bigcup_{i=1}^{\infty} O_i) \setminus (\bigcup_{i=1}^{\infty} E_i)) \leq \mu((\bigcup_{i=1}^{\infty} O_i \setminus E_i)) \leq \sum_{i=1}^{\infty} \mu(O_i \setminus E_i) < \epsilon$. Hence, $E$ satisfies (i). By Proposition 11.7, $\mu(E) = \lim_{n \in \mathbb{N}} \mu(\bigcup_{i=1}^{n} E_i)$. Then, $\exists n_0 \in \mathbb{N}$ such that $\mu(E) < \mu(\bigcup_{i=1}^{n_0} E_i) + \epsilon/2$. Let $F := \bigcup_{i=1}^{n_0} E_i$. Clearly, $F$ is compact and closed. By Claim 11.198.1, $F$ is a $G_\delta$. Then, $F \subseteq E$ and $\mu(E \setminus F) = \mu(E) - \mu(F) < \mu(\bigcup_{i=1}^{n_0} E_i) - \mu(\bigcup_{i=1}^{n_0} F_i) + \epsilon/2 = \mu((\bigcup_{i=1}^{n_0} E_i) \setminus (\bigcup_{i=1}^{n_0} F_i)) + \epsilon/2 \leq \mu((\bigcup_{i=1}^{n_0} (E_i \setminus F_i)) + \epsilon/2 \leq \sum_{i=1}^{n_0} \mu(E_i \setminus F_i) + \epsilon/2 < \epsilon$. Hence, $E$ satisfies (ii). Then, $E \in \mathcal{R}$. Thus, $\mathcal{R}$ is closed under countable unions.

$\forall E \in \mathcal{R}$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists O \in \mathcal{O}$ with $O$ being $\sigma$-compact and $E \subseteq O$ such that $\mu(O \setminus E) < \epsilon$; $\exists \bar{F} \in \mathcal{O}$ with $F$ being a compact $G_\delta$ and $F \subseteq E$ such that $\mu(E \setminus F) < \epsilon$. Then, $\bar{O}$ is a closed set, $\bar{E}$, and $\mu(\bar{E} \setminus E) = \mu(O \setminus E) < \epsilon$. By $\mathcal{X}$ being $\sigma$-compact, $\exists$ compact sets $(K_i)_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} K_i$. Without loss of generality, we may assume that $K_i \subseteq K_{i+1}$, $\forall i \in \mathbb{N}$. Note that $\epsilon > \mu(\bar{E} \setminus \bar{O}) = \mu(\bar{E} \setminus (\bar{O} \setminus \bigcup_{i=1}^{\infty} K_i)) = \mu(\bar{E} \setminus (\bigcup_{i=1}^{\infty} (\bar{O} \setminus K_i)) = \mu(\bigcap_{i=1}^{\infty} (E \setminus (\bar{O} \setminus K_i))) = \lim_{i \in \mathbb{N}} \mu(\bar{E} \setminus (\bar{O} \setminus K_i)).
where the last equality follows from Proposition 11.5. Then, \( \exists n \in \mathbb{N} \) such that \( \mu(\overline{E} \setminus (\overline{O} \cap K_n)) < \epsilon \). By Propositions 3.5 and 5.5, \( \overline{O} \cap K_n \) is closed and compact, then it is a \( G_\delta \) by Claim 11.198.1. Clearly, \( \overline{O} \cap K_n \subseteq \overline{E} \). Hence, \( E \) satisfies (ii). \( \overline{F} \in \mathcal{O} \) and is \( \sigma \)-compact by Claim 11.198.2. Note that \( \overline{E} \subseteq \overline{F} \) and \( \mu(\overline{F} \setminus \overline{E}) = \mu(\overline{E} \setminus \overline{F}) < \epsilon \). Then, \( E \) satisfies (i). Hence, \( \overline{E} \in \mathcal{R} \). \( \mathcal{R} \) is closed under set complements. Clearly, \( \emptyset \in \mathcal{R} \). This proves that \( \mathcal{R} \) is a \( \sigma \)-algebra.

\[ \forall E \in \mathcal{O} \] with \( E \) being a compact \( G_\delta \), \( \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R} \), then \( E \) satisfies (ii) trivially. By Lemma 11.197, there exists a continuous function \( \phi : X \to [0, 1] \subseteq \mathbb{R} \) such that \( E = \{ x \in X \mid \phi(x) = 1 \} \). \( \forall i \in \mathbb{N} \), let \( O_i := \{ x \in X \mid \phi(x) > 1 - 1/i \} \) \( \in \mathcal{O} \). Clearly \( E = \bigcap_{i=1}^{\infty} O_i \) and \( O_{i+1} \subseteq O_i \), \( \forall i \in \mathbb{N} \). By Proposition 11.5, \( \mu(E) = \lim_{i \to \infty} \mu(O_i) \). \( \exists n \in \mathbb{N} \) such that \( \mu(O_n) - \epsilon \leq \mu(E) \leq \mu(O_n) \). Clearly, we have \( E \subseteq O_n \). Then, \( \mu(\overline{O_n} \setminus E) = \mu(\overline{O_n}) - \mu(E) < \epsilon \). By Claim 11.198.2, \( O_n \) is \( \sigma \)-compact. Then, \( E \) satisfies (i) and \( E \in \mathcal{R} \).

Thus, \( \mathcal{R} \) is a \( \sigma \)-algebra and \( E := \{ E \in \mathcal{B}_R(\mathcal{X}) \mid E \text{ is a compact } G_\delta \} \subseteq \mathcal{R} \). By Proposition 11.196, \( \mathcal{R} = \mathcal{B}_R(\mathcal{X}) \). Then, \( \mathcal{X} \) is a finite metric measure space.

Thus, we have \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X})) \subseteq \mathcal{M}_{f^*}(\mathcal{X}) \). Clearly, \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X})) \supseteq \mathcal{M}_{f^*}(\mathcal{X}) \). Hence, \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X})) = \mathcal{M}_{f^*}(\mathcal{X}) \).

\[ \forall \tilde{\mu} \in \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X}), \mathcal{Y}) \text{ then } \mathcal{P} \circ \tilde{\mu} \in \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X})) = \mathcal{M}_{f^*}(\mathcal{X}) \text{.} \] This implies that \( \tilde{\mu} \in \mathcal{M}_{f^*}(\mathcal{X}, \mathcal{Y}) \). Hence, \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X}), \mathcal{Y}) \subseteq \mathcal{M}_{f^*}(\mathcal{X}, \mathcal{Y}) \). Clearly, \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X}), \mathcal{Y}) \supseteq \mathcal{M}_{f^*}(\mathcal{X}, \mathcal{Y}) \). Then, \( \mathcal{M}_f(X, \mathcal{B}_R(\mathcal{X}), \mathcal{Y}) = \mathcal{M}_{f^*}(\mathcal{X}, \mathcal{Y}) \).

This completes the proof of the theorem. \( \square \)

**Definition 11.199** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( \mathcal{Y} \) be a normed linear space, \( f : X \to \mathcal{Y} \). The support of \( f \) is the set \( \text{supp}(f) := \{ x \in X \mid f(x) \neq 0 \} \).

**Lemma 11.200** Let \( \mathcal{X} := (X, \mathcal{O}) \) be compact Hausdorff topological space, \( \mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathbb{R}) \), and \( f \in \mathcal{Z}^* \). \( f \in \mathcal{Z}^* \) is said to be a positive linear functional if \( f(z) \geq 0, \forall z \in P := \{ h \in \mathcal{Z} \mid h : X \to [0, \infty) \subseteq \mathbb{R} \} \). Then, \( f = f_+ - f_- \), where \( f_+, f_- \in \mathcal{Z}_* \) are positive linear functionals.

**Proof** \( \forall z \in P \), define \( f_+(z) := \sup_{0 \leq \phi \leq z} \exists_{\phi(x) \leq z(x), \forall x \in X} f(\phi) \). Then, \( 0 \leq f_+(z) \leq \| f \| \| z \| < \infty \) and \( f_+(z) \geq f(z), \forall z \in P \). Clearly, \( f_+(z) \).

(i) \( f_+(\alpha z) = \alpha f_+(z), \forall z \in P \) and \( \forall \alpha \in [0, \infty) \subseteq \mathbb{R} \).

\( \forall z_1, z_2 \in P, \forall \phi_i \in \mathcal{Z} \) with \( 0 \leq \phi_i(x) \leq z_i(x), \forall x \in X \), we have \( f(\phi_i) + f(\phi_2) = f(\phi_1 + \phi_2) \leq f_+(z_1 + z_2) \). Then, \( f_+(z_1) + f_+(z_2) \leq f_+(z_1 + z_2) \). On the other hand, \( \forall \phi \in \mathcal{Z} \) with \( 0 \leq \phi(x) \leq z_1(x) + z_2(x), \forall x \in X \), we have \( f(\phi) = f(\phi \wedge z_1) + f(\phi - \phi \wedge z_1) \leq f_+(z_1) + f_+(z_2) \). Then, \( f_+(z_1 + z_2) \leq f_+(z_1) + f_+(z_2) \). Therefore, \( f_+(z_1) + f_+(z_2) = f_+(z_1 + z_2), \forall z_1, z_2 \in P \).
11.10. **Dual of \(\mathcal{C}(X, Y)\) and \(\mathcal{C}_c(X, Y)\)**

\[\forall z \in Z, \text{ define } z_+ := z \vee 0 \in P \text{ and } z_- := (-z) \vee 0 \in P. \] Clearly, 
\[z = z_+ - z_. \] Define \(f_+(z) := f_+(z_+) - f_+(z_-) \in \mathbb{R}. \) Then, \(f_+ : Z \to \mathbb{R}\) is well-defined. We will show that \(f_+ \in Z^*. \)

\[\forall z \in Z, |f_+(z)| = |f_+(z_+) - f_+(z_-)| \leq f_+(z_+) + f_+(z_-) = f_+(z_+ + z_-) \leq \|f\|\|z_+ + z_-\| = \|f\|\|z\|.\]

\[\forall z_1, z_2 \in Z, \text{ let } z := z_1 + z_2. \text{ Then, } f_+(z_1) + f_+(z_2) = f_+(z_1 + z_2) = f_+(z_1 + z_2) = f_+(z_1) + f_+(z_2). \]

\[f_+(z_1 + z_2) = f_+(z_1) + f_+(z_2) - f_+(z_1) + f_+(z_2), \text{ where the first equality follows from the definition of } f_+; \text{ the second equality follows from (ii); the third equality follows from (ii); and the last equality follows from the definition of } f_+.\]

Note that, \(\forall x \in X,\)

\[ (z_1 + z_2 - z_+)(x) = z_1(x) \vee 0 + z_2(x) \vee 0 - (z_1(x) + z_2(x)) \vee 0 = \begin{cases} 0 & z_1(x) \geq 0 \text{ and } z_2(x) \geq 0 \\ -z_2(x) & z_1(x) \geq 0 > z_2(x) \geq -z_1(x) \\ z_1(x) & z_1(x) \geq 0 \geq -z_1(x) > z_2(x) \\ -z_1(x) & z_2(x) \geq 0 > z_1(x) \geq -z_2(x) \\ z_2(x) & z_2(x) \geq 0 > -z_2(x) > z_1(x) \\ 0 & z_1(x) < 0 \text{ and } z_2(x) < 0 \end{cases} \]

\[= (-z_1(x)) \vee 0 + (-z_2(x)) \vee 0 - (-z_1(x) - z_2(x)) \vee 0 = z_1(x) + z_2(x) - z_-(x) \]

Hence, (iii) \(f_+(z_1) + f_+(z_2) = f_+(z) = f_+(z_1 + z_2), \forall z_1, z_2 \in Z.\)

\[\forall z \in Z, \forall \alpha \in \mathbb{R}. \text{ If } \alpha = 0, \text{ then } f_+(\alpha z) = 0 = \alpha f_+(z). \text{ If } \alpha > 0, \text{ then } f_+(\alpha z) = f_+(\alpha z_+) - f_+(\alpha z_-) = \alpha f_+(z_+) - \alpha f_+(z_-) = \alpha f_+(z), \text{ where the first equality follows from the definition of } f_+; \text{ the second equality follows from (i); and the third equality follows from the definition of } f_+. \]

\[f_+(\alpha z) = f_+(\alpha z_+) - f_+(\alpha z_-) - \alpha f_+(z_+) + \alpha f_+(z_-) = \alpha f_+(z), \text{ where the first equality follows from the definition of } f_+; \text{ the second equality follows from (i); and the third equality follows from the definition of } f_+. \]

Hence, we have

(iv) \(f_+(\alpha z) = \alpha f_+(z), \forall z \in Z \text{ and } \forall \alpha \in \mathbb{R}. \)

Hence, \(f_+ \in Z^* \) and is a positive linear functional.

Let \(f_- := f_+ - f \in Z^*. \) Clearly, \(f_-(z) \geq 0, \forall z \in P. \) Hence, \(f_- \) is also a positive linear functional.

This completes the proof of the lemma. \(\square\)

**Theorem 11.201 (Riesz Representation Theorem)** Let \(X := (X, O)\) be a compact Hausdorff topological space, \(Y \) be a normed linear space over \(\mathbb{K}, \) \(Z := \mathcal{C}(X, Y), \) and \(Z := \{z \in Z \mid z = hy, h \in \mathcal{C}(X, \mathbb{R}), y \in Y\}. \) Assume that \(\text{span}(Z) = Z. \) Then, \(\forall f \in Z^*, \exists \mu \in \mathcal{M}_f(\mathcal{X}, \mathbb{Y}^*) \text{ such that } f(z) = \int_X \langle \mu(x), z(x) \rangle \rangle = \langle \mu, z \rangle, \forall z \in Z. \) Furthermore, \(Z^* = \mathcal{M}_f(\mathcal{X}, \mathbb{Y}^*) \) isometrically isomorphically.
Proof. By Proposition 5.14, $\mathcal{X}$ is normal. By Example 7.31, $\mathcal{Z}$ is a normed linear space over $\mathbb{K}$. Fix any $f \in \mathcal{Z}^*$. Define $\nu_0 : \mathcal{O} \to [0, \|f\|] \subseteq \mathbb{R}$ by $\nu_0(O) := \sup_{z \in \mathcal{Z}, \|z\| \leq 1} \sup \{ |f(z)| : \forall O \in \mathcal{O} \}$. Clearly, we have 
(i) $\nu_0(\emptyset) = 0$;
(ii) $0 \leq \nu_0(O) \leq \|f\| < \infty$, $\forall O \in \mathcal{O}$;
(iii) $\nu_0(O_1) \leq \nu_0(O_2)$, $\forall O_1, O_2 \in \mathcal{O}$ with $O_1 \subseteq O_2$.

For $(O_i)_{i=1}^\infty \subseteq \mathcal{O}$, let $O := \bigcup_{i=1}^\infty O_i \in \mathcal{O}$. $\forall z \in \mathcal{Z}$ with $K := \text{supp}(z) \subseteq O$ and $\|z\| \leq 1$, by Proposition 5.5, $K$ is compact. By Corollary 5.65 of Partition of Unity, $\exists m \in \mathbb{Z}_+$, $\exists n_1, \ldots, n_m \in \mathbb{N}$, which may be taken to be distinct, $\exists (\varphi_{n_i})_{i=1}^m \subseteq \mathcal{C}(X, \mathbb{R})$ such that $\varphi_{n_i} : X \to [0, 1] \subseteq \mathbb{R}$, $\text{supp}(\varphi_{n_i}) \subseteq O_{n_i}, \ i = 1, \ldots, m$, $\sum_{i=1}^m \varphi_{n_i}(x) = 1, \forall x \in K$, and $0 \leq \sum_{i=1}^m \varphi_{n_i}(x) \leq 1, \forall x \in X$. Then, $z = \sum_{i=1}^m z_i \varphi_{n_i}, \forall i \in \{1, \ldots, m\}$, $z \varphi_{n_i} \in \mathcal{Z}$, $\|z \varphi_{n_i}\| \leq 1$, and $\text{supp}(z \varphi_{n_i}) \subseteq \text{supp}(z) \cap \text{supp}(\varphi_{n_i}) \subseteq O_{n_i}$. Then, $|f(z)| = |\sum_{i=1}^m f(z \varphi_{n_i})| \leq \sum_{i=1}^m |f(z \varphi_{n_i})| \leq \sum_{i=1}^m \nu_0(O_{n_i}) \leq \sum_{i=1}^\infty \nu_0(O_i)$. By the arbitrariness of $z$, we have 
(iv) $\nu_0(O) = \nu_0(\bigcup_{i=1}^\infty O_i) \leq \sum_{i=1}^\infty \nu_0(O_i), \forall (O_i)_{i=1}^\infty \subseteq \mathcal{O}$.

For $O_1, O_2 \in \mathcal{O}$ with $O_1 \cap O_2 = \emptyset$, $\forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}$, $\exists z_i \in \mathcal{Z}$ with $\text{supp}(z_i) \subseteq O_i$ and $\|z_i\| \leq 1$, such that $|f(z_i)| > \nu_0(O_i) - \epsilon/2, \forall i \in \{1, 2\}$, when $f(z_i) \neq 0$, take $z_i := \frac{f(z_i)}{|f(z_i)|} \in \mathcal{Z}$; and when $f(z_i) = 0$, take $z_i := z_i \in \mathcal{Z}$. Then, $z_i \in \mathcal{Z}, \|z_i\| = \|z_i\| \leq 1$, $\text{supp}(z_i) \subseteq \text{supp}(z_i) \subseteq O_i$, $\text{supp}(z_1) \subseteq \text{supp}(z_1) \subseteq O_1$, and $f(z_i) = |f(z_i)| > \nu_0(O_i) - \epsilon/2$. Let $z := z_1 + z_2 \in \mathcal{Z}$. Clearly, $\|z\| \leq 1$ since $O_1 \cap O_2 = \emptyset$. Note that $\text{supp}(z) \subseteq \text{supp}(z_1) \cup \text{supp}(z_2) \subseteq \text{supp}(z_1) \cup \text{supp}(z_2) \subseteq O_1 \cup O_2$, where the equality follows from Proposition 3.3. Then, $\nu_0(O_1) + \nu_0(O_2) - \epsilon < f(z_1) + f(z_2) = f(z) \leq \nu_0(O_1 \cup O_2) \leq \nu_0(O_1) + \nu_0(O_2)$, where the last inequality follows from (iii). By the arbitrariness of $\epsilon$, we have 
(v) $\nu_0(O_1 \cup O_2) = \nu_0(O_1) + \nu_0(O_2), \forall O_1, O_2 \in \mathcal{O}$ with $O_1 \cap O_2 = \emptyset$.

By Proposition 11.193, there exists a finite topological measure space $\tilde{\mathcal{X}} := (X, \mathcal{B}_B(X), \nu)$ such that $\nu(O) = \nu_0(O), \forall O \in \mathcal{O}$ and $\nu(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \nu_0(O), \forall E \in \mathcal{B}_B(X)$. Furthermore, $\nu$ is unique among measures $\nu$ on $(X, \mathcal{B}_B(X))$ with $\nu_0 = \nu_0$. Clearly, $\nu \in \mathcal{M}_f(X)$.

Fix any $\tilde{F} \in \mathcal{O}$. Let $A_F := \{(U, h) \in \mathcal{O} \times \mathcal{O} : (X, \mathcal{B}_B(X), \nu) \mid F \subseteq U \subseteq \mathcal{U} \subseteq V, h : X \to [0, 1] \subseteq \mathbb{R}, \text{supp}(h) \subseteq V \}$. Define a relation $\prec$ on $A_F$ by $(U_1, V_1, h_1) \prec (U_2, V_2, h_2)$ if $V_1 \supseteq V_2$. By Proposition 3.35 and Urysohn's Lemma 3.55, $A_F := (A_F, \prec)$ is a directed system. $\forall (U, V, h) \in A_F$, the function $f_h : Y \to \mathbb{K}$, defined by $f_h(y) = f(hy), \forall y \in Y$, is a bounded linear functional since $f \in \mathcal{Z}^*$. Then, $f_h \in \mathcal{Y}^*$. This defines a net $(f_h)_{(U, V, h) \in A_F} \subseteq \mathcal{Y}^*$.
Claim 11.201.1 The net \((f_h)_{(U,V,h) \in \mathcal{A}_F} \subseteq Y^*\) is Cauchy.

Proof of claim: \(\forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \exists V \in \mathcal{O}\) with \(F \subseteq V\) such that \(\nu(V \setminus F) < \epsilon\). By Proposition 3.35, \(\exists U, \bar{U} \in \mathcal{O}\) such that \(F \subseteq U \subseteq \bar{U} \subseteq \bar{U} \subseteq U \subseteq V\). By Urysohn’s Lemma 3.55, \(\exists h \in C(\mathcal{X}, \mathbb{R})\) with \(h : \mathcal{X} \to [0,1] \subset \mathbb{R}\) such that \(h|_{\bar{U}} = 1\) and \(h|_{\overline{\mathcal{X}}_F} = 0\). Then, \(\text{supp}(h) \subseteq \overline{\mathcal{X}}_F\) and \((U,V,h) \in \mathcal{A}_F\) for all \((U,V,h) \in \mathcal{A}_F\). Fix \((U_1, V_1, h_1) \in \mathcal{A}_F\) with \((U,V,h) < (U_1, V_1, h_1)\), we have \(V_1 \subseteq V\) and \(\|f_h - f_{h_1}\| = \sup_{y \in Y, \|y\| \leq 1} |\langle f_h, y \rangle - \langle f_{h_1}, y \rangle| = \sup_{y \in Y, \|y\| \leq 1} |f(hy) - f(h_1y)| = \sup_{y \in Y, \|y\| \leq 1} |f((h - h_1)y)|\). Note that, \(\forall y \in Y\) with \(\|y\| \leq 1\), \((h - h_1)y \in \mathbb{Z}\), \(\|(h - h_1)y\| \leq 1\), and

\[
\begin{align*}
\text{supp}((h - h_1)y) & \subseteq \text{supp}((h - h_1)) \\
& \subseteq \left(\text{supp}(h) \cup \text{supp}(h_1)\right) \setminus \{x \in X \mid h(x) = 1 = h_1(x)\} \\
& \subseteq \left(\text{supp}(h) \cup \text{supp}(h_1)\right) \setminus (U \cap U_1) \\
& = \left(\text{supp}(h) \cup (U \cap U_1)\right) \cup \left(\text{supp}(h_1) \setminus (U \cap U_1)\right) \\
& \subseteq \left(\text{supp}(h) \setminus (U \cap U_1)\right) \cup \left(\text{supp}(h_1) \cup (U \cap U_1)\right) \\
& \subseteq \left(\text{supp}(h) \setminus (U \cap U_1)\right) \cup \left(\text{supp}(h_1) \setminus (U \cap U_1)\right)
\end{align*}
\]

where the third containment follows from the fact that \(h|_{\bar{U}} = 1\) and \(h|_{\overline{\mathcal{X}}_F} = 0\); and the first equality follows from Proposition 3.3. Then, \(\|f_h - f_{h_1}\| \leq \nu_0(V \setminus F) = \nu(V \setminus F) < \epsilon\). Hence, the net is Cauchy. This completes the proof of the claim. \(\square\)

By Propositions 4.47 and 7.72, \((vii) \lim_{(U,V,h) \in \mathcal{A}_F} f_h =: \bar{\mu}(F) \in Y^*, \forall F \in \mathcal{O}\).

Thus, we have defined a function \(\bar{\mu}\) that assigns a vector \(\bar{\mu}(F) \in Y^*\) to each closed subset \(F \subseteq \mathcal{X}\). We will next show that \(\bar{\mu}\) satisfies the assumption of Proposition 11.194.

Fix any \(F_1, F_2 \in \mathcal{O}\) with \(F_1 \cap F_2 = \emptyset\). Fix any \(\epsilon \in (0, \infty) \subset \mathbb{R}\). By the normality of \(\mathcal{X}\), \(\exists V_1, V_2 \in \mathcal{O}\) with \(V_1 \cap V_2 = \emptyset\) such that \(F_i \subseteq V_i, i = 1,2\). Without loss of generality, we may assume that \(\nu(V_i \setminus F) < \epsilon/4, i = 1,2\). Fix any \(i \in \{1,2\}\). By \(\lim_{(U,V,h) \in \mathcal{A}_{F_i}} f_h = \bar{\mu}(F_i),\)

\[\exists (U_i, V_i, h_i) \in \mathcal{A}_{F_i}, \text{ with } V_i \subseteq V_i \text{ such that } \left\|\bar{\mu}(F_i) - f_{h_i}\right\| < \epsilon/6.\]

By \(\lim_{(U,V,h) \in \mathcal{A}_{F_1 \cup F_2}} f_h = \bar{\mu}(F_1 \cup F_2),\)

\[\exists (U, V, h) \in \mathcal{A}_{F_1 \cup F_2} \text{ with } V \subseteq V_1 \cup V_2 \text{ such that } \left\|\bar{\mu}(F_1 \cup F_2) - f_h\right\| < \epsilon/6.\]

This implies that \(\left\|\bar{\mu}(F_1 \cup F_2) - f_h\right\| < \epsilon/2 + \sup_{y \in Y, \|y\| \leq 1} \left|\langle f_h, y \rangle - \langle f_{h_1}, y \rangle\right|.\]

Note that, \(\forall y \in Y\) with \(\|y\| \leq 1\), \(\hat{z} := (h - h_1 - h_2)y \in \mathbb{Z}\), \(\|\hat{z}\| \leq 1\), since \(\text{supp}(h_1) \cap \text{supp}(h_2) \subseteq \hat{V}_1 \cap \hat{V}_2 = \emptyset, \text{ and supp}(\hat{z}) \subseteq\)
(supp(\hat{h}) \cup supp(h_1) \cup supp(h_2)) \setminus ((\hat{U} \cap \hat{U}_1) \cup (\hat{U} \cap \hat{U}_2)). \text{ Then, supp}(\hat{z}) \subseteq supp(h) \cup supp(h_1) \cup supp(h_2) = supp(\hat{h}) \cup supp(h_1) \cup supp(h_2) \subseteq \hat{V} \cup \hat{V}_1 \cup \hat{V}_2 \subseteq V_1 \cup V_2. \text{ In addition, supp}(\hat{z}) \subseteq \left( \left( \hat{U} \cap \hat{U}_1 \cup \left( \hat{U} \cap \hat{U}_2 \right) \right) \right) \subseteq \left( \hat{U} \cap (\hat{U}_1 \cup \hat{U}_2) \right) \subseteq F_1 \cup F_2. \text{ This yields supp}(\hat{z}) \subseteq (V_1 \setminus V_2) \cap F_1 \cap F_2 = (V_1 \setminus F_1) \cup (V_2 \setminus F_2) = : \hat{O} \in \mathcal{O}. \text{ Then, we have} \|\hat{\mu}(F_1 \cup F_2) - \hat{\mu}(F_2)\| \leq \|\hat{\mu}(\hat{V}_1 \cap F_1) - \hat{\mu}(\hat{V}_2 \setminus F_2)\| < \epsilon/2 + \nu(\hat{O}) = \epsilon/2 + \nu(V_1 \setminus F_1) + \nu(V_2 \setminus F_2) < \epsilon. \text{ By the arbitrariness of} \epsilon, \text{ we have}

(viii) \hat{\mu}(F_1) + \hat{\mu}(F_2) = \hat{\mu}(F_1 \cup F_2), \forall F_1, F_2 \in \mathcal{O} \text{ with} F_1 \cap F_2 = \emptyset.

Fix any \tilde{F}_1, \tilde{F}_2, \hat{F} \in \mathcal{O} \text{ with} F_1 \triangle F_2 \subseteq O. \forall \epsilon \in (0, +\infty) \in \mathbb{R}, \forall i \in \{1, 2\}, \exists \tilde{V}_i \in \mathcal{O} \text{ with} F_i \subseteq \tilde{V}_i \text{ such that} \nu(\tilde{V}_i \setminus F_i) < \epsilon/2. \text{ By}

lim_{(U,V,h) \in \mathcal{A}_F} f = \hat{\mu}(F), \exists (\tilde{U}_1, \tilde{V}_1, \tilde{h}_1) \in \mathcal{A}_F, \nu(\tilde{V}_i \setminus V_i) < \epsilon/2 \text{ that}

\|\hat{\mu}(F_i) - \hat{\mu}(F_2)\| \leq \|\hat{\mu}(F_i) - f_{h_1}\| + \|f_{h_1} - f_{h_2}\| + \|f_{h_2} - \hat{\mu}(F_2)\| < \epsilon/2 + sup_{y \in \mathbb{B}(\mathcal{X}), \|y\| \leq 1} \|f((\tilde{h}_1 - \hat{h}_2) y)\|.

Note that, \forall y \in \mathbb{B}(\mathcal{X}) \text{ with} \|y\| \leq 1, \exists \epsilon > 0, \nu(\hat{O} \setminus \hat{O}_1) \leq 1, \text{ and supp}(\hat{z}) \subseteq (supp(\hat{h}) \cup supp(h_1) \cup supp(h_2)) \setminus (\hat{U}_1 \cap \hat{U}_2) \subseteq (supp(h_1) \cup supp(h_2)) \setminus (\hat{U}_1 \cap \hat{U}_2) \subseteq (V_1 \setminus V_2) \cap (F_1 \cap F_2) = (V_1 \setminus F_1) \cup (V_2 \setminus F_2) \cup (F_1 \setminus F_2) \cup (F_2 \setminus F_1) \cup (F_1 \setminus F_2) \cup (V_1 \setminus F_1) \cup (V_2 \setminus F_2) = \hat{O} \subseteq O. \text{ Thus,}

\|\hat{\mu}(F_1) - \hat{\mu}(F_2)\| < \epsilon/2 + \nu(\hat{O}) = \epsilon/2 + \nu(\hat{O}) < \epsilon/2 + \nu(V_1 \setminus F_1) + \nu(V_2 \setminus F_2) + \nu(\hat{O}) < \epsilon + \nu(\hat{O}). \text{ By the arbitrariness of} \epsilon, \text{ we have}

(ix) \|\hat{\mu}(F_1) - \hat{\mu}(F_2)\| \leq \nu(\hat{O}), \forall \tilde{F}_1, \tilde{F}_2, \hat{F} \in \mathcal{O} \text{ with} F_1 \cap F_2 \subseteq O.

By Proposition 11.194, \exists \mu \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*) \text{ such that} \mu(F) = \hat{\mu}(F), \forall \hat{F} \in \mathcal{O}. \text{ Furthermore,} \mathcal{P} \circ \mu \leq \nu. \text{ Then,} \|\mu\| = \mathcal{P} \circ \mu(X) \leq \|\mu\| \leq \|f\| < \infty. \text{ This defines a mapping} \Phi_{\mathbb{Z}} : \mathbb{Z}^* \rightarrow \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*) \text{ by} \Phi_{\mathbb{Z}}(f) = \mu, \forall f \in \mathbb{Z}^*.

The above can be expressed as

\(\Phi_{\mathbb{Z}}(f) \leq \|f\|, \forall f \in \mathbb{Z}^*.\)

\(\forall f_1, f_2 \in \mathbb{Z}^*, \forall \alpha \in \mathbb{K}, \text{ let} \mu_i = \Phi_{\mathbb{Z}}(f_i) \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*), \text{ i = 1, 2,} \quad f := f_1 + f_2 \in \mathbb{Z}^*, \text{ and} \quad \mu := \Phi_{\mathbb{Z}}(f) \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*). \forall \hat{F} \in \mathcal{O}, \mu(F) = \hat{\mu}(F) = lim_{(U,V,h) \in \mathcal{A}_F} f = lim_{(U,V,h) \in \mathcal{A}_F} (f_1 + f_2) = \hat{\mu}_1(F) + \hat{\mu}_2(F) = \mu_1(F) + \mu_2(F) = (\mu_1 + \mu_2)(F).

\text{By Proposition 11.192,} \mu_1 + \mu_2 \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*). \text{ Then, by Proposition 11.194,} \Phi_{\mathbb{Z}}(f_1 + f_2) = \mu = \mu_1 + \mu_2 = \Phi_{\mathbb{Z}}(f_1) + \Phi_{\mathbb{Z}}(f_2). \text{ Let} \hat{\mu} := \Phi_{\mathbb{Z}}(\alpha f_1) \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*). \forall \hat{F} \in \mathcal{O}, \text{ we have} \hat{\mu}(F) = \hat{\mu}(F) = lim_{(U,V,h) \in \mathcal{A}_F} \alpha f_1 = lim_{(U,V,h) \in \mathcal{A}_F} \alpha f_1 h = \alpha \hat{\mu}_1(F) = (\alpha \hat{\mu}_1)(F). \text{ By Proposition 11.192,} \alpha \hat{\mu}_1 \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*). \text{ Then, by Proposition 11.194,} \Phi_{\mathbb{Z}}(\alpha f_1) = \mu = \alpha \hat{\mu}_1 = \alpha \Phi_{\mathbb{Z}}(f_1). \text{ This shows that} \Phi_{\mathbb{Z}} \text{ is a bound linear operator.}

Fix any \mu \in \mathcal{M}_{f_1}(\mathcal{X}, \mathbb{Y}^*). \forall z \in \mathbb{Z}. \text{ By Proposition 11.37,} z \text{ is} \mathcal{B}_{\mathcal{B}}(\mathcal{X})-\text{measurable.} \text{ Clearly,} \mathcal{P} \circ z(x) \leq \|z\|, \forall x \in X. \text{ Then,} \int_X \mathcal{P} \circ z \, d\mathcal{P} \circ \mu \leq
We will show that $f(z) = \int_X \langle \mu(x), z(x) \rangle = \langle \langle \mu, z \rangle \rangle$, $\forall z \in \mathcal{Z}$, $\forall f \in \mathcal{Z}^*$, where $\mu = \Phi_z(f)$, in four steps. In the first step, we consider the special case: $\mathcal{Y} = \mathbb{R}$ and $f$ is a positive linear functional. $\forall F \subseteq \mathcal{O}$, $\forall (U, V, h) \in \mathcal{A}_F$, $f_h = f(h) \in [0, f\|] \subseteq \mathbb{R}$, since $h \in P := \{ z \in \mathcal{C}(X, \mathbb{R}) \mid z : X \to (0, \infty) \subseteq \mathbb{R} \} \subseteq \mathcal{C}(X, \mathbb{R}) = \mathbb{Z}$. Then, $\bar{\mu}(F) = \lim_{h \in \mathcal{F}(U, V, h) \in \mathcal{A}_F} f_h \in [0, f\|] \subseteq \mathbb{R}$.

Hence, by Proposition 11.194 and its proof, $\bar{\mu} : \mathcal{B}_B(X) \to [0, f\| \subseteq \mathbb{R}$ and $\mu \in \mathcal{M}_f(X)$. Fix any $z \in \mathcal{Z}$. We will distinguish three exhaustive cases: Case 1: $z = \partial z$; Case 2: $z \neq \partial z$ and $z \in P$; Case 3: $z \in \mathcal{Z}$. Case 1: $z = \partial z$. Then, $f(z) = 0 = \int_X z(x) \mu(x) = \langle \langle \mu, z \rangle \rangle$.

Case 2: $z \neq \partial z$ and $z \in P$. Let $\bar{z} := z/\|z\| \in P$. Then, $\bar{z} : X \to [0, 1] \subseteq \mathbb{R}$.

Fix any $n \in \mathbb{N}$ and fix any $i \in \{1, \ldots, n\}$. Define $\phi_i := ((n\bar{z} - i + 1) \cup 1 \subseteq P$, $O_i := \{ x \in X \mid n\bar{z}(x) > i - 1 \}$, and $F_i := \{ x \in X \mid n\bar{z}(x) \geq i - 1 \}$. Let $O_0 := \{ x \in X \mid n\bar{z}(x) > 0 \} = X$, $O_{i+1} := \{ x \in X \mid n\bar{z}(x) > n \} = 0$, $F_{i+1} := \{ x \in X \mid n\bar{z}(x) \geq n \}$, and $F_{i+2} := \{ x \in X \mid n\bar{z}(x) \geq n + 1 \} = 0$. Clearly, $F_{i+2} \subseteq O_{i+1} \subseteq O_{i+1} \subseteq F_{i+1} \subseteq O_{i+1} \subseteq F_{i+1} \subseteq O_{i-1}$, $\phi_i : X \to [0, 1] \subseteq \mathbb{R}$, $\phi_i | F_{i+1} = 1$, and $\text{supp}(\phi_i) \subseteq O_i \subseteq F_i \subseteq O_{i-1}$. Clearly $O_i \subseteq \mathcal{O}$ and $F_{i+1} \subseteq \mathcal{O}$, $i = 0, \ldots, n+1$, and $\bar{z} = \frac{1}{n} \sum_{i=1}^{n} \phi_i$. Fix any $i \in \{1, \ldots, n\}$. By the definition of $\bar{\mu}$ and the fact that $f$ is a positive linear functional, $\forall (U, V, h) \in \mathcal{A}_{F_{i+2}}$, with $V \subseteq O_{i+1}$, we have $h(x) \leq \chi_{O_{i+1}, X}(x) \leq \chi_{F_{i+1}, X}(x) \leq \phi_i(x), \forall x \in X$, and $f_h = f(h) \leq f(\phi_i)$. This implies that $\mu(F_{i+2}) \leq f(\phi_i)$. $\forall (U, V, h) \in \mathcal{A}_{F_i}$, we have $f(\phi_i) \leq \chi_{F_i, X}(x) \leq h(x), \forall x \in X$, and $f_h = f(h) \leq f(\phi_i)$.

This leads to $f(\phi_i) \leq \mu(F_i)$. By Propositions 11.75 and 11.83, we have $\mu(F_{i+2}) \leq \mu(F_{i+1}) = \int_X \chi_{F_{i+1}, X} d\mu \leq \int_X \phi_i d\mu = \mu(F_i)$. Then, $f(\phi_i) \leq \mu(F_i) = \mu(F_{i+2})$. Summing both side from 1 to $n$, we have $f(\bar{z}) = \int_X \bar{z} d\mu = \frac{1}{n} \sum_{i=1}^{n} f(\phi_i) - \int_X \bar{z} d\mu \leq \frac{1}{n} \sum_{i=1}^{n} \mu(F_i) = \mu(F_{i+2})$. The first inequality follows from (xi) and Proposition 11.92; and the last inequality follows from (x). By the arbitrariness of $n$, we have $f(\bar{z}) = \int_X \bar{z} d\mu$. By (xi) and Proposition 11.92, $f(z) = f(\|z\| \bar{z}) = \|z\| f(\bar{z}) = \int_X \bar{z} d\mu = \langle \langle \mu, z \rangle \rangle$. Hence, $f(z) = \int_X z d\mu = \langle \langle \mu, z \rangle \rangle$, $\forall z \in \mathcal{Z}$. This completes the first step.
In the second step, we consider the special case: $y = \mathbb{R}$. By Lemma 11.200, $f = f_+ - f_-$, where $f_+, f_- \in Z^*$ are positive linear functionals. Let $\mu_+ := \Phi_Z(f_+)$ and $\mu_- := \Phi_Z(f_-)$. By the first step, $f_+(z) = \int_X z \, d\mu_+$ and $f_-(z) = \int_X z \, d\mu_-$, $\forall z \in Z$. Then, by linearity of $\Phi_Z$, $\mu = \Phi_Z(f) = \Phi_Z(f_+ - f_-) = \Phi_Z(f_+) - \Phi_Z(f_-) = \mu_+ - \mu_-$. By Propositions 11.145 and 11.146, $f(z) = f_+(z) - f_-(z) = \int_X z \, d\mu_+ - \int_X z \, d\mu_- = \int_X z \, d\mu = \langle \langle \mu, z \rangle \rangle$. This completes the second step.

In the third step, we consider the special case when $y$ is a real normed linear space. $\forall y \in \mathbb{Y}$, define $\mu_y : B_B(X) \to \mathbb{R}$ by $\mu_y(E) = \langle \langle \mu(E), y \rangle \rangle$, $\forall E \in B_B(X)$. Then, $\mu_y \in M_{ft}(X, \mathbb{R})$ by Proposition 11.136. Define $f_y : C(X, \mathbb{R}) \to \mathbb{R}$ by $f_y(h) = f(hy)$, $\forall h \in C(X, \mathbb{R})$. By $f \in Z^*$, we have $f_y \in (C(X, \mathbb{R}))^*$. $\forall F \in O$, $\mu_y(F) = \langle \langle \mu(F), y \rangle \rangle = \langle \langle \lim_{(U,Y,F) \in A_F} f_h(y) \rangle \rangle = \lim_{(U,Y,F) \in A_F} \langle \langle f_h(y) \rangle \rangle = \langle \langle \Phi_{C(X,\mathbb{R})}(f)_y \rangle \rangle$. By Proposition 11.194, we have $\mu_y = \Phi_{C(X,\mathbb{R})}(f_y)$. By the second step, we have $f(hy) = f_y(y) = \int_X h \, d\mu_y = \int_X \langle \langle d\mu(x), h(x)y \rangle \rangle = \langle \langle \mu(hy), y \rangle \rangle$, $\forall h \in C(X, \mathbb{R})$, where the third equality follows from Proposition 11.146. $\forall z \in Z$, $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\exists \bar{z} \in \text{span}(Z)$ such that $\|z - \bar{z}\| < \epsilon/(2 \|f\| + 1)$. Then, $\bar{z} = \sum_{i=1}^n h_i y_i$, where $n \in Z_+$, $h_1, \ldots, h_n \in C(X, \mathbb{R})$, and $y_1, \ldots, y_n \in y$. Then, $|f(z) - \langle \langle \mu, z \rangle \rangle| \leq |f(z) - f(\bar{z})| + |f(\bar{z}) - \langle \langle \mu, \bar{z} \rangle \rangle| + |\langle \langle \mu, \bar{z} \rangle \rangle - \langle \langle \mu, z \rangle \rangle| \leq \|f\| \|z - \bar{z}\| + \left| \sum_{i=1}^n f(h_i y_i) - \int_X \langle \langle d\mu(x), \sum_{i=1}^n h_i(x) y_i \rangle \rangle \right| + \left| \int_X \langle \langle d\mu(x), \bar{z}(x) - z(x) \rangle \rangle \right| \leq \epsilon/2 + \left| \sum_{i=1}^n f(h_i y_i) - \int_X \langle \langle d\mu(x), h_i(x) y_i \rangle \rangle \right| + \left| \int_X \langle \langle d\mu(x), \bar{z}(x) - z(x) \rangle \rangle \right| \leq \epsilon/2 + 0 + \int_X |\bar{z} - z| \, d\mu(x) \leq \epsilon/2 + \frac{\|\bar{z} - z\| \|\mu\|}{2 \|f\| + 1} \|\mu\| \leq \epsilon/2 + \frac{\|\mu\| \|\mu\|}{2 \|f\| + 1} \|f\| < \epsilon$, where the second inequality follows from Proposition 7.72; the third inequality follows from Proposition 11.132; the fourth inequality follows from Proposition 11.132; the fifth inequality follows from Propositions 11.75 and 11.83; and the seventh inequality follows from (x). By the arbitrariness of $\epsilon$, we have $f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle$. This completes the third step.

In the fourth step, we consider the special case when $y$ is a complex normed linear space. Let $g := \text{Re} \circ f$. By Lemmas 7.40 and 7.81, $g$ is a bounded linear functional of the real normed linear space $Z_{\mathbb{R}} := (C(X, y_{\mathbb{R}}), \mathbb{R}, \| \cdot \|_{C(X, y_{\mathbb{R}})})$ and $f(z) = g(z) - ig(z)$, $\forall z \in Z_{\mathbb{R}}$. Let $\tilde{g} \in Z_{\mathbb{R}}$ be defined by $\tilde{g}(z) = g(iz)$, $\forall z \in Z_{\mathbb{R}}$, $\mu_{\tilde{g}} := \Phi_{Z_{\mathbb{R}}}(g) \in M_{ft}(X, y_{\mathbb{R}})$, and $\mu_{g} := \Phi_{Z_{\mathbb{R}}}(g) \in \text{M}_{ft}(X, y_{\mathbb{R}})$. $\forall F \in O$, $\mu(F) = \lim_{(U,Y,F) \in A_F} f_h = \lim_{(U,Y,F) \in A_F} (g_h - i\tilde{g}_h) F = \lim_{(U,Y,F) \in A_F} (g_h - i\tilde{g}_h) F = \Phi_{Z_{\mathbb{R}}}(g)(F) - i\Phi_{Z_{\mathbb{R}}}(\tilde{g})(F) = \mu_{gf} - i\mu_{gf}(F) = (\mu_{gf} - i\mu_{gf})(F)$. By the definition of $\tilde{g}$, we have $\mu_{\tilde{g}}(E)(\iy) = \mu_{\tilde{g}}(E)(\iy) \in \mathbb{C}, \forall E \in B_B(X)$, $\forall \iy \in y_{\mathbb{R}}$. Then, by Lemma 7.81, $\mu_{\tilde{g}}(E) - i\mu_{\tilde{g}}(E) \in y^*, \forall E \in B_B(X)$. By Proposition 11.136, $\mu_{gf} - i\mu_{gf} \in \text{M}_{ft}(X, y^*)$. By Proposition 11.194, we have $\mu = \mu_{gf} - i\mu_{gf}$. By the third step, $g(z) = \int_X \langle \langle d\mu_{gf}(x), z(x) \rangle \rangle$ and $\tilde{g}(z) = \int_X \langle \langle d\mu_{\tilde{g}}(x), z(x) \rangle \rangle$, $\forall z \in Z_{\mathbb{R}}$. This implies, $\forall z \in \mathbb{Z}, f(z) = g(z) - ig(iz) = g(z) - ig(z) = \int_X \langle \langle d\mu_{gf}(x), z(x) \rangle \rangle - i\int_X \langle \langle d\mu_{gf}(x), z(x) \rangle \rangle = \int_X \langle \langle d\mu_{\tilde{g}}(x), z(x) \rangle \rangle - i\int_X \langle \langle d\mu_{\tilde{g}}(x), z(x) \rangle \rangle = \int_X \langle \langle d\mu_{\tilde{g}}(x), z(x) \rangle \rangle$.
\[ \int_X \langle \langle d(\mu_f - i\mu_g) \rangle (x), z(x) \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle, \]
where the fourth equality follows from Propositions 11.145 and 11.146. This completes the fourth step.

Thus, we have the representation

\[(xii) \quad f(z) = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle, \quad \forall z \in \mathcal{Z}, \quad \forall f \in \mathcal{Z}^*, \text{ where } \mu = \Phi_z(f) \in \mathcal{M}_f(X, \mathcal{Y}^*).\]

Fix any \( \mu \in \mathcal{M}_f(X, \mathcal{Y}^*) \). Define \( f : \mathcal{Z} \to \mathcal{K} \) by \( f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle, \quad \forall z \in \mathcal{Z}. \) We will show that \( f \in \mathcal{Z}^*, \| f \| \leq \| \mu \|, \) and \( \Phi_z(f) = \mu. \quad \forall z \in \mathcal{Z}. \) By Proposition 5.7, \( z(X) \subseteq \mathcal{Y} \) is compact. Then, by Propositions 11.190, 7.35, and 7.17, \( H := \text{span}(z(X)) \subseteq \mathcal{Y} \) is a separable subspace. By \( (xi) \) and Proposition 11.132, \( f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \in \mathcal{K} \) is well-defined. By Proposition 11.132, \( f \in \mathcal{Z}^* \) and \( \| f \| = \sup_{z \in \mathcal{Z}, \| z \| \leq 1} \langle \langle \mu, z \rangle \rangle = \sup_{z \in \mathcal{Z}, \| z \| \leq 1} \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \leq \sup_{z \in \mathcal{Z}, \| z \| \leq 1} \| z \| \| \mu \| \leq \| \mu \|. \)

Let \( \tilde{\mu} := \Phi_z(f) \in \mathcal{M}_f(X, \mathcal{Y}^*). \) \( \forall F \in \mathcal{O}, \forall y \in \mathcal{Y}, \langle \langle \tilde{\mu}(F), y \rangle \rangle = \langle \langle \lim_{(U,V,h) \to A_F} f_{i} h, y \rangle \rangle = \lim_{(U,V,h) \to A_F} \langle \langle f_{i} h, y \rangle \rangle = \langle \langle \lim_{(U,V,h) \to A_F} f_{i} h(x(y)) \rangle \rangle, \quad \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \exists \epsilon \subset \mathcal{O} \text{ with } F \subseteq \mathcal{O} \text{ such that } \mathcal{P} \cap \mu(\mathcal{O} \setminus F) < \epsilon/(1 + \| y \|). \quad \forall (U,V,h) \in \mathcal{A}_F \text{ with } V \subseteq \mathcal{O}, \quad \int_{X} \langle \langle d\mu(x), h(x) y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle = \int_{\mathcal{P}} \langle \langle d\mu(x), h(x) y \rangle \rangle - \int_{\mathcal{F}} \langle \langle d\mu(x), h(x) y \rangle \rangle = \int_{\mathcal{O} \setminus \mathcal{F}} \langle \langle d\mu(x), h(x) y \rangle \rangle + \int_{\mathcal{O} \setminus \mathcal{F}} \langle \langle d\mu(x), h(x) y \rangle \rangle + 0 \leq \int_{\mathcal{O} \setminus \mathcal{F}} (\| y \| \mathcal{P} \cap \mu \leq \| y \| \mathcal{P} \cap \mu(\mathcal{O} \setminus F) < \epsilon, \text{ where the first equality follows from (xii) and Proposition 11.132; the second equality follows from Proposition 11.125; the first inequality follows from Proposition 11.132; and the second inequality follows from Proposition 11.83. This implies that } \| (\langle \langle \tilde{\mu}(F), y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle \| = \lim_{(U,V,h) \to A_F} \left| \int_{X} \langle \langle d\mu(x), h(x) y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle \right| \leq \epsilon. \text{ By the arbitrariness of } \epsilon \text{ and } y, \text{ we have } \tilde{\mu}(F) = \mu(F) \text{. By Proposition 11.194, } \tilde{\mu} = \mu. \text{ This shows that } \Phi_z \text{ is surjective.} \]

\( \forall f \in \mathcal{Z}^*, \text{ let } \mu := \Phi_z(f) \in \mathcal{M}_f(X, \mathcal{Y}^*). \) Then, by the above and (x), we have \( \| \mu \| = \| \Phi_z(f) \| \leq \| f \| \leq \| \mu \|. \) Hence, \( \| \Phi_z(f) \| = \| f \|. \) Then, \( \Phi_z \) is injective. This shows that \( \Phi_z \) is bijective, continuous, linear, and norm preserving. Hence, \( \Phi_z : \mathcal{Z}^* \to \mathcal{M}_f(X, \mathcal{Y}^*) \) is an isometrical isomorphism.

This completes the proof of the theorem.

\begin{definition}
Let \( m \in \mathbb{Z}_+, \quad I := [0, 1] \subset \mathbb{R}, \quad \mathcal{Y} \) be a normed linear space, \( f : I^m \to \mathcal{Y}, \) and \( n \in \mathbb{N}. \) The \( n \)th Bernstein function for \( f, \quad B_n : I^m \to \mathcal{Y}, \) is defined by, \( \forall x := (x_1, \ldots, x_m) \in I^m, \)

\[ B_n(x) = B_n(x; f) = \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right) \prod_{i=1}^{m} \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i} \]

where, for notational consistency, we have made the arbitrary assignment of \( 0^0 = 1 \) in this definition as well as the following theorem.
Theorem 11.203 (Bernstein Approximation Theorem) Let $m \in \mathbb{Z}_+$, $I \colonequals [0, 1] \subset \mathbb{R}$, $Y$ be a normed linear space, $f : I^m \to Y$ be continuous, and $(B_n)_{n=1}^\infty$ be the sequence of Bernstein functions for $f$. Then, $\lim_{n \in \mathbb{N}} \|f - B_n\|_{C(I^m, Y)} = 0$.

Proof $\forall n \in \mathbb{Z}_+$, $\forall y \in I$, we have

\[ 1 = (y + 1 - y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k (1 - y)^{n-k} \quad (11.6) \]

Then, $\forall n \in \mathbb{N}$, $\forall y \in I$, we have

\[ y = y \cdot 1 = y \left( \sum_{k=0}^{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right) \]
\[ = \sum_{k=0}^{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k} = \sum_{k=1}^{n} \binom{n-1}{k-1} y^k (1 - y)^{n-k} \]
\[ = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \quad (11.7) \]

where the second equality follows from (11.6). Similarly, $\forall n - 1 \in \mathbb{N}$, $\forall y \in I$,

\[ y^2 = y \left( \sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right) \]
\[ = \sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k} \]
\[ = \sum_{k=1}^{n} \frac{k-1}{n-1} \binom{n-1}{k-1} y^k (1 - y)^{n-k} = \sum_{k=1}^{n} \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k} \]
\[ = \sum_{k=0}^{n} \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k} \]

where the first equality follows from (11.7). This implies that

\[ (1 - \frac{1}{n})y^2 = \sum_{k=0}^{n} \frac{k^2 - k}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} \]
\[ = \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \]
\[ = \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} y \]
where the last equality follows from (11.7). Rearranging terms in the above yields, \( \forall y \in I, \forall n - 1 \in \mathbb{N} \),

\[
\frac{1}{n} y (1 - y) = \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} + y^2 - 2y^2 \\
= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} + y^2 \sum_{k=0}^{n} \binom{n}{k} y^k (1 - y)^{n-k} \\
- 2y \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \\
= \sum_{k=0}^{n} (y^2 - 2y \frac{k}{n} + \frac{k^2}{n^2} \binom{n}{k}) y^k (1 - y)^{n-k} \\
= \sum_{k=0}^{n} (y - \frac{k}{n})^2 \binom{n}{k} y^k (1 - y)^{n-k}
\]  

(11.8)

where the second equality follows from (11.6) and (11.7). \( \forall n \in \mathbb{Z}^+, \forall x := (x_1, \ldots, x_m) \in I^m \), we have

\[
1 = \prod_{i=1}^{m} (x_i + 1 - x_i)^n = \prod_{i=1}^{m} \left( \sum_{k=0}^{n} \binom{n}{k} x_i^k (1 - x_i)^{n-k} \right) \\
= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \prod_{i=1}^{m} \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i}
\]  

(11.9)

Then, \( \forall x := (x_1, \ldots, x_m) \in I^m \), \( \forall n - 1 \in \mathbb{N} \), we have

\[
\frac{1}{n} \sum_{i=1}^{m} x_i (1 - x_i) = \sum_{i=1}^{m} \sum_{k_i=0}^{n} (x_i - \frac{k_i}{n})^2 \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i} \\
= \sum_{i=1}^{m} \left( \sum_{k_i=0}^{n} (x_i - \frac{k_i}{n})^2 \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i} \right) \\
\cdot \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j} \\
= \sum_{i=1}^{m} \sum_{k_i=0}^{n} \cdots \sum_{k_m=0}^{n} (x_i - \frac{k_i}{n})^2 \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j} \\
= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \left( \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j} \right) \left( \sum_{i=1}^{m} (x_i - \frac{k_i}{n})^2 \right) \\
= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \left| x - \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right) \right|^2 \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j}
\]  

(11.10)
where the first equality follows from (11.9); and the second equality follows from (11.9).

Note that $I^m \subseteq \mathbb{R}^m$ is a compact metric space. By Propositions 5.22, 5.29, 7.21, 3.12, and 3.9, $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|f(x)\| \leq M, \forall x \in I^m$. By Proposition 5.39, $f$ is uniformly continuous. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x, \bar{x} \in I^m$ with $|x - \bar{x}| < \delta$, we have $\|f(x) - f(\bar{x})\| < \epsilon / 2$. Let $n_0 := \left[ \max \{ \delta^{-4}, m^2 M^2 / \epsilon^2, 2 \} \right] \in \mathbb{N}$. $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $\forall x := (x_1, \ldots, x_m) \in I^m$, we have

$$\|f(x) - B_n(x)\| = \left\| f(x) \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{m} \prod_{i=1}^{n} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i} \right\|
- \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{m} f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right) \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i} \right\|
\leq \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{m} \|f(x) - f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i}$$

where the first equality follows from (11.9) and Definition 11.202. Let $J := \{0, \ldots, n\}^m$ and $J_x := \{(k_1, \ldots, k_m) \in J \mid |x - \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)| < n^{-1/4}\}$. Then,

$$\|f(x) - B_n(x)\| \leq \sum_{(k_1, \ldots, k_m) \in J_x} \|f(x) - f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i}
+ \sum_{(k_1, \ldots, k_m) \in J \setminus J_x} \|f(x) - f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i}$$

Note that, $\forall (k_1, \ldots, k_m) \in J_x$, $\left| (x_1, \ldots, x_m) - \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right) \right| < n^{-1/4} \leq \delta$, then $\|f(x) - f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| < \epsilon / 2$. This implies that

$$\sum_{(k_1, \ldots, k_m) \in J_x} \|f(x) - f \left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i}
\leq \frac{\epsilon}{2} \sum_{(k_1, \ldots, k_m) \in J_x} \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i}
\leq \frac{\epsilon}{2} \sum_{(k_1, \ldots, k_m) \in J_x} \prod_{i=1}^{m} \left( n \atop k_i \right) x_i^{k_i} (1 - x_i)^{n-k_i} = \frac{\epsilon}{2}$$
where the equality follows from (11.9). Note also that, \( \forall (k_1, \ldots, k_m) \in J \setminus J_x, \| f(x) - f(\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \leq 2M = 2M \left[ \frac{x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})}{x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})} \right]^2 \leq 2M \sqrt{n} \| x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \|^2 \). Then,

\[
\sum_{(k_1, \ldots, k_m) \in J \setminus J_x} \| f(x) - f(\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \prod_{i=1}^{m} \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \leq 2M \sqrt{n} \sum_{(k_1, \ldots, k_m) \in J \setminus J_x} \| x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \|^2 \prod_{i=1}^{m} \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \leq 2M \sqrt{n} \sum_{(k_1, \ldots, k_m) \in J} \| x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \|^2 \prod_{i=1}^{m} \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} = 2M \sqrt{n} \sum_{n=1}^{m} x_i (1 - x_i) \leq \frac{Mm}{2 \sqrt{n}} \leq \epsilon/2
\]

Hence, we have \( \| f(x) - B_n(x) \| < \epsilon \). By the arbitrariness of \( x \in I^m \), we have \( \| f - B_n \|_{C(I^m, \mathbb{Y})} < \epsilon \). Therefore, \( \lim_{n \in \mathbb{N}} \| f - B_n \|_{C(I^m, \mathbb{Y})} = 0 \).

This completes the proof of the theorem. \( \square \)

**Theorem 11.204 (Riesz Representation Theorem)** Let \( m \in \mathbb{Z}_+ \), \( I_1, \ldots, I_m \subset \mathbb{R} \) be compact intervals, \( X := \prod_{i=1}^{m} I_i \subset \mathbb{R}^m \) with subset topology \( \mathcal{O} \), \( \mathcal{X} := (X, \mathcal{O}) \), \( \mathbb{Y} \) be a normed linear space, and \( \mathcal{Z} := C(\mathcal{X}, \mathbb{Y}) \).

Then, \( \mathcal{Z}^* = M_{ft}(\mathcal{X}, \mathbb{Y})^* = M_{ft}(\mathcal{X}, B_\mathbb{R}(\mathcal{X}), \mathbb{Y})^* \).

**Proof** By Tychonoff Theorem 5.47 and Proposition 4.37, \( \mathcal{X} \) is a compact metric space. By Proposition 5.5, \( \mathcal{X} \) is closed in \( \mathbb{R}^m \). By Proposition 4.39, \( \mathcal{X} \) is a complete metric space. Since \( \mathbb{R}^m \) is separable, by Proposition 4.38, \( \mathcal{X} \) is separable. Hence, \( \mathcal{X} \) is a separable compact complete metric space.

Let \( \mathcal{Z} := \{ z \in \mathcal{Z} \mid z = hy, h \in C(\mathcal{X}, \mathbb{R}), y \in \mathbb{Y} \} \). We will show that \( \mathcal{Z} = \text{span}(\mathcal{Z}) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \exists \theta_0 \in \{1, \ldots, m\} \) such that \( I_{\theta_0} = \emptyset \); Case 2: \( I_1, \ldots, I_m \) are nonempty. Case 1: \( \exists \theta_0 \in \{1, \ldots, m\} \) such that \( I_{\theta_0} = \emptyset \). Then, \( \mathcal{X} = \emptyset \) and \( \mathcal{Z} \) is the trivial Banach space with a singleton element. Clearly, \( Z = \mathcal{Z} \).

Hence, \( \mathcal{Z} = \text{span}(\mathcal{Z}) \).

Case 2: \( I_1, \ldots, I_m \) are nonempty. Without loss of generality, assume \( I_i = [a_i, b_i] \subset \mathbb{R} \) with \( a_i, b_i \in \mathbb{R} \) and \( a_i < b_i, i = 1, \ldots, m, \) and \( \exists \eta \in \mathbb{Z}_+ \) with \( m \leq m \) such that \( a_i < b_i, i = 1, \ldots, \eta, \) and \( a_i = b_i, i = \eta + 1, \ldots, m \). Let \( J := [0, 1] \subset \mathbb{R} \). Then, we may define a homeomorphism \( \psi : \mathcal{X} \to I^m \) by \( \psi(x) = (\frac{x_1 - a_1}{b_1 - a_1}, \ldots, \frac{x_m - a_m}{b_m - a_m}) \) for \( \forall x := (x_1, \ldots, x_m) \in X \). If \( f \in \mathcal{Z} \), let \( \hat{f} : I^m \to \mathbb{Y} \) be defined by \( \hat{f} := f \circ \psi^{-1} \). By Proposition 3.12, \( \hat{f} \in C(I^m, \mathbb{Y}) \). By Bernstein Approximation Theorem 11.203, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n \in \mathbb{N} \) such that \( \| f - B_n \|_{C(I^m, \mathbb{Y})} < \epsilon \), where \( B_n : I^m \to \mathbb{Y} \) is the nth Bernstein function.
for \( \tilde{f} \). Then, \( \| f - B_n \circ \psi \|_Z = \| (\tilde{f} - B_n) \circ \psi \|_Z = \| \tilde{f} - B_n \|_{C(f^m, Y)} < \epsilon \). Clearly, \( B_n \circ \psi \in \text{span} (Z) \). By the arbitrariness of \( \epsilon \), we have \( f \in \text{span} (Z) \).

By the arbitrariness of \( f \), we have \( Z = \text{span} (Z) \).

Hence, in both cases, we have \( Z = \text{span} (Z) \). By Riesz Representation Theorem 11.201, we have \( Z^* = M_{f_1}(X, Y^*) \). By Theorem 11.198, \( Z^* = M_{f_1}(X, Y^*) = M_f(X, B_0 (X), Y^*) \). This completes the proof of the theorem.

**Theorem 11.205 (Riesz Representation Theorem)** \( X := (X, \mathcal{O}) \) be a compact Hausdorff topological space, \( Y \) be a finite dimensional Banach space over \( \mathbb{K} \), and \( Z := C(X, Y) \). Then, \( Z^* = M_{f_1}(X, Y^*) \).

**Proof** Let \( Z := \{ z \in Z \mid z = hy, h \in C(X, \mathbb{R}), y \in Y \} \). We will show that \( Z = \text{span} (Z) \). Then, the theorem is a direct consequence of Riesz Representation Theorem 11.201.

Let the dimension of \( Y \) be \( m \in Z_+ \) and \( \{ y_1, \ldots, y_m \} \) be a basis of \( Y \). Then, there is an invertible bounded linear mapping \( \phi : Y \to \mathbb{K}^m \) such that \( y = \sum_{i=1}^m (\pi_i \circ \phi (y)) y_i \), \( \forall y \in Y \), where \( \pi_i : \mathbb{K}^m \to \mathbb{K} \) is the \( i \)-th coordinate projection function, \( i = 1, \ldots, m \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathbb{K} = \mathbb{R} \); Case 2: \( \mathbb{K} = \mathbb{C} \).

Case 1: \( \mathbb{K} = \mathbb{R} \). \( \forall f \in Z \), let \( g_i := \pi_i \circ \phi \circ f \in C(X, \mathbb{R}) \), \( i = 1, \ldots, m \). Then, \( f(x) = \sum_{i=1}^m (\pi_i \circ \phi (f(x))) y_i = \sum_{i=1}^m g_i(x) y_i \), \( \forall x \in X \). Hence, \( f \in \text{span} (Z) \). This case is proved.

Case 2: \( \mathbb{K} = \mathbb{C} \). \( \forall f \in Z \), let \( g_i := \text{Re} \circ \pi_i \circ \phi \circ f \in C(X, \mathbb{R}) \), \( h_i := \text{Im} \circ \pi_i \circ \phi \circ f \in C(X, \mathbb{R}) \), \( i = 1, \ldots, m \). Then, \( f(x) = \sum_{i=1}^m (\pi_i \circ \phi(f(x))) y_i = \sum_{i=1}^m (\text{Re} \circ \pi_i \circ \phi \circ f(x)) y_i + i \text{Im} \circ \pi_i \circ \phi \circ f(x) y_i = \sum_{i=1}^m (g_i(x) y_i + h_i(x) y_i) \), \( \forall x \in X \). Hence, \( f \in \text{span} (Z) \). This case is also proved.

This completes the proof of the theorem.

**Proposition 11.206** Let \( m \in Z_+ \), \( I_1, \ldots, I_m \subset \mathbb{R} \) be compact intervals, \( X := \prod_{i=1}^m I_i \subset \mathbb{R}^m \) with subset topology \( \mathcal{O} \), \( X := (X, \mathcal{O}) \), and \( Y \) be a separable normed linear space. Then, \( Z := C(X, Y) \) is separable.

**Proof** Let \( Z := \{ z \in Z \mid z = hy, h \in C(X, \mathbb{R}), y \in Y \} \). By Riesz Representation Theorem 11.204 and its proof, we have \( Z = \text{span} (Z) \). By Proposition 11.190, \( C(X, \mathbb{R}) \) is separable. Let \( D_1 \subset C(X, \mathbb{R}) \) be a countable dense subset. Let \( D_2 \subset Y \) be a countable dense subset. Then, \( Z_0 := \{ z \in Z \mid z = h_1 y_1, h_1 \in D_1, y_1 \in D_2 \} \) is a countable set. Clearly, \( \text{span} (Z_0) \supseteq Z \). Then, \( Z \) is separable. By Proposition 7.35, \( Z \) is separable. This completes the proof of the proposition.

**Theorem 11.207** Let \( X \) be a finite-dimensional normed linear space over \( \mathbb{K} \). Then, \( X \) is a second countable locally compact Hausdorff space.
11.10. DUAL OF \( \mathcal{C}(X, y) \) AND \( \mathcal{C}_c(X, y) \)

**Proof** Clearly, \( X \) is Hausdorff. By Propositions 11.42, \( \overline{B}_X(x_0, r) \) is compact, \( \forall x_0 \in X \) and \( \forall r \in \mathbb{R}_+ \). By Theorem 7.36, \( X \) is a Banach space. By Proposition 11.190, \( X \) is separable. Let \( D \subseteq X \) be a countable dense set. Clearly, \( \mathcal{M} := \{ B_X(x_0, r) \subseteq X \mid x_0 \in D, r \in \mathbb{Q} \cap \mathbb{R}_+ \} \) is a basis for the topology of \( X \), which is countable. Then, \( X \) is second countable. By Proposition 5.51, \( X \) is a locally compact space. Therefore, \( X \) is a second countable locally compact Hausdorff space.

**Proposition 11.208** Let \( m \in \mathbb{N} \), \( X \) be a \( m \)-dimensional Banach space over \( \mathbb{K}_1 \), \( X_c := (X, O_c) \) be the Alexandroff one-point compactification of \( X \), \( y \) be a normed linear space over \( \mathbb{K}_2 \), \( Z_c := \{ \tilde{z} \in C(X, y) \mid \tilde{z}(\omega) = \vartheta_y \} \), where \( \omega \in X_c \) is the point at infinity in \( X_c \), \( Z := \{ z \in C_b(X, y) \mid \exists \tilde{z} \in Z_c \exists z = \tilde{z}|X \} \). Then, \( Z = \{ z \in C_b(X, y) \mid \lim_{\| x \|_X \to \infty} \tilde{z}(x) = \vartheta_y \} =: \bar{Z} \), and \((Z, \mathcal{K}_2, \| \cdot \|)\) is a normed linear space over \( \mathbb{K}_2 \), where \( \| \cdot \| \) is the norm as in \( C_b(X, y) \). We will denote \((Z, \mathcal{K}_2, \| \cdot \|) := \mathcal{C}(X, y)\).

Furthermore, if \( y \) is a Banach space, then \( \mathcal{C}_c(X, y) \) is also a Banach space.

**Proof** By Theorem 11.207, \( X \) is a second countable locally compact Hausdorff space. Then, by Alexandroff One-Point Compactification Theorem 5.66, \( X_c \) is well-defined. By Example 9.71, \( C_b(X, y) \) is a normed linear space over \( \mathbb{K}_2 \). By Example 7.31 and Proposition 5.22, \( C(X_c, y) \) is a normed linear space over \( \mathbb{K}_2 \).

Fix any \( z \in Z \). \( z \in C_b(X, y) \) implies that \( \| \tilde{z} \|_{C_b(X, y)} \in \mathbb{R}_+ \). By the definition of \( Z \), \( \exists \tilde{z} \in Z_c \) such that \( \tilde{z} = \tilde{z}|X \). Then, \( \tilde{z}(\omega) = \vartheta_y \), \( \forall \vartheta \in \mathbb{R}_+ \), \( O := \tilde{z}_{\text{inv}}(B_{\vartheta} \vartheta_y, \varepsilon) \in O_c \), since \( \tilde{z} \) is continuous. Clearly, \( \omega \in O \). Then, \( X_c \setminus O \subseteq X \) is a compact set by Alexandroff One-Point Compactification Theorem 5.66. Thus, by Proposition 7.42, \( \exists M \in \mathbb{R}_+ \) such that \( X_c \setminus O \subseteq \overline{B}_X(\vartheta_X, M) \). Hence, \( \forall x \in X \) with \( \| x \|_X \geq M \), we have \( x \in O \cap X \), and then \( \| z(x) \|_y = \| \tilde{z}(x) \|_y < \varepsilon \). This shows that \( \lim_{\| x \|_X \to \infty} z(x) = \vartheta_y \). Therefore, \( z \in \bar{Z} \). By the arbitrariness of \( z \), we have \( Z \subseteq \bar{Z} \).

On the other hand, fix any \( z \in \bar{Z} \). Define \( \tilde{z} : X_c \to y \) by \( \tilde{z}(x) = \begin{cases} z(x) & x \in X \\ \vartheta_y & x = \omega \end{cases} \), \( \forall x \in X_c \). \( \forall x_0 \in X_c \), if \( x_0 \in X \), we have \( \tilde{z} \) is continuous at \( x_0 \). If \( x_0 = \omega \), then, by the definition of \( \bar{Z} \), \( \forall \vartheta \in \mathbb{R}_+ \), \( \exists M \in \mathbb{R}_+ \) such that \( \forall x \in X \) with \( \| x \|_X \geq M \), we have \( \| z(x) \|_y < \varepsilon \). Thus, let \( O := (X \setminus \overline{B}_X(\vartheta_X, M)) \cup \{ \omega \} \subseteq X_c \). By Alexandroff One-Point Compactification Theorem 5.66, \( O \in O_c \). Clearly, \( \forall x \in O \), we have \( \| \tilde{z}(x) \|_y < \varepsilon \). This proves that \( \omega \in O \subseteq \tilde{z}_{\text{inv}}(B_{\vartheta}(\vartheta_y, \varepsilon)) \). Hence, \( \tilde{z} \) is continuous at \( x_0 = \omega \). Thus, \( \tilde{z} \in Z_c \). Therefore, \( z \in \bar{Z} \). By the arbitrariness of \( z \), we have \( Z \subseteq \bar{Z} \).

Combining the above two paragraphs, we have \( Z = \bar{Z} \).

Clearly, \( \vartheta_{C_b(X, y)} \in \bar{Z} \neq \emptyset \). \( \forall z_1, z_2 \in Z \), \( \forall \vartheta \in \mathbb{K}_2 \), we have \( z_1 + z_2 \in C_b(X, y) \), since \( z_1, z_2 \in \bar{Z} \subseteq C_b(X, y) \) and \( \lim_{\| x \|_X \to \infty} (z_1(x) + z_2(x)) = \lim_{\| x \|_X \to \infty} z_1(x) + \lim_{\| x \|_X \to \infty} z_2(x) = \vartheta_y \) by the definition of \( \bar{Z} \) and \( Z = \bar{Z} \). This shows that \( z_1 + z_2 \in Z \). Furthermore, \( \alpha z_1 \in C_b(X, y) \),
and \( \lim_{\|x\| \to \infty} \alpha z_1(x) = \alpha \vartheta y = \vartheta y \). This shows that \( \alpha z_1 \in Z \). Hence, \( Z \subseteq C_b(X, Y) \) is a subspace, and therefore \( (Z, K_2, \|\cdot\|) \) is a normed linear subspace of \( C_b(X, Y) \).

Consider the case that \( Y \) is a Banach space. By Example 9.72, \( C_b(X, Y) \) is a Banach space. We will show that \( Z \) is a complete subspace of \( C_b(X, Y) \).

Then, \( (Z, K_2, \|\cdot\|) \) is a Banach subspace of \( C_b(X, Y) \).

Fix any Cauchy sequence \( (z_n)_{n=1}^\infty \subseteq Z \). Then, \( \lim_{n \in \mathbb{N}} z_n = z_0 \in C_b(X, Y) \) since \( Z \subseteq C_b(X, Y) \) and \( C_b(X, Y) \) is a Banach space. \( \forall \varepsilon \in \mathbb{R}_+ \), \( \exists N \in \mathbb{N}, \forall n, m \in \mathbb{N} \) with \( n, m \geq N \), we have \( \|z_n - z_m\|_{C_b(X, Y)} < \varepsilon /2 \).

By \( z_N \in Z \), we have \( \lim_{\|x\| \to \infty} z_N(x) = \vartheta y \). Then, \( \exists M \in \mathbb{R}_+ \), \( \forall x \in X \) with \( \|x\|_X > M \), we have \( \|z_N(x)\|_Y < \varepsilon /2 \). Then, \( \|z_0(x)\|_Y \leq \|z_0(x) - z_N(x)\|_Y + \|z_N(x)\|_Y < \|z_0 - z_N\|_{C_b(X, Y)} + \varepsilon /2 \leq \lim_{n \in \mathbb{N}} \|z_n - z_N\|_{C_b(X, Y)} + \varepsilon /2 \leq \varepsilon /2 \). Hence, \( \lim_{\|x\| \to \infty} z_0(x) = \vartheta y \) and \( z_0 \in Z \). By the arbitrariness of \( (z_n)_{n=1}^\infty \), we have \( Z \) is complete. Then, \( C_c(X, Y) \) is a Banach space.

\[ \square \]

**Theorem 11.209 (Riesz Representation Theorem)** Let \( m \in \mathbb{N} \) and \( n \in \mathbb{Z}_+ \), \( X \) be an \( m \)-dimensional Banach space over \( K_1 \), \( Y \) be an \( n \)-dimensional Banach space over \( K_2 \). Then, the following statements hold.

(i) \( (C_c(X, Y))^* = M_f(X, Y^*) = M_f(X, B_B(X), Y^*) \) isometrically isomorphically.

(ii) \( C_c(X, Y) \) is separable.

**Proof** (i) By Theorems 11.198 and 11.207, \( M_f(X, B_B(X), Y^*) = M_f(X, Y^*) \). By Proposition 11.192, \( M_f(X, Y^*) \) is a Banach space over \( K_2 \).

Let \( X_c := (X_c, O_c) \) be the Alexandroff one-point compactification of \( X \) as defined in Theorem 5.66. \( X_c \) is well-defined since \( X \) is a second countable locally compact Hausdorff space, by Theorem 11.207. By Riesz Representation Theorem 11.205, we have \( (C(X_c, Y))^* = M_f(X_c, Y^*) \). Let \( \mathcal{F} := \{ \emptyset, \{ \omega \} \} \), where \( \omega \in X_c \) is the point at infinity. Define \( \mu_1 : \mathcal{F} \to Y^* \) by \( \mu_1(\emptyset) = \vartheta y = \mu_1(\{ \omega \}) \). It is easy to see that \( \mu_1 \) is a finite \( Y^* \)-valued measure on the measurable space \( (\{ \omega \}, \mathcal{F}) \). Now, we need the following intermediate result.

**Claim 11.209.1** \( \forall \mu \in M_f(X, Y^*) \), we may define a unique finite \( Y^* \)-valued measure space \( (X_c, B_B(X_c), \mu_1) \) such that \( (X_c, B_B(X_c), \mu) \) and \( (\{ \omega \}, \mathcal{F}, \mu_1) \) are finite \( Y^* \)-valued measure subspaces of \( (X_c, B_B(X_c), \mu) \). Furthermore, \( \mu_1 \in M_f(X_c, Y^*) \).

**Proof of claim:** By Proposition 11.118, there exists a unique \( \sigma \)-finite \( Y^* \)-valued measure space \( (X_c, B_B(X_c), \mu) \) such that \( (X, B_B(X), \mu) \) and \( (\{ \omega \}, \mathcal{F}, \mu_1) \) are finite \( Y^* \)-valued measure subspaces of \( (X_c, B_B(X_c), \mu) \). This implies that \( \mathcal{P} \circ \mu_c(X_c) = \mathcal{P} \circ \mu_c(X) + \mathcal{P} \circ \mu_c(\{ \omega \}) = \mathcal{P} \circ \mu(X) + \mathcal{P} \circ \mu_1(X_c) \).
11.10. **Dual of \( C(X, \mathcal{Y}) \) and \( C_c(X, \mathcal{Y}) \)**

\[ \mu_1(\{\omega\}) = \mathcal{P} \circ \mu(\mathcal{X}) + 0 = \mathcal{P} \circ \mu(\mathcal{X}) < \infty, \]
where the first equality follows from the fact that \( \mathcal{P} \circ \mu \) is a measure; the second equality follows from Proposition 11.115; the third equality follows from the definition of \( \mu_1 \); and the inequality follows from \( \mu \) is finite. Hence, \( \tilde{\mu} \) is finite.

Thus, \( \tilde{\mu} \in M_f(\mathcal{X}, \mathcal{B}_B(\mathcal{X}^r), \mathcal{Y}^r) \). By Alexandroff One-Point Compactification Theorem 5.66, \( \mathcal{X}^r \) is a compact Hausdorff space. By Proposition 5.14, \( \mathcal{X}^r \) is normal. Clearly, \( \mathcal{X}^r \) is second countable. By Urysohn Metrization Theorem 4.53, \( \mathcal{X}^r \) is metrizable. Hence, \( \mathcal{X}^r \) is a compact Hausdorff metrizable topological space satisfying the second axiom of countability. By Theorem 11.198, \( M_f(\mathcal{X}^r, \mathcal{B}_B(\mathcal{X}^r), \mathcal{Y}^r) = M_{f_t}(\mathcal{X}^r, \mathcal{Y}^r) \). This completes the proof of the claim.

Fix any \( \mu \in M_{f_t}(\mathcal{X}, \mathcal{Y}^r) \). By Claim 11.209.1, let \( \tilde{\mu} \) be the finite \( \mathcal{Y}^r \)-valued measure on the measurable space \((\mathcal{X}^r, \mathcal{B}_B(\mathcal{X}^r))\) such that \((\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)\) and \((\{\omega\}, \mathcal{F}, \mu_1)\) are its finite \( \mathcal{Y}^r \)-valued measure subspaces. Then, \( \tilde{\mu} \in M_f(\mathcal{X}^r, \mathcal{Y}^r) = (C(\mathcal{X}^r, \mathcal{Y}))^r \). Thus, by Riesz Representation Theorem 11.205, we may define \( f : C(\mathcal{X}^r, \mathcal{Y}) \rightarrow K_2 \) by \( f(\tilde{z}) = f(x) = \int_{\mathcal{X}^r} \langle \langle d\mu(x), \tilde{z}(x) \rangle \rangle, \forall \tilde{z} \in C(\mathcal{X}^r, \mathcal{Y}). \) Then, \( f \) is a bounded linear functional on \( C(\mathcal{X}^r, \mathcal{Y}) \), and

\[
\left\| f \right\|_{(C(\mathcal{X}^r, \mathcal{Y}))^r} = \left\| \tilde{\mu} \right\|_{M_{f_t}(\mathcal{X}, \mathcal{Y}^r)}.
\]

This implies that \( f(\tilde{z}) = f(x) = \int_{\mathcal{X}^r} \langle \langle d\mu(x), \tilde{z}(x) \rangle \rangle = \int_{\mathcal{X}^r} \langle \langle d\mu(x), z(x) \rangle \rangle, \forall z \in C_c(\mathcal{X}, \mathcal{Y}) \) and \( z \in \mathcal{Z}_c \) is such that \( z = \tilde{z} |_{\mathcal{X}^r} \). Thus, \( f \) is a bounded linear functional on \( C(\mathcal{X}^r, \mathcal{Y}) \), and therefore on \( \mathcal{Z}_c \), which is a subspace of \( C(\mathcal{X}^r, \mathcal{Y}) \). Hence, \( \mu \) defines a bounded linear functional on \( C_c(\mathcal{X}, \mathcal{Y}) \) by \( f := \Phi(\mu) \), where \( f(z) = \int_{\mathcal{X}} \langle \langle d\mu(x), z(x) \rangle \rangle, \forall z \in C_c(\mathcal{X}, \mathcal{Y}). \) Therefore, \( \left\| f \right\|_{(C_c(\mathcal{X}, \mathcal{Y}))^r} = \sup_{z \in C_c(\mathcal{X}, \mathcal{Y})} \left\| \int_{\mathcal{X}} \langle \langle d\mu(x), z(x) \rangle \rangle \right\| \leq \sup_{z \in C_c(\mathcal{X}, \mathcal{Y})} \int_{\mathcal{X}} \| z(x) \| d\mathcal{P} \circ \mu(\mathcal{X}) \leq \mathcal{P} \circ \mu(\mathcal{X}) = \left\| \mu \right\|_{M_{f_t}(\mathcal{X}, \mathcal{Y}^r)} < \infty, \)

where the first equality follows from Proposition 7.63; the first inequality follows from Proposition 11.130; the second equality follows from Proposition 11.83; and the second equality follows from Proposition 11.142. Thus, we may define \( \Phi : M_{f_t}(\mathcal{X}, \mathcal{Y}^r) \rightarrow (C_c(\mathcal{X}, \mathcal{Y}))^r \) by \( f = \Phi(\mu), \forall \mu \in M_{f_t}(\mathcal{X}, \mathcal{Y}^r). \) Clearly, \( \Phi \) is a bounded linear operator.

On the other hand, fix any \( f \in (C_c(\mathcal{X}, \mathcal{Y}))^r \). We may define a functional \( \tilde{f} : \mathcal{Z}_c \rightarrow K_2 \) as follows. \( \forall z \in \mathcal{Z}_c \), let \( z : \mathcal{X} \rightarrow \mathcal{Y} \) be defined by \( z := \vec{z} |_{\mathcal{X}^r} \). Then, \( z \in C(\mathcal{X}, \mathcal{Y}) \). We define \( \tilde{f}(\vec{z}) := f(z) \).

It is easy to show that \( \tilde{f} \) is a bounded linear functional on \( \mathcal{Z}_c \) with \( \left\| \tilde{f} \right\|_{\mathcal{Z}_c^*} = \left\| f \right\|_{(C_c(\mathcal{X}, \mathcal{Y}))^r}. \) Clearly, \( \mathcal{Z}_c \) is a subspace of \( C(\mathcal{X}, \mathcal{Y}) \). By Hahn-Banach Theorem 7.83, \( \exists \tilde{f} \in (C(\mathcal{X}, \mathcal{Y}))^* = M_{f_t}(\mathcal{X}, \mathcal{Y}^r) \) such that
Proposition 11.125; and the last equality follows from Proposition 11.130; the second equality follows from Proposition 11.130; the second equality follows from Proposition 11.130.

This implies that \( \mu \) is bijective. By the preceding two paragraphs, if \( \Phi \) is not injective. Then, \( \Phi \) is linear. Then, \( \Phi \) is an isometrical isomorphism. This proves that \( \Phi(\mu) = \Phi(\mu) \) and \( \mu \neq \mu \). By Claim 11.209.1, we let \( \hat{\mu} \) be the finite \( y^* \)-valued measure on \( (X, B(X), \mu) \) and \( (w \in \mathcal{C}(X, y)^* \) and \( \mu \in M_f(\mathcal{C}(X, y)^*) \) such that \( \Phi(\mu) \) and \( \Phi(\mu) = \Phi(\mu) \) and \( \mu \neq \mu \). Let \( \hat{\mu} := \hat{\mu} - \mu \in M_f(\mathcal{C}(X, y)^*) \). Then, by linearity of \( \Phi \), \( \hat{\mu} \in M_f(\mathcal{C}(X, y)^*) \) and \( \mu \in M_f(\mathcal{C}(X, y)^*) \). Hence, \( \Phi \) is an isometry. By the expression for \( \Phi \), we can easily see that \( \Phi \) is linear. Then, \( \Phi \) is an isometrical isomorphism. This proves that \( M_f(\mathcal{C}(X, y)^*) = (\mathcal{C}(X, y)^*)^* \).

We will show that \( \Phi \) is injective by an argument of contradiction. Suppose \( \Phi \) is not injective. Then, \( \exists \mu, \hat{\mu} \in M_f(\mathcal{C}(X, y)^*) \) such that \( \Phi(\mu) = \Phi(\hat{\mu}) \) and \( \mu \neq \hat{\mu} \). Thus, \( \exists E \in B(X) \) such that \( \mu(E) \neq \hat{\mu}(E) \). Let \( \hat{\mu} := \hat{\mu} - \mu \in M_f(\mathcal{C}(X, y)^*) \). Then, by linearity of \( \Phi \), \( \Phi(\mu) = \Phi(\mu) \), and \( \mu \neq \hat{\mu}(E) \neq \hat{\mu}(E) \). By Claim 11.209.1, we let \( \hat{\mu} \) be the finite \( y^* \)-valued measure on \( (X, B(X), \mu) \) and \( \{w \} \), \( \mathcal{F}, \mu_1 \) are its finite \( y^* \)-valued measure subspaces. Therefore, \( \hat{\mu} \in M_f(\mathcal{C}(X, y)^*) \). Since \( \mu(E) \neq \mathcal{F} \), then \( \mu(E) = \hat{\mu}(E) \neq \mathcal{F} \), and \( \hat{\mu} \neq \mu \). This implies that \( \exists \hat{\mu} \in \mathcal{C}(X, y)^* \) such that \( \int_X (\hat{\mu}(x), \mathcal{F}(x)) = 0 \). Note that \( \int_X (\hat{\mu}(x), \mathcal{F}(x)) = \int_X (\hat{\mu}(x), \mathcal{F}(x)) = \int_X (\hat{\mu}(x), \mathcal{F}(x)) = \int_X (\hat{\mu}(x), \mathcal{F}(x)) = 0 \). By Theorem 11.207, \( \mathcal{C}(X, y)^* \)
is a second countable locally compact Hausdorff space. \( \forall n \in \mathbb{N}, \) by Proposition 5.62, there exists \( h_n : X \to [0, 1] \subseteq \mathbb{R} \) such that \( h_n(x) = 1, \forall x \in E_{n,1} := B_X (\varnothing_X, n) \) and \( h_n(x) = 0, \forall x \in E_{n,2} := B_X (\varnothing_X, n + 1) \). Clearly, \( \lim_{n \in \mathbb{N}} z(x)h_n(x) = z(x), \forall x \in X. \) By Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_X \langle (d\mu(x), z(x)) \rangle = \lim_{n \in \mathbb{N}} \int_X \langle (d\bar{\mu}(x), z(x)h_n(x)) \rangle \neq 0. \) Then, \( \exists n \in \mathbb{N} \) such that \( \Phi(\bar{\mu}) (zh_n) = \int_X \langle (d\bar{\mu}(x), z(x)h_n(x)) \rangle \neq 0, \) since \( zh_n \in C_c(X, y) \). This is a contradiction to \( \Phi(\bar{\mu}) = \varnothing(C_c(X, y)), \). This shows that hypothesis is impossible. Hence, \( \Phi \) is injective.

(ii) Let \( K_n := \overline{B}_X (\varnothing_X, n) \subseteq X, \forall n \in \mathbb{N}. \) Then, by Proposition 7.42, we have \( K_n \subseteq K_{n+1} \subseteq X \) are compact. By Proposition 5.62 and Theorem 11.207, \( \exists h_n : X \to [0, 1] \subseteq \mathbb{R} \) such that \( h_n(x) = 1, \forall x \in K_n; h_n(x) = 0, \forall x \in X \setminus K_n^c, \forall n \in \mathbb{N}. \) Fix any \( n \in \mathbb{N}. \) By Proposition 11.190, \( C(K_n, y) \) is separable. Let \( D_n \subseteq C(K_n, y) \) be a countable dense subset. Define \( \bar{D}_n := C_c(X, y) \) by \( \bar{D}_n := \{ f : X \to y \mid \exists g \in D_n \exists f(x) = \begin{cases} h_n(x)g(x) & x \in K_n \\ h_n(x)g \left( \frac{N_x}{\|x\|_X} \right) & x \in X \setminus K_n \end{cases}. \) Clearly, \( \bar{D}_n \) is countable. Define \( D := \bigcup_{n \in \mathbb{N}} \bar{D}_n \subseteq C_c(X, y). \) We will show that \( D \) is dense in \( C_c(X, y). \)

Fix any \( z \in C_c(X, y). \) Fix any \( \epsilon \in \mathbb{R}_+. \) Then, \( \exists N \in \mathbb{N} \) such that \( \|z(x)\|_y < \frac{\epsilon}{3}, \forall x \in X \setminus K_N^c. \) Since \( z|_{K_N} \in C(K_n, y). \) Then, \( \exists y \in D_N \) such that \( \|z|_{K_N} - g\|_{C(K_n, y)} < \frac{\epsilon}{3}. \) Let \( f \in D_n \subseteq D \) be such that \( f(x) = \begin{cases} h_N(x)g(x) & x \in K_N \\ h_N(x)g \left( \frac{N_x}{\|x\|_X} \right) & x \in X \setminus K_N \end{cases}. \) Then, \( \forall x \in K_N, \)

\[
\|f(x) - z(x)\|_y = \|g(x) - z(x)\|_y \leq \|z|_{K_N} - g\|_{C(K_n, y)} < \frac{\epsilon}{3}, \forall x \in X \setminus K_N, \|f(x) - z(x)\|_y \leq \|f(x)\|_y + \|z(x)\|_y = \|h_n(x)g \left( \frac{N_x}{\|x\|_X} \right)\|_y + \frac{\epsilon}{3} \leq \left( \frac{\epsilon}{3} \right) + \frac{\epsilon}{3} \leq \frac{\epsilon}{3} < \epsilon. \]

Hence, \( \|f - z\|_{C_c(X, y)} \leq \epsilon. \) This shows that \( D \) is dense in \( C_c(X, y). \) Therefore, \( C_c(X, y) \) is separable.

This completes the proof of the theorem. \( \square \)

\( C_c(\mathbb{R}^m, \mathbb{R}) \) is the continuous domain analogy of \( c_0. \)

**Definition 11.210** Let \( m \in \mathbb{N}, n \in \mathbb{Z}_+, X \) be an \( m \)-dimensional Banach space over \( \mathbb{K}_1, \) \( Y \) be an \( n \)-dimensional Banach space over \( \mathbb{K}_2, \) and \( (\mu_n)_{n=1}^\infty \subseteq M_f(X, y^*) = M_f(X, B(B(X), y^*). \) We will say that \( (\mu_n)_{n=1}^\infty \) converges to \( \mu_0 \in M_f(X, y^*) \) in distribution as \( n \to \infty, \) if \( \lim_{n \in \mathbb{N}} \mu_n = \mu_0 \) weak* (See Proposition 7.120 for details). In this case, we will denote \( \mu_n \Rightarrow \mu_0 \) as \( n \to \infty. \)
Chapter 12

Differentiation and Integration

12.1 Carathéodory Extension Theorem

Definition 12.1 Let $X$ be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on $X$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\nu : \mathcal{A} \to [0, \infty] \subseteq \mathbb{R}$ be a measure on the algebra $\mathcal{A}$. A $\mathcal{Y}$-valued measure on the algebra $\mathcal{A}$ is a function $\mu$ from $\mathcal{A}$ to $\mathcal{Y}$ that satisfies

(i) $\mu(\emptyset) = 0$;

(ii) $\forall A \in \mathcal{A}$ with $\nu(A) = \infty$, $\mu(A)$ is undefined;

(iii) $\forall A \in \mathcal{A}$ with $\nu(A) < \infty$, $\mu(A) \in \mathcal{Y}$ and all pairwise disjoint $(A_i)_{i=1}^\infty \subseteq \mathcal{A}$ with $A = \bigcup_{i=1}^\infty A_i$, we have $\sum_{i=1}^\infty \|\mu(A_i)\| < +\infty$ and $\mu(A) = \sum_{i=1}^\infty \mu(A_i)$;

(iv) $\forall A \in \mathcal{A}$ with $\nu(A) < \infty$, we have

$$\nu(A) = \sup_{n \in \mathbb{Z}^+, (A_i)_{i=1}^n \subseteq \mathcal{A}, A = \bigcup_{i=1}^n A_i, A_i \cap A_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(A_i)\|$$

Then, $\nu$ is said to be the total variation of $\mu$ and denoted by $\mathcal{P} \circ \mu$. $\mu$ is said to be finite if $\nu$ is finite; and $\mu$ is said to be $\sigma$-finite if $\nu$ is $\sigma$-finite.

Proposition 12.2 Let $X$ be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on $X$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\mu$ be a $\mathcal{Y}$-valued measure on the algebra $\mathcal{A}$, and $E \in \mathcal{A}$. Then, $\mathcal{A}_E := \{ C \in \mathcal{A} \mid C \subseteq E \}$ is an algebra on $E$, $\mu_E := \mu|_{\mathcal{A}_E}$ is a $\mathcal{Y}$-valued measure on the algebra $\mathcal{A}_E$, and $\mathcal{P} \circ \mu_E = (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$. 

543
Proof It is easy to show that $\mathcal{A}_E$ is an algebra on $E$. Note that $\mathcal{P} \circ \mu$ is a measure on the algebra $\mathcal{A}$. Define $\nu_E := (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$. Then, $\nu_E(\emptyset) = \mathcal{P} \circ \mu(\emptyset) = 0$. \forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}_E$ with $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_E$, we have $\nu_E(A) = \mathcal{P} \circ \mu(A) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) = \sum_{i=1}^{\infty} \nu_E(A_i)$. Then, $\nu_E$ is a measure on the algebra $\mathcal{A}_E$.

$\mu_E(\emptyset) = \mu(\emptyset) = \emptyset_Y$. \forall $A \in \mathcal{A}_E$ with $\nu_E(A) = \emptyset$, we have $\mathcal{P} \circ \mu(A) = \infty$, $\mu(A)$ is undefined, and therefore, $\mu_E(A)$ is undefined. \forall $A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\mathcal{P} \circ \mu(A) < \infty$ and $\mu(A) \in \emptyset_Y$. \forall pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}_E$ with $A = \bigcup_{i=1}^{\infty} A_i$, we have $\sum_{i=1}^{\infty} \|\mu_E(A_i)\| = \sum_{i=1}^{\infty} \|\mu(A_i)\| < \infty$ and $\sum_{i=1}^{\infty} \mu_E(A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \mu(A) = \mu_E(A)$. \forall $A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\nu_E(A) = \mathcal{P} \circ \mu(A) = \sup_{n \in \mathbb{Z}_+} (A_i)_{i=1}^{n} \subseteq \mathcal{A}_E, A = \bigcup_{i=1}^{n} A_i, A \cap A_j = \emptyset, \forall 1 \leq i < j \leq n \sum_{i=1}^{n} \|\mu(A_i)\| = \|\mu_E(A)\|$. Hence, $\mu_E$ is a $\emptyset_Y$-valued measure on the algebra $\mathcal{A}_E$ with $\mathcal{P} \circ \mu_E = \mathcal{P} \circ (\mu|_{\mathcal{A}_E}) = \nu_E = (\mathcal{P} \circ \mu)|_{\mathcal{A}_E}$.

This completes the proof of the proposition. $\square$

Proposition 12.3 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space. Define $\rho : \mathcal{B} \times \mathcal{B} \to [0, \infty) \subset \mathbb{R}$ by $\rho(E_1, E_2) = \mu(E_1 \triangle E_2)$, \forall $E_1, E_2 \in \mathcal{B}$. Then, $\rho$ defines a pseudo-metric on $\mathcal{B}$.

Proof \forall $E_1, E_2, E_3 \in \mathcal{B}$, $\rho(E_1, E_2) = \mu(E_1 \triangle E_2) \in [0, +\infty) \subset \mathbb{R}$ since $\mu$ is a finite measure on $(X, \mathcal{B})$. Clearly, $\rho(E_1, E_1) = \mu(\emptyset) = 0$ and $\rho(E_1, E_2) = \mu(E_1 \triangle E_2) = \mu(E_2 \triangle E_1) = \rho(E_2, E_1)$. Note that $\rho(E_1, E_2) + \rho(E_2, E_3) = \mu(E_1 \triangle E_2) + \mu(E_2 \triangle E_3) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1) + \mu(E_2 \setminus E_3) + \mu(E_3 \setminus E_2) + \mu(E_1 \setminus (E_2 \cup E_3)) + \mu((E_1 \setminus E_2) \cup (E_2 \setminus E_3)) \geq \mu(E_1 \setminus E_3) + \mu(E_3 \setminus E_1) = \mu(E_1 \triangle E_3) = \rho(E_1, E_3)$. Hence, $\rho$ defines a pseudo-metric on $\mathcal{B}$. This completes the proof of the proposition. $\square$

Theorem 12.4 (Carathéodory Extension Theorem) Let $\mathcal{A}$ be an algebra on a set $X$, $\emptyset_Y$ be a Banach space over $\mathbb{K}$, $\mu$ be a $\sigma$-finite $\emptyset_Y$-valued measure on the algebra $\mathcal{A}$, and $\mathcal{B}$ be the $\sigma$-algebra on $X$ generated by $\mathcal{A}$. Then, there is a unique $\emptyset_Y$-valued measure $\tilde{\mu}$ on $(X, \mathcal{B})$ such that $\tilde{\mu}|_{\mathcal{A}} = \mu$ and $(\mathcal{P} \circ \tilde{\mu})|_{\mathcal{A}} = \mathcal{P} \circ \mu$. Furthermore, $\tilde{\mu}$ is $\sigma$-finite.

Proof We first consider the special case where $\mu$ is finite. Then, $\mu : \mathcal{A} \to \emptyset_Y$ and $\mathcal{P} \circ \mu : \mathcal{A} \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$. By Carathéodory Extension Theorem 11.19, there is a unique finite measure $\nu : \mathcal{B} \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$ on the measurable space $(X, \mathcal{B})$ such that $\nu(A) = \mathcal{P} \circ \mu(A)$, $\forall A \in \mathcal{A}$, and $\nu(E) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, \bigcup_{i=1}^{\infty} A_i \supseteq E} \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i)$, $\forall E \in \mathcal{B}$. Define $\rho : \mathcal{B} \times \mathcal{B} \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$ by $\rho(E_1, E_2) = \nu(E_1 \triangle E_2)$, $\forall E_1, E_2 \in \mathcal{B}$. By Proposition 12.3, $\rho$ defines a pseudo-metric on $\mathcal{B}$. Define an equivalence relation $\equiv$ on $\mathcal{B}$ by $E_1 \equiv E_2$ if $\rho(E_1, E_2) = \nu(E_1 \triangle E_2) = 0$. By Lemma 4.48, $(\mathcal{B}/ \equiv, \tilde{\rho})$ is a metric space, where $\tilde{\rho} : (\mathcal{B}/ \equiv) \times (\mathcal{B}/ \equiv) \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$
is defined by \( \bar{\rho}([E_1], [E_2]) = \rho(E_1, E_2), \forall [E_1], [E_2] \in \mathcal{B}/ \equiv \) with \( E_1, E_2 \in \mathcal{B} \). Let \( \mathcal{A}_\equiv := \{ [A] \in \mathcal{B}/ \equiv \mid A \in \mathcal{A} \} \).

Claim 12.4.1 \( \mathcal{A}_\equiv \) is dense in \( \mathcal{B}/ \equiv \), that is \( \overline{\mathcal{A}_\equiv} = \mathcal{B}/ \equiv \).

**Proof of claim:** \( \forall [E] \in \mathcal{B}/ \equiv \) with \( E \in \mathcal{B}, \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists (A_i)_{i=1}^\infty \subseteq \mathcal{A} \) with \( E \subseteq \bigcup_{i=1}^\infty A_i \) such that \( \nu(E) \leq \sum_{i=1}^\infty \mathcal{P} \circ \mu(A_i) < \nu(E) + \epsilon/2 < +\infty \). Then, \( \exists n_0 \in \mathbb{N} \) such that \( \sum_{i=n_0}^\infty \mathcal{P} \circ \mu(A_i) < \epsilon/2 \). Let \( A := \bigcup_{i=1}^{n_0} A_i \). Note that \( E \setminus A \subseteq \bigcup_{i=1}^\infty A_i \setminus A \subseteq \bigcup_{i=n_0+1}^\infty A_i \). This implies that \( 0 \leq \nu(E \setminus A) \leq \nu(\bigcup_{i=n_0+1}^\infty A_i) \leq \sum_{i=n_0+1}^\infty \nu(A_i) = \sum_{i=n_0+1}^\infty \mathcal{P} \circ \mu(A_i) < \epsilon/2 \). Note also that \( A \setminus E \subseteq \bigcup_{i=1}^\infty A_i \setminus E \). This further implies that \( 0 \leq \nu(A \setminus E) \leq \nu(\bigcup_{i=1}^\infty A_i \setminus E) = \nu(\bigcup_{i=n_0+1}^\infty A_i \setminus E) \leq \sum_{i=n_0+1}^\infty \nu(A_i) - \nu(E) = \sum_{i=n_0+1}^\infty \mathcal{P} \circ \mu(A_i) - \nu(E) < \epsilon/2 \). Hence, \( \bar{\rho}([E], [A]) = \rho(E, A) = \nu(E \cap A) = \nu(E \setminus A) + \nu(A \setminus E) < \epsilon \). Hence, \( \overline{\mathcal{A}_\equiv} = \mathcal{B}/ \equiv \). This completes the proof of the claim.

Define \( g : \mathcal{A}_\equiv \to \mathcal{Y} \) by \( g([A]) = \mu(A), \forall [A] \in \mathcal{A}_\equiv \) with \( A \in \mathcal{A} \). We will show that \( g \) is well defined. \( \forall A_1, A_2 \in \mathcal{A} \) with \( A_1 \equiv A_2 \), we have \( \rho(A_1, A_2) = \nu(A_1 \triangle A_2) = \mathcal{P} \circ \mu(A_1 \triangle A_2) = 0 \). Then, \( \mathcal{P} \circ \mu(A_1 \setminus A_2) = 0 = \mathcal{P} \circ \mu(A_2 \setminus A_1) \). This implies that \( \mu(A_1 \setminus A_2) = \nu_y = \mu(A_2 \setminus A_1) \) and \( \mu(A_1) = \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) = \mu(A_1 \cap A_2) + \mu(A_2 \setminus A_1) \). Hence, \( g \) is well defined.

Claim 12.4.2 \( g \) is uniformly continuous.

**Proof of claim:** \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) let \( \delta = \epsilon \). \( \forall [A_1], [A_2] \in \mathcal{A}_\equiv \) with \( A_1, A_2 \in \mathcal{A} \) and \( \bar{\rho}([A_1], [A_2]) < \delta \), we have \( \| g([A_1]) - g([A_2]) \| = \| \mu(A_1) - \mu(A_2) \| = \| \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) - \mu(A_2 \cap A_1) - \mu(A_2 \setminus A_1) \| = \| \mu(A_1 \setminus A_2) + \mu(A_1 \setminus A_1) \| \leq \| \mu(A_1 \setminus A_2) \| + \| \mu(A_2 \setminus A_1) \| \leq \mathcal{P} \circ \mu(A_1 \setminus A_2) + \mathcal{P} \circ \mu(A_2 \setminus A_1) = \nu(A_1 \setminus A_2) = \nu(A_1 \setminus A_2) = \bar{\rho}([A_1], [A_2]) < \delta = \epsilon \). Hence, \( g \) is uniformly continuous. This completes the proof of the claim.

By Proposition 4.46 and the completeness of \( \mathcal{Y} \), there exists a unique continuous function \( \tilde{g} : \mathcal{B}/ \equiv \to \mathcal{Y} \) such that \( \tilde{g}|_{\mathcal{A}_\equiv} = g \), and \( \tilde{g} \) is uniformly continuous. Define \( \tilde{\mu} : \mathcal{B} \to \mathcal{Y} \) by \( \tilde{\mu}(E) = \tilde{g}(\nu(E)), \forall E \in \mathcal{B} \). We will show that \( \tilde{\mu} \) is the \( \mathcal{Y} \)-valued measure we seek. Clearly, \( \tilde{\mu}(A) = \tilde{g}(\mu(A)) = \mu(A), \forall A \in \mathcal{A} \). Then, \( \tilde{\mu}(\emptyset) = \mu(\emptyset) = \nu_y, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by the uniform continuity of \( \bar{\rho} \). \( \exists \delta(\epsilon) \in (0, \epsilon) \subset \mathbb{R} \) such that, \( \forall [E_1], [E_2] \in \mathcal{B}/ \equiv \) with \( E_1, E_2 \in \mathcal{B} \) and \( \bar{\rho}([E_1], [E_2]) = \rho(E_1, E_2) = \nu(E_1 \cap E_2) < \delta(\epsilon) \), we have \( \| \tilde{\mu}(E_1) - \tilde{\mu}(E_2) \| = \| \tilde{g}(\nu(E_1)) - \tilde{g}(\nu(E_2)) \| < \epsilon \).

Claim 12.4.3 \( \tilde{\mu} \) is finitely additive, that is, \( \tilde{\mu}(E_1 \cup E_2) = \tilde{\mu}(E_1) + \tilde{\mu}(E_2), \forall E_1, E_2 \in \mathcal{B} \) with \( E_1 \cap E_2 = \emptyset \).

**Proof of claim:** \( \forall E_1, E_2 \in \mathcal{B} \) with \( E_1 \cap E_2 = \emptyset \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) by Claim 12.4.1, \( \exists A_1, A_2 \in \mathcal{A} \) such that \( \bar{\rho}([A_1], [E_1]) = \rho(A_1, E_1) = \nu(A_1 \triangle E_1) < \delta(\epsilon)/2, i = 1, 2 \). Then, \( \| \tilde{\mu}(E_i) - \tilde{\mu}(A_i) \| < \epsilon/2, i = 1, 2 \). Note that
Claim 12.4.4. For all $E \in B$, $\| \tilde{\mu}(E) \| \leq \nu(E)$.

Proof of claim: Let $E \in B$. We can find $A \subseteq \mathbb{R}$ such that $\tilde{\mu}(E \cap \mathbb{R}) = \nu(E \cap \mathbb{R}) < \delta / 2$. Then, $\| \tilde{\mu}(E) - \tilde{\mu}(A) \| < \epsilon / 2$. Note that $\| \tilde{\mu}(A) \| = \| \mathcal{P} \circ \mu(A) \| = \nu(E \cap \mathbb{R})$. Therefore, $\| \tilde{\mu}(E) \| \leq \| \nu(E) \|$.

By the arbitrariness of $\epsilon$, we have $\| \tilde{\mu}(E) \| \leq \nu(E)$. This completes the proof of the claim. \qed
and $\tilde{\mu} : B \to Y$. Define $\tilde{g} : B/ \equiv \to Y$ by $\tilde{g}([E]) = \tilde{\mu}(E)$, $\forall [E] \in B/ \equiv$ with $E \in B$. We need to show that $\tilde{g}$ is well defined. $\forall E_1, E_2 \in B$ with $E_1 \equiv E_2$, we have $\nu(E_1 \triangle E_2) = 0$. This leads to $\nu(E_1 \setminus E_2) = 0 = \nu(E_2 \setminus E_1)$ and $\tilde{\mu}(E_1 \setminus E_2) = \tilde{\mu}(E_2 \setminus E_1)$. Then, $\tilde{\mu}(E_1) = \tilde{\mu}(E_1 \cap E_2) + \tilde{\mu}(E_1 \setminus E_2) = \tilde{\mu}(E_1 \cap E_2) = \tilde{\mu}(E_2 \setminus E_1) = \tilde{\mu}(E_2)$.

Hence, $\tilde{g}$ is well defined. Clearly, $\tilde{g}([A]) = \tilde{\mu}(A) = \mu(A) = g([A])$, $\forall [A] \in A_\equiv$ with $A \in A$. Hence, $\tilde{g}|_{A_\equiv}$ is continuous. Let $\delta = \varepsilon \in (0, \infty) \subset \mathbb{R}$. Then, $\forall [E_1], [E_2] \in B/ \equiv$ with $E_1, E_2 \in B$ and $\tilde{\mu}([E_1], [E_2]) = \nu(E_1 \triangle E_2) < \delta$, we have $\|\tilde{g}([E_1]) - \tilde{g}([E_2])\| = \|\tilde{\mu}(E_1) - \tilde{\mu}(E_2)\| = \|\tilde{\mu}(E_1 \cap E_2) + \tilde{\mu}(E_1 \setminus E_2) - \tilde{\mu}(E_2 \cap E_1) - \tilde{\mu}(E_2 \setminus E_1)\| \leq \|\tilde{\mu}(E_1 \setminus E_2)\| + \|\tilde{\mu}(E_2 \setminus E_1)\| \leq \nu\circ \tilde{\mu}(E_1 \setminus E_2) + \nu\circ \tilde{\mu}(E_2 \setminus E_1) = \nu\circ \tilde{\mu}(E_1 \setminus E_2) = \nu(E_1 \triangle E_2) < \delta = \varepsilon$. Hence, $\tilde{g}$ is uniformly continuous. By Proposition 4.46, $\tilde{g} = \tilde{g}$. Then, $\tilde{\mu}(E) = \tilde{g}([E]) = \tilde{g}([E]) = \tilde{\mu}(E)$, $\forall E \in B$. Hence, $\tilde{\mu} = \tilde{\mu}$. This establishes the uniqueness and completes the proof for the special case.

Next, we consider the general case where $\mu$ is $\sigma$-finite. Then, $\exists(X_n)_{n=1}^{\infty} \subseteq A$ such that $X = \bigcup_{n=1}^{\infty} X_n$ and $\nu\circ \mu(X_n) < \infty$, $\forall n \in \mathbb{N}$. Without loss of generality, we may assume that $(X_n)_{n=1}^{\infty}$ is pairwise disjoint. Fix any $n \in \mathbb{N}$. Let $A_n := \{ E \in A \mid E \subseteq X_n \}$ and $\mu_n := \mu|_{A_n}$. By Proposition 12.2, $\mu_n$ is a finite $\mathbb{R}$-valued measure on the algebra $A_n$ and $\nu\circ \mu_n = (\nu\circ \mu)|_{A_n}$. By the special case, there exists a unique $\mathbb{R}$-valued measure $\mu_n$ on $(X_n, \mathcal{B}_n)$, where $\mathcal{B}_n$ is the $\sigma$-algebra generated by $A_n$, such that $\mu_n(A) = \mu_n(A)$ and $\nu\circ \mu_n(A) = \nu\circ \mu_n(A)$, $\forall A \in A_n$. Furthermore, $\mu_n$ is finite. By Proposition 11.118, the generation process on $(X_n := (X_n, \mathcal{B}_n, \mu_n))_{n=1}^{\infty}$ yields a unique $\sigma$-finite $\mathbb{R}$-valued measure space $\mathcal{X} := (X, \mathcal{B}, \mu)$ on $X$, where $\mathcal{B} := \{ E \subseteq X \mid E \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N} \}$, $\forall E \in \mathcal{B}$, $\nu\circ \mu(E) = \sum_{n=1}^{\infty} \nu\circ \mu_n(E \cap X_n)$, $\mu(E)$ is undefined if $\nu\circ \mu(E) = \infty$, $\mu(E) = \sum_{n=1}^{\infty} \mu_n(E \cap X_n) \in Y$ if $\nu\circ \mu(E) < \infty$, and $X_n$ is the finite $\mathbb{R}$-valued measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$.

Claim 12.4.5 $B = \overline{B}$.

Proof of claim: $\forall A \in A$, $\forall n \in \mathbb{N}$, $A \cap X_n \subseteq A_n \subseteq B_n$. This implies that $A \subseteq \overline{B}$. Then, $A \subseteq B$. Since $B$ is a $\sigma$-algebra, then $B \subseteq B$. On the other hand, $\forall E \in B$, $\forall n \in \mathbb{N}$, we have $E \cap X_n \subseteq B_n \subseteq B$. Then, $E = \bigcup_{n=1}^{\infty} (E \cap X_n) \subseteq B$ and, hence, $B \subseteq B$. This shows that $B = B$ and completes the proof of the claim.

$\forall A \in A$, $\forall n \in \mathbb{N}$, we have $A \cap X_n \subseteq A_n \subseteq B_n$ and $\nu\circ \mu_n(A) = \sum_{n=1}^{\infty} \nu\circ \mu_n(A \cap X_n)$, $\nu\circ \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \nu\circ \mu_n(A \cap X_n) = \nu\circ \mu_n(A \cap X_n)$. Hence, $(\nu\circ \mu)|_{A} = \nu\circ \mu$. $\forall A \in A$ with $\nu\circ \mu(A) = \infty$, then $\nu\circ \mu(A) = \infty$ and $\nu\circ \mu(A) = \sum_{n=1}^{\infty} \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \mu_n(A \cap X_n) = \mu(A \cap X_n) = \mu(A) \in Y$. Hence, $\mu\circ \mu|_{A} = \mu$. This shows that $\mu$ is the $\sigma$-finite $\mathbb{R}$-valued measure on $(X, B)$ that we seek.

This completes the proof of the theorem. $\square$
Definition 12.5 Let $X$ be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on $X$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mu : \mathcal{A} \rightarrow \mathcal{Y}$. $\mu$ is said to be a $\mathcal{Y}$-valued pre-measure on the algebra $\mathcal{A}$ if

(i) $\mu(\emptyset) = 0$;

(ii) $\forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, we have

$$\sum_{i=1}^{\infty} \|\mu(A_i)\| < \infty$$

and $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i) \in \mathcal{Y}$.

Define $P \circ \mu : \mathcal{A} \rightarrow [0, \infty) \subset \mathbb{R}_e$ by, $\forall A \in \mathcal{A}$, $P \circ \mu(A) := \sup_{n \in \mathbb{Z}_+} \|A \cap \bigcup_{i=1}^{n} A_i\|$, $\forall_{1 \leq i \leq n} \mu(A_i) \in \mathcal{Y}$. $P \circ \mu$ is said to be the total variation of $\mu$.

$\mu$ is said to be finite if $P \circ \mu(X) < \infty$.

Proposition 12.6 Let $X$ be a set, $\mathcal{A} \subseteq 2^X$ be an algebra on $X$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, and $\mu : \mathcal{A} \rightarrow \mathcal{Y}$ be a $\mathcal{Y}$-valued pre-measure on the algebra $\mathcal{A}$. Then, $P \circ \mu : \mathcal{A} \rightarrow [0, \infty) \subset \mathbb{R}_e$ defines a measure on the algebra $\mathcal{A}$.

Proof Clearly, $P \circ \mu(\emptyset) = 0$. $\forall$ pairwise disjoint $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$, $\forall t \in \mathbb{N}$ with $P \circ \mu(A_t) > 0$, $\exists_{t_i \in \mathbb{Z}_+}$ and $\exists$ pairwise disjoint $(A_{i,j})_{j=1}^{n_i} \subseteq \mathcal{A}$ with $A_i = \bigcup_{j=1}^{n_i} A_{i,j}$ such that $t_i < \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \leq P \circ \mu(A_i)$. $\forall n \in \mathbb{N}$ with $P \circ \mu(A_t) = 0$, let $t_i = 0$, $n_i = 1$, $A_{i,1} = A_i$. Then, $0 = t_i \leq \|\mu(A_{i,1})\| \leq P \circ \mu(A_i) = 0$.

Claim 12.6.1 $0 \leq t := \sum_{i=1}^{\infty} t_i \leq P \circ \mu(A)$.

Proof of claim: Clearly, $t \geq 0$ since $t_i \geq 0, \forall t_i \in \mathbb{N}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $t < +\infty$; Case 2: $t = +\infty$.

Case 1: $t < +\infty$. Then, $t \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \geq 0$. Therefore, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $t - \epsilon < \sum_{i=1}^{N} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| \leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| + \|\mu(A \setminus \bigcup_{i=1}^{N} \bigcup_{j=1}^{n_i} A_{i,j})\| \leq P \circ \mu(A)$. By the arbitrariness of $\epsilon$, we have $t \leq P \circ \mu(A)$.

Case 2: $t = +\infty$. Then $t = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| = \infty$. $\forall M \in [0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ such that

$$M \leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\|$$

$$\leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \|\mu(A_{i,j})\| + \|\mu(A \setminus \bigcup_{i=1}^{N} \bigcup_{j=1}^{n_i} A_{i,j})\| \leq P \circ \mu(A)$$

This implies that $P \circ \mu(A) = +\infty = t$.

This completes the proof of the claim. \(\square\)
Hence, $\sum_{i=1}^{\infty} t_i \leq P \circ \mu(A)$. By the arbitrariness of $t_i$’s, we have $\sum_{i=1}^{\infty} P \circ \mu(A_i) \leq P \circ \mu(A)$.

On the other hand, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_j)^n_{j=1} \subseteq A$ with $A = \bigcup_{j=1}^{n} E_j$, we have

$$
\sum_{j=1}^{n} \| \mu(E_j) \| = \sum_{j=1}^{n} \mu \left( \bigcup_{i=1}^{\infty} (A_i \cap E_j) \right) = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \mu(A_i \cap E_j) \leq \sum_{i=1}^{\infty} P \circ \mu(A_i)
$$

By the arbitrariness of $n$ and $(E_j)^n_{j=1}$, we have $P \circ \mu(A) \leq \sum_{i=1}^{\infty} P \circ \mu(A_i)$.

Hence, $P \circ \mu(A) = \sum_{i=1}^{\infty} P \circ \mu(A_i)$. This implies that $P \circ \mu$ is a measure on the algebra $A$. This completes the proof of the proposition. \(\square\)

Clearly, $\mu : A \to \mathcal{Y}$ is a finite $\mathcal{Y}$-valued pre-measure on the algebra $A$ if, and only if, $\mu$ is a finite $\mathcal{Y}$-valued measure on the algebra $A$. In this case, the definitions of total variations of $\mu$ coincide considering $\mu$ as a $\mathcal{Y}$-valued pre-measure on the algebra $A$ or as a $\mathcal{Y}$-valued measure on the algebra $A$.

### 12.2 Change of Variable

**Definition 12.7** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, $\hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \hat{\nu})$ be a $\mathcal{Y}$-valued measure space, $\mathcal{Y}$ be a normed linear space, and $g : \mathcal{X} \to \hat{\mathcal{X}}$. $g$ is said to be an isomeasure between $\mathcal{X}$ and $\hat{\mathcal{X}}$ if $g$ is bijective; $\forall E \in \mathcal{B}, g(E) \in \mathcal{B}$; $\forall E \in \mathcal{B}, g_{\text{inv}}(E) \in \mathcal{B}$; $\forall E \in \mathcal{B}, P \circ \nu(E) = P \circ \mu(g_{\text{inv}}(E))$; $\forall E \in \text{dom}(\mu)$, we have $g(E) \in \text{dom}(\nu)$ and $\nu(g(E)) = \mu(E)$; and $\forall E \in \text{dom}(\nu)$, we have $g_{\text{inv}}(E) \in \text{dom}(\mu)$ and $\mu(E) = \nu(E) = \mu(g_{\text{inv}}(E))$.

We will say $\mathcal{X}$ and $\hat{\mathcal{X}}$ are isomeasurable if there is an isomeasure between $\mathcal{X}$ and $\hat{\mathcal{X}}$.

Two $\mathcal{Y}$-valued measure spaces are isomeasurable, then they are equivalent up to relabeling of elements. Thus, isomeasure preserves the completeness, finiteness, and $\sigma$-finiteness of the $\mathcal{Y}$-valued measure spaces.

**Definition 12.8** Let $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued topological measure space, $\hat{\mathcal{X}} := (\hat{\mathcal{X}}, \hat{\mathcal{B}}, \hat{\nu})$ be a $\mathcal{Y}$-valued topological measure space, $\mathcal{Y}$ be a normed linear space, and $g : \mathcal{X} \to \hat{\mathcal{X}}$. $g$ is said to be an homeomorphical isomeasure between $\mathcal{X}$ and $\hat{\mathcal{X}}$ if $g : \mathcal{X} \to \hat{\mathcal{X}}$ is a homeomorphism and $g$ is an isomeasure between $\mathcal{X}$ and $\hat{\mathcal{X}}$. We will say $\mathcal{X}$ and $\hat{\mathcal{X}}$ are homeomorphically isomeasurable if there is an homeomorphic isomeasure between $\mathcal{X}$ and $\hat{\mathcal{X}}$.

Two $\mathcal{Y}$-valued topological measure spaces are homeomorphically isomeasurable, then they are equivalent up to relabeling of elements.
Proposition 12.9  Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $(\bar{X}, \bar{\mathcal{B}})$ be a measurable space, and $g : X \to \bar{X}$ be bijective and such that $\forall E \in \mathcal{B}$, $g_{\text{inv}}(E) \in \mathcal{B}$ and $g(E) \in \bar{\mathcal{B}}$. Then, we may define the induced measure $g_\mu : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+$ by $g_\mu(\bar{E}) = \mu (g_{\text{inv}}(E))$, $\forall E \in \mathcal{B}$. Then, $\mathcal{X} := (\bar{X}, \bar{\mathcal{B}}, g_\mu)$ is a measure space and $g$ is an isomasure between $\mathcal{X}$ and $\mathcal{X}$.

Proof  This is straightforward, and therefore omitted. \qed

Proposition 12.10  Let $\mathcal{Y}$ be a normed linear space, $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, $(\bar{X}, \bar{\mathcal{B}})$ be a measurable space, and $g : X \to \bar{X}$ be bijective and such that $\forall E \in \mathcal{B}$, $g_{\text{inv}}(E) \in \mathcal{B}$ and $g(E) \in \bar{\mathcal{B}}$. Then, we may define the $\mathcal{Y}$-valued induced measure $g_\mu : \bar{\mathcal{B}} \to \mathcal{Y}$ by $g_\mu(g_{\text{inv}}(E)) = \mu (g_{\text{inv}}(E))$, $\forall E \in \mathcal{B}$ with $g_{\text{inv}}(E) \in \text{dom}(\mu)$; and $\forall E \in \bar{\mathcal{B}}$ with $g_{\text{inv}}(E) \in \mathcal{B} \setminus \text{dom}(\mu)$, $g_\mu(\bar{E})$ is undefined. Then, $\mathcal{X} := (\bar{X}, \bar{\mathcal{B}}, g_\mu)$ is a $\mathcal{Y}$-valued measure space, $\mathcal{P} \circ (g_\mu) = g_\mu(\mathcal{P} \circ \mu)$, and $g$ is an isomasure between $\mathcal{X}$ and $\mathcal{X}$.

Proof  Since $\mathcal{X}$ is a $\mathcal{Y}$-valued measure space, then $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ is a measure space. By Proposition 12.9, $(\bar{X}, \bar{\mathcal{B}}, g_\mu(\mathcal{P} \circ \mu))$ is a measure space and $g$ is an isomasure between $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ and $(\bar{X}, \bar{\mathcal{B}}, g_\mu(\mathcal{P} \circ \mu))$.

(i) $g_\mu(\emptyset) = \mu(g_{\text{inv}}(\emptyset)) = \mu(\emptyset) = \emptyset \in \mathcal{Y}$.

(ii) $\forall E \in \mathcal{B}$ with $g_\mu(\mathcal{P} \circ \mu)(\bar{E}) = +\infty$, we have $\mathcal{P} \circ \mu(g_{\text{inv}}(E)) = +\infty$ and $g_{\text{inv}}(E) \in \text{dom}(\mu)$. This implies that $g_\mu(E)$ is undefined.

(iii) $\forall E \in \mathcal{B}$ with $g_\mu(\mathcal{P} \circ \mu)(\bar{E}) < \infty$, we have $\mathcal{P} \circ \mu(g_{\text{inv}}(E)) < \infty$ and $g_{\text{inv}}(E) \in \text{dom}(\mu)$. This implies that $g_\mu(E) = \mu(g_{\text{inv}}(E)) \in \mathcal{Y}$. $\forall$ pairwise disjoint $\{\bar{E}_i\}_{i=1}^\infty \subseteq \bar{\mathcal{B}}$ with $\bar{E} = \bigcup_{i=1}^\infty \bar{E}_i$, $\forall i \in \mathbb{N}$, $\bar{E}_i \subseteq \bar{E}$, $g_{\text{inv}}(\bar{E}_i) \subseteq g_{\text{inv}}(\bar{E})$, $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E}_i)) \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E}))$, $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bigcup_{i=1}^\infty \bar{E}_i)) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) < \infty$, where the first inequality follows from Definition 11.108; the second equality follows from the fact that $\mathcal{P} \circ \mu$ is a measure; and the third equality follows from Proposition 2.5. Furthermore, $\sum_{i=1}^\infty g_\mu(\bar{E}_i) = \sum_{i=1}^\infty \mu(g_{\text{inv}}(\bar{E}_i)) = \mu\left(\bigcup_{i=1}^\infty g_{\text{inv}}(\bar{E}_i)\right) = \mu(g_{\text{inv}}(\bigcup_{i=1}^\infty \bar{E}_i)) = \mu(g_{\text{inv}}(\bar{E})) = g_\mu(\bar{E}) \in \mathcal{Y}$, where the second equality follows from Definition 11.108; and the third equality follows from Proposition 2.5.

(iv) $\forall \bar{E} \in \bar{\mathcal{B}}$ with $g_\mu(\mathcal{P} \circ \mu)(\bar{E}) < \infty$, by the previous paragraph, we have $\sup_{n \in \mathbb{Z}_+} \left(\bar{E}_i\right)_{i=1}^n \subseteq \bar{E}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n \sum_{i=1}^n \| g_\mu(\bar{E}_i) \| = s_{\bar{E}} \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = g_\mu(\mathcal{P} \circ \mu)(\bar{E}).$ $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n \in \mathbb{Z}_+$, $\exists$ pairwise disjoint $\{\bar{E}_i\}_{i=1}^n \subseteq \bar{\mathcal{B}}$ with $g_{\text{inv}}(\bar{E}) = \bigcup_{i=1}^n \bar{E}_i$, such that $\mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) - \epsilon < \sum_{i=1}^n \| \mu(\bar{E}_i) \| \leq \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = g_\mu(\mathcal{P} \circ \mu)(\bar{E})$. Then, $\left(\bar{E}_i\right)_{i=1}^n := \left(g(\bar{E}_i)\right)_{i=1}^n \subseteq \bar{\mathcal{B}}$ pairwise disjoint, $\bigcup_{i=1}^n \bar{E}_i = \mathcal{P} \circ \mu(g_{\text{inv}}(\bar{E})) = \bar{E}$, and $\sum_{i=1}^n \| g_\mu(\bar{E}_i) \| = \sum_{i=1}^n \| \mu(g_{\text{inv}}(\bar{E}_i)) \| = \sum_{i=1}^n \| \mu(\bar{E}_i) \| > g_\mu(\mathcal{P} \circ \mu)(\bar{E}) - \epsilon$. Hence, $s_{\bar{E}} \geq g_\mu(\mathcal{P} \circ \mu)(\bar{E})$ and $s_{\bar{E}} = g_\mu(\mathcal{P} \circ \mu)(\bar{E})$. \qed
Therefore, by Definition 11.108, \((\bar{X}, \mathcal{B}, g_\ast \mu)\) is a \(Y\)-valued measure space with \(P \circ (g_\ast \mu) = g_\ast (P \circ \mu)\). It is straightforward to show that \(g\) is an isomeasure between \(X\) and \(\bar{X}\). This completes the proof of the proposition.

\[\square\]

**Proposition 12.11** Let \(X := (X, \mathcal{O})\) and \(\bar{X} := (\bar{X}, \bar{O})\) be topological spaces, \(g : X \rightarrow \bar{X}\) be a homeomorphism, (\(Y\) be a normed linear space,) and \(\mathcal{X} := (X, \mathcal{B}(X), \mu)\) be a \((Y\)-valued\) topological measure space. Then, \(\bar{X} := (\bar{X}, \mathcal{B}(\bar{X}), \bar{g}_\ast \mu)\) is a \((Y\)-valued\) topological measure space and \(g\) is a homeomorphical isomeasure between \(\mathcal{X}\) and \(\bar{X}\).

**Proof** Clearly, \(g\) is bijective and \(g\) and \(g_\ast \mu\) are continuous. By Proposition 11.34, \(\forall E \in \mathcal{B}(X)\), we have \(g(E) \in \mathcal{B}(\bar{X})\); and \(\forall E \in \mathcal{B}(\bar{X})\), we have \(g_\ast \mu(E) \in \mathcal{B}(X)\). Then, by Propositions 12.9 and 12.10, the \((Y\)-valued\) induced measure \(g_\ast \mu\) is well defined, \((X, \mathcal{B}(X), \mu)\) is a \((Y\)-valued\) measure space, and \(g\) is an isomeasure between \(\mathcal{X}\) and \(\bar{X}\).

We will show that \(\bar{X}\) is a \((Y\)-valued\) topological measure space. \(\forall E \in \mathcal{B}(\bar{X}), \forall \bar{u} \in (0, \infty) \subset R\), we have \(g_\ast \mu(E) \in \mathcal{B}(X)\). Since \(\bar{X}\) is a \((Y\)-valued\) topological measure space, then \(\exists U \in \mathcal{O}\) with \(g_\ast \mu(E) \subset U\) such that \(\mu(U \setminus g_\ast \mu(E)) < \epsilon\). Let \(U := g(U) \in \mathcal{O}\). Then, \(E \subset g(U) = U\) and \(g_\ast \mu(U \setminus E) = g_\ast (g_\ast \mu(U \setminus E)) = \mu(U \setminus g_\ast \mu(E)) = \mu(U \setminus g_\ast \mu(E)) < \epsilon\). Note that \(U \setminus g_\ast \mu(E) = g_\ast (P \circ \mu)(U \setminus E) = P \circ \mu(g_\ast \mu(U \setminus E)) = P \circ \mu(g_\ast \mu(U) \setminus g_\ast \mu(E)) = P \circ \mu(U \setminus g_\ast \mu(E)) < \epsilon\). Where, the first equality follows from Proposition 12.10. Hence, \(\bar{X}\) is a \((Y\)-valued\) topological measure space.

Then, \(g\) is a homeomorphical isomeasure between \(\mathcal{X}\) and \(\bar{X}\).

\[\square\]

**Proposition 12.12** Let \(X := (X, \mathcal{B}, \mu)\) be a measure space, \(\bar{X} := (\bar{X}, \mathcal{B}, \nu)\) be a measure space, \(Y\) be a normed linear space, \(f : \bar{X} \rightarrow \bar{Y}\) be \(\mathcal{B}\)-measurable, and \(g : X \rightarrow \bar{X}\) be an isomeasure between \(X\) and \(\bar{X}\). Then, \(f_X(f \circ g) \, d\mu = f_X f \, d\nu\) whenever one of the integrals exists.

**Proof** Clearly, \(f \circ g\) is \(\mathcal{B}\)-measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\mu(X) < +\infty\); Case 2: \(\mu(X) = +\infty\).

Case 1: \(\mu(X) < +\infty\). \(\nu(\bar{X}) = \mu(g_\ast \mu(\bar{X})) = \mu(X) < +\infty\). Let \(\mathcal{Y}\) be the integration system on \(Y\). \(\forall \mathcal{R} \in \mathcal{Y}\), let \(F_{\mathcal{R}} := \sum_{\alpha \in \Lambda} y_{\alpha} \nu(f_\ast(U_{\alpha}))) \in \mathcal{Y}\) and \(\bar{F}_{\mathcal{R}} := \sum_{\alpha \in \Lambda} y_{\alpha} \mu((f \circ g)_\ast(U_{\alpha}))) \in \mathcal{Y}\). Note that \(\bar{F}_{\mathcal{R}} = \sum_{\alpha \in \Lambda} y_{\alpha} \mu(g_\ast(f_\ast(U_{\alpha}))) = F_{\mathcal{R}}\). Then, by Definition 11.70, we have \(\int_{\bar{X}} f \, d\nu = \lim_{\mathcal{R} \in \mathcal{Y}} \bar{F}_{\mathcal{R}} = \lim_{\mathcal{R} \in \mathcal{Y}} F_{\mathcal{R}} = \int_{X} (f \circ g) \, d\mu\) whenever one of the integrals exists.

Case 2: \(\mu(X) = +\infty\). \(\nu(\bar{X}) = \mu(g_\ast \mu(\bar{X})) = \mu(X) = +\infty\). Let \((\bar{F}_{\mathcal{A}})_{\mathcal{A} \in \mathcal{M}(\bar{X})}\) be the net for \(\int_{\bar{X}} f \, d\nu\) and \((F_{\mathcal{A}})_{\mathcal{A} \in \mathcal{M}(X)}\) be the net for \(\int_{X} (f \circ g) \, d\mu\) as defined in Definition 11.71. Without loss of generality,
Assume that $\int_{\tilde{X}} f \, d\nu$ exists. Then,

$$\int_{\tilde{X}} f \, d\nu = \lim_{A \in \mathcal{M}(\tilde{X})} \tilde{F}_A = \lim_{A \in \mathcal{M}(\tilde{X})} \int_{\tilde{A}} f \, d\nu$$

where the first two equalities follow from Definition 11.71; and the third equality follows from Case 1. Fix any open set $U (U \subseteq \mathbb{R}_e$ if $\mathcal{Y} = \mathbb{R}$ or $U \subseteq \mathcal{Y}$ if $\mathcal{Y} \neq \mathbb{R})$ with $\int_{\tilde{X}} f \, d\nu \in U$, \( \exists \tilde{A}_0 \in \mathcal{M}(\tilde{X}), \forall \tilde{A} \in \mathcal{M}(\tilde{X}) \) with $\tilde{A}_0 \subseteq \tilde{A}$, $F_{\tilde{A}} = \int_{\text{inv}(\tilde{A})} (f \circ g) \, d\mu_{\text{inv}(\tilde{A})} \in U$. \( \forall A \in \mathcal{M}(\tilde{X}) \) with $g_{\text{inv}}(\tilde{A}_0) \subseteq A$, we have $A \in \mathcal{B}$ and $\mu(A) < +\infty$. Then, $g(A) \in \mathcal{B}$ and $\mu(A) = \nu(g(A)) < +\infty$. Hence, $g(A) \in \mathcal{M}(\tilde{X})$ and $g(A) \supseteq g(g_{\text{inv}}(\tilde{A}_0)) = \tilde{A}_0$. Then, $F_A = \int_A (f \circ g) \, d\mu_A = \int_{g_{\text{inv}}(g(A))} (f \circ g) \, d\mu_{g_{\text{inv}}(g(A))} \in U$. Hence, $\int_{\tilde{X}} (f \circ g) \, d\mu = \lim_{A \in \mathcal{M}(\tilde{X})} F_A = \int_{\tilde{X}} f \, d\nu$.

This completes the proof of the proposition. \( \square \)

**Proposition 12.13** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\tilde{\mathcal{X}} := (\tilde{X}, \tilde{\mathcal{B}}, \nu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $W$ be a separable subspace of $\mathcal{Y}$, $f : \tilde{X} \to W$ be absolutely integrable over $\tilde{X}$, and $g : X \to \tilde{X}$ be an isomorphism between $\mathcal{X}$ and $\tilde{\mathcal{X}}$. Then, $f \circ g$ is absolutely integrable over $\mathcal{X}$ and $\int_{\tilde{X}} (f \circ g) \, d\mu = \int_{\tilde{X}} f \, d\nu$. \( \square \)

**Proof** Clearly, $f \circ g$ is $\mathcal{B}$-measurable. By the assumption, $0 \leq \int_{\tilde{X}} (P \circ f) \, d\nu < +\infty$. By Proposition 12.12, $\int_{\tilde{X}} (P \circ f \circ g) \, d\mu = \int_{\tilde{X}} (P \circ f) \, d\nu \in [0, +\infty) \subseteq \mathbb{R}$. Hence, $f \circ g$ is absolutely integrable over $\tilde{X}$. By Proposition 11.92, $f$ is integrable over $\tilde{X}$ and hence, $\int_{\tilde{X}} f \, d\nu < +\infty$. Then, $f : \tilde{X} \to W$ is integrable. \( \square \)

**Proposition 12.14** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\tilde{\mathcal{X}} := (\tilde{X}, \tilde{\mathcal{B}}, \nu)$ be a $\mathcal{Y}$-valued measure space, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : X \to \mathcal{W}$ is $\mathcal{B}$-measurable, and $g : X \to \tilde{X}$ be an isomorphism between $\mathcal{X}$ and $\tilde{\mathcal{X}}$. Then, $\int_{\tilde{X}} f \, d\nu = \int_{\tilde{X}} (f \circ g) \, d\mu$, whenever one of the integrals exists.

**Proof** Clearly, $f \circ g$ is $\mathcal{B}$-measurable. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu(X) < +\infty$; Case 2: $\mathcal{P} \circ \mu(X) = +\infty$.

Case 1: $\mathcal{P} \circ \mu(X) < +\infty$. $\mathcal{P} \circ \nu(\tilde{X}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\tilde{X})) = \mathcal{P} \circ \mu(X) < +\infty$. Let $\mathcal{A}(\mathcal{W})$ be the integration system on $\mathcal{W}$. $\forall A \in \mathcal{A}(\mathcal{W})$, let $F_A := \sum_{\alpha \in A} w_{\alpha}(f_{\text{inv}}(U_{\alpha})) \in \mathcal{Z}$ and $F_R := \sum_{\alpha \in \Lambda} w_{\alpha}(g_{\text{inv}}(f_{\text{inv}}(U_{\alpha}))) \in \mathcal{Z}$. Then, $F_R = \sum_{\alpha \in \Lambda} w_{\alpha}(g_{\text{inv}}(f_{\text{inv}}(U_{\alpha}))) = F_R$. Then, by Definition 11.119, we have $\int_{\tilde{X}} f \, d\nu = \lim_{R \subseteq \mathcal{A}(\mathcal{W})} F_R = \lim_{R \subseteq \mathcal{A}(\mathcal{W})} F_R = \int_{\tilde{X}} (f \circ g) \, d\mu$, whenever one of the integrals exists.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $\mathcal{P} \circ \nu(\tilde{X}) = \mathcal{P} \circ \mu(g_{\text{inv}}(\tilde{X})) = \mathcal{P} \circ \mu(X) = +\infty$. Let $\{ F_A \}_{A \in \mathcal{M}(\tilde{X})}$ be the net for $\int_{\tilde{X}} f \, d\nu$ and $\{ F_A \}_{A \in \mathcal{M}(\mathcal{X})}$ be the net for
12.2. CHANGE OF VARIABLE

If \( f : X \to \bar{W} \) is absolutely integrable over \( X \), and \( g : X \to \bar{X} \) is an isomorphism between \( \bar{X} \) and \( X \). Then, \( \int_{\bar{X}} (f \circ g) \, d\mu \) as defined in Definition 11.120. Without loss of generality, assume that \( \int_{\bar{X}} f \, d\nu \) exists. Then,

\[
\int_{\bar{X}} f \, d\nu = \lim_{A \in \mathcal{M} (\bar{X})} \bar{F}_A = \lim_{A \in \mathcal{M} (\bar{X})} \int_{\bar{A}} f \, d\nu_{\bar{A}}
\]

where the first two equalities follow from Definition 11.120; and the third equality follows from Case 1. Fix any open set \( U \subseteq \mathbb{R} \) if \( \mathcal{Z} = \mathbb{R} \) or \( U \subseteq \mathcal{Z} \) if \( \mathcal{Z} \neq \mathbb{R} \) with \( \int_X f \, d\nu \in U \), \( \exists \bar{A}_0 \in \mathcal{M} (\bar{X}) \), \( \forall \bar{A} \in \mathcal{M} (\bar{X}) \) with \( \bar{A}_0 \subseteq \bar{A} \), \( \bar{F}_\bar{A} = \int_{\text{inv}(\bar{A})} (f \circ g) \, d\mu_{\text{inv}(\bar{A})} \in U \). Hence, \( f \circ g \) is absolutely integrable over \( \bar{X} \) and \( \bar{X} \) is a Banach space over \( \bar{Y} \).

**Proposition 12.15** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( K \), \( \hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \nu) \) be a \( \mathcal{Y} \)-valued measure space, \( \mathcal{Z} \) be a Banach space over \( K \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B} (\mathcal{Y}, \mathcal{Z}) \), \( f : \hat{X} \to \mathcal{W} \) be absolutely integrable over \( \hat{X} \), and \( g : X \to \hat{X} \) be an isomorphism between \( X \) and \( \hat{X} \). Then, \( f \circ g \) is absolutely integrable over \( X \), and \( \int_{\hat{X}} (f \circ g) \, d\mu_{\text{inv}} \in \mathcal{Z} \).

**Proof** Cleariy, \( f \circ g \) is \( \mathcal{B} \)-measurable. By Definition 12.7, \( g \) is an isomorphism between \( (X, \mathcal{B}, \mathcal{P} \circ \mu) \) and \( (\hat{X}, \hat{\mathcal{B}}, \mathcal{P} \circ \nu) \). By Proposition 12.12, \( \int_X (\mathcal{P} \circ f) \, d\mathcal{P} = \int_{\hat{X}} (f \circ g) \, d\mathcal{P} \circ \nu \in \mathbb{R} \), where the last step follows from the assumption that \( f \) is absolutely integrable over \( \hat{X} \). Hence, \( f \circ g \) is absolutely integrable over \( X \). By Proposition 12.14, \( \int_{\hat{X}} (f \circ g) \, d\mu_{\text{inv}} = \int_{\hat{X}} (f \circ g) \, d\mu \in \mathcal{Z} \).

**Theorem 12.16 (Change of Variable)** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( K \)-valued measure space, \( \hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \nu) \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a separable Banach space over \( K \), \( g : \hat{X} \to X \) be an isomorphism between \( \hat{X} \) and \( X \), \( \mathcal{Z} \) be a Banach space over \( K \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B} (\mathcal{Y}, \mathcal{Z}) \), \( f : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable, and \( \frac{d(g_{\mu})}{d\nu} = h : X \to \mathcal{Y} \). Then, \( f \) is absolutely integrable over \( \hat{X} \) if, and only if, \( (\mathcal{P} \circ f \circ g_{\mu})(\mathcal{P} \circ h) \) is integrable over \( X \) and \( \int_X f \, d\nu = \int_{\hat{X}} (f \circ g_{\mu}) h \, d\mu \in \mathcal{Z} \).

**Proof** By Proposition 12.15, \( f \) is absolutely integrable over \( \hat{X} \) if, and only if, \( f \circ g_{\mu} \) is absolutely integrable over \( \hat{X} \). By (ii) of Proposition 11.168, \( f \circ g_{\mu} \) is absolutely integrable over \( \hat{X} \) if, and only if, \( (\mathcal{P} \circ f \circ g_{\mu})(\mathcal{P} \circ h) \) is integrable over \( X \) and \( \int_X f \, d\nu = \int_{\hat{X}} (f \circ g_{\mu}) h \, d\mu \in \mathcal{Z} \).
and only if, \((P \circ f \circ g_{\text{inv}})(P \circ h)\) is integrable over \(X, \mathcal{B}, P \circ \mu\). In this case, 
\((f \circ g_{\text{inv}})h\) is absolutely integrable over \(X\) and, by Proposition 12.15 and (ii) 
of Proposition 11.168, \(\int_X f \, d\nu = \int_X (f \circ g_{\text{inv}}) \, d(\gamma, \nu) = \int_X (f \circ g_{\text{inv}})h \, d\mu \in \mathbb{Z}\). 
This completes the proof of the theorem. \(\Box\)

12.3 Product Measure

**Definition 12.17** Let \(X\) be a set. \(\mathcal{A} \subseteq X^2\) is said to be a \(\pi\)-system on \(X\) if: 
(i) \(\emptyset, X \in \mathcal{A}\); and (ii) \(A, B \in \mathcal{A}\), we have \(A \cap B \in \mathcal{A}\). \(\mathcal{D} \subseteq X^2\) is said to be a monotone class on \(X\) if 
(i) \(\emptyset, X \in \mathcal{D}\); 
(ii) \(A, B \in \mathcal{D}\) with \(A \subseteq B\), we have \(B \setminus A \in \mathcal{D}\); 
(iii) \(\forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{D}\) with \(A_i \subseteq A_{i+1}\), \(\forall i \in \mathbb{N}\), we have \(\bigcup_{i=1}^{\infty} A_i \in \mathcal{D}\).

Let \(X\) be a set and \(\mathcal{A} \subseteq X^2\). Then, there exists the smallest monotone class on \(X\) that contains \(\mathcal{A}\), which is the intersection of all monotone classes on \(X\) that contains \(\mathcal{A}\).

**Proposition 12.18** Let \(X\) be a set. \(\mathcal{B} \subseteq X^2\) is a \(\sigma\)-algebra on \(X\) if, and 
only if, \(\mathcal{B}\) is a \(\pi\)-system on \(X\) and is a monotone class on \(X\).

**Proof** “Necessity” Let \(\mathcal{B}\) be a \(\sigma\)-algebra on \(X\). Then, \(\emptyset, X \in \mathcal{B}\). \(\forall A, B \in \mathcal{B}\), we have \(A \cap B \in \mathcal{B}\). Hence, \(\mathcal{B}\) is a \(\pi\)-system on \(X\). \(\forall A, B \in \mathcal{B}\) with \(A \subseteq B\), \(B \setminus A \in \mathcal{B}\). \(\forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{B}\) with \(A_i \subseteq A_{i+1}\), \(\forall i \in \mathbb{N}\), we have \(\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}\). Hence, \(\mathcal{B}\) is a monotone class on \(X\).

“Sufficiency” Let \(\mathcal{B}\) be a monotone class on \(X\) and a \(\pi\)-system on \(X\). Then, \(\emptyset, X \in \mathcal{B}\). \(\forall A \subseteq \mathcal{B}\), we have \(A \subseteq X\) and \(X \setminus A = \tilde{A} \in \mathcal{B}\), since \(\mathcal{B}\) is a monotone class on \(X\). \(\forall A_1, A_2 \in \mathcal{B}\), we have \(A_1 \cup A_2 = \tilde{A_1} \cap \tilde{A_2}\). By the above, \(A_1, A_2 \in \mathcal{B}\). Then, \(A_1 \cap A_2 \in \mathcal{B}\) since \(\mathcal{B}\) is a \(\pi\)-system on \(X\). This further implies that \(A_1 \cup A_2 \in \mathcal{B}\). Hence, \(\mathcal{B}\) is an algebra on \(X\). \(\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{B}\), \(\forall n \in \mathbb{N}\), let \(A_n := \bigcup_{i=1}^{n} E_i \in \mathcal{B}\), since \(\mathcal{B}\) is an algebra. Then, \(A_n \subseteq A_{n+1}\), \(\forall n \in \mathbb{N}\). By \(\mathcal{B}\) being a monotone class on \(X\), we have \(\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} E_i \in \mathcal{B}\). This shows that \(\mathcal{B}\) is a \(\sigma\)-algebra.

This completes the proof of the proposition. \(\Box\)

**Lemma 12.19 (Monotone Class Lemma)** Let \(X\) be a set, \(\mathcal{A} \subseteq X^2\) be a \(\pi\)-system on \(X\), \(\mathcal{B}\) be the \(\sigma\)-algebra on \(X\) generated by \(\mathcal{A}\), and \(\mathcal{D}\) be the smallest monotone class on \(X\) that contains \(\mathcal{A}\). Then, \(\mathcal{B} = \mathcal{D}\).

**Proof** By Proposition 12.18, \(\mathcal{B}\) is a monotone class on \(X\). Hence, \(\mathcal{D} \subseteq \mathcal{B}\).

Define \(\mathcal{D}_1 := \{ E \in \mathcal{D} \mid E \cap A \in \mathcal{D}, \forall A \in \mathcal{A} \}\). Since \(\mathcal{A}\) is a \(\pi\)-system on \(X\), then \(\mathcal{A} \subseteq \mathcal{D}_1\). Clearly \(\mathcal{D}_1 \subseteq \mathcal{D}\). We will show that \(\mathcal{D}_1\) is a monotone class on \(X\). Then, \(\mathcal{D} \subseteq \mathcal{D}_1\) and \(\mathcal{D} = \mathcal{D}_1\).
12.3. PRODUCT MEASURE

Clearly, \( \emptyset \in A \subseteq D_1 \) and \( X \in A \subseteq D_1 \). \( \forall E_1, E_2 \in D_1 \) with \( E_1 \subseteq E_2 \), we have \( E_1, E_2 \in D \) and \( E_2 \setminus E_1 \in D \). \( \forall A \in A \), we have \( A \cap E_1 \in D \) and \( A \cap E_2 \in D \). Clearly, \( A \cap E_1 \subseteq A \cap E_2 \). Since \( D \) is a monotone class on \( X \), we have \( D \ni (A \cap E_2) \setminus (A \cap E_1) = A \cap (E_2 \setminus E_1) \). By the arbitrariness of \( A \), we have \( E_2 \setminus E_1 \in D_1 \). \( \forall (E_i)_{i=1}^{\infty} \subseteq D_1 \) with \( E_i \subseteq E_{i+1}, \forall i \in \mathbb{N} \), we have \( \bigcup_{i=1}^{\infty} E_i \in D \) since \( D \) is a monotone class on \( X \). \( \forall A \subseteq A, \forall i \subseteq \mathbb{N}, E_i \subseteq D_1 \) implies that \( A \cap E_i \in D \). Clearly, \( A \cap E_i \subseteq A \cap E_{i+1}, \forall i \subseteq \mathbb{N} \).

Then, \( D \ni \bigcup_{i=1}^{\infty} (A \cap E_i) = A \cap (\bigcup_{i=1}^{\infty} E_i) \), since \( D \) is a monotone class on \( X \). By the arbitrariness of \( A \), we have \( \bigcup_{i=1}^{\infty} E_i \in D_1 \). This shows that \( D_1 \) is a monotone class on \( X \).

Define \( D_2 := \{ E \subseteq D : E \cap B \in D, \forall B \subseteq D \} \). Since \( D_1 = D \), then \( A \subseteq D_2 \). Clearly, \( D_2 \subseteq D \). We will show that \( D_2 \) is a monotone class on \( X \). Then, \( D \subseteq D_2 \) and \( D = D_2 \).

Clearly, \( \emptyset \in A \subseteq D_2 \) and \( X \in A \subseteq D_2 \). \( \forall E_1, E_2 \in D_2 \) with \( E_1 \subseteq E_2 \), we have \( E_2 \setminus E_1 \in D \). \( \forall B \in D \), we have \( E_1 \cap B \subseteq E_2 \cap B \subseteq D \). \( E_1 \subseteq E_2 \cap B \subseteq D \). \( \forall E_1, E_2 \in D_2 \) with \( E_1 \subseteq E_{i+1}, \forall i \in \mathbb{N} \), we have \( \bigcup_{i=1}^{\infty} E_i \subseteq D \) since \( D \) is a monotone class on \( X \). \( \forall B \subseteq D \), \( \forall i \subseteq \mathbb{N} \), \( E_i \subseteq D_2 \) implies that \( B \cap E_i \in D \). Clearly, \( B \cap E_i \subseteq B \cap E_{i+1}, \forall i \subseteq \mathbb{N} \). Then, \( D \ni \bigcup_{i=1}^{\infty} (B \cap E_i) = B \cap (\bigcup_{i=1}^{\infty} E_i) \), since \( D \) is a monotone class on \( X \). By the arbitrariness of \( B \), we have \( \bigcup_{i=1}^{\infty} E_i \subseteq D_2 \). This shows that \( D_2 \) is a monotone class on \( X \).

Therefore, \( D_2 = D \). This implies, by the definition of \( D_2 \), that \( D \) is a \( \pi \)-system on \( X \). By Proposition 12.18, we have \( D \) is \( \sigma \)-algebra. Then, \( B \subseteq D \). Hence, \( B = D \). This completes the proof of the lemma.

\[ \square \]

**Proposition 12.20** Let \( X \) be a set, \( \mathcal{C} \) be a semialgebra on \( X \), \( \mathcal{Y} \) be a normed linear space, \( A \) be the algebra on \( X \) generated by \( \mathcal{C} \), \( \nu : \mathcal{C} \rightarrow [0, \infty] \subseteq \mathbb{R}_+ \), and \( \mu \) be a function from \( \mathcal{C} \) to \( \mathcal{Y} \). Assume that

(i) \( \mu(\emptyset) = \emptyset \text{ and } \nu(\emptyset) = 0 \);

(ii) \( \forall C \subseteq \mathcal{C}, \forall n \subseteq \mathbb{Z}_+, \forall pairwise \ disjoint (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^{n} C_i \), we have \( \nu(C) = \sum_{i=1}^{n} \nu(C_i) \);

(iii) \( \forall C \subseteq \mathcal{C}, \forall pairwise \ disjoint (C_i)_{i=1}^{\infty} \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^{\infty} C_i \), we have \( \nu(C) \leq \sum_{i=1}^{\infty} \nu(C_i) \);

(iv) \( \forall C \subseteq \mathcal{C} \) with \( \nu(C) = \infty \), \( \mu(C) \) is undefined;

(v) \( \forall C \subseteq \mathcal{C} \) with \( \nu(C) < \infty \), \( \mu(C) \) is undefined, \( \forall n \subseteq \mathbb{Z}_+, \forall pairwise \ disjoint (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^{n} C_i \), we have \( \mu(C) = \sum_{i=1}^{n} \mu(C_i) \);

(vi) \( \forall C \subseteq \mathcal{C} \) with \( \nu(C) < \infty \), we have

\[ \nu(C) = \sup_{n \subseteq \mathbb{Z}_+, (C_i)_{i=1}^{n} \subseteq \mathcal{C}, C = \bigcup_{i=1}^{n} C_i, C_i \cap C_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(C_i) \| \]
Then, \( \nu \) admits a unique extension to a measure \( \bar{\nu} \) on the algebra \( \mathcal{A} \); and \( \mu \) admits a unique extension to a \( \mathcal{Y} \)-valued measure \( \bar{\mu} \) on the algebra \( \mathcal{A} \) with \( (\mathcal{P} \circ \bar{\mu})|_{\mathcal{E}} = \nu \). Furthermore, \( \mathcal{P} \circ \bar{\mu} = \bar{\nu} \).

**Proof**

By Proposition 11.31, \( \mathcal{A} = \{ A \subseteq X \mid A \text{ is the finite disjoint union of sets in } \mathcal{C} \} \). Define \( \bar{\nu} : \mathcal{A} \rightarrow [0, \infty] \subset \mathbb{R} \) by \( \bar{\nu}(A) = \sum_{i=1}^{n} \nu(C_i) \), \( \forall A = \bigcup_{i=1}^{n} C_i \in \mathcal{A} \), where \( n \in \mathbb{Z}_{+} \) and \( (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) is pairwise disjoint.

By (ii), (iii), and Proposition 11.32 and its proof, \( \bar{\nu} \) is the unique measure on the algebra \( \mathcal{A} \) which is an extension of \( \nu \). Hence, \( \bar{\nu} \) is the measure on the algebra \( \mathcal{A} \) that we seek. Define \( \bar{\mu} \) to be a function from \( \mathcal{A} \) to \( \mathcal{Y} \) by: \( \bar{\mu}(A) \) is undefined if \( \bar{\nu}(A) = \infty \); and \( \bar{\mu}(A) = \sum_{i=1}^{n} \mu(C_i) \in \mathcal{Y} \) if \( \bar{\nu}(A) < \infty \), \( A = \bigcup_{i=1}^{n} C_i \), \( n \in \mathbb{Z}_{+} \), and \( (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) is pairwise disjoint. By (i) and (v), \( \bar{\mu} \) is well-defined. Clearly, \( \bar{\mu}|_{\mathcal{E}} = \mu \) and hence \( \bar{\mu} \) is an extension of \( \mu \).

We will show that \( \bar{\mu} \) is a \( \mathcal{Y} \)-valued measure on the algebra \( \mathcal{A} \) with \( \mathcal{P} \circ \bar{\mu} = \bar{\nu} \). Then, \( \bar{\mu} \) is the \( \mathcal{Y} \)-valued measure on the algebra \( \mathcal{A} \) that we seek. By (i), \( \bar{\mu}(\emptyset) = 0_{\mathcal{Y}} \). \( \forall A \in \mathcal{A} \) with \( \bar{\nu}(A) = \infty \), \( \bar{\mu}(A) \) is undefined. \( \forall A \in \mathcal{A} \) with \( \bar{\nu}(A) < \infty \), \( \forall \text{ pairwise disjoint } (A_i)_{i=1}^{n} \subseteq \mathcal{A} \) with \( A = \bigcup_{i=1}^{n} A_i \), we have \( \bar{\mu}(A) \in \mathcal{Y} \), \( \bar{\nu}(A_i) < \infty \), and \( \bar{\mu}(A_i) \in \mathcal{Y} \), \( \forall i \in \mathbb{N} \). Since \( A \in \mathcal{A} \), then \( A = \bigcup_{i=1}^{n} C_i \), where \( n_0 \in \mathbb{Z}_{+} \) and \( (C_i)_{i=1}^{n_0} \subseteq \mathcal{C} \) is pairwise disjoint. Fix any \( i \in \mathbb{N} \), \( A_i \in \mathcal{A} \) implies that \( A_i = \bigcup_{j=1}^{n_i} C_{i,j} \), where \( n_i \in \mathbb{Z}_{+} \) and \( (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C} \) is pairwise disjoint. This leads to \( \sum_{i=1}^{\infty} \| \bar{\mu}(A_i) \| = \sum_{i=1}^{\infty} \| \sum_{j=1}^{n_i} \mu(C_{i,j}) \| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(C_{i,j}) \| \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) = \bar{\nu}(\bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} C_{i,j}) = \bar{\nu}(A) < \infty \), where the first equality follows from the definition of \( \bar{\mu} \); the second inequality follows from (vi); the second equality follows from \( \bar{\nu}|_{\mathcal{E}} = \nu \); and the third equality follows from the fact that \( \bar{\nu} \) is a measure on the algebra \( \mathcal{A} \). This implies that \( \forall \epsilon \in (0, +\infty] \subset \mathbb{R} \), \( \exists N \in \mathbb{N} \) such that \( \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(C_{i,j}) < \epsilon \), \( \forall n \in \mathbb{N} \) with \( N \leq n \), let \( \bar{A} := A \setminus \left( \bigcup_{i=1}^{n} \bigcup_{j=1}^{n_i} C_{i,j} \right) \in \mathcal{A} \). Clearly, \( \bar{A} = \bigcup_{i=1}^{n} \bigcup_{j=1}^{n_i} C_{i,j} \) and \( \bar{\nu}(\bar{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) < \epsilon \). By \( \bar{A} \in \mathcal{A} \), we have \( \bar{A} = \bigcup_{i=1}^{n} \bar{C}_i \), where \( \bar{n} \in \mathbb{Z}_{+} \) and \( (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C} \) is pairwise disjoint. Then, \( \| \bar{\mu}(A) - \sum_{i=1}^{\infty} \mu(C_{i,j}) \| = \| \sum_{k=1}^{n_0} \mu(C_k) - \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mu(C_{i,j}) \| = \| \sum_{k=1}^{n_0} \mu(C_k \cap (\bigcup_{i=1}^{n} \bigcup_{j=1}^{n_i} C_{i,j} \cup \bar{A})) - \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mu(C_{i,j} \cap (\bigcup_{k=1}^{n_0} C_k)) \| = \| \sum_{k=1}^{n_0} \mu(C_k \cap \bar{C}_i) \| + \| \sum_{k=1}^{n_0} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mu(C_k \cap C_{i,j}) \| - \| \sum_{i=1}^{n} \sum_{j=1}^{n_i} \mu(C_{i,j} \cap \bar{C}_i) \| = \| \sum_{k=1}^{n_0} \mu(C_k \cap \bar{C}_i) \| = \| \sum_{i=1}^{n} \mu(\bar{C}_i) \| = \sum_{i=1}^{n} \mu(\bar{C}_i) \| = \| \sum_{i=1}^{n} \nu(C_i) \| = \| \sum_{i=1}^{n} \bar{\nu}(C_i) \| = \| \nu(\bigcup_{i=1}^{n} C_i) \| = \| \bar{\nu}(A) \| < \epsilon \), where the first equality follows from the definition of \( \bar{\mu} \); the fourth equality follows from (v); the sixth equality follows from (vi); the second inequality follows from \( \bar{\nu}|_{\mathcal{E}} = \nu \); the seventh equality follows from \( \bar{\nu}|_{\mathcal{E}} = \nu \); the eighth equality follows from the fact that \( \bar{\nu} \) is a measure on the algebra \( \mathcal{A} \). Hence, \( \bar{\mu}(A) = \sum_{i=1}^{\infty} \mu(A_i) \).
Then, \( \sum_{i=1}^{n} \| \tilde{\mu}(A_i) \| = \sum_{i=1}^{n} \left\| \sum_{j=1}^{n_i} \mu(C_{i,j}) \right\| \leq \sum_{i=1}^{n} \sum_{j=1}^{n_i} \| \mu(C_{i,j}) \| \leq \sum_{i=1}^{n} \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \tilde{\nu}(C_{i,j}) = \sum_{i=1}^{n} \tilde{\nu}(A_i) = \tilde{\nu}(A) \), where the first equality follows from the definition of \( \tilde{\mu} \); the second inequality follows from (vi); the second equality follows from \( \tilde{\nu}|_\mathcal{C} = \nu \); and the last two equalities follow from the fact that \( \tilde{\nu} \) is a measure on the algebra \( \mathcal{A} \). Hence, \( s_A := \sup \{ \| \tilde{\mu}(A\cap C) \| : A \in \mathcal{A}, C \in \mathcal{C}, \forall C \in \mathcal{C}, \forall j, \sum_{i=1}^{n} \| \tilde{\mu}(A_i) \| \leq \tilde{\nu}(A) \} \). On the other hand, \( \forall e \in (0, +\infty) \in \mathbb{R} \), \( \forall i \in \{1, \ldots, n\} \), by (vi), \( \exists n_i \in \mathbb{Z}_{++} \), \( \exists \text{ pairwise disjoint } (C_{i,j})_{j=1}^{n} \subseteq \mathcal{C} \) with \( \tilde{\nu}(C_{i,j}) = \nu(C_i) < \sum_{j=1}^{n_i} \| \mu(C_{i,j}) \| + \epsilon/(1+n_0) = \sum_{j=1}^{n_i} \| \tilde{\mu}(C_{i,j}) \| + \epsilon/(1+n_0) \). Then, \( \tilde{\nu}(A) = \sum_{i=1}^{n} \tilde{\nu}(C_i) < \sum_{i=1}^{n} \sum_{j=1}^{n_i} \| \tilde{\mu}(C_{i,j}) \| + \epsilon \leq s_A + \epsilon. \) By the arbitrariness of \( \epsilon \), we have \( \tilde{\nu}(A) \leq s_A. \) Hence, \( s_A = \tilde{\nu}(A) \). This shows that \( \tilde{\mu} \) is a \( y \)-valued measure on the algebra \( \mathcal{A} \) with \( \mathcal{P} \circ \tilde{\mu} = \tilde{\nu} \).

To show the uniqueness of \( \tilde{\mu} \), let \( \hat{\mu} \) be another \( y \)-valued measure on the algebra \( \mathcal{A} \) with \( \hat{\mu}|_\mathcal{C} = \mu \) and \( (\mathcal{P} \circ \hat{\mu})|_\mathcal{C} = \nu \). Since \( \mathcal{P} \circ \hat{\mu} \) is a measure on the algebra \( \mathcal{A} \) that is an extension of \( \nu \), then \( \mathcal{P} \circ \hat{\mu} = \hat{\nu} = \mathcal{P} \circ \tilde{\mu} \) by Proposition 11.32. \( \forall A \in \mathcal{A} \) with \( \tilde{\nu}(A) = \infty \), then \( \tilde{\nu}(A) = \mu(A) \) and \( \tilde{\mu}(A) \), as both are undefined. \( \forall A \in \mathcal{A} \) with \( \tilde{\nu}(A) < \infty \), then \( A = \bigcup_{j=1}^{n} C_j \), where \( n \in \mathbb{Z}_{++} \) and \( (C_{j})_{j=1}^{n} \subseteq \mathcal{C} \) is pairwise disjoint. Then, \( \tilde{\mu}(A) = \sum_{j=1}^{n} \tilde{\mu}(C_j) = \sum_{j=1}^{n} \| \mu(C_j) \| = \hat{\mu}(A) \), where the first equality follows from the fact that \( \tilde{\mu} \) is a \( y \)-valued measure on the algebra \( \mathcal{A} \); the second equality follows from \( \tilde{\mu}|_\mathcal{C} = \mu \); and the last equality follows from the definition of \( \hat{\mu} \). Hence, \( \hat{\mu} = \hat{\mu} \). Therefore, \( \hat{\mu} \) is unique.

This completes the proof of the proposition. \( \square \)

**Proposition 12.21** Let \( m \in \mathbb{N} \), \( X_j := (X_j, \mathcal{B}_j, \mu_j) \) be a finite \( K \)-valued measure space, \( \forall j \in \{1, \ldots, m\} \), \( \mathcal{C}_m := \{ \prod_{j=1}^{m} E_j \subseteq \prod_{j=1}^{m} X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \ldots, m\} \} \), which is the set of measurable rectangles in \( \prod_{j=1}^{m} X_j \), \( \mathcal{B}_m \) be the \( \sigma \)-algebra on \( \prod_{j=1}^{m} X_j \) generated by \( \mathcal{C}_m \), \( \mu : \mathcal{C}_m \to \mathbb{K} \) be defined by \( \mu(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mu_j(E_j) \), \( \forall \prod_{j=1}^{m} E_j \in \mathcal{C}_m \), and \( \nu : \mathcal{C}_m \to (0, \infty) \subseteq \mathbb{R} \) be defined by \( \nu(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mathcal{P} \circ \mu_j(E_j) \), \( \forall \prod_{j=1}^{m} E_j \in \mathcal{C}_m \). Then, \( \mathcal{C}_m \) is a subalgebra on \( \prod_{j=1}^{m} X_j \), \( \mu \) and \( \nu \) satisfy the assumptions of Proposition 12.20. Therefore, there exists a unique \( K \)-valued measure \( \lambda := \prod_{j=1}^{m} \mu_j \) on the measurable space \( (\prod_{j=1}^{m} X_j, \mathcal{B}_m, \mathcal{P}) \) such that \( \lambda(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mu_j(E_j) \) and \( \mathcal{P} \circ \lambda(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mathcal{P} \circ \mu_j(E_j) \), \( \forall \prod_{j=1}^{m} E_j \in \mathcal{C}_m \). Furthermore, \( \lambda \) is finite. The \( K \)-valued measure space \( \mathcal{X} := (\prod_{j=1}^{m} X_j, \mathcal{B}_m, \lambda) \) is said to be the product \( K \)-valued measure space of \( X_1, \ldots, X_m \), and denoted by \( \prod_{j=1}^{m} X_j \). \( \lambda \) is said to be the product \( K \)-valued measure of \( \mu_1, \ldots, \mu_m \) and denoted by \( \prod_{j=1}^{m} \mu_j \).

**Proof** We will prove the result using mathematical induction on \( m \).
1° \( m = 1 \). The result is trivial and \( \prod_{j=1}^{1} X_1 = X_1 \) and \( \prod_{j=1}^{1} \mu_1 = \mu_1 \).

2° Assume that the result holds \( \forall m \in \{1, \ldots, k - 1\} \) with \( k \in \mathbb{N} \) and \( k \geq 2 \).

3° Consider the case where \( m = k \). Let \( \lambda := \prod_{j=1}^{k-1} \mu_j \) by the inductive assumption.

\[
\forall \prod_{j=1}^{k} E_{j,1}, \prod_{j=1}^{k} E_{j,2} \in \mathcal{E}_k, \text{ we have } (\prod_{j=1}^{k} E_{j,1}) \cap (\prod_{j=1}^{k} E_{j,2}) = \prod_{j=1}^{k} (E_{j,1} \cap E_{j,2}) \in \mathcal{E}_m. \text{ Note that } \left( \prod_{j=1}^{k} E_{j,1} \right) = \left( \left( \prod_{j=1}^{k-1} E_{j,1} \right) \times E_{k,1} \right) \cup \left( \left( \prod_{j=1}^{k-1} E_{j,1} \right) \times \bar{E}_{k,1} \right) \cup \left( \left( \prod_{j=1}^{k-1} E_{j,1} \right) \times \bar{E}_{k,1} \right). \text{ By inductive assumption, } \prod_{j=1}^{k-1} E_{j,1} \in \mathcal{E}_{k-1} \text{ and } \mathcal{E}_{k-1} \text{ is a semi-algebra on } \prod_{j=1}^{k-1} X_j, \text{ then } \left( \prod_{j=1}^{k-1} E_{j,1} \right) \text{ is a finite disjoint union of sets in } \mathcal{E}_{k-1}. \text{ Hence, } \left( \prod_{j=1}^{k-1} E_{j,1} \right) \text{ is a finite disjoint union of sets in } \mathcal{E}_k. \text{ Clearly, } \emptyset = \prod_{j=1}^{k} \emptyset \in \mathcal{E}_k \neq \emptyset. \text{ Hence, } \mathcal{E}_k \text{ is a semi-algebra on } X_1 \times X_2.

Next, we will show that \( \mu \) and \( \nu \) satisfy (i) – (vi) of Proposition 12.20.

\( \mu(\emptyset) = \mu(\prod_{j=1}^{k} \emptyset) = \prod_{j=1}^{k} \mu_j(\emptyset) = 0 \) and \( \nu(\emptyset) = \nu(\prod_{j=1}^{k} \emptyset) = \prod_{j=1}^{k} \mathcal{P} \circ \mu_j(\emptyset) = 0 \). Hence, (i) of Proposition 12.20 holds.

(iv) of Proposition 12.20 holds trivially since \( \nu(\prod_{j=1}^{k} E_{j}) = \prod_{j=1}^{k} \mathcal{P} \circ \mu_j(E_{j}) \leq \prod_{j=1}^{k} \mathcal{P} \circ \mu_j(X_j) < \infty, \forall \prod_{j=1}^{k} E_{j} \in \mathcal{E}_k. \)

\[
\forall \prod_{j=1}^{k} E_{j} \in \mathcal{E}_k, \forall \text{ pairwise disjoint } \left( \prod_{j=1}^{k} E_{j,i} \right) \subset \mathcal{E}_k \text{ with } \prod_{j=1}^{k} E_{j} = \bigcup_{i=1}^{\infty} (\prod_{j=1}^{k} E_{j,i}). \text{ Fix any } x_k \in E_k. \forall (x_1, \ldots, x_{k-1}) \in \prod_{j=1}^{k-1} E_{j}, \exists i \in \mathbb{N} \text{ such that } (x_1, \ldots, x_k) \in \prod_{j=1}^{k} E_{j,i}. \text{ Then, } \prod_{j=1}^{k-1} E_{j} = \bigcup_{i\in\mathbb{N}} x_k \in E_k, \prod_{j=1}^{k-1} E_{j,i} \text{ and the sets in the union are pairwise disjoint.}

Then, \[
\sum_{i\in\mathbb{N}} \chi_{E_k, x_k}(x_k) \prod_{j=1}^{k} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i\in\mathbb{N}, x_k \in E_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i\in\mathbb{N}, x_k \in E_k} \mathcal{P} \circ \lambda \left( \bigcup_{i\in\mathbb{N}, x_k \in E_k} \prod_{j=1}^{k-1} E_{j,i} \right) = \mathcal{P} \circ \lambda \left( \prod_{j=1}^{k-1} E_{j} \right) = \mathcal{P} \circ \mu_k(E_{k,i}) \subset [0, \infty) \subset \mathbb{R}, \text{ where the last four equalities follow from the inductive assumption.} \forall x_k \in X_k \setminus E_k, \text{ we have } \chi_{E_k, x_k}(x_k) \cdot \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j}) = 0 = \sum_{i\in\mathbb{N}} \chi_{E_k, x_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}). \text{ Then, } \forall x_k \in X_k, \chi_{E_k, x_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j}) = \sum_{i\in\mathbb{N}} \chi_{E_k, x_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}). \text{ By Monotone Convergence Theorem 11.81 and Proposition 11.75, } \nu(\prod_{j=1}^{k} E_{j}) = \prod_{j=1}^{k} \mathcal{P} \circ \mu_j(E_{j}) = \int_{X_k} \chi_{E_k, x_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) d\mathcal{P} \circ \mu_k = \int_{X_k} \sum_{i\in\mathbb{N}} \chi_{E_k, x_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) d\mathcal{P} \circ \mu_k = \sum_{i=1}^{\infty} (\prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i})) \mathcal{P} \circ \mu_k(E_{k,i}) = \sum_{i=1}^{\infty} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i=1}^{\infty} \nu(\prod_{j=1}^{k} E_{j,i}). \text{ Hence, (ii) and (iii) of Proposition 12.20 hold.} \)
\[ \forall \prod_{j=1}^k E_j \in \mathcal{E}_k, \forall n \in \mathbb{Z}_+, \forall \text{pairwise disjoint} \left( \prod_{j=1}^k E_{j,i} \right)_{i=1}^n \subseteq \mathcal{E}_k \]

with \( \prod_{j=1}^k E_j = \bigcup_{i=1}^n \prod_{j=1}^k E_{j,i} \). Fix any \( x_k \in E_k \). \( \forall (x_1, \ldots, x_{k-1}) \in \prod_{j=1}^{k-1} E_j, \exists i \in N := \{1, \ldots, n\} \) such that \( (x_1, \ldots, x_k) \in \prod_{j=1}^k E_{j,i} \). Then, \( \prod_{j=1}^{k-1} E_j = \bigcup_{i \in N, x_k \in E_{k,i}} \prod_{j=1}^{k-1} E_{j,i} \) and the sets in the union are pairwise disjoint. Then, \( \sum_{i \in N} \chi_{E_{k,i}}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N, x_k \in E_{k,i}} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N, x_k \in E_{k,i}} \lambda \left( \prod_{j=1}^{k-1} E_{j,i} \right) = \lambda \left( \bigcup_{i \in N} \prod_{j=1}^{k-1} E_{j,i} \right) = \lambda \left( \prod_{j=1}^{k-1} E_j \right) \]

where the last four equalities follow from the inductive assumption. \( \forall x_k \in X_k \setminus E_k \), we have \( \chi_{E_k \setminus X_k}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = 0 = \sum_{i \in N} \chi_{E_{k,i}}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \). Then, \( \forall x_1 \in X_1, \chi_{E_{k,i}}(x_k) \cdot \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N} \chi_{E_{k,i}}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \) by Propositions 11.125 and 11.132, we have \( \mu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_j(E_{j,i}) = \int_{X_k} \chi_{E_k \setminus X_k} \cdot \prod_{j=1}^{k-1} \mu_j(E_{j,i}) d\mu_k \]

\( = \int_{X_k} (\sum_{i \in N} \chi_{E_{k,i}}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i})) d\mu_k \)

\( = \sum_{i=1}^n \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i=1}^n \mu(\prod_{j=1}^k E_{j,i}) \). Hence, (v) of Proposition 12.20 holds.

\[ \forall \prod_{j=1}^k E_j \in \mathcal{E}_k, \forall n \in \mathbb{Z}_+, \forall \text{pairwise disjoint} \left( \prod_{j=1}^k E_{j,i} \right)_{i=1}^n \subseteq \mathcal{E}_k \]

with \( \prod_{j=1}^k E_j = \bigcup_{i=1}^n \prod_{j=1}^k E_{j,i} \), we have \( \sum_{i=1}^n \mu(\prod_{j=1}^k E_{j,i}) \leq \sum_{i=1}^n \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \leq \sum_{i=1}^n \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \leq \sum_{i=1}^n \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \) by the inequality follows from Definition 11.108; and the last equality follows from (ii). Then, \( \nu(\prod_{j=1}^k E_j) \geq s \prod_{j=1}^k E_j \) := \( \sup_{n \in \mathbb{Z}_+} (\sum_{i=1}^n \prod_{j=1}^k E_{j,i})^{1/n} \)

On the other hand, \( \forall \epsilon \in (0, 1] \subseteq \mathbb{R}, \forall j \in \{1, \ldots, k\}, \exists n_j \in \mathbb{Z}_+, \exists \text{pairwise disjoint} (E_{j,i})_{i=1}^{n_j} \subseteq B_j \) with \( E_j = \bigcup_{i=1}^{n_j} E_{j,i} \) such that \( \sum_{i=1}^{n_j} \mu_j(E_{j,i}) \geq \mathcal{P} \circ \mu_j(E_{j,i}) \)\). Then, \( \prod_{j=1}^k E_j = \bigcup_{i_1=1}^{n_1} \cdots \bigcup_{i_k=1}^{n_k} \prod_{j=1}^{n_j} E_{j,i_j} \) and the sets in the union are pairwise disjoint and in \( \mathcal{E}_k \). Then, \( \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} \mu(\prod_{j=1}^k E_{j,i_j}) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} \prod_{j=1}^{n_j} \mu_j(E_{j,i_j}) \)

\( \geq \left( 1 - \frac{\epsilon}{\mathcal{P}(\prod_{j=1}^k E_j)} \right) \prod_{j=1}^k \mu_j(E_{j,i_j}) \) by the arbitrariness of \( \epsilon \), we have \( s \prod_{j=1}^k E_j \geq \nu(\prod_{j=1}^k E_j) \).

Hence, (vi) of Proposition 12.20 holds.

Let \( A_k \) be the algebra on \( \prod_{j=1}^k X_j \) generated by \( \mathcal{E}_k \). By Proposition 12.20, \( \nu \) admits a unique extension to a measure \( \bar{\nu} \) on the algebra \( A_k \); and \( \mu \) admits a unique extension to a \( \mathbb{K} \)-valued measure \( \bar{\mu} \) on the
algebra \( \mathcal{A}_k \) with \( (\mathcal{P} \circ \hat{\mu})|_{\mathcal{C}_k} = \nu \). Furthermore, \( \mathcal{P} \circ \hat{\mu} = \hat{\nu} \). Clearly, \( \hat{\mu} \) is finite. By Carathéodory Extension Theorem 12.4, there is a unique \( \mathbb{K} \)-valued measure \( \lambda \) on the measurable space \( (\prod_{j=1}^k X_j, \mathcal{B}_k) \) such that \( \lambda|_{\mathcal{A}} = \hat{\mu} \) and \( (\mathcal{P} \circ \lambda)|_{\mathcal{A}} = \mathcal{P} \circ \hat{\mu} = \hat{\nu} \). Then, \( \lambda|_{\mathcal{C}_k} = \mu \) and \( (\mathcal{P} \circ \lambda)|_{\mathcal{C}_k} = \nu \). This implies that \( \mathcal{P} \circ \lambda(\prod_{j=1}^k X_j) = \nu(\prod_{j=1}^k X_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(X_j) < \infty \). Hence, \( \lambda \) is finite.

This completes the induction process.

This completes the proof of the proposition. \( \square \)

**Proposition 12.22** Let \( k \in \mathbb{N} \), \( X_j := (X_j, \mathcal{B}_j, \mu_j) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( \forall j \in \{1, \ldots, k\} \), \( \mathcal{C} := \left\{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \ldots, k\} \right\} \) be the set of measurable rectangles, \( \mathcal{B} \) be the \( \sigma \)-algebra on \( \prod_{j=1}^k X_j \) generated by \( \mathcal{C} \). Then, there is a unique \( \mathbb{K} \)-valued measure \( \lambda := \prod_{j=1}^k \mu_j \) on the measurable space \( (\prod_{j=1}^k X_j, \mathcal{B}) \) such that \( \lambda(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_j(E_j) \) and \( \mathcal{P} \circ \lambda(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j), \forall E_j \in \text{dom}(\mu_j), \forall j \in \{1, \ldots, k\} \). Furthermore, \( \lambda \) is \( \sigma \)-finite. The \( \mathbb{K} \)-valued measure space \( \mathcal{X} := \left( \prod_{j=1}^k X_j, \mathcal{B}, \lambda \right) \) is said to be the product \( \mathbb{K} \)-valued measure space of \( X_1, \ldots, X_k \) and denoted by \( \prod_{j=1}^k X_j \). \( \lambda \) is said to be the product \( \mathbb{K} \)-valued measure of \( \mu_1, \ldots, \mu_k \) and denoted by \( \prod_{j=1}^k \mu_j \).

**Proof** Fix any \( j \in \{1, \ldots, k\} \). Since \( X_j \) is \( \sigma \)-finite, then \( \exists (X_j,n)_{n=1}^\infty \subseteq \mathcal{B}_j \) such that \( X_j = \bigcup_{n=1}^\infty X_j,n \) and \( \mathcal{P} \circ \mu_j(X_j,n) < \infty, \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( (X_j,n)_{n=1}^\infty \) is pairwise disjoint. Let \( X_{j,n} := (X_j,n, \mathcal{B}_{j,n}, \mu_{j,n}) \) be the finite \( \mathbb{K} \)-valued measure subspace of \( X_j, \forall n \in \mathbb{N} \). Let \( \mathcal{C}_{n_1,\ldots,n_k} := \left\{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_j \mid E_j \in \mathcal{B}_{j,n_j}, \forall j = 1, \ldots, k \right\} \) and \( \mathcal{B}_{n_1,\ldots,n_k} \) be the \( \sigma \)-algebra on \( \prod_{j=1}^k X_j \) generated by \( \mathcal{C}_{n_1,\ldots,n_k} \). By Proposition 12.21, there is a unique \( \mathbb{K} \)-valued measure \( \lambda_{n_1,\ldots,n_k} = \prod_{j=1}^k \mu_{j,n_j} \) on \( (\prod_{j=1}^k X_{j,n_j}, \mathcal{B}_{n_1,\ldots,n_k}) \) such that \( \lambda_{n_1,\ldots,n_k}(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mu_{j,n_j}(E_j) \) and \( \mathcal{P} \circ \lambda_{n_1,\ldots,n_k}(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_{j,n_j}(E_j), \forall E_j \in \mathcal{C}_{n_1,\ldots,n_k} \). Furthermore, \( \lambda_{n_1,\ldots,n_k} \) is finite. Denote the finite \( \mathbb{K} \)-valued measure space \( (\prod_{j=1}^k X_{j,n_j}, \mathcal{B}_{n_1,\ldots,n_k}, \lambda_{n_1,\ldots,n_k}) =: \mathcal{X}_{n_1,\ldots,n_k} \). By Proposition 11.118, the generation process on \( (\mathcal{X}_{n_1,\ldots,n_k})_{n_j \in \mathbb{N}, j \in \{1, \ldots, k\}} \) yields a unique \( \sigma \)-finite \( \mathbb{K} \)-valued measure space \( \mathcal{X} := \left( \prod_{j=1}^k X_j, \mathcal{B}, \lambda \right) \) on \( \prod_{j=1}^k X_j \) such that \( \mathcal{X}_{n_1,\ldots,n_k} \) is the finite \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X}, \forall n_1, \ldots, n_k \in \mathbb{N} \), where \( \mathcal{B} := \left\{ E \subseteq \prod_{j=1}^k X_j \mid E \cap (\prod_{j=1}^k X_{j,n_j}) \in \mathcal{B}_{n_1,\ldots,n_k}, \forall n_j \in \mathbb{N}, \forall j \in \{1, \ldots, k\} \right\}, \mathcal{P} \circ \lambda(E) = \sum_{n_1=1}^\infty \cdots \sum_{n_k=1}^\infty \mathcal{P} \circ \lambda_{n_1,\ldots,n_k}(E \cap (\prod_{j=1}^k X_{j,n_j})), \forall E \in \mathcal{B}, \lambda(E) = \sum_{n_1=1}^\infty \cdots \sum_{n_k=1}^\infty \lambda_{n_1,\ldots,n_k}(E \cap (\prod_{j=1}^k X_{j,n_j})), \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \lambda(E) < \infty \).
Claim 12.22.1 \( \mathcal{B} = \overline{\mathcal{B}} \).

**Proof of claim:** \( \forall \prod_{j=1}^{k} E_j \in \mathcal{E}, \forall n_1, \ldots, n_k \in \mathbb{N}, E_j \cap X_{j,n_j} \in \mathcal{B}_{j,n_j}, \forall j \in \{1, \ldots, k\}, \) by Proposition 11.115. Then, \( \prod_{j=1}^{k} E_j \cap \prod_{j=1}^{k} X_{j,n_j} \in \mathcal{E}_{n_1, \ldots, n_k} \subseteq \overline{\mathcal{B}}_{n_1, \ldots, n_k} \). This implies that \( \prod_{j=1}^{k} E_j \in \overline{\mathcal{B}} \). Then, \( \mathcal{E} \subseteq \overline{\mathcal{B}} \). Since \( \overline{\mathcal{B}} \) is a \( \sigma \)-algebra, then \( \mathcal{B} \subseteq \overline{\mathcal{B}} \).

On the other hand, \( \forall n_1, \ldots, n_k \in \mathbb{N}, \forall \prod_{j=1}^{k} E_j \in \mathcal{E}_{n_1, \ldots, n_k}, \) we have \( E_j \in \mathcal{B}_{j,n_j} \subseteq \mathcal{B}_j, \forall j \in \{1, \ldots, k\}. \) Thus, \( \prod_{j=1}^{k} E_j \in \mathcal{E} \subseteq \mathcal{B}. \) Then, \( \mathcal{E}_{n_1, \ldots, n_k} \subseteq \mathcal{B} \) and \( \overline{\mathcal{B}}_{n_1, \ldots, n_k} \subseteq \mathcal{B}. \) This further implies that \( \mathcal{B} \subseteq \mathcal{B}. \) Hence, \( \mathcal{B} = \overline{\mathcal{B}}. \) This completes the proof of the claim. \( \square \)

\[ \forall j \in \{1, \ldots, k\}, \forall E_j \in \text{dom}(\mu_j), \text{we have} \infty > \mathcal{P} \circ \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_j, n_j(E_j \cap X_{j,n_j}), \mathbb{K} \ni \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mu_j(E_j \cap X_{j,n_j}) = \sum_{n_j=1}^{\infty} \mu_j, n_j(E_j \cap X_{j,n_j}). \] Then,

\[
\mathcal{P} \circ \lambda \left( \prod_{j=1}^{k} E_j \right) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^{k} E_j \cap \prod_{j=1}^{k} X_{j,n_j} \right)
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^{k} (E_j \cap X_{j,n_j}) \right)
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^{k} \mathcal{P} \circ \mu_{j, n_j} (E_j \cap X_{j,n_j})
= \prod_{j=1}^{k} \left( \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_{j, n_j} (E_j \cap X_{j,n_j}) \right) = \prod_{j=1}^{k} \mathcal{P} \circ \mu_j (E_j)
< \infty
\]

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. We also have

\[
\lambda \left( \prod_{j=1}^{k} E_j \right) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^{k} E_j \cap \prod_{j=1}^{k} X_{j,n_j} \right)
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^{k} (E_j \cap X_{j,n_j}) \right)
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^{k} \mu_{j, n_j} (E_j \cap X_{j,n_j})
= \prod_{j=1}^{k} \left( \sum_{n_j=1}^{\infty} \mu_{j, n_j} (E_j \cap X_{j,n_j}) \right) = \prod_{j=1}^{k} \mu_j (E_j) \in \mathbb{K}
\]

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. Hence, \( \mathcal{X} \) is the \( \sigma \)-finite \( \mathbb{K} \)-valued
measure space that we seek.

This completes the proof of the proposition. \(\square\)

**Proposition 12.23** Let \(m \in \mathbb{N}\), \(\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)\) be a \(\sigma\)-finite \(\mathbb{K}\)-valued measure space, \(\mathcal{X} := (X, \mathcal{B}, \lambda) := \prod_{i=1}^m \mathcal{X}_i\) be the \(\sigma\)-finite product \(\mathbb{K}\)-valued measure space of \(\mathcal{X}_1, \ldots, \mathcal{X}_m\). Then, \(\mathcal{X} := \prod_{i=1}^m \mathcal{X}_i, \mathcal{B} \lambda)\) is the \(\mathbb{K}\)-valued measure space of \(\mathcal{X}\) if, and only if, \(\mathcal{X} = \prod_{i=1}^m \mathcal{X}_i\).

**Proof** Let \(\mathcal{X} := \prod_{i=1}^m \mathcal{X}_i, \mathcal{B} \lambda)\) be the \(\sigma\)-finite \(\mathbb{K}\)-valued measure subspace of \(\mathcal{X}\) and \(\tilde{\mathcal{X}} := \prod_{i=1}^m \tilde{X}_i := (\prod_{i=1}^m \tilde{X}_i, \tilde{\mathcal{B}}, \tilde{\lambda})\) be the \(\sigma\)-finite product \(\mathbb{K}\)-valued measure space of \(\tilde{X}_1, \ldots, \tilde{X}_m\). All we need to show is that \(\mathcal{X} = \tilde{\mathcal{X}}\).

Let \(\tilde{X}_{i,1} := \tilde{X}_i\) and \(\tilde{X}_{i,2} := X_i \setminus \tilde{X}_i, i = 1, \ldots, m\). Then, \(\mathcal{E}_{n_1, \ldots, n_m} := \{\prod_{i=1}^m B_i \subseteq \prod_{i=1}^m \tilde{X}_{i, n_i} \mid B_i \subseteq \mathcal{B}_i, \forall i \in \{1, \ldots, m\}\}\) be the set of measurable rectangles on \(\prod_{i=1}^m \tilde{X}_{i, n_i}\). \(\tilde{\mathcal{B}}_{n_1, \ldots, n_m}\) be the \(\sigma\)-algebra generated by \(\mathcal{E}_{n_1, \ldots, n_m}\) on \(\prod_{i=1}^m \tilde{X}_{i, n_i}, \forall n_1, \ldots, n_m \in \{1, 2\}\), and \(\mathcal{E} := \{\prod_{i=1}^m B_i \subseteq \prod_{i=1}^m \tilde{X}_i \mid B_i \subseteq \mathcal{B}_i, \forall i \in \{1, \ldots, m\}\}\) be the set of measurable rectangles on \(\prod_{i=1}^m \tilde{X}_i\). Then, by Proposition 12.22, \(\mathcal{B}\) is the \(\sigma\)-algebra generated by \(\mathcal{E}\) on \(\prod_{i=1}^m \tilde{X}_i\).

**Claim 12.23.1** \(\mathcal{B} = \{E \subseteq \prod_{i=1}^m X_i \mid E \cap (\prod_{i=1}^m \tilde{X}_{i, n_i}) \subseteq \tilde{B}_{n_1, \ldots, n_m}, \forall n_i \in \{1, 2\}, \forall i \in \{1, \ldots, m\}\} =: \tilde{\mathcal{B}}\).

**Proof of claim:** \(\forall \prod_{i=1}^m B_i \subseteq \mathcal{B}\), we have \(B_i \subseteq \mathcal{B}_i, i = 1, \ldots, m\), which implies that \(B_i \subseteq B_i \cap \tilde{X}_{i, n_i} \subseteq \tilde{X}_{i, n_i}, n_i = 1, 2, i = 1, \ldots, m\). Hence, \(\prod_{i=1}^m B_i \subseteq \tilde{\mathcal{B}}\). By the arbitrariness of \(\prod_{i=1}^m B_i\), we have \(\mathcal{B} \subseteq \tilde{\mathcal{B}}\).

Now, we will show that \(\tilde{\mathcal{B}}\) is a \(\sigma\)-algebra on \(\prod_{i=1}^m \tilde{X}_i\). Clearly, \(\emptyset = \prod_{i=1}^m \emptyset \subseteq \mathcal{E} \subseteq \mathcal{B}\) and \(\prod_{i=1}^m X_i \in \mathcal{E} \subseteq \mathcal{B}\). \(\forall (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}, \forall n_i = 1, 2, \forall i = 1, \ldots, m, \forall n \in \mathbb{N}\), by the definition of \(\tilde{\mathcal{B}}\), \(E_n \cap (\prod_{i=1}^m \tilde{X}_{i, n_i}) \subseteq \prod_{i=1}^m \tilde{X}_{i, n_i}\), which implies \((\bigcup_{n=1}^{\infty} E_n) \cap (\prod_{i=1}^m \tilde{X}_{i, n_i}) = \bigcup_{n=1}^{\infty} (E_n \cap (\prod_{i=1}^m \tilde{X}_{i, n_i})) \subseteq \prod_{i=1}^m \tilde{X}_{i, n_i}\). Since \(\tilde{\mathcal{B}}_{n_1, \ldots, n_m}\) is a \(\sigma\)-algebra, \((\bigcup_{n=1}^{\infty} E_n) \subseteq \tilde{\mathcal{B}}_{n_1, \ldots, n_m}\). Hence, \((\bigcup_{n=1}^{\infty} E_n) \subseteq \mathcal{B}\). \(\forall E \subseteq \mathcal{B}, \forall n_i = 1, 2, \forall i = 1, \ldots, m, \forall (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}\), then \((\prod_{i=1}^m \tilde{X}_i) \setminus E \cap (\prod_{i=1}^m \tilde{X}_{i, n_i}) \subseteq \prod_{i=1}^m \tilde{X}_i \setminus \mathcal{B}_{n_1, \ldots, n_m}\). This implies that \((\prod_{i=1}^m \tilde{X}_i) \setminus E \subseteq \mathcal{B}\). Hence, \(\tilde{\mathcal{B}}\) is a \(\sigma\)-algebra on \(\prod_{i=1}^m \tilde{X}_i\).

Since \(\mathcal{B}\) is the \(\sigma\)-algebra on \(\prod_{i=1}^m \tilde{X}_i\) generated by \(\mathcal{E}\) and \(\mathcal{E} \subseteq \mathcal{B}\), then \(\mathcal{B} \subseteq \tilde{\mathcal{B}}\).

On the other hand, \(\tilde{\mathcal{B}}_{n_1, \ldots, n_m} \subseteq \mathcal{E}\), \(n_i = 1, 2, i = 1, \ldots, m\), implies that \(\tilde{\mathcal{B}}_{n_1, \ldots, n_m} \subseteq \mathcal{B}\). \(\forall E \subseteq \mathcal{B}\), we have \(E \cap (\prod_{i=1}^m \tilde{X}_{i, n_i}) \subseteq \tilde{\mathcal{B}}_{n_1, \ldots, n_m} \subseteq \mathcal{B}\), \(n_i = 1, 2, i = 1, \ldots, m\), which implies that \(E = \bigcup_{n_1=1}^{2} \cdots \bigcup_{n_m=1}^{2} (E \cap (\prod_{i=1}^m \tilde{X}_{i, n_i})) \subseteq \mathcal{B}\). Hence, \(\mathcal{B} \subseteq \mathcal{B}\). Therefore, \(\mathcal{B} = \tilde{\mathcal{B}}\). This completes the proof of the claim. \(\square\)
12.3. PRODUCT MEASURE

Note that \( \tilde{B} = \tilde{B}_{1, \ldots, l} = \{ E \subseteq \prod_{i=1}^{m} X_i \mid E \in \tilde{B} \} = \{ E \subseteq \prod_{i=1}^{m} \tilde{X}_i \mid E \in \tilde{B} \} = \tilde{B} \), where the first equality follows from Proposition 12.22; the second equality follows from the definition of \( \tilde{B} \); the third equality follows from Claim 12.23.1; and the last equality follows from Proposition 11.115.

Finally, we will show that \( \check{\lambda} = \check{\lambda} \) by Propositions 12.22 and 11.115. By Proposition 11.115 and Claim 12.23.1, \( \lambda \) is a \( K \)-valued measure on \( (\prod_{i=1}^{m} X_i, \mathcal{B}) \), \( \check{\lambda} = \lambda_{|G} = \lambda_{|G} \), and \( \mathcal{P} \circ \check{\lambda} = (\mathcal{P} \circ \lambda)|_G \). By Proposition 12.22, \( \check{\lambda} \) is the unique \( K \)-valued measure on \( (\prod_{i=1}^{m} X_i, \mathcal{B}) \) such that \( \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mu_i(E_i) \) and \( \mathcal{P} \circ \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \mu_i(E_i), \forall E_i \in \text{dom}(\mu_i), \forall i \in \{1, \ldots, m\} \). Then, \( \forall i \in \{1, \ldots, m\}, \forall E_i \in \text{dom}(\mu_i) \), by Proposition 11.115, \( \mu_i = \mu_{|G} \) and \( \mathcal{P} \circ \mu_i = (\mathcal{P} \circ \mu_i)|_G \). Hence, \( E_i \in \text{dom}(\mu_i) \). This implies that \( \check{\lambda}(\prod_{i=1}^{m} E_i) = \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mu_i(E_i) \) and \( \mathcal{P} \circ \check{\lambda}(\prod_{i=1}^{m} E_i) = \mathcal{P} \circ \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \mu_i(E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \mu_i(E_i) \).

By Proposition 12.22, \( \check{\lambda} \) is the unique \( K \)-valued measure on \( (\prod_{i=1}^{m} \tilde{X}_i, \tilde{B}) \) such that \( \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \tilde{\mu}_i(E_i) \) and \( \mathcal{P} \circ \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \tilde{\mu}_i(E_i), \forall E_i \in \text{dom}(\tilde{\mu}_i), \forall i \in \{1, \ldots, m\} \). Hence, we have \( \lambda = \check{\lambda} \).

Hence, \( \hat{X} = \hat{X} \). This completes the proof of the proposition. 

**Proposition 12.24** Let \( \Gamma \) be a finite nonempty index set, \( \Lambda_\beta \) be a finite nonempty index set, \( \forall \beta \in \Gamma \), with \( \Lambda_\beta \)’s being pairwise disjoint, \( \Lambda := \bigcup_{\beta \in \Gamma} \Lambda_\beta, X_\alpha := (X_\alpha, \mathcal{B}_\alpha, \mu_\alpha) \) be a \( \sigma \)-finite \( K \)-valued measure space, \( \forall \alpha \in \Lambda, \hat{X}_\beta := \prod_{\alpha \in \Lambda_\beta} X_\alpha := (\prod_{\alpha \in \Lambda_\beta} X_\alpha, \hat{B}_\beta, \check{\mu}_\beta) \) be the \( \sigma \)-finite product \( K \)-valued measure space, \( \forall \beta \in \Gamma, \hat{X} := \prod_{\beta \in \Gamma} \hat{X}_\beta := (\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \hat{B}, \check{\mu}) \) be the \( \sigma \)-finite product \( K \)-valued measure space, and \( \hat{X} := \prod_{\alpha \in \Lambda} X_\alpha := (\prod_{\alpha \in \Lambda} X_\alpha, \hat{B}, \check{\mu}) \) be the \( \sigma \)-finite product \( K \)-valued measure space. Then, \( \hat{X} \) and \( \hat{X} \) are isomeric.

**Proof** Define \( \Psi : \hat{X} \to \hat{X} \) by, \( \forall x \in \hat{X}, \forall \alpha \in \Lambda, \exists! \beta_\alpha \in \Gamma \exists! \alpha \in \Lambda_\beta, \pi_\alpha(\Psi(x)) = \pi_\alpha(\beta_\alpha(\pi_\beta(\alpha))) \). By the proof of Proposition 3.30, \( \Psi \) is bijective with inverse \( \Psi_{\text{inv}} : \hat{X} \to \hat{X} \).

Let \( \hat{E}_\beta := \{ \prod_{\alpha \in \Lambda_\beta} E_\alpha \subseteq \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda_\beta \}, \forall \beta \in \Gamma, \hat{E} := \{ \prod_{\beta \in \Gamma} E_\beta \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\beta \in \hat{B}_\beta, \forall \beta \in \Gamma \}, \check{E} := \{ \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda \}, \text{ and } \hat{E} := \{ \prod_{\alpha \in \Lambda} E_\alpha \subseteq \prod_{\alpha \in \Lambda} X_\alpha \mid E_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda \}. \) By Proposition 12.22, \( \hat{B}_\beta \) is the \( \sigma \)-algebra generated by \( \hat{E}_\beta \) on \( \prod_{\alpha \in \Lambda_\beta} X_\alpha, \forall \beta \in \Gamma \), \( \hat{B} \) is the \( \sigma \)-algebra generated by \( \check{E} \) on \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \), and \( \hat{B} \) is the \( \sigma \)-algebra generated by \( \hat{E} \) on \( \prod_{\alpha \in \Lambda} X_\alpha \). Clearly, \( \hat{E} \subseteq \hat{E} \).

Fix any \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \in \hat{E} \), we will show that \( \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) = \prod_{\alpha \in \Lambda} E_\alpha. \forall x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha, \forall \alpha \in \Lambda, \) we have \( \pi_\alpha(\Psi(x)) = \)
By the previous paragraph, we have $\bar{\alpha}$.

Hence, $\Psi(x) \in \prod_{\alpha \in \Lambda} E_{\alpha}$. This implies that $\Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha}) \subseteq \prod_{\alpha \in \Lambda} E_{\alpha}$. On the other hand, $\forall y \in \prod_{\alpha \in \Lambda} E_{\alpha}, \forall \beta \in \Gamma, \forall \alpha \in \Lambda_{\beta}$, we have $\pi_{\alpha}(\pi_{\beta}^{(T)}(\Psi(\psi_{\text{inv}}(y)))) = \pi_{\alpha}(y) \in E_{\alpha}$. Then, $\Psi_{\text{inv}}(y) \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha}$. This leads to $\Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_{\alpha}) \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha}$. This is equivalent to $\prod_{\alpha \in \Lambda} E_{\alpha} \subseteq \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha})$. Therefore, $\Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha}) = \prod_{\alpha \in \Lambda} E_{\alpha} \in \tilde{c}$. Conversely, $\forall \prod_{\alpha \in \Lambda} E_{\alpha} \in \tilde{c}$, we have $\Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_{\alpha}) = \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha} \in \tilde{c}$.

Let $\tilde{B}$ be the $\sigma$-algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}$ generated by $\tilde{c}$. We will show that $\tilde{B} = \{ E \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \mid \Psi(E) \in \tilde{B} \} =: \hat{B}$. Since $\tilde{c} \subseteq \tilde{B}$, by the previous paragraph, we have $\hat{c} \subseteq \hat{B}$. It is easy to show that $\hat{B}$ is a $\sigma$-algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}$ since $\tilde{B}$ is a $\sigma$-algebra. Then, $\hat{B} \subseteq \hat{B}$. On the other hand, let $\hat{B} := \{ E \subseteq \prod_{\alpha \in \Lambda} X_{\alpha} \mid \Psi_{\text{inv}}(E) \in \hat{B} \}$. Since $\hat{B}$ is a $\sigma$-algebra on $\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}$, then $\hat{B}$ is a $\sigma$-algebra on $\prod_{\alpha \in \Lambda} X_{\alpha}$. Since $\hat{c} \subseteq \hat{B}$, by the previous paragraph, then $\hat{c} \subseteq \hat{B}$. By the definition of $\hat{B}$, we have $B \subseteq \hat{B}$. Since $\hat{B} \subseteq \hat{B}$, then, by the definitions of $B$ and $\hat{B}$, $\hat{B} \subseteq B$. Thus, we have $B = \hat{B} = \hat{B}$. We can now prove the following result.

Claim 12.24.1 $\hat{c} \subseteq \hat{B}$.

Proof of claim: Without loss of generality, assume that $\Gamma = \{ \beta_{1}, \ldots, \beta_{n} \}$ for some $n \in \mathbb{N}$. We will prove the following statement:

1. $\forall i \in \{1, \ldots, n+1\}, \forall E_{\beta_{j}} := \prod_{\alpha \in \Lambda_{\beta_{j}}} E_{\alpha}, j = i, \ldots, n, \forall E_{\beta_{j}} \in \hat{B}_{\beta_{j}}, j = 1, \ldots, i-1$, we have $\prod_{j=1}^{n} E_{\beta_{j}} \in \hat{c}$.

by mathematical induction on $i$.

1\circ $i = 1$. $\forall E_{\beta_{j}} := \prod_{\alpha \in \Lambda_{\beta_{j}}} E_{\alpha} \in \hat{c}_{\beta_{j}}, j = 1, \ldots, n$, we have $\prod_{j=1}^{n} E_{\beta_{j}} = \prod_{j=1}^{n} \prod_{\alpha \in \Lambda_{\beta_{j}}} E_{\alpha} \in \hat{c} \subseteq \hat{B}$.

This case is proved.

2\circ Assume that the statement holds for $i = k \in \{1, \ldots, n\}$.

3\circ Consider the case $i = k + 1$. Fix any $E_{\beta_{j}} := \prod_{\alpha \in \Lambda_{\beta_{j}}} E_{\alpha} \in \hat{c}_{\beta_{j}}, j = k+1, \ldots, n$, and any $E_{\beta_{k}} \in \hat{B}_{\beta_{k}}, j = 1, \ldots, k-1$. Define $F := \{ E_{\beta_{k}} \in \hat{B}_{\beta_{k}} \mid \prod_{j=1}^{k-1} E_{\beta_{j}} \in \hat{B} \}$. By the inductive assumption, we have $\hat{c}_{\beta_{k}} \subseteq F$. Then, $\emptyset = \prod_{\alpha \in \Lambda_{\beta_{k}}} E_{\alpha} \in \hat{c}_{\beta_{k}} \subseteq F$ and $X_{\beta_{k}} := \prod_{\alpha \in \Lambda_{\beta_{k}}} X_{\alpha} \in \hat{c}_{\beta_{k}} \subseteq F$. $\forall E_{\beta_{k}} \in F$, we have $E_{\beta_{k}} \in \hat{B}_{\beta_{k}}$ and $\prod_{j=1}^{n} E_{\beta_{j}} \in \hat{B}$. By $X_{\beta_{k}} \in \hat{c}_{\beta_{k}} \subseteq F$, we have $(\prod_{j=1}^{k-1} E_{\beta_{j}}) \times X_{\beta_{k}} \times (\prod_{j=k+1}^{n} E_{\beta_{j}}) \in \hat{B}$. This implies that $(\prod_{j=1}^{k-1} E_{\beta_{j}}) \times X_{\beta_{k}} \times (\prod_{j=k+1}^{n} E_{\beta_{j}}) \in \hat{B}$. Then, $X_{\beta_{k}} \setminus E_{\beta_{k}} = \hat{c}$.
Therefore, the statement holds for $B_1$ and $F$ is a $\sigma$-algebra on $\prod_{\alpha \in \Lambda_\beta} X_\alpha$. Since $\hat{B}_\beta \supseteq F \supseteq \bar{\sigma}_{\beta_k}$, then $F = \hat{B}_\beta$. Therefore, the statement holds for $i = k + 1$.

This completes the induction process and the statement holds for $i = 1, \ldots, n + 1$.

By the above statement for $i = n + 1$ case, we have $\forall E_{\beta, j} \in \hat{B}_{\beta}, j = 1, \ldots, n, \prod_{j=1}^n E_{\beta, j} \in \hat{B}$. Hence, $\hat{C} \subseteq \hat{B}$. This completes the proof of the claim.

By Claim 12.24.1 and the definition of $\hat{B}$, we have $\hat{B} \subseteq \hat{C}$. Since $\hat{C} \subseteq \hat{C}$, we have $\hat{B} \subseteq \hat{B}$. Therefore, $\hat{B} = \hat{B}$. This further implies that, $\forall E \in \hat{B}, \Psi_{\text{inv}}(E) \in \hat{B} = \hat{B} = \hat{B}$ and, $\forall E \in \hat{B} = \hat{B}, \Psi(E) \in \hat{B}$.

By Proposition 12.10, $\Psi$ is an isomeasure between $\hat{X}$ and $(\prod_{\alpha \in \Lambda} X_\alpha, \hat{B}, \Psi, \mu)$. $\forall \alpha \in \Lambda, \forall E_\alpha \in \text{dom}(\mu_\alpha)$, we have

$$P \circ (\Psi \mu) \left( \prod_{\alpha \in \Lambda} E_\alpha \right) = \Psi_{\text{inv}} \left( \prod_{\alpha \in \Lambda} P \circ \mu \right) \left( \prod_{\alpha \in \Lambda} E_\alpha \right) = \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_\alpha$$

and

$$\Psi \mu \left( \prod_{\alpha \in \Lambda} E_\alpha \right) = \mu \left( \Psi_{\text{inv}} \left( \prod_{\alpha \in \Lambda} E_\alpha \right) \right) = \mu \left( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_\alpha \right)$$

By Proposition 12.22, $\mu$ is the unique $K$-valued measure on $(\prod_{\alpha \in \Lambda} X_\alpha, \hat{B})$ such that $\mu(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\alpha \in \Lambda} \mu_\alpha(E_\alpha)$ and $P \circ \mu(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\alpha \in \Lambda} P \circ \mu_\alpha(E_\alpha), \forall E_\alpha \in \text{dom}(\mu_\alpha), \forall \alpha \in \Lambda$. Therefore, $\mu = \Psi \mu$. This implies that $\Psi$ is an isomeasure between $X$ and $\hat{X}$. This completes the proof of the proposition. \qed

**Proposition 12.25** Let $X_i := (X_i, B_i, \mu_i)$ be a $\sigma$-finite measure space, $i = 1, 2$, and $X := (X_1 \times X_2, B, \mu) := X_1 \times X_2$ be the product measure space. Then, $\forall E \in B, \forall x_1 \in X_1, \forall x_2 \in X_2, \forall E_{x_1} := \{ x_2 \in X_2 \mid (x_1, x_2) \in E \} \in B_2$; and, $\forall x_2 \in X_2, \forall E_{x_2} := \{ x_1 \in X_1 \mid (x_1, x_2) \in E \} \in B_1$. 
Proof. Let $\mathcal{E} := \{ E \in \mathcal{B} \mid E_{x_1} \in \mathcal{B}_2, \forall x_1 \in X_1 \} \subseteq \mathcal{B}$. \( \forall E \in \mathcal{C} := \{ E_1 \times E_2 \subseteq X_1 \times X_2 \mid E_1 \in \mathcal{B}_1, E_2 \in \mathcal{B}_2 \} \), we have $E \in \mathcal{E}$. Then, $\mathcal{C} \subseteq \mathcal{E}$. We will show that $\mathcal{E}$ is a $\sigma$-algebra on $X_1 \times X_2$. Then, by Proposition 12.22, we have $\mathcal{B} \subseteq \mathcal{E}$ and $\mathcal{B} = \mathcal{E}$.

Clearly, $\emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. \( \forall E \in \mathcal{E} \), \( \forall x_1 \in X_1 \), we have \((X_1 \times X_2) \setminus E\) and $E_{x_1} \in \mathcal{B}_2$ since $\mathcal{B}_2$ is a $\sigma$-algebra. By the arbitrariness of $x_1$, we have $(X_1 \times X_2) \setminus E \in \mathcal{E}$. \( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E} \), let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_1$. \( \forall x_1 \in X_1 \), we have $E_{x_1} = \bigcup_{i=1}^{\infty} E_{i,x_1} \in \mathcal{B}_2$ since $\mathcal{B}_2$ is a $\sigma$-algebra. Then, $E \in \mathcal{E}$. Therefore, $\mathcal{E}$ is a $\sigma$-algebra and $\mathcal{B} = \mathcal{E}$. By the definition of $\mathcal{E}$, we have $\forall E \in \mathcal{B}, \forall x_1 \in X_1, E_{x_1} \in \mathcal{B}_2$. By symmetry, $\forall E \in \mathcal{B}, \forall x_2 \in X_2, E_{x_2} \in \mathcal{B}_1$. This completes the proof of the proposition.

Proposition 12.26 Let $\mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i)$ be a $\sigma$-finite measure space, $i = 1, 2$, and $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2$ be the product measure space. Then, $\forall \mu \in \mathcal{B}$, the function $g_\mu : X_1 \to [0, \infty] \subseteq \mathcal{R}$, defined by $g_\mu(x_1) = \mu_2(E_{x_1})$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$-measurable, where $E_{x_1}$ is as defined in Proposition 12.25. By symmetry, $\forall \mu \in \mathcal{B}$, the function $h_\mu : X_2 \to [0, \infty] \subseteq \mathcal{R}$, defined by $h_\mu(x_2) = \mu_1(E_{x_2})$, $\forall x_2 \in X_2$, is $\mathcal{B}_2$-measurable, where $E_{x_2}$ is as defined in Proposition 12.25.

Proof. By Proposition 12.25, $\forall x_1 \in X_1, E_{x_1} \in \mathcal{B}_2$. Then, $g_\mu(x_1) \in [0, \infty] \subseteq \mathcal{R}$. Hence, the function $g_\mu$ is well-defined.

First, consider the special case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are finite. Let $\mathcal{E} := \{ E \in \mathcal{B} \mid g_\mu \text{ is } \mathcal{B}_1\text{-measurable} \}$, and $\mathcal{C}$ be the set of measurable rectangles in $X_1 \times X_2$. \( \forall E \in \mathcal{C}, \forall x_1 \in X_1, \mu_2(E_{x_1}) \subseteq \mathcal{B}_1\text{-measurable} \). Clearly, $\mathcal{E}$ is a $\sigma$-algebra on $X_1 \times X_2$. We will show that $\mathcal{E}$ is a monotone class on $X_1 \times X_2$. Then, by Monotone Class Lemma 12.19, we have $\mathcal{E} = \mathcal{B}$.

Clearly, $\emptyset = \emptyset \times \emptyset \in \mathcal{C} \subseteq \mathcal{E}$ and $X_1 \times X_2 \in \emptyset \subseteq \mathcal{E}$. \( \forall E_1, E_2 \in \mathcal{E} \) with $E_1 \subseteq E_2$, then $g_\mu, E : X_1 \to [0, \mu_2(X_2)] \subseteq \mathcal{R}$ is $\mathcal{B}_1$-measurable, $i = 1, 2$. Let $E := E_2 \setminus E_1$, \( \forall x_1 \in X_1 \), we have $g_\mu(x_1) = \mu_2(E_{x_1}) = \mu_2(E_{x_1} \setminus E_{x_1}) = \mu_2(E_{x_1}) - \mu_2(E_{x_1}) = g_\mu(x_1) - g_\mu(x_1) \subseteq \mathcal{R}$. By Propositions 7.23, 11.39, and 11.38, $g_\mu$ is $\mathcal{B}_1$-measurable. Then, $E = E_2 \setminus E_1 \in \mathcal{E}$.

Next, consider the general case when $\mathcal{X}_1$ and $\mathcal{X}_2$ are $\sigma$-finite. Since $X_1$ is $\sigma$-finite, then $\exists (X_{1,n})_{n=1}^{\infty} \subseteq \mathcal{B}_1$ such that $X_1 = \bigcup_{n=1}^{\infty} X_{1,n}$ and
involved in the equalities are finite. This shows that $E_\mu$.

Lemma 12.19, we have $E$.

Proof First, consider the special case when $X_1, X_2 := (X_{1,m}, \mathcal{B}_{1,m}, \mu_{1,m})$ be the finite measure subspace of $X_1, X_2 := (X_{2,m}, \mathcal{B}_{2,m}, \mu_{2,m})$ be the finite measure subspace of $X_1, X_2 := (X_{1,m}, \mathcal{B}_{1,m}, \lambda_{1,m}) := X_1 \times X_2$, which is the finite measure subspace of $X$ by Proposition 12.22 and its proof. $orall E \in \mathcal{B}, \forall n, m \in \mathbb{N}$, let $E_{n,m} := E \cap (X_{1,n} \times X_{2,m}) \in \mathcal{B}_{n,m}$. By the special case, the function $g_{n,m} : X_1 \times \to [0, \infty) \in \mathbb{R}$, defined by $g_{n,m}(x_1) = \mu_{2,m}((E_{n,m}) \cap X_{2,m})$, $\forall x_1 \in X_{1,n}$, is $B_{1,n}$-measurable. Define $g_{n,m} : X_1 \to [0, \infty) \in \mathbb{R}$ by $g_{n,m}(x_1) = \mu_{2}(E_{n,m} \cap X_{1,n})$, $\forall x_1 \in X_1$. Then, $g_{n,m}(x_1) = \begin{cases} g_{n,m}(x_1) & \forall x_1 \in X_{1,n} \\ 0 & \forall x_1 \in X_1 \setminus X_{1,n} \end{cases}$, $\forall x_1 \in X_1$. By Proposition 11.41, $g_{n,m}$ is $B_1$-measurable. Note that $g_{E}(x_1) = \mu_{2}(E_{x_1}) = \mu_{2}(\bigcup_{n=1}^{\infty} (E_{n,n} \cap x_1)) = \lim_{n \to \infty} \mu_{2}(E_{n,n} \cap x_1)$, $\forall x_1 \in X_1$, where the second equality follows from the fact that $E = \bigcup_{n=1}^{\infty} E_{n,n}$; and the third equality follows form Proposition 11.7. By Proposition 11.48 and the fact that $[0, \infty) \subseteq \mathbb{R}$ is metrizable, $g_{E}$ is $B_1$-measurable. By symmetry, $\forall E \in \mathcal{B}, h_{E}$ is $B_2$-measurable.

This completes the proof of the proposition.

Lemma 12.27 Let $X_i := (X_i, \mathcal{B}_i, \mu_i)$ be a $\sigma$-finite measure space, $i = 1, 2$, and $X := X_1 \times X_2, \mathcal{B}, \mu := X_1 \times X_2$ be the product measure space. Then, $\forall E \in \mathcal{B}$, we have $\mu(E) = \int_{X_1 \times X_2} g_{E} \, d\mu_1 + \int_{X_2} h_{E} \, d\mu_2$, where the function $g_{E} : X_1 \to [0, \infty) \subseteq \mathbb{R}$, as defined in Proposition 12.26, is $B_1$-measurable; and the function $h_{E} : X_2 \to [0, \infty) \subseteq \mathbb{R}$, as defined in Proposition 12.26, is $B_2$-measurable.

Proof First, consider the special case when $X_1$ and $X_2$ are finite. Then $\mathcal{X}$ is finite. $\forall E \in \mathcal{B}, 0 \leq \mu(E) \leq \mu(X_1 \times X_2) = \mu_1(X_1) \mu_2(X_2) < \infty$.

By Proposition 12.26, the function $g_{E} : X_1 \to [0, \mu_2(X_2)) \subseteq \mathbb{R}$ is $B_1$-measurable. Let $E := \{ E \in \mathcal{B} \mid \mu(E) = \int_{X_1} g_{E} \, d\mu_1 \} \subseteq \mathcal{B}$. Let $\mathcal{C}$ be the set of measurable rectangles in $X_1 \times X_2$, which is clearly a $\pi$-system.

$\forall E := E_1 \times E_2 \in \mathcal{C}, g_{E} = \mu_2(E_2)\chi_{E_1 \times X_1}$. Then, by Proposition 11.75, $\mu(E) = \mu_1(E_1)\mu_2(E_2) = \int_{X_1} g_{E} \, d\mu_1$. This shows that $E \in \mathcal{E}$. Then $\mathcal{C} \subseteq \mathcal{E}$.

We will shown that $\mathcal{C}$ is a monotone class on $X_1 \times X_2$. By Monotone Class Lemma 12.19, we have $\mathcal{E} = \mathcal{B}$.

Clearly, $\emptyset \in \mathcal{E}$ and $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E_1, E_2 \in \mathcal{E}$ with $E_1 \subseteq E_2$, let $E := E_2 \setminus E_1 \in \mathcal{B}$. By Proposition 12.26, $g_{E_1}, g_{E_2}$ and $g_{E}$ are $B_1$-measurable. As shown in the proof of Proposition 12.26, $g_{E} = g_{E_2} - g_{E_1}$. Then, $\mu(E) = \mu(E_2) - \mu(E_1) = \int_{X_1} g_{E_2} \, d\mu_1 - \int_{X_1} g_{E_1} \, d\mu_1 = \int_{X_1} g_{E} \, d\mu_1$, where the first equality follows from the fact that $\mu$ is a measure; the second equality follows from the definition of $\mathcal{E}$; the last equality follows from Proposition 11.83; and all equality follows from the fact all quantity involved in the equalities are finite. This shows that $E = E_2 \setminus E_1 \in \mathcal{E}$.
∀ (E_i)_{i=1}^\infty \subseteq \mathcal{E} with E_i \subseteq E_{i+1}, \forall i \in \mathbb{N}, let E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}. \forall i \in \mathbb{N}, by Proposition 12.26, g_{E_i} is \mathcal{B}_1-measurable. \forall x_1 \in X_1, we have g_{E_i}(x_1) \leq g_{E_{i+1}}(x_1) since E_i \subseteq E_{i+1}. By Proposition 12.26, g_E is \mathcal{B}_1-measurable.

As shown in the proof of Proposition 12.26, g_E(x_1) = \lim_{i \in \mathbb{N}} g_{E_i}(x_1), \forall x_1 \in X_1. This implies that \mu(E) = \lim_{i \in \mathbb{N}} \mu(E_i) = \lim_{i \in \mathbb{N}} \int_{X_i} g_E \, d\mu_1 = \int_{X_1} g_E \, d\mu_1, where the first equality follows from Proposition 11.7; the second equality follows from the definition of \mathcal{E}; and the last equality follows from Monotone Convergence Theorem 11.81. Hence, E = \bigcup_{i=1}^\infty E_i \in \mathcal{E}.

Hence, \mathcal{E} is a monotone class on X_1 \times X_2 and \mathcal{E} = \mathcal{B}. Therefore, \mu(E) = \int_{X_1} g_E \, d\mu_1, \forall E \in \mathcal{B}. By symmetry, \mu(E) = \int_{X_2} h_E \, d\mu_2, \forall E \in \mathcal{B}. This completes the proof for the special case.

Next, consider the general case when X_1 and X_2 are \sigma-finite. Since X_1 is \sigma-finite, then \exists (X_1^\infty)_{n=1} \subseteq \mathcal{B}_1 such that X_1 = \bigcup_{n=1}^\infty X_1^n and \mu_1(X_1^\infty) < \infty, \forall n \in \mathbb{N}. Without loss of generality, we may assume that X_1 \subseteq X_1^{n+1}, \forall n \in \mathbb{N}. Since X_2 is \sigma-finite, then \exists (X_2^{\infty})_{m=1} \subseteq \mathcal{B}_2 such that X_2 = \bigcup_{m=1}^\infty X_2^m and \mu_2(X_2^{\infty}) < \infty, \forall m \in \mathbb{N}. Without loss of generality, we may assume that X_2 \subseteq X_2^{m+1}, \forall m \in \mathbb{N}, \forall n, m \in \mathbb{N}, let X_1,n := (X_1 \times X_2^m, \mathcal{B}_1 \times \mathcal{B}_2, \mu_{1,n}) := X_1 \times X_2^m, which is the finite measure subspace of \mathcal{X} by Proposition 12.22. \forall E \in \mathcal{B}, \forall n, m \in \mathbb{N}, let E_{n,m} := E \cap (X_1 \times X_2^m) \in \mathcal{B}_n,m. By the special case, the function g_{n,m} : X_1,n \to [0, \infty) \subseteq \mathbb{R}, defined by g_{n,m}(x_1) = \mu_2((E_{n,m})_{x_1} \cap X_2^m), \forall x_1 \in X_1,n, is \mathcal{B}_1,n-measurable and \mu(E_{n,m}) = \lambda_{n,m}(E_{n,m}) = \int_{X_1,n} g_{n,m} \, d\mu_{1,n}. Define \tilde{g}_{n,m} : X_1 \to [0, \infty) \subseteq \mathbb{R} by \tilde{g}_{n,m}(x_1) = \mu_2((E_{n,m})_{x_1}), \forall x_1 \in X_1.

Then, \tilde{g}_{n,m}(x_1) = \begin{cases} g_{n,m}(x_1) & \forall x_1 \in X_1,n \\ 0 & \forall x_1 \in X_1 \setminus X_1,n \end{cases}, \forall x_1 \in X_1. By Proposition 11.41, \tilde{g}_{n,m} is \mathcal{B}_1-measurable. By Definition 11.79, \mu(E_{n,m}) = \int_{X_1} \tilde{g}_{n,m} \, d\mu_1. By Proposition 12.26 and its proof, we have g_{E_i}(x_1) = \lim_{i \in \mathbb{N}} g_{n,m}(x_1), \forall x_1 \in X_1, and g_E is \mathcal{B}_1-measurable. Clearly, \tilde{g}_{n,m}(x_1) \leq g_{n+1,m+1}(x_1), \forall x_1 \in X_1, \forall n \in \mathbb{N}, since E_{n,n} \subseteq E_{n+1,n+1}. Then, \mu(E) = \lim_{n \in \mathbb{N}} \mu(E_{n,m}) = \lim_{n \in \mathbb{N}} \int_{X_1} \tilde{g}_{n,m} \, d\mu_1 = \int_{X_1} g_E \, d\mu_1, where the first equality follows from Proposition 11.7; and the last equality follows from Monotone Convergence Theorem 11.81.

By symmetry, \mu(E) = \int_{X_2} h_E \, d\mu_2, \forall E \in \mathcal{B}.

This completes the proof of the proposition. \qed

**Proposition 12.28** Let X_i := (X_i, B_i, \mu_i) be a \sigma-finite measure space, i = 1, 2, \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2 be the product measure space, \mathcal{Y} := (\mathcal{Y}, \mathcal{O}) be a topological space, and f : X_1 \times X_2 \to \mathcal{Y} be \mathcal{B}-measurable. Then,

(i) \forall x_1 \in X_1, f_{x_1} : X_2 \to \mathcal{Y}, defined by f_{x_1}(x_2) = f(x_1, x_2), \forall x_2 \in X_2, is \mathcal{B}_2-measurable;

(ii) \forall x_2 \in X_2, f_{(x_2)} : X_1 \to \mathcal{Y}, defined by f_{(x_2)}(x_1) = f(x_1, x_2), \forall x_1 \in X_1, is \mathcal{B}_1-measurable.
12.3. PRODUCT MEASURE

Proof (i) \( \forall \) open set \( O \in \mathcal{O} \), \( f_{\text{inv}}(O) \in \mathcal{B} \) since \( f \) is \( \mathcal{B} \)-measurable. \( \forall x_1 \in X_1 \), by Proposition 12.25, \( \{ f_{\text{inv}}(O) \}_{x_1} := \{ x_2 \in X_2 \mid (x_1, x_2) \in f_{\text{inv}}(O) \} \in \mathcal{B}_2 \). Clearly, \( f_{x_1, \text{inv}}(O) = \{ f_{\text{inv}}(O) \}_{x_1} \in \mathcal{B}_2 \). Then, \( f_{x_1} \) is \( \mathcal{B}_2 \)-measurable.

(ii) By symmetry, (ii) holds. This completes the proof of the proposition.

\[ \square \]

Theorem 12.29 (Tonelli) Let \( X_i := (X_i, \mathcal{B}_i, \mu_i) \) be a \( \sigma \)-finite measure space, \( i = 1, 2 \), \( X := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2 \) be the product measure space, and \( f : X_1 \times X_2 \to [0, \infty] \subset \mathbb{R}_e \) be \( \mathcal{B} \)-measurable. Define \( f_i : X_2 \to \mathbb{R} \) and \( f(x_2) : X_1 \to \mathbb{R} \) as in Proposition 12.28. Then, the following statements hold.

(i) \( \int X_1 \times X_2 f \, d\mu = \int X_1 p \, d\mu_1 \), where \( p : X_1 \to [0, \infty] \subset \mathbb{R}_e \), defined by \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \), \( \forall x_1 \in X_1 \), is \( \mathcal{B}_1 \)-measurable.

(ii) \( \int X_1 \times X_2 f \, d\mu = \int X_2 q \, d\mu_2 \), where \( q : X_2 \to [0, \infty] \subset \mathbb{R}_e \), defined by \( q(x_2) = \int_{X_1} f(x_2) \, d\mu_1 \), \( \forall x_2 \in X_2 \), is \( \mathcal{B}_2 \)-measurable.

Proof (i) By Proposition 11.67, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^\infty \), \( \varphi_n : X \to [0, \infty] \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), such that \( 0 \leq \varphi_n(x_1, x_2) \leq f(x_1, x_2) \), \( \forall (x_1, x_2) \in X_1 \times X_2 \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \in \mathbb{N}} \varphi_n(x_1, x_2) = f(x_1, x_2) \), \( \forall (x_1, x_2) \in X_1 \times X_2 \).

By Proposition 12.28 and Definition 11.79, \( p \) is well-defined. Define \( p_n : X_1 \to [0, \infty] \subset \mathbb{R}_e \) by \( p_n(x_1) = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \), \( \forall x_1 \in X_1 \), \( \forall n \in \mathbb{N} \). By Proposition 12.28, \( p_n \)'s are well-defined. By Proposition 11.83, \( p_n(x_1) \leq p_{n+1}(x_1) \leq p(x_1) \), \( \forall x_1 \in X_1 \), \( \forall n \in \mathbb{N} \). By Monotone Convergence Theorem 11.81, \( p(x_1) = \lim_{n \in \mathbb{N}} p_n(x_1) \), \( \forall x_1 \in X_1 \). Fix any \( n \in \mathbb{N} \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m_n} a_i \chi_{E_i} \times X_2 \), where \( m_n \in \mathbb{Z}^+ \), \( a_1, \ldots, a_{m_n} \in [0, \infty] \subset \mathbb{R} \) are distinct, \( E_1, \ldots, E_{m_n} \in \mathcal{B} \) are nonempty, pairwise disjoint, and \( \mu(E_i) < \infty \), \( \forall i \in \{1, \ldots, m_n\} \). Fix any \( i \in \{1, \ldots, m_n\} \), define \( g_i : X_1 \to [0, \infty] \subset \mathbb{R}_e \) by \( g_i(x_1) = \mu_2(E_{i,x_1}) \), \( \forall x_1 \in X_1 \), where \( (E_{i,x_1}) := \{ x_2 \in X_2 \mid (x_1, x_2) \in E_i \} \).

By Proposition 12.26, \( g_i \) is \( \mathcal{B}_1 \)-measurable. By Lemma 12.27, \( \mu(E_i) = \int_{X_1} g_i \, d\mu_1 \). By Proposition 11.75, \( \int_{X_1 \times X_2} \varphi_n \, d\mu = \sum_{i=1}^{m_n} a_i \mu(E_i) = \sum_{i=1}^{m_n} a_i \int_{X_1} \chi_{E_i} \times X_2 \, d\mu_1 = \int_{X_1} \left( \sum_{i=1}^{m_n} a_i g_i \right) \, d\mu_1 =: \alpha \), where the third equality follows from Proposition 11.83. By Propositions 3.89, 11.38, and 11.39, \( \tilde{p}_n : X_1 \to [0, \infty] \subset \mathbb{R}_e \) is \( \mathcal{B}_1 \)-measurable. Note that \( p_n(x_1) = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \int_{X_1} \left( \sum_{i=1}^{m_n} a_i \chi_{E_i,x_1 \times X_2} \right) \, d\mu_2 = \sum_{i=1}^{m_n} a_i \mu_2(E_{i,x_1}) = \sum_{i=1}^{m_n} a_i g_i(x_1) = \tilde{p}_n(x_1) \), \( \forall x_1 \in X_1 \), where the first equality follows from the definition of \( p_n \); the second equality follows from the canonical representation of \( \varphi_n \); the third equality follows from Proposition 11.83; the fourth equality follows from the fact that \( X_2 \) is \( \sigma \)-finite, Propositions 11.75 and 11.7, and the Monotone Convergence Theorem 11.81; the fifth equality follows from the definition
of \( g_i \)'s; and the last equality follows from the definition of \( \tilde{p}_n \). Then, \( p_n \) is \( \mathcal{B}_1 \)-measurable. By Proposition 11.48 and the fact that \([0, \infty] \subset \mathbb{R}_e\) is metrizable, \( p \) is \( \mathcal{B}_1 \)-measurable. Then, \( \int_{X_1 \times X_2} \varphi_n \, d\mu = \int_{X_1} p_n \, d\mu_1 = \int_{X_1} p_n \, d\mu_1 [0, \infty] \subset \mathbb{R}_e \). By Monotone Convergence Theorem 11.81, \( \int_{X_1 \times X_2} f \, d\mu = \lim_{n \to \infty} \int_{X_1 \times X_2} \varphi_n \, d\mu = \lim_{n \to \infty} \int_{X_1} p_n \, d\mu_1 = \int_{X_1} p \, d\mu_1 \).

Hence, (i) holds. By symmetry, (ii) holds. This completes the proof of the theorem. 

\[ \square \]

**Theorem 12.30 (Fubini)** Let \( X_i := (X_i, \mathcal{B}_i, \mu_i) \) be a \( \sigma \)-finite measure space, \( i = 1, 2 \), \( X := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2 \) be the product measure space, \( \mathcal{Y} \) be a separable Banach space, and \( f : X_1 \times X_2 \to \mathcal{Y} \) be absolutely integrable over \( X \). Define \( f_{x_1} : X_2 \to \mathcal{Y} \) and \( f(x_2) : X_1 \to \mathcal{Y} \) as in Proposition 12.28. Then, the following statements hold.

(i) There exists \( \bar{p} : X_1 \to \mathcal{Y} \) and exists \( \bar{U} \in \mathcal{B}_1 \) such that \( \mu_1(X_1 \setminus \bar{U}) = 0 \); \( \forall x_1 \in \bar{U}, f_{x_1} \) is absolutely integrable over \( X_2 \) and \( \bar{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y} \); \( \forall x_1 \in X_1 \setminus \bar{U}, \bar{p}(x_1) = 0 \); \( \bar{p} \) is absolutely integrable over \( X_1 \); and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_1} \bar{p} \, d\mu_1 \).

(ii) There exists \( \bar{q} : X_2 \to \mathcal{Y} \) and exists \( \bar{V} \in \mathcal{B}_2 \) such that \( \mu_2(X_2 \setminus \bar{V}) = 0 \); \( \forall x_2 \in \bar{V}, f(x_2) \) is absolutely integrable over \( X_1 \) and \( \bar{q}(x_2) = \int_{X_2} f(x_2) \, d\mu_1 \in \mathcal{Y} \); \( \forall x_2 \in X_2 \setminus \bar{V}, \bar{q}(x_2) = 0 \); \( \bar{q} \) is absolutely integrable over \( X_2 \); and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_2} \bar{q} \, d\mu_2 \).

(iii) \( \forall p : X_1 \to \mathcal{Y} \) satisfying: \( p \) is \( \mathcal{B}_1 \)-measurable; and \( \exists U \in \mathcal{B}_1 \) with \( \mu_1(X_1 \setminus U) = 0 \) such that, \( \forall x_1 \in U, f_{x_1} \) is absolutely integrable over \( X_2 \) and \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y} \); we have \( p \) is absolutely integrable over \( X_1 \) and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \).

(iv) \( \forall q : X_2 \to \mathcal{Y} \) satisfying: \( q \) is \( \mathcal{B}_2 \)-measurable; and \( \exists V \in \mathcal{B}_2 \) with \( \mu_2(X_2 \setminus V) = 0 \) such that, \( \forall x_2 \in V, f(x_2) \) is absolutely integrable over \( X_1 \) and \( q(x_2) = \int_{X_1} f(x_2) \, d\mu_1 \in \mathcal{Y} \); we have \( q \) is absolutely integrable over \( X_2 \) and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2 \).

**Proof**

(i) Since \( f \) is absolutely integrable over \( X \), then \( \mathcal{P} \circ f \) is \( \mathcal{B}_1 \)-measurable and \( \int_{X_1 \times X_2} \mathcal{P} \circ f \, d\mu < \infty \). By Tonelli Theorem 12.29, \( g : X_1 \to [0, \infty] \subset \mathbb{R}_e \), defined by \( g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) \, d\mu_2(x_2), \forall x_1 \in X_1 \), is \( \mathcal{B}_1 \)-measurable, and \( \int_{X_1 \times X_2} \mathcal{P} \circ f \, d\mu = \int_{X_1} g \, d\mu_1 \). By Proposition 11.82, the function \( \bar{g} : X_1 \to [0, \infty] \subset \mathbb{R}_e \), defined by \( \bar{g}(x_1) = \begin{cases} g(x_1) & g(x_1) < \infty \\ 0 & g(x_1) = \infty \end{cases} \), \( \forall x_1 \in X_1 \), is \( \mathcal{B}_1 \)-measurable, \( g = \bar{g} \) a.e. in \( X_1 \), and \( \int_{X_1 \times X_2} \mathcal{P} \circ f \, d\mu = \int_{X_1} \bar{g} \, d\mu_1 < \infty \). Let \( U := \{ x_1 \in X_1 \mid g(x_1) < \infty \} \). Then, \( U \in \mathcal{B}_1 \) and \( \mu_1(X_1 \setminus U) = 0 \); \( \forall x_1 \in U \), we have \( g(x_1) = \int_{X_2} \mathcal{P} \circ f(x_1, x_2) \, d\mu_2(x_2) = \int_{X_1 \times X_2} \mathcal{P} \circ f_{x_1} \, d\mu_2 < \infty \). Then, \( f_{x_1} \) is absolutely integrable over \( X_2 \). By
12.3. PRODUCT MEASURE

Propositions 12.28 and 11.92, \( \int_{X \times Y} f \, d\mu_2 \in \mathcal{Y}, \forall x_1 \in U. \) Define \( \tilde{p} : X_1 \to \mathcal{Y} \) by 
\[
\tilde{p}(x_1) = \begin{cases} 
\int_{X_2} f_{x_1} \, d\mu_2 & x_1 \in U \\
0 & x_1 \in X_1 \setminus U,
\end{cases} \quad \forall x_1 \in X_1.
\]

By Proposition 11.66, \( \exists \) a sequence of simple functions \( (\varphi_n)_{n=1}^\infty \), \( \varphi_n : \mathcal{X} \to \mathcal{Y}, \forall n \in \mathbb{N} \), such that \( \|\varphi_n(x_1, x_2)\| \leq \|f(x_1, x_2)\|, \forall (x_1, x_2) \in X_1 \times X_2, \forall n \in \mathbb{N} \), and \( \lim_{n \in \mathbb{N}} \varphi_n = f \) a.e. in \( \mathcal{X} \). Let \( A := \{(x_1, x_2) \in X_1 \times X_2 \mid (\varphi_n(x_1, x_2))_{n=1}^\infty \) does not converge to \( f(x_1, x_2)\}. \) Then, \( A \in \mathcal{B} \) and \( \mu(A) = 0 \). Define \( g_A : X_1 \to [0, \infty] \subset \mathbb{R} \) by \( g_A(x_1) = \mu_2(A_{x_1}), \forall x_1 \in X_1 \), where \( A_{x_1} := \{x_2 \in X_2 \mid (x_1, x_2) \in A\} \) is as defined in Proposition 12.25, and \( \hat{g}_A : X_1 \to [0, \infty] \subset \mathbb{R} \) by \( \hat{g}_A(x_1) = \begin{cases} 
g_A(x_1) & g_A(x_1) < \infty \\
0 & g_A(x_1) = \infty
\end{cases} \), \( \forall x_1 \in X_1 \).

By Lemma 11.82, \( \hat{g}_i \) is \( \mathcal{B}_1 \)-measurable and \( \mu(E_i) = \int_{X_1} \hat{g}_i \, d\mu_1 \). Define 
\( \hat{g}_i : X_1 \to [0, \infty] \subset \mathbb{R} \) by \( \hat{g}_i(x_1) = \begin{cases} 
g_i(x_1) & g_i(x_1) < \infty \\
0 & g_i(x_1) = \infty
\end{cases} \), \( \forall x_1 \in X_1 \).

By Proposition 11.82, \( \hat{g}_i \) is \( \mathcal{B}_1 \)-measurable, \( g_i = \hat{g}_i \) a.e. in \( X_1 \), and \( \mu(E_i) = \int_{X_1} \hat{g}_i \, d\mu_1 \). Then, \( \int_{X_1 \times X_2} \varphi \, d\mu = \sum_{i=1}^m y_i \mu(E_i) = \sum_{i=1}^m y_i \int_{X_1} \hat{g}_i \, d\mu_1 = \int_{X_1} \sum_{i=1}^m y_i \hat{g}_i \, d\mu_1 = \int_{X_1} \tilde{p} \, d\mu_1 \), where the first equality follows from Proposition 11.75; and the third equality follows from Proposition 11.92.

By Propositions 7.23, 11.38, and 11.39, \( \tilde{p}_n : X_1 \to \mathcal{Y} \) is \( \mathcal{B}_1 \)-measurable and, by Proposition 11.92, \( \tilde{p}_n \) is absolutely integrable over \( X_1 \). Let \( A_n := \bigcup_{i=1}^m \{x_1 \in X_1 \mid g_i(x_1) = \infty\} \). Then, \( A_n \in \mathcal{B} \) and \( \mu_1(A_n) = 0 \). \( \forall x_1 \in X_1 \setminus A_n, \) \( \tilde{p}_n(x_1) = \sum_{i=1}^m y_i g_i(x_1) = \sum_{i=1}^m y_i g_i(x_1) = \sum_{i=1}^m y_i \mu_2(E_{x_1}) = \sum_{i=1}^m y_i \int_{X_2} \chi_{E_{x_1}} \, d\mu_2 = \int_{X_2} \sum_{i=1}^m y_i \chi_{E_{x_1}} \, d\mu_2 = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \in \mathcal{Y} \), where the first equality follows from the definition of \( \tilde{p}_n \); the second equality follows from the definitions of \( \hat{g}_i \)’s; the third equality follows from the definitions of \( \hat{g}_i \)’s; the fourth equality follows from Proposition 11.75; and the fifth equality follows from Proposition 11.92. Let \( A_2 := \bigcup_{n=1}^\infty A_n \). Then, \( A_2 \in \mathcal{B} \) and \( \mu_1(A_2) = 0 \). \( \forall x_1 \in U \setminus (A_1 \cup A_2) =: \bar{U}, \) we have 
\[
\tilde{p}(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \lim_{n \in \mathbb{N}} \tilde{p}_n(x_1) \in \mathcal{Y}.
\]
Clearly, $\tilde{U} \in \mathcal{B}_1$ and $0 \leq \mu_1(X_1 \setminus \tilde{U}) = \mu_1((X_1 \setminus U) \cup A_1 \cup A_2) \leq \mu_1(X_1 \setminus U) + \mu_1(A_1) + \mu_1(A_2) = 0$. Fix any $n \in \mathbb{N}$. Define $\tilde{p}_n : X_1 \to \mathcal{Y}$ by

$$\tilde{p}_n(x_1) = \begin{cases} p_n(x_1) & x_1 \in \tilde{U} \\ \vartheta_y & x_1 \in X_1 \setminus \tilde{U} \end{cases}, \forall x_1 \in X_1.$$ By Proposition 11.41, $\tilde{p}_n$ is $\mathcal{B}_1$-measurable. By Lemma 11.43, $\tilde{p}_n = \tilde{p}$ a.e. in $X_1$. Define $\tilde{p} : X_1 \to \mathcal{Y}$ by

$$\tilde{p}(x_1) = \begin{cases} \tilde{p}(x_1) & x_1 \in \tilde{U} \\ \vartheta_y & x_1 \in X_1 \setminus \tilde{U} \end{cases}, \forall x_1 \in X_1.$$ Then, $\tilde{p}(x_1) = \lim_{n \to \infty} \tilde{p}_n(x_1)$, $\forall x_1 \in X_1$. By Proposition 11.48, $\tilde{p}$ is $\mathcal{B}_1$-measurable. $\forall x_1 \in \tilde{U}$, $\tilde{p}(x_1) = \tilde{p}(x_1) = \int_{X_2} f_{x_1} d\mu_2$ and $f_{x_1}$ is absolutely integrable over $X_2$.

Note that $\int_{X_1 \times X_2} f d\mu = \lim_{n \to \infty} \int_{X_1 \times X_2} \varphi_n d\mu = \lim_{n \to \infty} \int_{X_1} \tilde{p}_n d\mu_1 = \lim_{n \to \infty} \int_{X_1} \tilde{p}_n d\mu_1$, where the first equality follows from Lebesgue Dominated Convergence Theorem 11.91; the second equality follows from the definition of $\tilde{p}_n$; and the third equality follows from Proposition 11.92.

(iii) Fix any $p : X_1 \to \mathcal{Y}$ satisfying: $p$ is $\mathcal{B}_1$-measurable; and $\exists U \in \mathcal{B}_1$ with $\mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, $f_{x_1}$ is absolutely integrable over $X_2$ and $p(x_1) = \int_{X_2} f_{x_1} d\mu_2$. By (i), we have $U_0 := \tilde{U} \cap U \in \mathcal{B}_1$ and $\mu_1(X_1 \setminus U_0) = \mu_1((X_1 \setminus U) \cup (X_1 \setminus U)) = 0$. $\forall x_1 \in U_0$, $p(x_1) = \int_{X_2} f_{x_1} d\mu_2 = \tilde{p}(x_1)$. By Lemma 11.43, $p = \tilde{p}$ a.e. in $X_1$. By Proposition 11.83, we have $p$ is absolutely integrable over $X_1$ since $\tilde{p}$ is so. By Proposition 11.92, $\int_{X_1} \tilde{p} d\mu_1 = \int_{X_1} \tilde{p} d\mu_1 \leq \int_{X_1 \times X_2} f d\mu \in \mathcal{Y}$. Hence, (iii) holds.

By symmetry, (iv) also holds.

When the range of the function $f$ is a $\sigma$-compact subset of a separable Banach space, then the set $U$ and $V$ of the previous theorem takes a simpler form. This is established in the following theorem.

**Theorem 12.31 (Fubini)** Let $\mathcal{X}_1 := (X_1, \mathcal{B}_1, \mu_1)$ be a $\sigma$-finite measure space, $i = 1, 2$, $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2$ be the product measure space, $\mathcal{Y}$ be a separable Banach space, $U \subseteq \mathcal{Y}$ be a $\sigma$-compact conic segment, $f : X_1 \times X_2 \to U$ be absolutely integrable over $\mathcal{X}$. Define $f_{x_1} : X_2 \to \mathcal{Y}$ and $\int_{\mathcal{X}_2} f_{x_1} d\mu = \int_{X_1 \times X_2} f d\mu \in \mathcal{Y}$. Hence, (iii) holds.

By symmetry, (iv) holds. This completes the proof of the theorem. $\square$

When the range of the function $f$ is a $\sigma$-compact subset of a separable Banach space, then the set $U$ and $V$ of the previous theorem takes a simpler form. This is established in the following theorem.
12.3. Product Measure

(ii) Define \( V \) := \( \{ x_2 \in X_2 \mid f(x_2) \) is absolutely integrable over \( X_1 \} \), and then, by Lebesgue Dominated Convergence Theorem 11.91, 

\[ \forall \| p \|_X, \ \exists q \in \mathcal{B}_2, \mu_2(X_2 \setminus V) = 0, \ \text{and} \ \text{is absolutely integrable over} \ X_2, \ \text{and} \ \int_{X_1 \times X_2} f \ d\mu = \int_{X_2} q \ d\mu_2 \in \mathcal{Y}. \]

**Proof**

(i) Since \( f \) is absolutely integrable over \( X \), then \( f \) and \( P \circ f \) are \( \mathcal{B} \)-measurable and \( \int_{X_1 \times X_2} P \circ f \ d\mu < \infty \). By Tonelli Theorem 12.29, \( g : X_1 \to [0, \infty] \subset \mathbb{R} \), defined by \( g(x_1) = \int_{X_2} P \circ f(x_1, x_2) \ d\mu_2(x_2), \ \forall x_1 \in X_1 \), is \( \mathcal{B}_1 \)-measurable and \( \int_{X_1 \times X_2} P \circ f \ d\mu = \int_{X_1} g \ d\mu_1 \). By Proposition 11.82, the function \( g : X_1 \to [0, \infty] \subset \mathbb{R} \), defined by \( g(x_1) = \left\{ \begin{array}{ll} g(x_1) & \text{if } x_1 \in X_1, \\
(x_1) & \text{if } x_1 = \infty, \end{array} \right. \) and \( \int_{X_1 \times X_2} P \circ f \ d\mu = \int_{X_1} g \ d\mu_1 \). Let \( U := \{ x_1 \in X_1 \mid g(x_1) < \infty \} \). Then, \( \bar{U} = U \), \( \mathcal{B}_1 \)-measurable, and \( \int_{X_1 \times X_2} P \circ f \ d\mu = \int_{X_1} g \ d\mu_1 \). By Lemma 11.68, \( \exists \) a sequence of simple functions \( (\varphi_n)_{n=1}^{\infty} \), \( \varphi_n : X_1 \to U \), \( \forall n \in \mathbb{N} \), such that \( \| \varphi_n(x_1, x_2) \| < \| f(x_1, x_2) \| \), \( \forall (x_1, x_2) \in X_1 \times X_2 \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \varphi_n(x_1, x_2) = f(x_1, x_2) \), \( \forall (x_1, x_2) \in X_1 \times X_2 \). Fix any \( x_1 \in X_1 \), we have \( \lim_{n \to \infty} \varphi_n(x_1, x_2) = f(x_1, x_2) \), \( \forall x_2 \in X_2 \). Then, by Lebesgue Dominated Convergence Theorem 11.91, \( p(x_1) = \lim_{n \to \infty} \int_{X_2} \varphi_n(x_1, x_2) \ d\mu_2(x_2), \forall x_1 \in U \).

Fix any \( n \in \mathbb{N} \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m} y_i \chi_{E_i}, x_1 \in X_1 \), where \( m \in \mathbb{Z}^+, y_1, \ldots, y_m \in U \) are distinct and none equals to \( \vartheta_y \), \( E_1, \ldots, E_m \in \mathcal{B} \) are nonempty, pairwise disjoint, and \( \mu(E_i) < \infty, i = 1, \ldots, m \). Fix any \( i \in \{1, \ldots, m\} \). Define \( g_i : X_1 \to [0, \infty] \subset \mathbb{R} \) by \( g_i(x_1) = \mu_2(E_{i,x_1}), \forall x_1 \in X_1 \), where \( E_{i,x_1} := \{ x_2 \in X_2 \mid (x_1, x_2) \in E_i \} \). By Lemma 12.27, \( g_i \) is \( \mathcal{B}_1 \)-measurable and \( \mu(E_i) = \int_{X_1} g_i \ d\mu_1 \). Define \( \hat{g}_i : X_1 \to [0, \infty] \subset \mathbb{R} \) by \( \hat{g}_i(x_1) = \left\{ \begin{array}{ll} g_i(x_1) & \text{if } g_i(x_1) < \infty, \\
(x_1) & \text{if } g_i(x_1) = \infty, \end{array} \right. \) and \( \int_{X_1 \times X_2} g_i \ d\mu_1 = \int_{X_1} \hat{g}_i \ d\mu_1 \). Then, by Proposition 11.82, \( \hat{g}_i \) is \( \mathcal{B}_1 \)-measurable, \( \hat{g}_i \) a.e. in \( X_1 \), and \( \mu(E_i) = \int_{X_1} \hat{g}_i \ d\mu_1 \).
= \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \in Y, \forall x_1 \in U, \text{where the first equality follows from}
the definition of \( \hat{\rho}_n \); the second equality follows from the preceding discussion; the third equality follows from the definitions of \( g_i \)'s; the fourth equality follows from Proposition 11.75; and the fifth equality follows from Proposition 11.92. Define \( \hat{\rho}_n : X_1 \to Y \) by \( \hat{\rho}_n(x_1) = \left\{ \begin{array}{ll} \hat{\rho}_n(x_1) & x_1 \in U \\ \vartheta_y & x_1 \in X_1 \setminus U \end{array} \right\}, \forall x_1 \in X_1. \ By Proposition 11.41, \( \hat{\rho}_n \) is \( B_1 \)-measurable. Then, \( \hat{\rho}_n = \hat{\rho}_n \) a.e. in \( X_1 \).

Note that, \( \forall x_1 \in U, \) we have

\[
\| \hat{\rho}_n(x_1) \| = \| \hat{\rho}_n(x_1) \| = \left\| \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \right\| \leq \int_{X_2} \mathcal{P} \circ \varphi_n(x_1, x_2) \, d\mu_2(x_2) \leq \int_{X_2} \mathcal{P} \circ f(x_1, x_2) \, d\mu_2(x_2) = \tilde{g}(x_1) < \infty
\]

where the second equality follows from the previous paragraph; the first inequality follows from Proposition 11.92; the second inequality follows from the property of \( \varphi_n; \) and the last equality follows from the definition of \( U. \)

Note also that, \( \forall x_1 \in X_1 \setminus U, \) we have \( \| \hat{\rho}_n(x_1) \| = 0 = \tilde{g}(x_1). \) Then, \( \| \hat{\rho}_n(x_1) \| \leq \tilde{g}(x_1), \forall x_1 \in X_1. \) Hence, \( \hat{\rho}_n \) is absolutely integrable over \( X_1 \).

By Proposition 11.92, \( \int_{X_1 \times X_2} \varphi_n \, d\mu = \int_{X_1} \hat{\rho}_n \, d\mu_1 = \int_{X_1} \hat{\rho}_n \, d\mu_1, \) where the first equality follows from the definition of \( \hat{\rho}_n \).

By the result of the third preceding paragraph, \( p(x_1) = \lim_{n \in \mathbb{N}} \int_{X_1} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \lim_{n \in \mathbb{N}} \hat{\rho}_n(x_1) = \lim_{n \in \mathbb{N}} \hat{\rho}_n(x_1), \forall x_1 \in U, \) where the second equality follows from the result of the second preceding paragraph; the last equality follows from the definition of \( \hat{\rho}_n \). Note that \( p(x_1) = \vartheta_y = \lim_{n \in \mathbb{N}} \hat{\rho}_n(x_1), \forall x_1 \in X_1 \setminus U. \) Then, \( p(x_1) = \lim_{n \in \mathbb{N}} \hat{\rho}_n(x_1), \forall x_1 \in X_1. \) This implies that, by Proposition 11.51, \( p \) is \( B_1 \)-measurable. Then, we have \( \int_{X_1 \times X_2} f \, d\mu = \lim_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi_n \, d\mu = \lim_{n \in \mathbb{N}} \int_{X_1} \hat{\rho}_n \, d\mu_1 = \int_{X_1} \hat{\rho}_n \, d\mu_1, \) where the first equality follows from Lebesgue Dominated Convergence Theorem 11.91; the second equality follows from the result of the previous paragraph; and the last equality follows from the Lebesgue Dominated Convergence Theorem 11.91. Hence, (i) holds.

By symmetry, (ii) also holds. This completes the proof of the theorem.

\[ \square \]

**Proposition 12.32** Let \( m \in \mathbb{N}, X_i := (X_i, \mathcal{O}_i) \) be a second countable topological space, \( X := (X_i, \mathcal{B}_i, \mu_i) \) be a second countable \( \sigma \)-finite topological measure space, \( i = 1, \ldots, m, \mathcal{X} := (\prod_{i=1}^m X_i, \mathcal{O}) := \prod_{i=1}^m X_i \) be the product topological space, and \( (\prod_{i=1}^m X_i, \mathcal{B}, \mu) := \prod_{i=1}^m (X_i, \mathcal{B}_i, \mu_i) \) be the product measure space. Then, \( \mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu) \) is a second countable \( \sigma \)-finite topological measure space, which is said to be the product topological measure space of \( X_1, \ldots, X_m, \) denoted by \( \mathcal{X} = \prod_{i=1}^m X_i. \)

**Proof** By Proposition 12.22, \( \mathcal{B} \) is the \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \) generated by \( \mathcal{E} := \{ \prod_{i=1}^m B_i \subseteq \prod_{i=1}^m X_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \ldots, m\} \}. \) By \( X_i \),
Let $U$ be a topological measure space, we have $\mathcal{B}_i = \mathcal{B}_\mathcal{B}(\mathcal{X}_i)$, $i = 1, \ldots, m$. Let $\bar{\mathcal{X}}_1 := \mathcal{X}_1$ and $\bar{\mathcal{X}}_i := \bar{\mathcal{X}}_{i-1} \times \mathcal{X}_i := ((\cdots (X_1 \times X_2) \times \cdots) \times X_i, \mathcal{O}_i)$, $i = 2, \ldots, m$, $(X_1, \mathcal{B}_1, \bar{\mu}_1) := (X_1, \mathcal{B}_1, \mu_1)$ and $(\cdots (X_1 \times X_2) \times \cdots) \times X_i, \mathcal{B}_i, \bar{\mu}_i) := (\cdots (X_1 \times X_2) \times \cdots) \times X_{i-1}, \mathcal{B}_{i-1}, \bar{\mu}_{i-1}) \times (X_i, \mathcal{B}_i, \mu_i)$, $i = 2, \ldots, m$. By Proposition 3.28, $\bar{\mathcal{X}}_i$, $i = 1, \ldots, m$, are second countable topological spaces, and by Proposition 11.24, $\mathcal{B}_m = \mathcal{B}_\mathcal{B}(\bar{\mathcal{X}}_m)$, $i = 1, \ldots, m$. By Propositions 3.30 and 12.24, $\bar{\mathcal{X}}_m$ is homeomorphic to $\mathcal{X}$ and $(\cdots (X_1 \times X_2) \times \cdots) \times X_m, \mathcal{B}_m, \mu_m)$ is isomeric to $(\prod_{i=1}^m X_i, \mathcal{B}, \mu)$ with the same bijective mapping. Then, we have $\mathcal{B} = \mathcal{B}_\mathcal{B}(\mathcal{X})$. By Proposition 3.28, $\mathcal{X}$ is second countable.

First, consider the special case when $\mathcal{X}_i$, $i = 1, \ldots, m$, are finite. Let $\mathcal{E} := \{ E \in \mathcal{B} \mid \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, (i) \exists \mathcal{U} \in \mathcal{O}$ with $E \subseteq \mathcal{U}$ such that $\mu(\mathcal{U} \setminus E) < \varepsilon; (ii) \exists \mathcal{F} \in \mathcal{O}$ with $F \subseteq E$ such that $\mu(\mathcal{E} \setminus F) < \varepsilon \}$. We will show that $\mathcal{E}$ is a $\sigma$-algebra on $\prod_{i=1}^m \mathcal{X}_i$ and contains $\mathcal{E}$. Then, $\mathcal{E} = \mathcal{B}$ and $\mathcal{X}$ is a second countable finite topological measure space.

Let $\forall \varepsilon := \prod_{i=1}^m B_i \in \mathcal{E}$, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, fix any $i \in \{1, \ldots, m\}$, then $B_i \in \mathcal{B}_i$. Since $\mathcal{X}_i$ is a topological measure space, then $\exists O_i \in \mathcal{O}_i$ with $B_i \subseteq O_i$ such that $\mu_i(O_i \setminus B_i) < \frac{\varepsilon}{1 + m \prod_{j=1}^m \mu_j(B_j)}$. By Proposition 11.27, $\exists X_i \setminus F_i \in \mathcal{O}_i$ with $F_i \subseteq B_i$ such that $\mu_i(B_i \setminus F_i) < \frac{\varepsilon}{1 + m \prod_{j=1}^m \mu_j(B_j)}$. Let $U := \prod_{i=1}^m O_i \in \mathcal{O}$ and $F := \prod_{i=1}^m F_i$. Clearly, $(\prod_{i=1}^m X_i) \setminus F = \bigcup_{i=1}^m ((\prod_{j=1}^{i-1} X_j) \times (X_i \setminus F_i) \times (\prod_{j=i+1}^m X_j)) \in \mathcal{O}$. Then, $F \subseteq E \subseteq U$.

$\mu(U \setminus E) = \mu\left(\prod_{i=1}^m O_i \setminus \left(\prod_{i=1}^m B_i\right)\right)$

$= \mu\left(\prod_{i=1}^m \left(\prod_{j=1}^{i-1} O_j \times (O_i \setminus B_i) \times \prod_{j=i+1}^m O_j\right)\right)$

$\leq \sum_{i=1}^m \mu_i(\prod_{j=1}^{i-1} O_j \times (O_i \setminus B_i) \times \prod_{j=i+1}^m O_j)$

$= \sum_{i=1}^m \mu_i(O_i \setminus B_i) \prod_{j=1, j \neq i}^m \mu_j(O_j)$

$\leq \sum_{i=1}^m \frac{\varepsilon}{1 + m \prod_{j=1, j \neq i}^m (1 + \mu_j(B_j))} \prod_{j=1, j \neq i}^m (1 + \mu_j(B_j)) < \varepsilon$

where the first inequality follows from Proposition 11.6; and the third equality follows from Proposition 12.21, and

$\mu(E \setminus F) = \mu\left(\prod_{i=1}^m B_i \setminus \left(\prod_{i=1}^m F_i\right)\right)$
where the first inequality follows from Proposition 11.6; and the third equality follows from Proposition 12.21. Hence, \( E \notin \mathcal{E} \). This shows that \( \mathcal{E} \subseteq \mathcal{E} \).

Let \( E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \), where \( E_i \subseteq U \) such that \( \mu(U \setminus E_i) < \epsilon/n \). Then, \( \mu(U \setminus E) = \mu(U \setminus E_1) + \mu(U \setminus E_2) + \cdots + \mu(U \setminus E_n) + \mu(U \setminus E) < \epsilon/n + \cdots + \epsilon/n + \epsilon/n = \epsilon \).

Hence, \( \mu \) satisfies (i). Next, consider the case when \( X_i \) is \( \sigma \)-finite. Fix any \( i = 1, \ldots, m \). By \( X_i \) being \( \sigma \)-finite, \( \exists (X_{i,n})_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X_i = \bigcup_{n=1}^{\infty} X_{i,n} \) and \( \mu_i(X_{i,n}) < \epsilon/n \), \( \forall n \in \mathbb{N} \). By \( X_i \) being a topological measure space, without loss of generality, we may assume that \( X_{i,n} \subseteq X_{i,n+1} \) and \( X_{i,n} \in \mathcal{O}, \forall n \in \mathbb{N} \). By Proposition 11.29, let \( X_{i,n} := ((X_{i,n}, \mathcal{O}_{i,n}), \mathcal{B}_{i,n}, \mu_{i,n}) \) be the finite topological measure space of \( X_i \), \( \forall n \in \mathbb{N} \). By the special case, let \( \tilde{X}_{i,n} := \prod_{j=1}^{m} X_{i,n} \) be the second countable finite product topological measure space. By Proposition 12.23, \( \prod_{j=1}^{m} X_{i,n} \) is the finite

- \( \mathcal{E} \subseteq \mathcal{E} \) 
- \( \mathcal{E} \subseteq \mathcal{E} \) 
- \( \mathcal{E} \subseteq \mathcal{E} \)
12.3. PRODUCT MEASURE

measure subspace of \( \prod_{i=1}^{m} X_{i, B, \mu} \). \( \forall E \in B, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N} \), let \( E_{n} := E \cap (\prod_{i=1}^{m} X_{i, n}) \in \mathcal{B}_{n} \). Since \( \hat{X}_{n} \) is a topological measure space, \( \exists U_{n} \in \mathcal{O}_{n} \) with \( E_{n} \subseteq U_{n} \) such that \( \hat{\mu}_{n}(U_{n} \setminus E_{n}) < 2^{-n}\epsilon \). Since \( X_{i, n} \in \mathcal{O}_{i} \), \( i = 1, \ldots, m \), then \( \prod_{i=1}^{m} X_{i, n} \in \mathcal{O} \), which further implies that \( U_{n} \in \mathcal{O} \). Let \( U := \bigcup_{n=1}^{\infty} U_{n} \in \mathcal{O} \). Then, \( E = \bigcup_{n=1}^{\infty} E_{n} \subseteq \bigcup_{n=1}^{\infty} U_{n} = U \) and \( \mu(U \setminus E) = \mu(\bigcup_{n=1}^{\infty}(U_{n} \setminus E)) \leq \sum_{n=1}^{\infty} \mu(U_{n} \setminus E) \leq \sum_{n=1}^{\infty} \hat{\mu}_{n}(U_{n} \setminus E_{n}) < \epsilon \), where the first inequality follows from Proposition 11.6; the second inequality follows from Proposition 11.4; and the equality follows from Proposition 11.13. By the arbitrariness of \( E \), we have \( X \) is a \( \sigma \)-finite topological measure space. It is second countable since \( X \) is second countable. This completes the proof of the proposition. \( \square \)

Proposition 12.33 Let \( m \in \mathbb{N} \), \( X_{i} := (X_{i}, \mathcal{O}_{i}) \) be a second countable topological space, \( X_{i} := (X_{i}, \mathcal{B}_{i}, \mu_{i}) \) be a second countable \( \sigma \)-finite \( K \)-valued topological measure space, \( i = 1, \ldots, m \), \( X := (\prod_{i=1}^{m} X_{i}, \mathcal{O}) := \prod_{i=1}^{m} X_{i} \) be the product topological space, and \( (\prod_{i=1}^{m} X_{i}, B, \mu) := \prod_{i=1}^{m} (X_{i}, B_{i}, \mu_{i}) \) be the product \( K \)-valued measure space. Then, \( X := (X, \mathcal{B}, \mu) \) is a second countable \( \sigma \)-finite \( K \)-valued topological measure space, which is said to be the product \( K \)-valued topological measure space of \( X_{1}, \ldots, X_{m} \), denoted by \( X = \prod_{i=1}^{m} X_{i} \).

Proof Let \( \hat{X}_{i} := (X_{i}, B_{i}, P \circ \mu_{i} =: \nu_{i}), i = 1, \ldots, m \), and \( \hat{X} := (\prod_{i=1}^{m} X_{i}, B, P \circ \mu =: \nu) \). By Proposition 12.22, \( \hat{X}_{i}, i = 1, \ldots, m \), and \( \hat{X} \) are \( \sigma \)-finite measure spaces and \( \hat{X} = \prod_{i=1}^{m} \hat{X}_{i} \). Since \( X_{i} \) is a second countable \( \sigma \)-finite \( K \)-valued topological measure space, then, by Definition 11.191, \( \hat{X}_{i} := (X_{i}, B_{i}, \nu_{i}) \) is a second countable \( \sigma \)-finite topological measure space, \( i = 1, \ldots, m \). By Proposition 12.32, \( \hat{X} := (X, B, \nu) \) is a second countable \( \sigma \)-finite topological measure space. Then, \( \hat{X} \) is a second countable \( \sigma \)-finite \( K \)-valued topological measure space. This completes the proof of the proposition. \( \square \)

Proposition 12.34 Let \( m \in \mathbb{N} \), \( X_{i} := (X_{i}, B_{i}, \mu_{i}) \) be a \( \sigma \)-finite measure space, \( \nu_{i} \) be a \( \sigma \)-finite \( K \)-valued measure on \( (X_{i}, B_{i}) \) with \( \frac{d\nu_{i}}{d\mu_{i}} = f_{i} : X_{i} \to K \) a.e. in \( X_{i} \), \( i = 1, \ldots, m \), and \( f : \prod_{i=1}^{m} X_{i} \to K \) be defined by \( f(x_{1}, \ldots, x_{m}) = \prod_{i=1}^{m} f_{i}(x_{i}), \forall (x_{1}, \ldots, x_{m}) \in \prod_{i=1}^{m} X_{i} \). Then, \( \frac{d\nu}{d\mu} = f \) a.e. in \( X \), where \( \nu := \prod_{i=1}^{m} \nu_{i}, \mu := \prod_{i=1}^{m} \mu_{i} \), and \( X := \prod_{i=1}^{m} X_{i}, \mathcal{B}, \mu \) := \( \prod_{i=1}^{m} \mathcal{B}_{i} \).

Proof By Proposition 12.22, \( X \) is a \( \sigma \)-finite measure space. By Propositions 7.23, 11.38, and 11.39, \( f \) is \( B \)-measurable. By Proposition 11.116 and Definition 11.166, \( f \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( K \)-valued measure \( \hat{\nu} \) on \( (\prod_{i=1}^{m} X_{i}, \mathcal{B}) \) with respect to \( \mu \). By Proposition 11.167, \( \frac{d\nu}{d\mu} = f \) a.e. in \( X \).

First, consider the special case when \( \nu_{i}, i = 1, \ldots, m, \) are finite \( K \)-valued measures. Let \( E := \{ E \in B \mid \nu(E) = \hat{\nu}(E) \} \) and \( \mathcal{C} := \{ \prod_{i=1}^{m} B_{i} \subseteq \)
\[ \prod_{i=1}^{m} X_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \ldots, m\} \], where \( \mathcal{C} \) is the set of measurable rectangles in \( \prod_{i=1}^{m} X_i \), \( \forall E := \prod_{i=1}^{m} B_i \in \mathcal{C} \), we have \( \nu(E) = \prod_{i=1}^{m} \nu_i(B_i) \) by Proposition 12.21. By repeated application of Tonelli Theorem 12.29 and Proposition 12.24, we have

\[
\int_{\prod_{i=1}^{m} X_i} \mathcal{P} \circ f \, d\mu = \left( \int_{\prod_{i=1}^{m-1} X_i} \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d(\prod_{i=1}^{m-1} \mu_i) \right) \left( \int_{X_m} \mathcal{P} \circ f_m \, d\mu_m \right)
= \left( \int_{\prod_{i=1}^{m-1} X_i} \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d(\prod_{i=1}^{m-1} \mu_i) \right) \mathcal{P} \circ \nu_m(X_m) = \cdots = \prod_{i=1}^{m} \mathcal{P} \circ \nu_i(X_i)
< \infty
\]

where the first equality follows from Tonelli Theorem 12.29; and the second equality follows from Definition 11.166. Hence, \( f \) is absolutely integrable over \( \mathcal{X} \) and \( \tilde{\nu} \) is a finite \( \mathbb{K} \)-valued measure by Proposition 11.116. By Proposition 11.83, \( f \chi_{\prod_{i=1}^{m} B_i, \prod_{i=1}^{m} X_i} \), is absolutely integrable over \( \mathcal{X} \). By repeated application of Fubini Theorem 12.30,

\[
\tilde{\nu}(E) = \int_{\prod_{i=1}^{m} B_i} f \, d\mu = \int_{\prod_{i=1}^{m} X_i} f \chi_{\prod_{i=1}^{m} B_i, \prod_{i=1}^{m} X_i} \, d(\prod_{i=1}^{m} \mu_i)
= \left( \int_{\prod_{i=1}^{m-1} X_i} \left( \prod_{i=1}^{m-1} f_i \right) \chi_{\prod_{i=1}^{m-1} B_i, \prod_{i=1}^{m-1} X_i} \, d(\prod_{i=1}^{m-1} \mu_i) \right) \left( \int_{B_m} f_m \, d\mu_m \right)
= \left( \int_{\prod_{i=1}^{m-1} X_i} \left( \prod_{i=1}^{m-1} f_i \right) \chi_{\prod_{i=1}^{m-1} B_i, \prod_{i=1}^{m-1} X_i} \, d(\prod_{i=1}^{m-1} \mu_i) \right) \nu_m(B_m)
= \cdots = \prod_{i=1}^{m} \nu_i(B_i) = \nu(E)
\]

where the first equality follows from Definition 11.166; the second equality follows from Proposition 11.92; the third equality follows from Fubini Theorem; and the fourth equality follows from Definition 11.166. This implies that \( E \in \mathcal{E} \). By the arbitrariness of \( E \), we have \( \mathcal{C} \subseteq \mathcal{E} \).

Clearly, \( \mathcal{C} \) is a \( \pi \)-system on \( \prod_{i=1}^{m} X_i \). We will show that \( \mathcal{E} \) is a monotone class on \( \prod_{i=1}^{m} X_i \). Then, by Monotone Class Lemma 12.19, we have \( \mathcal{E} = \mathcal{B} \). This further implies that \( \nu = \tilde{\nu} \) and \( \frac{d\mu}{d\nu} = f \) a.e. in \( \mathcal{X} \), which completes the proof of the special case.

Note that \( \emptyset = \prod_{i=1}^{m} \emptyset \subseteq \mathcal{E} \subseteq \mathcal{C} \subseteq \mathcal{E} \) and \( \prod_{i=1}^{m} X_i \subseteq \mathcal{C} \subseteq \mathcal{E} \). \( \forall E_1, E_2 \in \mathcal{E} \) with \( E_1 \subseteq E_2 \), we have \( \nu(E_2 \setminus E_1) = \nu(E_2) - \nu(E_1) = \tilde{\nu}(E_2) - \tilde{\nu}(E_1) = \tilde{\nu}(E_2 \setminus E_1) \), where the first equality follows from the fact that \( \nu \) is a finite \( \mathbb{K} \)-valued measure; the second equality follows from the fact that \( E_1, E_2 \in \mathcal{E} \); and the third equality follows from fact that \( \tilde{\nu} \) is a finite \( \mathbb{K} \)-valued measure. Hence, \( E_2 \setminus E_1 \in \mathcal{E} \).

\( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E} \) with \( E_i \subseteq E_{i+1} \), \( \forall i \in \mathbb{N} \), let \( E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \). Then,

\[
\nu(E) = \lim_{i \in \mathbb{N}} \nu(E_i) = \lim_{i \in \mathbb{N}} \tilde{\nu}(E_i) = \tilde{\nu}(E), \quad \text{where the first equality}
\]
follows from Proposition 11.112 and the fact that $\nu$ is finite; the second equality follows from the fact that $E_i \in \mathcal{E}$, $\forall i \in \mathbb{N}$; and the third equality follows from Proposition 11.112 and the fact that $\nu$ is finite. Then, $E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{E}$.

Hence, $\mathcal{E}$ is a monotone class on $\prod_{i=1}^{m} X_i$. This completes the proof of the special case.

Next, consider the general case when $\nu_i$, $i = 1, \ldots, m$, are $\sigma$-finite. Then, there exists $(X_{i,n})_{n=1}^{\infty} \subseteq \mathcal{B}_i$ such that $X_i = \bigcup_{n=1}^{\infty} X_{i,n}$ and $\mathcal{P} \circ \nu_i(X_{i,n}) < \infty$, $\forall n \in \mathbb{N}$, $i = 1, \ldots, m$. Without loss of generality, we may assume that $(X_{i,n})_{n=1}^{\infty}$ is pairwise disjoint, $i = 1, \ldots, m$. For any $i = 1, \ldots, m$ and $\forall n \in \mathbb{N}$, let $X_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \mu_{i,n})$ be the $\sigma$-finite measure subspace of $X_i$, $\tilde{X}_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \nu_{i,n})$ be the finite $\mathbb{K}$-valued measure subspace of $X_i$, $\mathcal{X}_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \bar{\mu}_{i,n})$ be the finite $\mathbb{K}$-valued measure space, and $\tilde{X}_{i,n,m} := (\prod_{i=1}^{m} X_{i,n}, \mathcal{B}_{i_{1},n_{1}} \times \cdots \times \mathcal{B}_{i_{m},n_{m}}, \bar{\nu}_{1,n_{1},\ldots,n_{m}})$ be the finite product $\mathbb{K}$-valued measure space. By Proposition 12.23, $\tilde{X}_{i_{1},n_{1}} \times \cdots \times \tilde{X}_{i_{m},n_{m}}$ is the measure subspace of $\mathcal{X}$ and $\tilde{X}_{i_{1},n_{1}} \times \cdots \times \tilde{X}_{i_{m},n_{m}}$ is the finite $\mathbb{K}$-valued measure subspace of $\tilde{X} := (\prod_{i=1}^{m} X_i, \mathcal{B}, \nu)$. By Definition 11.166, we have $\frac{d\nu_{1,n_{1}}}{d\mu_{1,n_{1}}} = f_{1}|_{X_{1,n_{1}}}$ a.e. in $X_{1,n_{1}}$, $i = 1, \ldots, m$. By the special case, we have $\frac{d\tilde{\nu}_{1,n_{1},\ldots,n_{m}}}{d\bar{\mu}_{1,n_{1},\ldots,n_{m}}} = f|_{\prod_{i=1}^{m} X_{i,n_{i}}}$ a.e. in $\tilde{X}_{i_{1},n_{1}} \times \cdots \times \tilde{X}_{i_{m},n_{m}}$, $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$, let $E_{1,n_{1}} \times \cdots \times E_{m,n_{m}} := E \cap (\prod_{i=1}^{m} X_{i,n_{i}}) \in \mathcal{B}_{1,n_{1}} \times \cdots \times \mathcal{B}_{m,n_{m}}$. Then, we have

$$\mathcal{P} \circ \nu(E) = \sum_{n_{1}=1}^{\infty} \cdots \sum_{n_{m}=1}^{\infty} \mathcal{P} \circ \nu(E_{1,n_{1}} \times \cdots \times E_{m,n_{m}})$$

$$= \sum_{n_{1}}^{\infty} \cdots \sum_{n_{m}}^{\infty} \mathcal{P} \circ \tilde{\nu}_{1,n_{1},\ldots,n_{m}}(E_{1,n_{1}} \times \cdots \times E_{m,n_{m}})$$

$$= \sum_{n_{1}}^{\infty} \cdots \sum_{n_{m}}^{\infty} \int_{E_{1,n_{1}} \times \cdots \times E_{m,n_{m}}} \mathcal{P} \circ (f|_{\prod_{i=1}^{m} X_{i,n_{i}}}) d\bar{\mu}_{1,n_{1},\ldots,n_{m}}$$

$$= \sum_{n_{1}=1}^{\infty} \cdots \sum_{n_{m}=1}^{\infty} \int_{E_{1,n_{1}} \times \cdots \times E_{m,n_{m}}} \mathcal{P} \circ f d\mu = \int_{E} \mathcal{P} \circ f d\mu < \infty$$

where the first equality follows from the fact that $\mathcal{P} \circ \nu$ is a measure; the second equality follows from Proposition 11.115; the third equality follows from the fact that $\frac{d\tilde{\nu}_{1,n_{1},\ldots,n_{m}}}{d\bar{\mu}_{1,n_{1},\ldots,n_{m}}} = f|_{\prod_{i=1}^{m} X_{i,n_{i}}}$ a.e. in $\tilde{X}_{i_{1},n_{1}} \times \cdots \times \tilde{X}_{i_{m},n_{m}}$, and the last two equalities follow from Proposition 11.83. We also have

$$\nu(E) = \sum_{n_{1}=1}^{\infty} \cdots \sum_{n_{m}=1}^{\infty} \nu(E_{1,n_{1}} \times \cdots \times E_{m,n_{m}}) = \sum_{n_{1}=1}^{\infty} \cdots \sum_{n_{m}=1}^{\infty} \tilde{\nu}_{1,n_{1},\ldots,n_{m}}(E_{1,n_{1}} \times \cdots \times E_{m,n_{m}})$$

$$= \sum_{n_{1}=1}^{\infty} \cdots \sum_{n_{m}=1}^{\infty} \int_{E_{1,n_{1}} \times \cdots \times E_{m,n_{m}}} f|_{\prod_{i=1}^{m} X_{i,n_{i}}} d\bar{\mu}_{1,n_{1},\ldots,n_{m}}$$
Theorem 12.35 (Fubini) Let \( \mathcal{X}_i := (X_i, \mathcal{B}_i, \mu_i) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( i = 1, 2, \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := \mathcal{X}_1 \times \mathcal{X}_2 \) be the product \( \mathbb{K} \)-valued measure space, \( \mathcal{Y} \) be a separable Banach space over \( \mathbb{K} \), and \( f : X_1 \times X_2 \to \mathcal{Y} \) be absolutely integrable over \( \mathcal{X} \). Define \( f_{x_1} : X_2 \to \mathcal{Y} \) and \( f_{(x_2)} : X_1 \to \mathcal{Y} \) as in Proposition 12.28. Then, the following statements hold.

(i) There exists \( \tilde{p} : X_1 \to \mathcal{Y} \) and exists \( \tilde{U} \in \mathcal{B}_1 \) such that \( \mathcal{P} \circ \mu_1(X_1 \setminus \tilde{U}) = 0 \); \( \forall x_1 \in \tilde{U}, f_{x_1} \) is absolutely integrable over \( X_2 \) and \( \tilde{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y} \); \( \forall x_1 \in X_1 \setminus \tilde{U}, \tilde{p}(x_1) = 0 \); \( \tilde{p} \) is absolutely integrable over \( \mathcal{X}_1 \); and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_1} \tilde{p} \, d\mu_1 \in \mathcal{Y} \).

(ii) There exists \( \tilde{q} : X_2 \to \mathcal{Y} \) and exists \( \tilde{V} \in \mathcal{B}_2 \) such that \( \mathcal{P} \circ \mu_2(X_2 \setminus \tilde{V}) = 0 \); \( \forall x_2 \in \tilde{V}, f_{(x_2)} \) is absolutely integrable over \( \mathcal{X}_1 \) and \( \tilde{q}(x_2) = \int_{X_1} f_{(x_2)} \, d\mu_1 \in \mathcal{Y} \); \( \forall x_2 \in X_2 \setminus \tilde{V}, \tilde{q}(x_2) = 0 \); \( \tilde{q} \) is absolutely integrable over \( \mathcal{X}_2 \); and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_2} \tilde{q} \, d\mu_2 \in \mathcal{Y} \).

(iii) \( \forall p : X_1 \to \mathcal{Y} \) satisfying: \( p \) is \( \mathcal{B}_1 \)-measurable; and \( \exists U \in \mathcal{B}_1 \) with \( \mathcal{P} \circ \mu_1(X_1 \setminus U) = 0 \) such that, \( \forall x_1 \in U, f_{x_1} \) is absolutely integrable over \( \mathcal{X}_2 \) and \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y} \), we have \( p \) is absolutely integrable over \( \mathcal{X}_1 \) and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \in \mathcal{Y} \).

(iv) \( \forall q : X_2 \to \mathcal{Y} \) satisfying: \( q \) is \( \mathcal{B}_2 \)-measurable; and \( \exists V \in \mathcal{B}_2 \) with \( \mathcal{P} \circ \mu_2(X_2 \setminus V) = 0 \) such that, \( \forall x_2 \in V, f_{(x_2)} \) is absolutely integrable over \( \mathcal{X}_1 \) and \( q(x_2) = \int_{X_1} f_{(x_2)} \, d\mu_1 \in \mathcal{Y} \), we have \( q \) is absolutely integrable over \( \mathcal{X}_2 \) and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2 \in \mathcal{Y} \).

Proof

(i) Let \( \mathcal{X}_i := (X_i, \mathcal{B}_i, \mathcal{P} \circ \mu_i =: \nu_i), i = 1, 2 \), and \( \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mathcal{P} \circ \mu =: \nu) \). By Proposition 12.22, \( \mathcal{X}_1, \mathcal{X}_2, \) and \( \mathcal{X} \) are \( \sigma \)-finite measure spaces and \( \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \). Fix any \( i = 1, 2 \). By Radon-Nikodym Theorem 11.171, we may define functions \( \frac{d\nu_i}{d\nu} =: g_i : \mathcal{X}_i \to \mathbb{K}, \forall E \in \mathcal{B}_i, \nu_i(E) = \mathcal{P} \circ \mu_i(E) = \int_E \mathcal{P} \circ g_i \, d\nu \). By Proposition 11.167, we have \( \mathcal{P} \circ g_1 = 1 \) a.e. in \( \mathcal{X}_1 \). By Proposition 12.34, we have \( \frac{d\nu}{d\nu} = g_1g_2 \) a.e. in \( \mathcal{X} \).
Since $f$ is absolutely integrable over $\mathcal{X}$, then, by Proposition 11.168, $\int_{X_1 \times X_2} f \, dm = \int_{X_1 \times X_2} (f g_1) \, d\nu$ and $(P \circ f) (P \circ (g_1 g_2)) = (P \circ f) (P \circ g_1) (P \circ g_2)$ is integrable over $\mathcal{X}$. Define $\tilde{f} := fg_1g_2 : X_1 \times X_2 \to Y$. By Fubini Theorem 12.30, there exists $\tilde{p} : X_1 \to Y$ and exists $\tilde{U} \in \mathcal{B}_1$ such that $\nu_1(X_1 \setminus \tilde{U}) = 0$, $\forall x_1 \in \tilde{U}$, $\tilde{f}$ is absolutely integrable over $\mathcal{X}_2$ and $\tilde{p}(x_1) = \int_{X_2} f_{x_1} \, d\nu_2$; $\forall x_1 \in X_1 \setminus \tilde{U}$, $\tilde{p}(x_1) = \vartheta_Y$; $\tilde{p}$ is absolutely integrable over $\mathcal{X}_1$; and $\int_{X_1 \times X_2} \tilde{f} \, d\nu = \int_{X_1} \tilde{p} \, d\nu_1$. Let $A_1 := \{x_1 \in X_1 \mid g_1(x_1) = 0\}$. Clearly, $A_1 \in \mathcal{B}_1$ and $\nu_1(A_1) = 0$. Let $\tilde{U} := \tilde{U} \setminus A_1 \in \mathcal{B}_1$, then, $P \circ \mu_1(X_1 \setminus \tilde{U}) = \nu_1(X_1 \setminus \tilde{U}) = \nu_1((X_1 \setminus \tilde{U}) \cup A_1) = 0$. $\forall x_1 \in \tilde{U}$, $\tilde{f}$ is absolutely integrable over $\mathcal{X}_2$, which implies that $f_{x_1} g_1(x_1) g_2$ is absolutely integrable over $\mathcal{X}_2$, which further implies that $(P \circ f_{x_1})(P \circ g_2)$ is integrable over $\mathcal{X}_2$. By Proposition 11.168, $f_{x_1}$ is absolutely integrable over $\mathcal{X}_1$ and $\int_{X_1 \times X_2} f_{x_1} \, d\mu_2 = \int_{X_1} (f_{x_1} g_2) \, d\nu_2$. Then, $\tilde{p}(x_1) = g_1(x_1) \int_{X_2} f_{x_1} \, d\mu_2$.

Define $\tilde{p} : X_1 \to Y$ by $\tilde{p}(x_1) = \begin{cases} \int_{X_2} f_{x_1} \, d\mu_2 & x_1 \in \tilde{U} \\ \vartheta_y & x_1 \in X_1 \setminus \tilde{U} \end{cases}$. By Propositions 11.41, 11.38, and 11.39, $\tilde{p}$ is $\mathcal{B}_1$-measurable. Since $\tilde{p}$ and $g_1 \tilde{p}$ are $\mathcal{B}_1$-measurable, then, by Lemma 11.43, $\tilde{p} = g_1 \tilde{p}$ a.e. in $X_1$. Since $\tilde{p}$ is absolutely integrable over $\mathcal{X}_1$, then, $P \circ (g_1 \tilde{p}) = (P \circ \tilde{p})(P \circ g_1)$ is integrable over $\mathcal{X}_1$. By Proposition 11.168, $\tilde{p}$ is absolutely integrable over $\mathcal{X}_1$ and $\int_{X_1} (g_1 \tilde{p}) \, d\mu_1 = \int_{X_1} \tilde{p} \, d\mu_1$.

By Proposition 11.92, we have $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1 \times X_2} \tilde{f} \, d\nu = \int_{X_1} \tilde{p} \, d\nu_1 = \int_{X_1} \tilde{p} \, d\mu_1$. Hence, (i) holds.

By symmetry, (ii) holds.

(iii) Fix any $p : X_1 \to Y$ satisfying: $p$ is $\mathcal{B}_1$-measurable; and $\exists U \in \mathcal{B}_1$ with $P \circ \mu_1(X_1 \setminus U) = 0$ such that, $\forall x_1 \in U$, $f_{x_1}$ is absolutely integrable over $\mathcal{X}_2$ and $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in Y$. By (i), we have $U_0 := \tilde{U} \cap U \in \mathcal{B}_1$ and $P \circ \mu_1((X_1 \setminus \tilde{U}) \cup (X_1 \setminus U)) = 0$. $\forall x_1 \in U_0$, $p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 = \tilde{p}(x_1) \in Y$. By Lemma 11.43, $p = \tilde{p}$ a.e. in $X_1$. By Proposition 11.83, we have $p$ is absolutely integrable over $\mathcal{X}_1$ since $\tilde{p}$ is so. By Proposition 11.92, $\int_{X_1} p \, d\mu_1 = \int_{X_1} \tilde{p} \, d\mu_1 = \int_{X_1 \times X_2} f \, d\mu \in Y$. Hence, (iii) holds.

By symmetry, (iv) holds. This completes the proof of the theorem. □

12.4 Functions of Bounded Variation

Definition 12.36 Let $m \in \mathbb{N}$, $R^m$ be endowed with the usual positive cone, $Y$ be a normed linear space over $K$, $D \subseteq R^m$, and $F : D \to Y$. $F$ is said to be continuous on the right at $x_0 \in D$ if $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in B_{R^m} (x_0, \delta) \cap D$ with $x_0 \preceq x$, we have $\|F(x) - F(x_0)\| < \varepsilon$. $F$ is said to be continuous on the right if it is continuous on the right at $x$, $\forall x \in D$. $F$ is said to be continuous on the right on $E \subseteq D$ if it is continuous on the
Definition 12.38 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone $P$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $x_0 \in \Omega \subseteq \mathbb{R}^m$, and $F : \Omega \rightarrow \mathcal{Y}$. Then, the following statements are equivalent.

(i) $F$ is continuous on the right at $x_0$.

(ii) If $x_0$ is an accumulation point of $\Omega \cap (x_0 + P)$, then $\lim_{h \to a_m} F(x_0 + h) = F(x_0)$.

(iii) If $x_0$ is an accumulation point of $\Omega \cap (x_0 + P)$, then $\forall (h_n)_{n=1}^{\infty} \subseteq (P \cap (\Omega - x_0)) \setminus \{0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, we have $\lim_{h \in \mathbb{R}_+} F(x_0 + h_n) = F(x_0)$.

Proof (i) $\Rightarrow$ (iii). Let $F$ be continuous on the right at $x_0$ and $x_0$ be an accumulation point of $\Omega \cap (x_0 + P)$. By (i), $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}_+}(x_0, \delta) \cap \Omega$ with $x_0 \triangleq x$, we have $\|F(x) - F(x_0)\| < \epsilon$. Hence, $\forall (h_n)_{n=1}^{\infty} \subseteq (P \cap (\Omega - x_0)) \setminus \{0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, we have $\lim_{h \in \mathbb{R}_+} F(x_0 + h_n) = F(x_0)$. This implies that $x_0 + h_n \in \mathcal{B}_{\mathbb{R}_+}(x_0, \delta) \cap \Omega$, $x_0 + h_n \supseteq x_0$, and $\|F(x_0 + h_n) - F(x_0)\| < \epsilon$. Hence, $\lim_{h \in \mathbb{R}_+} F(h) = F(x_0)$.

(ii) $\Rightarrow$ (i). We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0$ is not an accumulation point of $\Omega \cap (x_0 + P)$; Case 2: $x_0$ is an accumulation point of $\Omega \cap (x_0 + P)$. Case 1: $x_0$ is not an accumulation point of $\Omega \cap (x_0 + P)$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_{\mathbb{K}}(x_0, \delta) \cap ((\Omega \cap (x_0 + P)) \setminus \{x_0\}) = \emptyset$ and $\mathcal{B}_{\mathbb{R}_+}(x_0, \delta) \cap \Omega \cap (x_0 + P) = \{x_0\}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}_+}(x_0, \delta) \cap \Omega$ with $x_0 \triangleq x$, we have $x = x_0$ and $\|F(x) - F(x_0)\| = 0 < \epsilon$. Hence, $F$ is continuous on the right at $x_0$.

Case 2: $x_0$ is an accumulation point of $\Omega \cap (x_0 + P)$. By (ii), $\lim_{h \in \mathbb{R}_+} F(x_0 + h) = F(x_0)$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall h \in \mathcal{B}_{\mathbb{K}}(0_m, \delta) \cap \Omega \cap (\Omega - x_0) \setminus \{0_m\}$, we have $\|F(x_0 + h) - F(x_0)\| < \epsilon$. Then, $\forall x \in \mathcal{B}_{\mathbb{R}_+}(x_0, \delta) \cap \Omega \cap (x_0 + P)$, we have $x-x_0 \in \mathcal{B}_{\mathbb{R}_+}(0_m, \delta) \cap (\Omega - x_0) \cap P$ and $\|F(x) - F(x_0)\| < \epsilon$. Hence, $F$ is continuous on the right at $x_0$.

In both cases, $F$ is continuous on the right at $x_0$.

This completes the proof of the proposition.  

Definition 12.38 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone. $x_1, x_2 \in \mathbb{R}^m$ with $x_1 \lesssim x_2$, the semi-open rectangle with corners $x_1$ and $x_2$ is $r_{x_1, x_2} := \left\{ x \in \mathbb{R}^m \mid x_1 < x \lesssim x_2 \right\}$. The closed rectangle
with corners $x_1$ and $x_2$ is $\mathbf{r}_{x_1,x_2} := \{ x \in \mathbb{R}^m \mid x_1 \leq x \leq x_2 \}$. The open rectangle with corners $x_1$ and $x_2$ is $\mathbf{r}^0_{x_1,x_2} := \{ x \in \mathbb{R}^m \mid x_1 < x < x_2 \}$.

$\Omega \subseteq \mathbb{R}^m$ is said to be a region if $\exists N \subseteq \mathbb{N}$ and $(\hat{x}_i)_{i \in N}, (\check{x}_i)_{i \in N} \subseteq \Omega$ with $\check{x}_i \searrow \hat{x}_i$ and $\mathbf{r}_{\check{x}_i,\hat{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \mathbf{r}_{\check{x}_i,\hat{x}_i}$, $(\check{x}_i,\hat{x}_i)_{i \in N}$ is pairwise disjoint, $P(\Omega) = \bigcup_{i \in N} \mathbf{r}_{\check{x}_i,\hat{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\mathbf{r}_{x_1,x_2} \subseteq \Omega$, we have $\mathbf{r}_{x_1,x_2} \subseteq P(\Omega)$. $P(\Omega)$ is uniquely defined for each region $\Omega$ (see below) and is called the principal of $\Omega$.

Clearly, any region $\Omega \subseteq \mathbb{R}^m$ we have $\Omega \in B_\mathbb{B}(\mathbb{R}^m)$ and $P(\Omega) \in B_\mathbb{B}(\mathbb{R}^m)$.

**Proposition 12.39** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone. Then, the following holds.

(i) If $\Omega \subseteq \mathbb{R}^m$ is a region, then, $P(\Omega) \subseteq \Omega$ is uniquely defined, $\Omega \in B_\mathbb{B}(\mathbb{R}^m)$, and $P(\Omega) \in B_\mathbb{B}(\mathbb{R}^m)$.

(ii) If $\Omega_1, \Omega_2 \subseteq \mathbb{R}^m$ are regions with $\Omega_1 \subseteq \Omega_2$, then $P(\Omega_1) \subseteq P(\Omega_2)$.

(iii) If $\Omega_1, \Omega_2, \Omega_1 \cup \Omega_2 \subseteq \mathbb{R}^m$ are regions, then $P(\Omega_1) \cup P(\Omega_2) \subseteq P(\Omega_1 \cup \Omega_2)$.

(iv) If $\Omega_1, \Omega_2, \Omega_1 \cap \Omega_2 \subseteq \mathbb{R}^m$ are regions, then $P(\Omega_1) \cap P(\Omega_2) \supseteq P(\Omega_1 \cap \Omega_2)$.

(v) Let $\Omega \in B_\mathbb{B}(\mathbb{R}^m)$ be a rectangle, that is $\exists$ intervals $I_1, \ldots, I_m$ such that $\Omega = \prod_{i=1}^{m} I_i$, $\forall i \in \{1, \ldots, m\}$: if $I_i = [a_i, b_i] \subseteq \mathbb{R}$ with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, let $\hat{I}_i := (a_i, b_i) \subseteq \mathbb{R}$; if $I_i = [a_i, \infty) \subseteq \mathbb{R}$ with $a_i \in \mathbb{R}$, let $\hat{I}_i := [a_i, \infty) \subseteq \mathbb{R}$; if $I_i = (\infty, b_i) \subseteq \mathbb{R}$ with $b_i \in \mathbb{R}$, let $\hat{I}_i := (\infty, b_i) \subseteq \mathbb{R}$; in all other cases, let $\hat{I}_i := I_i$. Then, $\Omega$ is a region and $P(\Omega) = \prod_{i=1}^{m} \hat{I}_i$ is a rectangle.

(vi) $\forall x_1, x_2 \in \mathbb{R}^m$ with $x_1 \leq x_2$, we have $\mathbf{r}_{x_1,x_2}$ and $\mathbf{r}_{x_1,x_2}$ are regions with $P(\mathbf{r}_{x_1,x_2}) = P(\mathbf{r}_{x_1,x_2}) = \mathbf{r}_{x_1,x_2}$.

(vii) Let $M \subseteq \mathbb{N}$ and $(\hat{x}_j)_{j \in M}, (\check{x}_j)_{j \in M} \subseteq \mathbb{R}^m$ with $\hat{x}_j \leq \check{x}_j, \forall j \in M$.

Then, $\Omega := \bigcup_{j \in M} \mathbf{r}_{\hat{x}_j,\check{x}_j} \subseteq \mathbb{R}^m$ is a region and $P(\Omega) = \Omega$.

(viii) If $\Omega_1 \subseteq \mathbb{R}^l$ and $\Omega_2 \subseteq \mathbb{R}^{m-l}$ are regions, where $l \in \{1, \ldots, m-1\}$, then $\Omega_1 \times \Omega_2 \subseteq \mathbb{R}^m$ is a region with $P(\Omega_1 \times \Omega_2) = P(\Omega_1) \times P(\Omega_2)$.

**Proof**

(i) Let $j = 1, 2$, and $\exists N_j \subseteq \mathbb{N}$ and $(\hat{x}_{j,i})_{i \in N_j}, (\check{x}_{j,i})_{i \in N_j} \subseteq \Omega$ with $\check{x}_{j,i} \leq \hat{x}_{j,i}$ and $\mathbf{r}_{\check{x}_{j,i},\hat{x}_{j,i}} \subseteq \Omega, \forall i \in N_j$, such that $\Omega = \bigcup_{i \in N_j} \mathbf{r}_{\check{x}_{j,i},\hat{x}_{j,i}}$.

$(\mathbf{r}_{\check{x}_{j,i},\hat{x}_{j,i}})_{i \in N_j}$ is pairwise disjoint, $\Omega_j := \bigcup_{i \in N_j} \mathbf{r}_{\check{x}_{j,i},\hat{x}_{j,i}}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\mathbf{r}_{x_1,x_2} \subseteq \Omega$, we have $\mathbf{r}_{x_1,x_2} \subseteq \Omega_j$. We will show that $\Omega_1 = \Omega_2$. Then, the result follows.
∀i ∈ N, we have \( x_{1,i} \leq x_{1,i} \) and \( r_{x_{1,i},x_{1,i}} \subseteq \Omega \). This implies that \( r_{x_{1,i},x_{1,i}} \subseteq \Omega_n \). Hence, \( \Omega_1 = \bigcup_{i \in N_1} r_{x_{1,i},x_{1,i}} \subseteq \Omega_2 \). By symmetry, we have \( \Omega_2 \subseteq \Omega_1 \).

(ii) By \( \Omega_1 \) being a region, \( \exists N_1 \subseteq N \) and \((\bar{x}_{1,i})_{i \in N_1}, (\bar{x}_{1,i})_{i \in N_1} \subseteq \Omega_1 \subseteq \Omega_2 \) with \( \bar{x}_{1,i} \leq \bar{x}_{1,i} \) and \( r_{\bar{x}_{1,i},\bar{x}_{1,i}} \subseteq \Omega_1 \subseteq \Omega_2 \), \( \forall i \in N_1 \), such that \( \Omega_1 = \bigcup_{i \in N_1} r_{\bar{x}_{1,i},\bar{x}_{1,i}} \) and \( P(\Omega_1) = \bigcup_{i \in N_1} r_{\bar{x}_{1,i},\bar{x}_{1,i}} \). By \( \Omega_2 \) being a region, we have \( r_{x_{1,i},x_{1,i}} \subseteq P(\Omega_2) \), \( \forall i \in N_1 \). Hence, \( P(\Omega_1) \subseteq P(\Omega_2) \).

(iii) By (ii), \( P(\Omega_1 \cup \Omega_2) \supseteq P(\Omega_j), \) \( i = 1, 2 \). Then, the result follows.

(iv) By (ii), \( P(\Omega_1 \cap \Omega_2) \subseteq P(\Omega_1), \) \( i = 1, 2 \). Then, the result follows.

(v) This is straightforward, and is therefore omitted.

(vi) This follows immediately from (v).

(vii) We will first prove the result under the additional assumption that \( (r_{x_{j,i},x_{j,i}})_{j \in M} =:= (\Omega_j)_{j \in M} \) is pairwise disjoint. \( \forall j \in M \), by (vi), \( \Omega_j \) is a region and \( P(\Omega_j) = \Omega_j \). Then, \( \exists N_1 \subseteq N \) and \((\bar{x}_{j,i})_{i \in N_1}, (\bar{x}_{j,i})_{i \in N_1} \subseteq \Omega_j \) with \( \bar{x}_{j,i} \leq \bar{x}_{j,i} \), \( \forall i \in N_1 \), such that \( \Omega_j = \bigcup_{i \in N_1} r_{\bar{x}_{j,i},\bar{x}_{j,i}} \) is pairwise disjoint, \( P(\Omega_j) = \Omega_j = \bigcup_{i \in N_1} r_{\bar{x}_{j,i},\bar{x}_{j,i}} \). \( \forall i \in N_1 \), \( r_{x_{j,i},x_{j,i}} \subseteq \Omega_j \), \( \forall i \in N_1 \), \( r_{x_{j,i},x_{j,i}} \subseteq P(\Omega_j) \), \( \forall i \in N_1 \). Hence, \( P(\Omega_j) \subseteq P(\Omega) \).

Next, we prove the general case when \( (\Omega_j)_{j \in M} \) is not necessarily pairwise disjoint. Without loss of generality, take \( M = N \). Then, \( \Omega = \bigcup_{n=1}^{\infty} \left( \Omega_n \setminus \bigcup_{i=1}^{n-1} \Omega_i \right) \). Since for open semi-rectangles \( A, B \), \( A \setminus B \) can be expressed as a finite union of pairwise disjoint semi-open rectangles. Then, \( \Omega = \bigcup_{n=1}^{\infty} \Omega_n \), where \( \Omega_n = r_{\bar{x}_n,\bar{x}_n} \) for some \( \bar{x}_n, \bar{x}_n \in R^m \) with \( \bar{x}_n \equiv \bar{x}_n, \forall n \in N \), and the sets in the union are pairwise disjoint. By the special case, we have \( \Omega \) is a region and \( P(\Omega) = \Omega \).

(viii) By \( \Omega_j \) being a region, \( \exists N_j \subseteq N \) and \((\bar{x}_{j,i})_{i \in N_j}, (\bar{x}_{j,i})_{i \in N_j} \subseteq \Omega_j \) with \( \bar{x}_{j,i} \leq \bar{x}_{j,i} \) and \( r_{\bar{x}_{j,i},\bar{x}_{j,i}} \subseteq \Omega_j \), \( \forall i \in N_j \), such that \( \Omega_j = \bigcup_{i \in N_j} r_{\bar{x}_{j,i},\bar{x}_{j,i}} \) is pairwise disjoint, \( P(\Omega_j) = \bigcup_{i \in N_j} r_{\bar{x}_{j,i},\bar{x}_{j,i}} \). \( \forall i \in N_j \), \( r_{x_{j,i},x_{j,i}} \subseteq \Omega_j \), \( \forall i \in N_j \), \( r_{x_{j,i},x_{j,i}} \subseteq P(\Omega_j) \), \( j = 1, 2 \).
Proposition 12.39, $\Omega$ is a region with $P(\Omega) = \Omega$. This completes the proof of the proposition.

**Proposition 12.40** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{O}_{\mathbb{R}^m}$ with the subset topology $\mathcal{O}$. Then, $\Omega$ is a region with $P(\Omega) = \Omega$.

**Proof** By $\mathbb{R}$ being second countable and Proposition 3.28, $\mathbb{R}^m$ is second countable. By Proposition 4.4, $\mathbb{R}^m$ is separable, which implies that, by Proposition 4.38, $(\Omega, \mathcal{O})$ is separable and second countable. Let $D \subseteq \Omega$ be a countable dense subset and define $B := \{ r_0 \mid x \in D, \delta \in \mathbb{Q}, \delta > 0, x - \delta \leq x \leq x + \delta \subseteq \Omega \}$. It is easy to show, by Definition 3.17, that $B$ is a countable basis for $\mathcal{O}$. Then, $\Omega = \bigcup_{B \in B} B =: \bigcup_{i \in \mathbb{N}} B_i = \bigcup_{i \in \mathbb{N}} \bigcup_{x \in \mathbb{R}} B_i = \bigcup_{i \in \mathbb{N}} \bigcup_{x \in \mathbb{R}} B_i$, where $N \subseteq \mathbb{N}$. By (vii) of Proposition 12.39, $\Omega$ is a region with $P(\Omega) = \Omega$. This completes the proof of the proposition.

**Definition 12.41** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{O}$ be a region, and $F : \Omega \to \mathbb{Y}$. For each $x$, $x \in \mathbb{R}^m$ with $x_i \leq x$, $\forall J \subseteq \bar{J} := \{1, \ldots, m\}$, the value of $r_{x_1, x_2}$ with any of $J$ coordinates equal to that of $x_1$ and the rest of coordinates equal to that of $x_2$ is $r_{x_1, x_2} \in \mathbb{R}^m$ such that $\pi_i(x) = \{ \pi_i(x_1) \mid \forall i \in J \). We define the set of vertices of $r_{x_1, x_2}$ with $i$ coordinates equal to that of $x_1$ to be $\text{VRect}(x_1, x_2) := \{ x \in \mathbb{R}^m \mid J \subseteq \bar{J}, \text{card}(J) = i \}$, if $x_1 < x_2$. When $r_{x_1, x_2} \subseteq \Omega$, the increment of $F$ on $r_{x_1, x_2}$ is $\Delta_F(r_{x_1, x_2}) := \sum_{i=0}^m \sum_{x \in \text{VRect}(x_1, x_2)} (-1)^i F(x)$, $x_1 < x_2$.

The total variation of $F$ on the semi-open rectangle $r_{x_1, x_2}$ is

$$T_F(r_{x_1, x_2}) := \sup_{x \in \mathbb{R}} \sum_{i=1}^n \| \Delta_F(r_{x_i, x_{i-1}}) \|$$

where $T_F(r_{x_1, x_2}) \in [0, \infty) \subset \mathbb{R}_e$. $F$ is said to be of locally bounded variation if

(i) $F$ is continuous on the right;

(ii) $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1, x_2} \subseteq \Omega$, we have $T_F(r_{x_1, x_2}) < \infty$;
(iii) \( \forall x_1, x_2 \in \Omega \) with \( x_1 \neq x_2 \) and \( \mathbf{r}_{x_1, x_2} \subseteq \Omega, \ G : \mathbf{r}_{x_1, x_2} \to [0, \infty) \subset \mathbb{R} \)
defined by \( G(x) = T_F(\mathbf{r}_{x_1, x}) \), \( \forall x \in \mathbf{r}_{x_1, x_2} \), is continuous on the right.

The total variation of \( F \) is defined by \( T_F := \sum_{i \in \mathbb{N}} T_F(\mathbf{r}_{\tilde{x}_i, \tilde{x}_i}) \) if the sum is independent of the choice of \( (\tilde{x}_i)_{i \in \mathbb{N}} \) and \( (\tilde{x}_i)_{i \in \mathbb{N}} \) as prescribed in Definition 12.38. \( F \) is said to be of bounded variation if \( F \) is of locally bounded variation and \( T_F < \infty \).

**Definition 12.42** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \mathcal{Y} \) be a normed linear space, \( \Omega \in \mathcal{B}_\mathbb{R}^m \) be a region with the subset topology, \( \mu \) be a \( \mathcal{Y} \)-valued measure on \( (\mathbb{P}(\Omega), \mathcal{B}_\mathbb{R}(\mathbb{P}(\Omega))) \), and \( F : \Omega \to \mathcal{Y} \). \( F \) is said to be a cumulative distribution function of \( \mu \) if \( \forall x_1, x_2 \in \Omega \) with \( \mathbf{r}_{x_1, x_2} \subseteq \Omega \) and \( x_1 \leq x_2 \), we have \( \mu(\mathbf{r}_{x_1, x_2}) = \Delta_F(\mathbf{r}_{x_1, x_2}) \).

Note that, for \( \sigma \)-finite \( \mathcal{Y} \)-valued measures, the cumulative distribution function is not unique and may not exist. For finite \( \mathcal{Y} \)-valued measures, the function \( F : \Omega \to \mathcal{Y} \) defined by \( F(x) = \mu(\mathbb{P}(\Omega) \cap \{ \tilde{x} \in \mathbb{R}^m \mid \tilde{x} \leq x \}) \), \( \forall x \in \Omega \), is a cumulative distribution function of \( \mu \).

**Proposition 12.43** Let \( m \in \mathbb{N} \), \( \forall i \in \{1, \ldots, m\} \), \( X_i \) be a set, \( \mathcal{C}_i \) be a semi-algebra on \( X_i \), \( \mathcal{B}_i \) be the \( \sigma \)-algebra on \( X_i \) generated by \( \mathcal{C}_i \), \( \mathcal{C} := \{ \prod_{i=1}^m C_i \subseteq \prod_{i=1}^m X_i \mid C_i \in \mathcal{C}_i, \ i = 1, \ldots, m \} \), \( \mathcal{B} \) be the \( \sigma \)-algebra on \( X := \prod_{i=1}^m X_i \) generated by \( \mathcal{C} \), and \( \mathcal{M} := \{ \prod_{i=1}^m B_i \subseteq X \mid B_i \in \mathcal{B}_i, \ i = 1, \ldots, m \} \). Then, \( \mathcal{M} \subseteq \mathcal{B} \) and \( \mathcal{B} \) is the \( \sigma \)-algebra on \( X \) generated by \( \mathcal{M} \).

**Proof** First, we will prove the following intermediate result.

**Claim 12.43.1** \( \forall C_i \in \mathcal{C}_i, \ i = 1, \ldots, m-1, \forall B_m \in \mathcal{B}_m \), we have \( (\prod_{i=1}^{m-1} C_i) \times B_m \in \mathcal{B} \).

**Proof of claim:** Fix any \( C_i \in \mathcal{C}_i, \ i = 1, \ldots, m-1 \). Define \( \mathcal{F} := \{ E \subseteq X_m \mid (\prod_{i=1}^{m-1} C_i) \times E \in \mathcal{B} \} \). By the definition of \( \mathcal{B} \), we have \( \mathcal{C}_m \subseteq \mathcal{F} \). Clearly, \( \emptyset \in \mathcal{C}_m \subseteq \mathcal{F} \). Then, by \( \mathcal{C}_m \) being a semi-algebra on \( X_m \), we have \( X_m \setminus X_m \setminus \emptyset = \bigcup_{i=1}^n C_m, \) where \( n \in \mathbb{Z}_+ \) and \( (C_m)_{i=1}^n \subseteq \mathcal{C}_m \) is pairwise disjoint. Then, \( (\prod_{i=1}^{m-1} C_i) \times C_m \subseteq \mathcal{B}, \ i = 1, \ldots, m, \) and hence \( (\prod_{i=1}^{m-1} C_i) \times X_m = \bigcup_{i=1}^n ((\prod_{i=1}^{m-1} C_i) \times C_m, \) \( \subseteq \mathcal{B} \). This implies that \( X_m \in \mathcal{F}, \ \forall E \subseteq \mathcal{F} \), we have \( (\prod_{i=1}^{m-1} C_i) \times E \in \mathcal{B} \). Then, by \( \mathcal{B} \) being a semi-algebra and \( (\prod_{i=1}^{m-1} C_i) \times X_m \subseteq \mathcal{B} \), we have \( (\prod_{i=1}^{m-1} C_i) \times (X_m \setminus E) = (\prod_{i=1}^{m-1} C_i) \times (X_m \setminus E) \subseteq \mathcal{B} \). Thus, \( X_m \setminus E \in \mathcal{F}, \ \forall E \subseteq \mathcal{F} \), we have \( (\prod_{i=1}^{m-1} C_i) \times E_j \in \mathcal{B}, \forall j \in \mathbb{N} \). Then, by \( \mathcal{B} \) being a semi-algebra, we have \( (\prod_{i=1}^{m-1} C_i) \times (\bigcup_{j=1}^\infty E_j) = \bigcup_{j=1}^\infty ((\prod_{i=1}^{m-1} C_i) \times E_j) \subseteq \mathcal{B} \). Thus, \( \bigcup_{j=1}^\infty E_j \in \mathcal{F} \). This shows that \( \mathcal{F} \) is a semi-algebra. Hence, \( \mathcal{B}_m \subseteq \mathcal{F} \) since \( \mathcal{B}_m \) is the semi-algebra generated by \( \mathcal{C}_m \). This completes the proof of the claim. □

Fix any \( B_i \in \mathcal{B}_i, \ i = 1, \ldots, m \). Clearly, \( \emptyset \in \mathcal{C}_i, \ i = 1, \ldots, m-1 \). Fix any \( i = 1, \ldots, m-1 \). Since \( \mathcal{C}_i \) is a semi-algebra on \( X_i \), we have
12.4. FUNCTIONS OF BOUNDED VARIATION

\[ X_i = X_i \setminus \emptyset = \bigcup_{j=1}^{n_i} C_{i,j}, \text{ where } n_i \in \mathbb{Z}_+ \text{ and } (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{E}_i \text{ is pairwise disjoint. By Claim 12.43.1, } (\prod_{i=1}^{m-1} C_{i,j}) \times B_m \in \mathcal{B}, i = 1, \ldots, m, i = 1, \ldots, m - 1. \text{ Then, } (\prod_{i=1}^{m-1} X_i) \times B_m = (\prod_{i=1}^{m-1} C_{i,j}) \times B_m = \bigcup_{j=1}^{n_i} \cdots \bigcup_{j_{m-1}=1}^{n_{m-1}} \prod_{i=1}^{m-1} C_{i,j} \times B_m \in \mathcal{B}. \text{ By symmetry, we have } (\prod_{i=1}^{k-1} X_i) \times B_k \times (\prod_{i=k+1}^{m} X_i) \in \mathcal{B}, k = 1, \ldots, m. \text{ Then, } \prod_{i=1}^{m} B_i = \bigcap_{k=1}^{m} \left( (\prod_{i=1}^{k-1} X_i) \times B_k \times (\prod_{i=k+1}^{m} X_i) \right) \in \mathcal{B}. \text{ By the arbitrariness of } B_i, \text{ we have } \mathcal{E} \subseteq M \subseteq \mathcal{B}. \text{ Since } \mathcal{B} \text{ is the } \sigma\text{-algebra on } X \text{ generated by } \mathcal{E}, \text{ then } \mathcal{B} \text{ is the } \sigma\text{-algebra on } X \text{ generated by } M. \]

Proposition 12.44 Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( \bar{x}, \bar{x} \in \mathbb{R}^m \) with \( \bar{x} \preceq \bar{x} \), \( X := r_{\bar{x}, \bar{x}} \subseteq \mathbb{R}^m \) with subset topology \( \mathcal{O} \), \( \mathcal{X} := (X, \mathcal{O}) \) the topological space, \( \mathcal{E} := \left\{ r_{x_1, x_2} \left| x_1 \preceq x_2 \right. \right\} \), and \( \mathcal{B} \) be the \( \sigma\)-algebra on \( X \) generated by \( \mathcal{E} \). Then, \( \mathcal{B} = \mathcal{B}_{\mathcal{B}}(\mathcal{X}) \). This completes the proof of the proposition.

Proposition 12.45 Let \( \mathcal{X} := (X, \mathcal{O}) \) be a topological space, \( (X_i)_{i=1}^{\infty} \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{X}) \) be pairwise disjoint with \( X = \bigcup_{i=1}^{\infty} X_i \), \( X_i := (X_i, \mathcal{O}_i) \) be the topological subspace of \( \mathcal{X} \), \( \forall i \in \mathbb{N} \), and \( \mathcal{B} := \{ E \subseteq X \mid E \cap X_i \in \mathcal{B}_{\mathcal{B}}(\mathcal{X}_i), \forall i \in \mathbb{N} \} \). Then, \( \mathcal{B} = \mathcal{B}_{\mathcal{B}}(\mathcal{X}) \).

Proposition 12.46 Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \subseteq \mathbb{R}^m \), \( \gamma \) be a normed linear space, and \( F : \Omega \to \gamma \). Then, \( \forall \bar{x}, \bar{x} \in \Omega \) with \( \bar{x} \preceq \bar{x} \) and \( r_{\bar{x}, \bar{x}} \subseteq \Omega \), \( \forall n \in \mathbb{Z}_+ \), \( (x_i)_{i=1}^{n} \subseteq \Omega \) with \( x_i \preceq x_i, r_{x_i, x_i} \subseteq \Omega \), \( \forall i \in \{1, \ldots, n\} \), \( r_{\bar{x}, \bar{x}} = \bigcup_{i=1}^{n} r_{x_i, x_i} \), and the sets in the union being pairwise disjoint, we have \( \Delta_F(r_{\bar{x}, \bar{x}}) = \sum_{i=1}^{n} \Delta_F(r_{x_i, x_i}) \). This case is proved.
\( n = 1 \). This case is trivial and proved.

\( n = 2 \). Without loss of generality, assume that \( \bar{x} < \bar{x} \) and \( \bar{x} < \bar{x} \), \( i = 1, 2 \). Otherwise, the result follows from \( n = 1 \) or \( n = 0 \) cases, since \( \Delta F(r_{x_1,x_2}) = \emptyset \) if \( x_1 \leq x_2 \) and \( x_1 \not< x_2 \). Then, we have \( (\bar{x}_i)_{i=1}^n = (\bar{x}_i)_{i=1}^n \subseteq \mathcal{R}_{\bar{x},\bar{x}} \). Without loss of generality, assume \( \bar{x} \in r_{x_2, \bar{x}_2} \). Then, \( \bar{x} = \bar{x}_2 \) and \( \bar{x} \not< \bar{x}_2 = \bar{x} \). Clearly, \( \bar{x} \not< \bar{x}_2 \) since \( \mathcal{R}_{\bar{x},\bar{x}} = \bigcup_{i=1}^2 \mathcal{R}_{\bar{x},\bar{x}} \) and the sets in the union are nonempty and pairwise disjoint. Then, \( \exists j_0 \in \{1, \ldots, m\} \) such that \( \pi_{j_0}(\bar{x}) > \pi_{j_0}(\bar{x}_2) > \pi_{j_0}(\bar{x}) \). Without loss of generality, assume that \( j_0 = m \).

Define \( x_C \in \mathcal{T}_{\bar{x},\bar{x}} \) be such that \( \pi_j(x_C) = \begin{cases} \pi_\bar{x} & j = 1, \ldots, m-1 \\ \pi_{\bar{x}_2} & j = m \end{cases} \) and \( x_D \in \mathcal{T}_{\bar{x},\bar{x}} \) be such that \( \pi_j(x_D) = \begin{cases} \pi_\bar{x} & j = 1, \ldots, m-1 \\ \pi_{\bar{x}_2} & j = m \end{cases} \). Clearly, we have \( x_D \in r_{\bar{x}_1, \bar{x}_1} \setminus r_{\bar{x}_2, \bar{x}_2} \). Then, \( \bar{x}_1 = \bar{x}, \bar{x}_1 = x_D, \bar{x}_2 = x_C, \) and \( \bar{x}_2 = \bar{x} \).

By Definition 12.41, we have

\[
\sum_{i=1}^2 \Delta F(r_{\bar{x}_i,\bar{x}_i}) = \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_D}} (-1)^i F(s)
\]

\[
+ \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s)
\]

\[
+ \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \mathcal{VRec}_{1,x,x_D}} (-1)^i F(s) + \Delta F(r_{\bar{x},\bar{x}})
\]

where the first two equalities follow from Definition 12.41; the third equality follows from the fact that, \( \forall i \in \{0, \ldots, m\}, \forall s \in \mathcal{VRec}_{i,x,x_D} \) with \( \pi_m(s) = \pi_m(x_D) \) if, and only if, \( s \in \mathcal{VRec}_{i+1,x,x_D} \) with \( \pi_m(s) = \pi_m(x_C) \), \( \forall s \in \mathcal{VRec}_{i,x,x_D} \) with \( \pi_m(s) = \pi_m(x) \) if, and only if, \( s \in \mathcal{VRec}_{i+1,x,x_D} \) with \( \pi_m(s) = \pi_m(x) \) if, and only if, \( s \in \mathcal{VRec}_{i,x,x_D} \) with \( \pi_m(s) = \pi_m(x) \); and the fourth equality follows from the fact that the second and fourth summations cancel each other out. This case is proved.

2 Assume that the result holds for \( n \leq k-1 \), where \( k \in \mathbb{N} \) and \( k \geq 3 \).
Consider the case when \( n = k \in \{3,4,\ldots\} \). Without loss of generality, assume that \( \bar{x} \in \hat{x} \) and \( \hat{x} \in \hat{x}_i \), \( i = 1,\ldots,n \). Otherwise, the result follows from \( 1^\circ \) or \( 2^\circ \), since \( \Delta_F(x_1, x_2) = \emptyset \) if \( x_1 \equiv x_2 \) and \( x_1 \not\equiv x_2 \). Then, we have \( (\hat{x}_i)_{i=1}^n, (\hat{x}_i)_{i=1}^n \subseteq \mathfrak{T}_{\mathfrak{x}, \bar{x}} \). Without loss of generality, assume \( \bar{x} \in \hat{x}_{n,x} \). Then, \( \hat{x} = \hat{x}_n \) and \( \hat{x} \leq \hat{x}_n = \hat{x} \).

Clearly, \( \hat{x} \not\equiv \hat{x}_i \) since \( \hat{r}_{\mathfrak{x}, \hat{x}} = \bigcup_{i=1}^n \{ r_{\mathfrak{x}, \hat{x}} \cap r_{\mathfrak{x}, \hat{x}_i} \} \) and the sets in the union are nonempty and pairwise disjoint. Then, \( \exists j_0 \in \{1,\ldots,m\} \) such that \( \pi_{j_0}(\hat{x}) > \pi_{j_0}(\hat{x}_n) > \pi_{j_0}(\hat{x}) \).

Without loss of generality, assume that \( j_0 = m \).

Define \( x_C \in \mathfrak{T}_{\mathfrak{x}, \hat{x}} \) be such that \( \pi_j(x_C) = \left\{ \begin{array}{ll} \pi_j(\hat{x}) & j = 1,\ldots,m-1 \\ \pi_m(\hat{x}_n) & j = m \end{array} \right. \) and \( x_D \in \mathfrak{T}_{\mathfrak{x}, \hat{x}} \) be such that \( \pi_j(x_D) = \left\{ \begin{array}{ll} \pi_j(\hat{x}) & j = 1,\ldots,m-1 \\ \pi_m(\hat{x}_n) & j = m \end{array} \right. \).

Then, \( r_{\mathfrak{x}, \hat{x}} = r_{\mathfrak{x}, x_D} \cup r_{\mathfrak{x}, x_C} = (\bigcup_{i=1}^n (r_{\mathfrak{x}, x_D} \cap r_{\mathfrak{x}, \hat{x}_i})) \cup (\bigcup_{i=1}^n (r_{\mathfrak{x}, x_C} \cap r_{\mathfrak{x}, \hat{x}_i})) \).

Clearly, \( r_{\mathfrak{x}, x_D} \cap r_{\mathfrak{x}, \hat{x}_i} = \emptyset \). Note that \( x_D \in \mathfrak{T}_{\mathfrak{x}, \hat{x}} \cap \mathfrak{T}_{\mathfrak{x}, \hat{x}_i} \).

Without loss of generality, assume that \( x_D \in \hat{x}_{n-1}, x_{n-1} \). Then, \( \hat{x} \equiv (\hat{x}_n)_{n-1} \equiv (\hat{x}_n)_{n-1} \equiv \hat{x} \).

Then, we must have \( x_D = \hat{x}_{n-1} \) since \( r_{\hat{x}_{n-1}, \hat{x}_{n-1}} \cap r_{\hat{x}_{n-1}, \hat{x}_n} = \emptyset \).

Thus, \( r_{\mathfrak{x}, x_D} \cap r_{\mathfrak{x}, \hat{x}_{n-1}} = \emptyset \). Then, \( r_{\mathfrak{x}, x_D} = \bigcup_{i=1}^n (r_{\mathfrak{x}, x_D} \cap r_{\mathfrak{x}, \hat{x}_i}) \) and \( r_{\mathfrak{x}, x_C} = \bigcup_{i=1}^n (r_{\mathfrak{x}, x_C} \cap r_{\mathfrak{x}, \hat{x}_i}) \).

By the inductive assumption, we have \( \Delta_F(r_{\mathfrak{x}, \hat{x}}) = \Delta_F(r_{\mathfrak{x}, x_D}) + \Delta_F(r_{\mathfrak{x}, x_C}) = \Delta_F(r_{\mathfrak{x}, x_D} \cap r_{\mathfrak{x}, \hat{x}_i}) + \sum_{i=1}^n \Delta_F(r_{\mathfrak{x}, x_C} \cap r_{\mathfrak{x}, \hat{x}_i}) \).

This case is proved.

Therefore, the result holds for any \( n \in \mathbb{N} \). This completes the proof of the proposition.

\( \square \)

**Proposition 12.47** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone; \( \Omega \subseteq \mathbb{R}^m \), and \( F : \Omega \to \mathfrak{Y} \). Assume that \( F \) satisfies (i) and (ii) of Definition 12.41. Then, \( \forall \hat{x}, \bar{x} \in \Omega \) with \( \hat{x} \equiv \hat{x} \) and \( \bar{x} \equiv \bar{x} \), \( \forall x_i \in \mathbb{Z}_+, \exists (\hat{x}_i)_{i=1}^n \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \) with \( \hat{x}_i \equiv \hat{x}_i \), and \( \forall i \in \{1,\ldots,n\}, \exists (\hat{x}_i)_{i=1}^n \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \), \( r_{\mathfrak{x}, \hat{x}_i} = \bigcup_{i=1}^n r_{\mathfrak{x}, \hat{x}_i} \), and \( x_{C,D} \in \mathfrak{T}_{\mathfrak{x}, \hat{x}} \).

Proof: \( \forall \epsilon \in (0,\infty) \subseteq \mathbb{R}, \forall i \in \{1,\ldots,n\}, \exists n_i \in \mathbb{Z}_+, \exists (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \). \( r_{\mathfrak{x}, \hat{x}_i} \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \), \( \forall \epsilon \in (0,\infty) \subseteq \mathbb{R}, \forall i \in \{1,\ldots,n\}, \exists (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \).

Let \( \epsilon_1 \) and \( \epsilon_2 \) be the union being pairwise disjoint, \( \forall i \in \{1,\ldots,n\}, \exists n_i \in \mathbb{Z}_+, \exists (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \).

By the inductive assumption, \( \sum_{j=1}^{n_i} \Delta_F(r_{\mathfrak{x}, \hat{x}_{i,j}}) = \sum_{j=1}^{n_i} \Delta_F(r_{\mathfrak{x}, \hat{x}_{i,j}}) \) and the sets in the union being pairwise disjoint, we have \( \epsilon_1, \epsilon_2 \).

Proof: \( \forall \epsilon \in (0,\infty) \subseteq \mathbb{R}, \forall i \in \{1,\ldots,n\}, \exists n_i \in \mathbb{Z}_+, \exists (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq \mathfrak{T}_{\mathfrak{x}, \hat{x}} \).
Proposition 12.48 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), and \( F : \Omega \to \mathbb{R} \) be of locally bounded variation. Assume that \( \Delta_F(r_{x_1,x_2}) \geq 0 \), \( \forall x_1, x_2 \in \Omega \) with \( r_{x_1,x_2} \subseteq \Omega \) and \( x_1 \leq x_2 \). Then, there exists a unique \( \sigma \)-finite measure \( \mu \) on the measurable space \( (\mathbb{P}(\Omega), \mathcal{B}_B(\mathbb{P}(\Omega))) \) such that \( F \) is a cumulative distribution function of \( \mu \). Furthermore, \( T_F(r_{x_1,x_2}) = \Delta_F(r_{x_1,x_2}) = \mu(r_{x_1,x_2}) \), \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), and \( T_F = \mu(\mathbb{P}(\Omega)) \in [0, \infty) \subseteq \mathbb{R}_e \).

Proof: Since \( \Omega \) is a region, then \( \exists N \subseteq \mathbb{N} \), \( \exists (\tilde{x}_i) \in N \), \( (\tilde{x}_i) \in N \subseteq \Omega \) with \( \tilde{x}_i \leq \tilde{x} \) and \( r_{\tilde{x}_i, \tilde{x}} \subseteq \Omega \), \( \forall i \in N \), such that \( \Omega = \bigcup_{i \in N} r_{\tilde{x}_i, \tilde{x}} \). By (i) of Definition 12.41, \( \mathbb{P}(r_{\tilde{x}_i, \tilde{x}}) \) is pairwise disjoint, \( \mathbb{P}(\Omega) = \bigcup_{i \in N} \mathbb{P}(r_{\tilde{x}_i, \tilde{x}}) \), and \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq \mathbb{P}(\Omega) \).

Claim 12.48.1 \( \forall \tilde{\tilde{x}}, \tilde{x} \in \Omega \) with \( r_{\tilde{x}, \tilde{x}} \subseteq \Omega \) and \( \tilde{x} \leq \tilde{x} \), let \( \mathcal{C} := \{r_{x_1,x_2} \mid \tilde{x} \leq x_1 \leq x_2 \leq \tilde{x} \} \), which is the collection of semi-open rectangles contained in \( r_{\tilde{x}, \tilde{x}} \). Clearly, \( \mathcal{C} \) is semialgebra on \( r_{\tilde{x}, \tilde{x}} \). Define \( \mu : \mathcal{C} \to [0, \infty) \subseteq \mathbb{R} \) by \( \mu(r_{x_1,x_2}) = \Delta_F(r_{x_1,x_2}) \), \( \forall r_{x_1,x_2} \in \mathcal{C} \). Then, \( \mu \) admits a unique extension to a measure \( \hat{\mu} \) on the algebra on \( r_{\tilde{x}, \tilde{x}} \), \( \mathcal{A} \), generated by \( \mathcal{C} \).

Proof of claim: We will prove the claim by Proposition 11.32. By Proposition 12.46, (i) of Proposition 11.32 is satisfied. \( \forall C := r_{\tilde{x}, \tilde{x}} \in \mathcal{C} \), \( \forall \text{ pairwise disjoint } \left( C_j := r_{\tilde{x}_j, \tilde{x}_j} \right)_{j=1}^\infty \subseteq \mathcal{C} \) with \( C = \bigcup_{j=1}^\infty C_j \), we will show that \( \mu(C) \leq \sum_{j=1}^\infty \mu(C_j) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( C = \emptyset \); Case 2: \( C \neq \emptyset \).

Case 1: \( C = \emptyset \). Then, \( \mu(C) = 0 \). Clearly \( C_j = \emptyset \), \( \forall j \in \mathbb{N} \). Then, \( \mu(C) = 0 = \sum_{j=1}^\infty \mu(C_j) \). This case is proved.

Case 2: \( C \neq \emptyset \). Then, \( \hat{x} < \tilde{x} \). By (i) of Definition 12.41, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists \tilde{y} \in r_{\tilde{x}, \tilde{x}} \) such that \( \Delta_F(r_{\tilde{y}, \hat{x}}) > \Delta_F(r_{\tilde{x}, \tilde{x}}) - \epsilon/2 \). Since \( r_{\tilde{y}, \hat{x}} \subseteq r_{\tilde{x}, \tilde{x}} \), by the assumption and (i) of Proposition 11.32, we have \( \Delta_F(r_{\tilde{y}, \hat{x}}) \leq \Delta_F(r_{\tilde{x}, \tilde{x}}) \).

Fix any \( j \in \mathbb{N} \). \( C_j = r_{\tilde{x}_j, \tilde{x}_j} \subseteq \overline{C_j} \subseteq \overline{C} \subseteq \Omega \). By (i) of Definition 12.41, \( \exists \tilde{y}_j \in \mathbb{R}^m \) with \( \tilde{x}_j < \tilde{y}_j \) such that \( \Delta_F(r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) < \Delta_F(r_{\tilde{x}_j, \tilde{x}_j}) + 2^{-j-1}\epsilon \). Since \( \mathbb{C}_j \subseteq r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}} \), by the assumption and (i) of Proposition 11.32, we have \( \Delta_F(r_{\tilde{x}_j, \tilde{x}_j}) \leq \Delta_F(r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) \).

Obviously, \( r_{\tilde{y}_j, \tilde{x}} \subseteq r_{\tilde{x}, \tilde{x}} = C = \bigcup_{j=1}^\infty C_j = \bigcup_{j=1}^\infty r_{\tilde{x}_j, \tilde{x}_j} \subseteq \bigcup_{j=1}^\infty (r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) \). Since \( \mathbb{C}_{j=1}^\infty (r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) \subseteq \bigcup_{j=1}^\infty (r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) \). By (i) of Proposition 11.32, \( \mu(C) = \Delta_F(r_{\tilde{x}, \tilde{x}}) < \Delta_F(r_{\tilde{y}, \hat{x}}) + \epsilon/2 < \sum_{j=1}^\infty \Delta_F(r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) + \epsilon/2 < \sum_{j=1}^\infty (\Delta_F(r_{\tilde{x}_j, \tilde{y}_j} \cap r_{\tilde{x}, \tilde{x}}) + 2^{-j-1}\epsilon) + \epsilon/2 < \sum_{j=1}^\infty \mu(C_j) + \epsilon \leq \sum_{j=1}^\infty \mu(C_j) + \epsilon \).
By the arbitrariness of \( \epsilon \), we have \( \mu(C) \leq \sum_{j=1}^{\infty} \mu(C_j) \). This case is also proved.

In both cases, we have \( \mu(C) \leq \sum_{j=1}^{\infty} \mu(C_j) \). Then, (ii) of Proposition 11.32 holds. Then, by Proposition 11.32, \( \mu \) admits a unique extension to a measure \( \hat{\mu} \) on the algebra \( \mathcal{A} \). This completes the proof of the claim. \( \square \)

Fix any \( i \in N \). By Claim 12.48.1 and Carathéodory Extension Theorem 11.19, there exists a unique finite measure \( \hat{\mu}_i \) on the measurable space \( (\bar{x}_i, \bar{\xi}_i, B_i) \) that is an extension of \( \hat{\mu} \), where \( B_i \) is the \( \sigma \)-algebra on \( \bar{x}_i, \bar{\xi}_i \) generated by \( \mathcal{C}_i := \left\{ x_{1,2} \mid \bar{x}_i \leq x_1 \leq x_2 \leq \bar{x}_i \right\} \). By Proposition 12.44, \( B_i = B_{\mathbb{R}}(\bar{x}_i, \bar{\xi}_i) \). Let \( \mathcal{X}_i := (\bar{x}_i, \bar{\xi}_i, B_i, \hat{\mu}_i) \) be the finite measure space. By Proposition 11.118, the generation process on \( (\mathcal{X}_i)_{i=1}^{\infty} \) yields a unique \( \sigma \)-finite measure space \( \mathcal{X} := (\mathcal{P}(\Omega), B, \hat{\mu}) \) such that \( \mathcal{X}_i \) is the finite measure subspace of \( \mathcal{X} \), \( \forall i \in N \).

By Proposition 12.45, \( B = B_{\mathbb{R}}(\mathcal{P}(\Omega)) \). Then, \( \mathcal{X} = (\mathcal{P}(\Omega), B_{\mathbb{R}}(\mathcal{P}(\Omega)), \hat{\mu}) \).

Finally, we will show that \( T_F = \hat{\mu}(\mathcal{P}(\Omega)) \in [0, \infty] \subseteq \mathbb{R} \). Note that

\[
T_F = \sum_{i \in N} T_F(\bar{x}_i, \bar{\xi}_i) = \sum_{i \in N} \Delta_F(\bar{x}_i, \bar{\xi}_i) = \sum_{i \in N} \hat{\mu}(\bar{x}_i, \bar{\xi}_i) = \hat{\mu}(\mathcal{P}(\Omega)) \in [0, \infty] \subseteq \mathbb{R}.
\]

The second equality follows from the assumption of the proposition and Definition 12.41; the third equality follows from the fact that \( \hat{\mu} \) is a measure; the fourth equality follows from the fact that \( \hat{\mu} \) is a measure. Clearly, \( T_F \) is well-defined since it is independent of \( (\bar{x}_i)_{i \in N} \) and \( (\bar{\xi}_i)_{i \in N} \), by the uniqueness of \( \hat{\mu} \). This completes the proof of the proposition. \( \square \)

**Proposition 12.49** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in B_{\mathbb{R}}(\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), and \( F : \Omega \to \mathbb{R} \) satisfy

(i) \( F \) is continuous on the right on \( \Omega \);

(ii) \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( x_{1,2} \subseteq \Omega \), we have \( \Delta_F(x_{1,2}) \geq 0 \).

Then,

1. \( F \) is of locally bounded variation;

2. there exists a unique \( \sigma \)-finite measure \( \hat{\mu} \) on the measurable space \( (\mathcal{P}(\Omega), B_{\mathbb{R}}(\mathcal{P}(\Omega))) \) such that \( F \) is a cumulative distribution function of \( \hat{\mu} \).
3. \( T_F(r_{x_1,x_2}) = \Delta F(r_{x_1,x_2}) = \hat{\mu}(r_{x_1,x_2}), \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \tau_{x_1,x_2} \subseteq \Omega \), and \( T_F = \hat{\mu}(P(\Omega)) \in [0, \infty) \subset \mathbb{R} \).

**Proof** By (i), (ii) of Definition 12.41 is satisfied. \( \forall \bar{x}, \bar{x} \in \Omega \) with \( \bar{x} \leq \bar{x} \) and \( \tau_{\bar{x},\bar{x}} \subseteq \Omega \), \( \forall n \in \mathbb{Z}_+, \forall \bar{x}_1, \bar{x}_1, \ldots, \bar{x}_n, \bar{x}_n \in \tau_{\bar{x},\bar{x}} \) with \( \bar{x}_i \leq \bar{x}_i, \forall i \in \{1, \ldots, n\} \), \( r_{\bar{x},\bar{x}} = \bigcup_{i=1}^n r_{\bar{x}_i,\bar{x}_i} \), and the sets in the union being pairwise disjoint, we have \( \sum_{i=1}^n \sum_i^n \Delta F(r_{\bar{x}_i,\bar{x}_i}) = \Delta F(r_{\bar{x},\bar{x}}) = \Delta F(r_{\bar{x},\bar{x}}) \in [0, \infty) \subset \mathbb{R} \), where the first equality follows from (ii); and the second equality follows from Proposition 12.46. Then, \( T_F(r_{\bar{x},\bar{x}}) = \Delta F(r_{\bar{x},\bar{x}}) \in [0, \infty) \subset \mathbb{R} \). Thus, (ii) of Definition 12.41 is satisfied.

\( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \tau_{x_1,x_2} \subseteq \Omega \), define \( \hat{G} : \tau_{x_1,x_2} \to [0, \infty) \subset \mathbb{R} \) by \( \hat{G}(x) = T_F(r_{x_1,x_2}), \forall x \in \tau_{x_1,x_2} \). By the previous paragraph, \( \hat{G}(x) = \Delta F(r_{x_1,x_2}), \forall x \in \tau_{x_1,x_2} \). Then, \( \forall x \in \tau_{x_1,x_2} \) with \( x \neq x_2 \), \( \hat{G}(x) = \Delta F(r_{x_1,x_2}) = \lim_{h \to 0} \Delta F(r_{x_1,x_2+h}) = \lim_{h \to 0} G(x+h) \), where the second equality follows from Definition 12.41 and (i). Hence, by Proposition 12.37, (iii) of Definition 12.41 is satisfied.

This shows that \( F \) is of locally bounded variation. Hence, the first statement holds. The second and third statements of the proposition follows directly from Proposition 12.48. This completes the proof of the proposition. □

**Theorem 12.50** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}(\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), \( \mathcal{Y} \) be a Banach space, \( F : \Omega \to \mathcal{Y} \) be of locally bounded variation. Then, there exists a unique \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \hat{\mu} \) on the measurable space \((P(\Omega), \mathcal{B}(P(\Omega)))\) such that \( F \) is a cumulative distribution function of \( \hat{\mu} \). Furthermore,

\( \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2}) = T_F(r_{x_1,x_2}), \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \tau_{x_1,x_2} \subseteq \Omega \), and \( \mathcal{P} \circ \hat{\mu}(P(\Omega)) = T_F \).

**Proof** Since \( \Omega \in \mathcal{B}(\mathbb{R}^m) \) is a region, then \( \exists N \subseteq \mathbb{N} \) and \( (\bar{x}_i)_{i \in N} \), \( (\bar{x}_i)_{i \in N} \subseteq \Omega \) with \( \bar{x}_i \leq \bar{x}_i \) and \( \tau_{\bar{x}_i,\bar{x}_i} \subseteq \Omega \), \( \forall i \in N \), such that \( \Omega = \bigcup_{i \in N} \tau_{\bar{x}_i,\bar{x}_i}, (\bar{x}_i,\bar{x}_i)_{i \in N} \) is pairwise disjoint, \( P(\Omega) := \bigcup_{i \in N} r_{\bar{x}_i,\bar{x}_i} \), and \( \forall \bar{x}_1, \bar{x}_2 \in \Omega \) with \( \bar{x}_1 \leq \bar{x}_2 \) and \( \tau_{\bar{x}_1,\bar{x}_2} \subseteq \Omega \), we have \( r_{\bar{x}_1,\bar{x}_2} \subseteq P(\Omega) \). We need the following result.

**Claim 12.50.1** Fix any \( \bar{x}, \bar{x} \in \Omega \) with \( \bar{x} \leq \bar{x} \) and \( \tau_{\bar{x},\bar{x}} \subseteq \Omega \) be endowed with subset topology, then, there exists a unique finite \( \mathcal{Y} \)-valued measure \( \hat{\mu} \) on the measurable space \((\Omega, \mathcal{B}(\Omega))\) such that \( \hat{\mu}(r_{x_1,x_2}) = \Delta F(r_{x_1,x_2}) \) and \( \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2}) = T_F(r_{x_1,x_2}), \forall x_1 \leq x_2 \leq \bar{x} \). Furthermore, \( \hat{\mu} \in \mathcal{M}_{ft}(\hat{\Omega}, \mathcal{Y}) \).

**Proof of claim:** Let \( \hat{\Omega} := \tau_{\bar{x},\bar{x}}, \hat{F} := F|_{\hat{\Omega}} : \hat{\Omega} \to \mathcal{Y} \), and \( \hat{G} : \hat{\Omega} \to [0, \infty) \subset \mathbb{R} \) be defined by \( \hat{G}(x) = T_F(r_{\bar{x},\bar{x}}), \forall x \in \hat{\Omega} \). Note that \( \hat{G} \) is real-valued since \( F \) is of locally bounded variation. By (iii)
of Definition 12.41, $\tilde{G}$ is continuous on the right. By Proposition 12.47, $\Delta_G(r_{x_1,x_2}) = T_F(x_{1,x_2}) \geq 0$, $\forall x_1, x_2 \in \hat{\Omega}$ with $x_1 \lessgtr x_2$. By Proposition 12.49, $\tilde{G}$ is of locally bounded variation, there exists a unique $\sigma$-finite measure $\tilde{\nu}$ on the measurable space $(\hat{\Omega}, B_{\hat{\Omega}})$ such that $\tilde{G}$ is a cumulative distribution function of $\tilde{\nu}$, and $\tilde{\nu}(\hat{\Omega}) = T_G(\hat{\Omega}) < \infty$. Hence, $\tilde{\nu}$ is finite.

Let $\mathcal{C} := \{ r_{x_1,x_2} \mid \bar{x} \lessgtr x_1 \lessgtr x_2 \leq \bar{x} \}$ be the collection of semi-open rectangles contained in $\hat{\Omega}$. Clearly, $\mathcal{C}$ is a semialgebra on $\hat{\Omega}$. Define $\nu : \mathcal{C} \to [0, \infty) \subset \mathbb{R}$ by $\nu := \tilde{\nu}|_{\mathcal{C}}$ and $\mu : \mathcal{C} \to \mathbb{Y}$ by $\mu(r_{x_1,x_2}) = \Delta_F(r_{x_1,x_2}) \in \mathbb{Y}$, $\forall r_{x_1,x_2} \in \mathcal{C}$. It is easy to show that (i) – (v) of Proposition 12.20 are satisfied by $\mu$ and $\nu$. $\forall r_{x_1,x_2} \in \mathcal{C}$, we have

$$\nu(r_{x_1,x_2}) = \mu(r_{x_1,x_2}) = T_F(r_{x_1,x_2})$$

Hence, (vi) of Proposition 12.20 is satisfied. By Proposition 12.20, $\nu$ admits a unique extension to a measure $\hat{\nu}$ on the algebra on $\hat{\Omega}$, $\mathcal{A}$, generated by $\mathcal{C}$; and $\mu$ admits a unique extension to a $\mathbb{Y}$-valued measure $\hat{\mu}$ on the algebra $\mathcal{A}$ with $\mathcal{P} \circ \hat{\mu} = \nu$ and $\mathcal{P} \circ \hat{\mu}|_{\mathcal{C}} = \nu$. Clearly, $\hat{\nu}$ is finite. By Carathéodory Extension Theorem 12.4, there is a unique $\hat{\mathcal{Y}}$-valued measure $\hat{\mu}$ on $(\hat{\Omega}, B)$ such that $\hat{\mu}|_{\mathcal{A}} = \hat{\mu}$ and $(\mathcal{P} \circ \hat{\mu})|_{\mathcal{A}} = \hat{\nu}$, where $B$ is the $\sigma$-algebra on $\hat{\Omega}$ generated by $\mathcal{A}$. Then, $B$ equals to the $\sigma$-algebra on $\hat{\Omega}$ generated by $\mathcal{C}$. By Proposition 12.44, $\hat{B} = B_{\hat{\Omega}}(\hat{\Omega})$. By Carathéodory Extension Theorem 11.19, we have $\mathcal{P} \circ \hat{\mu} = \hat{\nu}$. Then, $\hat{\mu}$ is a finite $\hat{\mathcal{Y}}$-valued measure on $(\hat{\Omega}, B_{\hat{\Omega}}(\hat{\Omega}))$. By Theorem 11.198, $\hat{\mu} \in M_{f}(\hat{\Omega}, \hat{\mathcal{Y}})$. This completes the proof of the claim. \( \square \)

$\forall i \in N$, let $X_i := (x_i, \bar{x}_i) \subset \Omega$ be the unique finite $\hat{\mathcal{Y}}$-valued metric measure space as defined in Claim 12.50.1. By Proposition 11.118, the generation process on $(X_i)_{i \in N}$ yields a unique $\sigma$-finite $\hat{\mathcal{Y}}$-valued measure space $\hat{\mathcal{X}} := (\hat{\mathcal{P}}(\hat{\Omega}), B, \hat{\mu})$ such that $X_i$ is the finite $\hat{\mathcal{Y}}$-valued measure subspace of $\hat{\mathcal{X}}$, $\forall i \in N$. By Proposition 12.45, we have $\hat{B} = B_{\hat{\Omega}}(\hat{\mathcal{P}}(\hat{\Omega}))$.

Next, we will show that $F$ is a cumulative distribution function of $\hat{\mu}$. Fix any $x_1, x_2 \in \hat{\Omega}$ with $x_1 \lessgtr x_2$ and $\tau_{x_1,x_2} \subseteq \hat{\Omega}$. Let $(r_{x_1,x_2}, B_{\hat{\mathcal{P}}}(r_{x_1,x_2}), \hat{\mu})$ be the finite $\hat{\mathcal{Y}}$-valued metric measure space as prescribed in Claim 12.50.1. Note that $\mathcal{P} \circ \hat{\mu}(x_{1,x_2}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2} \cap \tau_{x_1,x_2}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2} \cap \tau_{x_1,x_2}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2} \cap \tau_{x_1,x_2}) = \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2}) = T_F(r_{x_1,x_2}) < \infty$, where the first equality follows from the fact $\mathcal{P} \circ \hat{\mu}$ is a measure; the second equality follows from Proposition 11.115; the third equality follows from Claim 12.50.1; the fourth equality follows from Claim 12.50.1; the fifth equality follows from the fact that $\mathcal{P} \circ \hat{\mu}$ is a measure; and the sixth equality and the inequality follow from Claim 12.50.1.
This implies that \( \hat{\mu}(r_{x_1,x_2}) = \sum_{i \in N} \hat{\mu}(r_{x_1,x_2} \cap r_{\tilde{x}_i,\xi_i}) = \sum_{i \in N} \Delta_F(r_{x_1,x_2} \cap r_{\tilde{x}_i,\xi_i}) = \Delta_F(r_{x_1,x_2}) \in \mathcal{Y} \), where the first equality follows from the fact that \( \hat{\mu} \) is a \( \mathcal{Y} \)-valued measure; and the second equality follows from Proposition 11.115; the third equality follows from Claim 12.50.1; the fourth equality follows from Claim 12.50.1; the fifth equality follows from the fact that \( \hat{\mu} \) is a finite \( \mathcal{Y} \)-valued measure; and the last equality follows from Claim 12.50.1. By the arbitrariness of \( x_1 \) and \( x_2 \), \( F \) is a cumulative distribution function of \( \hat{\mu} \).

Clearly, \( T_F = \sum_{i \in N} T_F(r_{\tilde{x}_i,\xi_i}) = \sum_{i \in N} P \circ \hat{\mu}(r_{\tilde{x}_i,\xi_i}) = P \circ \hat{\mu}(P(\Omega)) \), where the first equality follows from Definition 12.41; the second equality follows from the preceding argument; and the last equality follows from the fact \( P \circ \hat{\mu} \) is a measure. \( T_F \) is well-defined since it is independent of \( (\tilde{x}_i)_{i \in N} \) and \( (\xi_i)_{i \in N} \), by the uniqueness of \( \hat{\mu} \).

This completes the proof of the theorem.

**Proposition 12.51.** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone \( P, \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), \( \mathcal{Y} \) be a normed linear space over \( \mathbb{K} \), and \( \mu \in \mathcal{M}(P(\Omega), \mathcal{B}_B(P(\Omega)), \mathcal{Y}) \). Define \( F : \Omega \to \mathcal{Y} \) and \( G : \Omega \to [0, \infty) \subset \mathbb{R} \) by \( F(x) = \mu(\{ \tilde{x} \in P(\Omega) \mid \tilde{x} \leq x \}) \) and \( G(x) = P \circ \mu(\{ \tilde{x} \in P(\Omega) \mid \tilde{x} \leq x \}) \), \( \forall x \in \Omega \). Then,

(i) \( F \) is a cumulative distribution function of \( \mu \), \( F \) is of locally bounded variation, \( T_F(r_{x_1,x_2}) \leq P \circ \mu(r_{x_1,x_2}) \), \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), and \( T_F \leq P \circ \mu(P(\Omega)) \) if it is well-defined;

(ii) \( G \) is a cumulative distribution function of \( P \circ \mu \) and \( G \) is of bounded variation.

(iii) If, in addition, \( \mathcal{Y} \) is a Banach space, then, \( T_F(r_{x_1,x_2}) = P \circ \mu(r_{x_1,x_2}) = \Delta_G(r_{x_1,x_2}) \), \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), and \( T_F = P \circ \mu(P(\Omega)) \). Therefore, \( F \) is of bounded variation.

**Proof** Since \( \Omega \) is a region, then \( \exists N \subseteq \mathbb{N} \) and \( (\tilde{x}_i)_{i \in N}, (\xi_i)_{i \in N} \subseteq \Omega \) with \( \tilde{x}_i \leq \tilde{x}_i \) and \( r_{\tilde{x}_i,\xi_i} \subseteq \Omega \), \( \forall i \in N \), such that \( \Omega = \bigcup_{i \in N} r_{\tilde{x}_i,\xi_i} \), \( (r_{\tilde{x}_i,\xi_i})_{i \in N} \) is pairwise disjoint, \( P(\Omega) := \bigcup_{i \in N} r_{\tilde{x}_i,\xi_i} \), and \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq P(\Omega) \).

\[
\forall x_1, x_2 \in \Omega \text{ with } x_1 \leq x_2 \text{ and } r_{x_1,x_2} \subseteq \Omega, \text{ we have}
\]

\[
\Delta_F(r_{x_1,x_2}) = \left\{ \sum_{i=0}^m \sum_{x \in \mathbb{V} \cap x_1,x_2} \left( -1 \right)^i F(\tilde{x}) \mid x \leq x_2 \text{ and } x_1 \neq x_2 \right\} x_1 \leq x_2 \text{ and } x_1 \neq x_2
\]

\[
= \mu(P(\Omega) \cap r_{x_1,x_2}) = \mu(r_{x_1,x_2})
\]
Hence, $F$ is a cumulative distribution function of $\mu$.

By Proposition 11.4, $G(x_1) \leq G(x_2)$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. By an argument that is similar to the previous paragraph, we have $G$ is a cumulative distribution function of $P \circ \mu$ and $\Delta_G(r_{x_1,x_2}) = P \circ \mu(r_{x_1,x_2}) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1,x_2} \subseteq \Omega$. Since $\mu$ is finite, by Proposition 11.5, we have, $\forall x \in \Omega$, $G(x) = \lim_{n \to \infty} P \circ \mu(x)$, $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $P \circ \mu(x)$ is continuous on the right. Then, $G$ is continuous on the right. By Proposition 12.49, $G$ is of locally bounded variation. Then, $T_G = \sum_{i \in I} T_G(r_{x_i}) = \sum_{i \in I} \Delta_G(r_{x_i}) = \sum_{i \in I} P \circ \mu(r_{x_i}) = \sum_{i \in I} P \circ \mu(\Omega)$, where the first equality follows from Definition 12.41; the second equality follows from the fact that $\Delta_G(r_{x_1,x_2}) = P \circ \mu(r_{x_1,x_2}) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1,x_2} \subseteq \Omega$; the third equality follows from the fact that $G$ is cumulative distribution function of $P \circ \mu$; and the last two steps follow from the fact that $P \circ \mu$ is a finite measure. Clearly, $T_G$ is well defined since it is independent of $(x_i)_{i \in I}$ and $(x_i)_{i \in I}$. Hence, $G$ is of bounded variation. Thus, (ii) holds.

$\forall x \in \Omega$, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, by $G$ being continuous on the right, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall 0_m \leq h < \delta 1_m$ with $x + h \in \Omega$, we have $G(x) \leq G(x + h) - G(x)$ $+ \varepsilon$. Then, $\|F(x+h) - F(x)\| = \|\mu(\{x \in P(\Omega) \mid x \leq x + h\}) - \mu(\{x \in P(\Omega) \mid x \leq x\})\| = \|\mu(\{x \in P(\Omega) \mid x \leq x + h\}) - \mu(\{x \in P(\Omega) \mid x \leq x\})\| = \|P \circ \mu(\{x \in P(\Omega) \mid x \leq x + h\}) - P \circ \mu(\{x \in P(\Omega) \mid x \leq x\})\| = G(x + h) - G(x) < \varepsilon$, where the second equality and the first inequality follow from the fact that $\mu$ is a finite $\mu$-measurable value; and the third equality follows from Fact 11.3. Hence, $F$ is continuous on the right. Then, $F$ satisfies (i) of Definition 12.41.
∀ \sigma \in P \circ \text{follows from the second to last paragraph; and the second equality and the}
T \text{ we have }
that \ t \text{ive cone } T \text{ with (ii), further concludes that }
P(\Omega) \text{ is well-defined, and is finite, then } F \text{ is of locally bounded variation. When } F \text{ is well-defined,}
we have }
G \text{ is cumulative distribution function of } \mu \text{; the third inequality follows from Proposition 11.4; and the fourth}
equality follows from Proposition 11.47; the second inequality follows from }
\exists \in (0, \infty) \subset \mathbb{R} \text{, } \forall \in (0, \infty) \subset \mathbb{R} \text{, } \forall 0 \leq h < \delta 1_m \text{ with}
x + h \in \Omega, \text{ we have } G(x) \leq G(x + h) < G(x) + \epsilon. \text{ Then, } \forall 0 \leq h < \delta 1_m \text{ with}
x + h \in \mathbb{R}, 0 \leq G(x + h) - G(x) = T_F(r_{x_1}, x_2) - T_F(x_1, x_2) = \sum_{(\hat{x}, \hat{z}) \in M} P \circ \mu(\hat{x}, \hat{z}) = P \circ \mu(x_1 + h, x_2) \leq P \circ \mu(x_1, x_2) - P \circ \mu(x_1, x_2 + h) \text{ where the first inequality follows from the monotonicity of}
G; M := \{(\hat{x}, \hat{z}) \in \mathbb{R} \times \mathbb{R} | i \in \{0, \ldots, m - 1\}, J \subseteq J_m := \{1, \ldots, m\}, \text{ card } (J) = i, \pi_k(\hat{x}) = \pi_k(x_1) \forall k \in J, \pi_k(\hat{z}) = \pi_k(x) \forall k \in J_m \setminus J, \pi_k(\hat{x}) = \pi_k(x), \forall k \in J, \pi_k(\hat{z}) = \pi_k(x + h) \forall k \in J_m \setminus J \}; \text{ the second equality follows from Proposition 12.47; the second inequality follows from the last paragraph; the third equality follows from the fact that } P \circ \mu \text{ is a measure; the third inequality follows from Proposition 11.4; and the fourth}
equality follows from the fact that } P \circ \mu \text{ is a finite measure. Hence, } G \text{ is}
continuous on the right. Thus, } F \text{ satisfies (iii) of Definition 12.41.}

Therefore, } F \text{ is of locally bounded variation. When } F \text{ is well-defined, we have }
T_F = \sum_{i \in N} T_F(r_{\hat{x}_i, \hat{z}_i}) \leq \sum_{i \in N} P \circ \mu(\hat{x}_i, \hat{z}_i) = P \circ \mu(\Omega) < \infty, \text{ where the first equality follows from Definition 12.41; the first inequality follows from the second to last paragraph; and the second equality and the second inequality follows from the fact that } P \circ \mu \text{ is a finite measure. Thus, (i) holds.}

(iii) } \forall \text{ a Banach space. By Theorem 12.50 and (i), } \mu \text{ is the unique}
\sigma\text{-finite } \forall \text{-valued measure on the measurable space } (P(\Omega), \mathcal{B}(P(\Omega))) \text{ such that } F \text{ is a cumulative distribution function of } \mu. \text{ Theorem 12.50, coupled with (ii), further concludes that } T_F(r_{x_1, x_2}) = P \circ \mu(r_{x_1, x_2}) = \Delta_G(r_{x_1, x_2}), \forall x_1, x_2 \in \Omega \text{ with } x_1 \leq x_2 \text{ and } r_{x_1, x_2} \subseteq \Omega, \text{ and } T_F = P \circ \mu(\Omega) < \infty. \text{ Since } T_F \text{ is well-defined and is finite, then } F \text{ is of bounded variation.}

This completes the proof of the proposition. \qed

Proposition 12.52 Let } m \in \mathbb{N}, \mathbb{R}^m \text{ be endowed with the usual positive cone } P, \Omega \in \mathcal{B}(\mathbb{R}^m) \text{ be a rectangle with the subset topology } \mathcal{O}, \forall \text{
be a normed linear space over $\mathbb{K}$, $\mu$ be a $\sigma$-finite $\mathcal{Y}$-valued measure on $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)))$ such that $\mathcal{P} \circ \mu(x_1, x_2) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$.

By Proposition 12.39, $\mathcal{P}(\Omega) = \prod_{i=1}^{m} I_i$ and $\Omega$ is a region.

Define $F: \Omega \to \mathcal{Y}$ by, if $\Omega \neq \emptyset$ then fix $x_0 \in \Omega$, $F(x) = \emptyset y$, $\forall x \in \Omega$ with $\pi_{i_0}(x) = \pi_{i_0}(x_0)$ for some $i_0 \in \hat{J} := \{1, \ldots, m\}$; and $\forall x \in \Omega$ with $\pi_i(x) \neq \pi_i(x_0)$, $\forall i \in \hat{J}$, $F(x)$ is such that $\Delta F(\bar{x}, \hat{x}) = \mu(\bar{x}, \hat{x})$, where $\bar{x} := x \wedge x_0$ and $\hat{x} := x \vee x_0$. Then, $F$ is a cumulative distribution function of $\mu$, which is called the cumulative distribution function of $\mu$ with origin $x_0$. Define $G: \Omega \to \mathbb{R}$ to be the cumulative distribution function of $\mathcal{P} \circ \mu$ with origin $x_0$. Then, the following statements hold.

(i) $F$ and $G$ are of locally bounded variation.

(ii) If, in addition, $\mathcal{Y}$ is a Banach space, then $T_F(\pi, x_2) = \mathcal{P} \circ \mu(\pi, x_2)$

\[
\Delta_G(\pi, x_2) = \forall \pi_1, x_2 \in \Omega \text{ with } x_1 \leq x_2 \text{ and } \pi_1, x_2 \in \Omega, \text{ and } T_F = \mathcal{P} \circ \mu(\mathcal{P}(\Omega)).
\]

**Proof**

(i) By the assumption on $\mu$, we have $\mathcal{P} \circ \mu(x_1, x_2) < \infty$,

$\forall \pi_1, x_2 \in \Omega$ with $x_1 \leq x_2$. Then, $\pi_1, x_2 \in \text{dom}(\mu)$. This implies that $F(x) \in \mathcal{Y}$ and $G(x) \in \mathbb{R}$, $\forall x \in \Omega$. Hence, $F$ and $G$ are well defined.

If $\Omega = \emptyset$, then the result holds trivially. This case is proved. Consider the case $\Omega \neq \emptyset$.

**Claim 12.52.1** $F$ is a cumulative distribution function of $\mu$.

**Proof of claim:**

$\forall \pi_1, x_2 \in \Omega$ with $x_1 \leq x_2$. If $x_1 \neq x_2$, then by Definition 12.41, we have $\mu(\pi, x_2) = \mu(\emptyset) = \emptyset y = \Delta F(\pi, x_2)$. On the other hand, if $x_1 < x_2$, then let $n := \text{card} (\{ i \in \hat{J} \mid (\pi_i(x_1) - \pi_i(x_0)) (\pi_i(x_2) - \pi_i(x_0)) < 0 \}) \in \{0, \ldots, m\}$. We will show that $\mu(\pi, x_2) = \Delta F(\pi, x_2)$ by mathematical induction on $n$.

1° Consider the case $n = 0$. Define $J := \{ i \in \hat{J} \mid \pi_i(x_1) < \pi_i(x_0) \text{ or } \pi_i(x_2) < \pi_i(x_0) \}$ and $\bar{n} := \text{card}(J)$.

$\forall j \in \{0, \ldots, m\}$, $\forall x \in \text{VRect}_{j, x_2}$, if $\pi_i(x) \neq \pi_i(x_0)$, $\forall i \in \hat{J}$, then, $\bar{n} = \text{card}(\{ i \in \hat{J} \mid \pi_i(x) < \pi_i(x_0) \})$, by the definition of $F$ and Definition 12.41, we have $\Delta F(\pi, x_2) = \mu(\pi, x_2) = (-1)^n F(x)$. On the other hand, if $\pi_i(x) = \pi_i(x_0)$ for some $i_0 \in \hat{J}$, then, by the definition of $F$ and Definition 12.41, $\Delta F(\pi, x_2) = \emptyset y = \mu(\emptyset) = \mu(\pi, x_2) = F(x) = (-1)^n F(x)$. Hence, we have $\mu(\pi, x_2) = (-1)^n F(x)$, $\forall x \in \text{VRect}_{j, x_2}$, $\forall j \in \{0, \ldots, m\}$.

Then, we have

\[
\mu(\pi, x_1) = \mu(\{ x \in \mathcal{P}(\Omega) \mid x_1 \leq x \leq x_2 \})
\]

\[
= \mu(\{ x \in \mathcal{P}(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2) \leq \pi_i(x_0), \forall i \in J; \pi_i(x_0) \leq \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \hat{J} \text{ \setminus \} J}\})
\]

\[
= \mu(\{ x \in \mathcal{P}(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J; \}
\]
\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \hat{J} \setminus J \} \]

\[ -\mu \left( \bigcup_{i_0=1}^{m} \{ x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \{i_0\} \} \right) \]

\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\hat{J} \setminus J) \setminus \{i_0\}; \]

\[ \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \{i_0\}; \]

\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\hat{J} \setminus J) \cap \{i_0\} \}

\[ = \sum_{j=0}^{m} (-1)^j \sum_{J \subseteq J, \text{card}(J) = j} \mu(\{ x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \hat{J} \}) \]

\[ \forall i \in J \setminus \hat{J}; \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\hat{J} \setminus J) \setminus \hat{J}; \]

\[ \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \hat{J}; \]

\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\hat{J} \setminus J) \cap \hat{J} \} \]

\[ \pi_i(x_j) \]

\[ \left\{ \begin{array}{l}
\pi_i(x_1) \quad \forall i \in (\hat{J} \setminus J) \cup ((\hat{J} \setminus J) \cap \hat{J}) = J \triangle \hat{J} \\
\pi_i(x_2) \quad \forall i \in ((\hat{J} \setminus J) \setminus \hat{J}) \cup (\hat{J} \setminus J) = \hat{J} \setminus (J \triangle J), \forall i \in \hat{J}. \end{array} \right. \]

Then, \( \{ x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \hat{J}; \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\hat{J} \setminus J) \setminus \hat{J}; \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \hat{J}; \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\hat{J} \setminus J) \cap \hat{J} \} = r_{\bar{\hat{J}}, \bar{x}}. \)

By the definition of \( x_j \), we have \( x_j \in \text{VRect}_{\text{card}(J \triangle J),x_1,x_2} \) and \( \text{card}(J \triangle \hat{J}) = \text{card}(J) + \text{card}(\hat{J}) = n + j \equiv 0 (\text{mod } 2) \), \( \forall J \subseteq \hat{J} \) with \( \text{card}(\hat{J}) = j \in \{0, \ldots, m\} \). This leads to the conclusion that

\[ \mu(r_{x_1,x_2}) = \sum_{j=0}^{m} (-1)^j \sum_{J \subseteq J, \text{card}(J) = j} \mu(r_{\bar{x},\bar{x}}) \]

\[ = \sum_{j=0}^{m} (-1)^j \sum_{J \subseteq J, \text{card}(J) = j} (-1)^{\text{card}(J \triangle J)} F(x_j) \]

\[ = \sum_{j=0}^{m} \sum_{J \subseteq J, \text{card}(\hat{J}) = j} (-1)^{\text{card}(J \triangle \hat{J})} F(x_j) \]

\[ = \sum_{i=0}^{m} \sum_{\bar{x} \in \text{VRect}_{i,x_1,x_2}} (-1)^{\text{card}(J \triangle \hat{J})} F(r_{x_1,x_2}) \]

where the second equality follows from the previous paragraph. This case is proved.

2° Assume that \( \mu(r_{x_1,x_2}) = \Delta_F(r_{x_1,x_2}), \forall n \leq k - 1 \in \{0, \ldots, m - 1\} \).

3° Consider the case when \( n = k \in \hat{J} \). Since \( n > 0 \), then \( \exists i_0 \in \hat{J} \) such that \( \pi_{i_0}(x_1) < \pi_{i_0}(x_0) < \pi_{i_0}(x_2) \). Without loss of generality, assume that
By Proposition 4.16, we have $\lim_{h \to_m} \pi_i(x) = \pi_i(\Omega)$, $\forall i \in J$, $\exists x \in \Omega$ where the first inequality follows from the fact that $\mu(\Omega) = 1$. Then, $\lim_{h \to_m} \mu(x) \leq \mu(\Omega) = 1$, $\forall x \in \Omega \setminus \{1\}$ and $\mu(x) = \mu(x) = \Delta x$, $\forall x \in \Omega \setminus \{1\}$, respectively. We need the following intermediate result.

**Claim 12.5.2** $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$, if $x_2$ is an accumulation point of $\Omega \cap (x + P)$, then $\lim_{h \to_m} \mu(x_1, x_2) = 0$. $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $x_1 \neq x_2$, then $\lim_{h \to_m} \mu(x_1, x_2) = 0$.

**Proof of claim:** $\forall (H_n)_{n=1}^{\infty} \subseteq P \setminus \Omega \setminus \{0_m\}$ with $\lim_{n \in N} h_n = 0_m$, we have $\lim_{n \in N} h_n = \lim_{n \in N} \bigwedge_{i=n}^{\infty} h_i = 0_m$, $0_m \leq h_n = 0_m$, and $h_n \in P \cap (\Omega \setminus \{0_m\}) \setminus \{0_m\}, \forall n \in N$. Then, $0 \leq \lim_{n \in N} \mu(x_1, x_2) \leq \lim_{n \in N} \mu(x_1, x_2) = 0$, where the first inequality follows from the fact that $\mu(x_1, x_2)$ is a measure; the second inequality follows from Proposition 11.4; and the first equality follows from Proposition 11.5. Then, $\lim_{n \in N} \mu(x_1, x_2) = 0$. By Proposition 4.16, we have $\lim_{h \to_m} h \subseteq \{x_1, x_2, \ldots, x_n\}, h \neq 0_m$. This proves the first statement.

$\forall h \in \Omega$, let $J := \bigcup_{i=n}^{\infty} \pi_i(x) < \pi_i(x_0)$ and $n := \text{card}(J)$. Let $x$ be an accumulation point of $\Omega \cap (x + P)$. Then, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that, $\forall h \in P \cap (\Omega \setminus h)$ with $h \neq 0_m$ and $h < \delta 1_m$, we have $x + h \in \Omega$.
and \( \{ i \in \bar{J} \mid \pi_i(x+h) < \pi_i(x) \} = J \). Then, \( 0 \leq |G(x+h) - G(x)| = (-1)^n (P \circ \mu(r_{x_h}, \hat{x}_h) - P \circ \mu(r_{\hat{x}_h}, \hat{x}_h)) \leq |P \circ \mu(r_{x_h}, \hat{x}_h) - P \circ \mu(r_{\hat{x}_h}, \hat{x}_h)| + |P \circ \mu(r_{x_h}, \hat{x}_h) - P \circ \mu(\hat{x}_h, \hat{x}_h) - P \circ \mu(\hat{x}_h, \hat{x}_h) \leq P \circ \mu(r_{\hat{x}_h \setminus x_h}, \hat{x}_h) + P \circ \mu(r_{x_h \setminus \hat{x}_h}, \hat{x}_h), \) where \( x_h := x + h; \) the first equality follows from the definition of \( G; \) the second inequality follows from the fact that \( x \leq x_h \) and \( \hat{x} \leq \hat{x}_h \leq \hat{x} \equiv \hat{x}_h; \) the second equality follows from Fact 11.3; and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have \( 0 \leq \lim_{h \to 0^m} |G(x+h) - G(x)| \leq \lim_{h \to 0^m} \|P \circ \mu(r_{\hat{x}_h \setminus x_h}, \hat{x}_h) + \mathcal{P} \circ \mu(r_{x_h \setminus \hat{x}_h}, \hat{x}_h)\| = 0. \) Hence, by Proposition 12.37, \( G \) is continuous on the right. By Proposition 12.49, \( G \) is of locally bounded variation.

\( \forall x \in \Omega, \) let \( x \) be an accumulation point of \( \Omega \cap (x + P), \) and let \( J, \) \( n, \) \( \delta, \) and \( x_h \) be as defined in the previous paragraph. \( \forall h \in P \cap (\Omega - x) \) with \( h \neq 0_m \) and \( h < \delta 1_m, \) we have \( 0 \leq \|F(x+h) - F(x)\| = (-1)^n (\mu(r_{x_h}, \hat{x}_h) - \mu(r_{\hat{x}_h}, \hat{x}_h), \| \mu(r_{x_h}, \hat{x}_h) - \mu(r_{\hat{x}_h}, \hat{x}_h)\| + \| \mu(r_{x_h}, \hat{x}_h) - \mu(r_{\hat{x}_h}, \hat{x}_h)\| = \| \mu(r_{x_h}, \hat{x}_h) - \mu(r_{\hat{x}_h}, \hat{x}_h)\| + \| \mu(r_{x_h}, \hat{x}_h) - \mu(r_{\hat{x}_h}, \hat{x}_h)\| \leq \mathcal{P} \circ \mu(r_{x_h}, \hat{x}_h \setminus \hat{x}_h, x_h) + \mathcal{P} \circ \mu(r_{\hat{x}_h}, \hat{x}_h \setminus \hat{x}_h, x_h), \) where the first equality follows from the definition of \( F; \) the second inequality follows from the fact that \( x \equiv x_h \) and \( \hat{x} \equiv \hat{x}_h \equiv \hat{x} \equiv \hat{x}_h; \) the second equality and the third inequality follow from the theorem on \( \mu; \) and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have \( 0 \leq \lim_{h \to 0^m} \|F(x+h) - F(x)\| \leq \lim_{h \to 0^m} \mathcal{P} \circ \mu(r_{x_h}, \hat{x}_h \setminus \hat{x}_h, x_h) + \mathcal{P} \circ \mu(r_{\hat{x}_h}, \hat{x}_h \setminus \hat{x}_h, x_h) = 0. \) Hence, by Proposition 12.37, \( F \) is continuous on the right and \( F \) satisfies (i) of Definition 12.41.

\( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2, \) we have

\[
T_F(x_1, x_2) = \sup_{n \in \mathbb{Z}^+} \sum_{\substack{\eta \in \mathcal{D}_n, \eta_i \geq 0, \eta_i \neq 0 \forall i \leq n \cup \mathcal{I}_i \geq 1 \cup \mathcal{I}_i < 1 \leq n}} \| \Delta_F(r_{x_i}, \hat{x}_i) \| \leq \sup_{n \in \mathbb{Z}^+} \sum_{\substack{\eta \in \mathcal{D}_n, \eta_i \geq 0, \eta_i \neq 0 \forall i \leq n \cup \mathcal{I}_i \geq 1 \cup \mathcal{I}_i < 1 \leq n}} \| \mathcal{P} \circ \mu(r_{x_i}, \hat{x}_i) \| \leq \sup_{n \in \mathbb{Z}^+} \sum_{\substack{\eta \in \mathcal{D}_n, \eta_i \geq 0, \eta_i \neq 0 \forall i \leq n \cup \mathcal{I}_i \geq 1 \cup \mathcal{I}_i < 1 \leq n}} \| \mathcal{P} \circ \mu(r_{x_i}, \hat{x}_i) \| = \mathcal{P} \circ \mu(r_{x_1}, x_2) < \infty
\]

where the second equality follows from the fact that \( F \) is a cumulative distribution function of \( \mu; \) the first inequality follows from Definition 11.108; the third equality follows from the fact that \( \mathcal{P} \circ \mu \) is a measure; and the last inequality follows from the assumption on \( \mu. \) Then, \( F \) satisfies (ii) of Definition 12.41.
12.4. FUNCTIONS OF BOUNDED VARIATION

∀x_1, x_2 ∈ Ω with x_1 < x_2, define ̃G : T_{x_1,x_2} → [0, ∞) ⊂ R by ̃G(x) = T_F (r_{x_1,x}), ∀x ∈ T_{x_1,x_2}. ∀x ∈ T_{x_1,x_2} with x ̸= x_2, we have

0 ≤ \lim_{h \to 0} \frac{\bar{G}(x+h) - \bar{G}(x)}{h} = \lim_{h \to 0} \frac{T_F (r_{x_1,x+h}) - T_F (r_{x_1,x})}{h}

= \lim_{h \to 0} \frac{\sum_{T_F (r_{\bar{x},\bar{x}}) \leq \lim_{h \to 0} \sum_{(\bar{x},\bar{x}) \in M} P \circ \mu(\bar{x},\bar{x})}{h}

= \lim_{h \to 0} \frac{\sum_{T_F (r_{\bar{x},\bar{x}}) \leq \lim_{h \to 0} \sum_{(\bar{x},\bar{x}) \in M} P \circ \mu(\bar{x},\bar{x})}{h}

where the first inequality follows from the monotonicity of ̃G; the first equality follows from the definition of ̃G; M := \{(x, x) ∈ T_{x_1,x} \times T_{x_1,x+h} \mid i ∈ \{0, \ldots, m - 1\}, ĵ ∈ ĵ \text{ card}(E) = i, \pi_k(x) = \pi_k(x_1), \forall k ∈ ĵ, \pi_k(\bar{x}) = \pi_k(\bar{x}) = \pi_k(x+h), \forall k ∈ ĵ \}\; \text{with (i), further concludes that}

T_F (r_{x_1,x_2}) = P \circ \mu(\bar{x},\bar{x}) = \Delta_G(r_{x_1,x_2}),

∀x_1, x_2 ∈ Ω with x_1 ≤ x_2 and T_{x_1,x_2} ⊆ Ω, and T_F = P \circ \mu(Ω).

This completes the proof of the proposition. □

Lemma 12.53 Let m ∈ N, R^m be endowed with the usual positive cone P, Ω ∈ B_B (R^m) be a region with the subset topology O, and F : Ω → R. Assume that the following holds.

(i) F is continuous on the right.

(ii) F is monotonically nondecreasing, i.e., F(x_1) ≤ F(x_2), ∀x_1, x_2 ∈ Ω with x_1 ≤ x_2.

Then, F is B_B (R^m)-measurable.

Proof Consider the special case when Ω is a rectangle. ∀α ∈ R, let B := \{x ∈ Ω \mid F(\bar{x}) < α\}. We will show that B ∈ O ⊆ B_B (Ω) ⊆ B_B (R^m). Then, by Proposition 11.35, F is B_B (R^m)-measurable.

∀x ∈ B, F(x_1) < α. Since F is continuous on the right, then ∃δ ∈ (0, ∞) ⊂ R such that F(\bar{x}) < α, ∀\bar{x} ∈ Ω ∩ (x + P) with \bar{x} < x + δ1_m. ∀\bar{x} ∈ \{\bar{x} ∈ Ω \mid \bar{x} < x + δ1_m\} := O ∈ O, since \bar{x}, x ∈ Ω and Ω is a rectangle, we have x ∨ \bar{x} ∈ Ω ∩ (x + P) with x ∨ \bar{x} < x + δ1_m. Then, F(\bar{x}) ≤ F(x ∨ \bar{x}) < α.
This implies that \( \bar{x} \in B \). By the arbitrariness of \( \bar{x} \), we have \( O \subseteq B \). By the arbitrariness of \( x \), we have \( B \in \mathcal{B}(\Omega) \subseteq \mathcal{B}(\mathbb{R}^m) \).

Now, consider the general case when \( \Omega \) is a region. Then, \( \exists \mathbb{N} \subseteq \mathbb{N} \), \( \exists (\bar{x}_i)_{i \in \mathbb{N}}, (\bar{x}_i)_{i \in \mathbb{N}} \subseteq \Omega \) with \( \bar{x}_i \triangleq \bar{x}_i \) and \( r_{\bar{x}_i, \bar{x}_i} \subseteq \Omega \), \( \forall i \in \mathbb{N} \), such that \( \Omega = \bigcup_{i \in \mathbb{N}} r_{\bar{x}_i, \bar{x}_i} \). Let \( B := \{ x \in \Omega \mid F(x) < \alpha \} = \bigcup_{i \in \mathbb{N}} (B \cap r_{\bar{x}_i, \bar{x}_i}) =: \bigcup_{i \in \mathbb{N}} B_i \). By the special case, \( B_i \in \mathcal{B}(\mathbb{R}^m) \), \( \forall i \in \mathbb{N} \). Then, \( B \in \mathcal{B}(\mathbb{R}^m) \). By Proposition 11.35, \( F \) is \( \mathcal{B}(\mathbb{R}^m) \)-measurable.

This completes the proof of the lemma. \( \Box \)

**Proposition 12.54** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}(\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), \( \mathcal{Y} \) be a separable normed linear space over \( \mathbb{K} \) with \( \mathcal{Y}^* \) being separable, \( \mu \in \mathcal{M}_f(\mathcal{P}(\Omega), \mathcal{B}(\mathcal{P}(\Omega)), \mathcal{Y}) \), and \( F : \Omega \to \mathcal{Y} \) be the cumulative distribution function of \( \mu \) as defined by \( F(x) = \mu(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) \), \( \forall x \in \Omega \), as shown in Proposition 12.51. Then, \( F \) is \( \mathcal{B}(\mathbb{R}^m) \)-measurable.

**Proof** We will prove the result in three steps. In the first step, we consider the special case \( \mathcal{Y} = \mathbb{R} \). By Jordan Decomposition Theorem 11.162, there exists a unique pair of mutually singular finite measures \( \mu_+ \) and \( \mu_- \) on the measurable space \( (\mathcal{P}(\Omega), \mathcal{B}(\mathcal{P}(\Omega))) \) such that \( \mu = \mu_+ - \mu_- \). Furthermore, \( \mathcal{P} \circ \mu = \mu_+ + \mu_- \). Let \( F_+ : \Omega \to [0, \infty) \subset \mathbb{R} \) and \( F_- : \Omega \to [0, \infty) \subset \mathbb{R} \) be the cumulative distribution functions of \( \mu_+ \) and \( \mu_- \), respectively, as defined by \( F_+(x) = \mu_+(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) \) and \( F_-(x) = \mu_-(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \succeq x \}) \), \( \forall x \in \Omega \). By Proposition 12.51, \( F_+ \) and \( F_- \) are of locally bounded variation, and hence, are continuous on the right. By their definitions, \( F_+(x_1) \leq F_+(x_2) \) and \( F_-(x_1) \leq F_-(x_2) \), \( \forall x_1, x_2 \in \Omega \) with \( x_1 \preceq x_2 \). Hence, \( F_+ \) and \( F_- \) are monotonically nondecreasing. By Lemma 12.53, \( F_+ \) and \( F_- \) are \( \mathcal{B}(\mathbb{R}^m) \)-measurable. Clearly, \( F(x) = \mu(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) = \mu_+(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) - \mu_-(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \succeq x \}) = F_+(x) - F_-(x) \), \( \forall x \in \Omega \). By Propositions 11.38, 11.39, and 7.23, \( F \) is \( \mathcal{B}(\mathbb{R}^m) \)-measurable. This completes the first step.

In the second step, we consider the special case \( \mathcal{Y} = \mathbb{C} \). By Proposition 11.164, \( \mu = \mu_r + i\mu_i \), where \( \mu_r \) and \( \mu_i \) are finite \( \mathbb{R} \)-valued measures on \( (\mathcal{P}(\Omega), \mathcal{B}(\mathcal{P}(\Omega))) \). Let \( F_r : \Omega \to \mathbb{R} \) and \( F_i : \Omega \to \mathbb{R} \) be defined by \( F_r(x) = \mu_r(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) \) and \( F_i(x) = \mu_i(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) \), \( \forall x \in \Omega \). By Proposition 12.51, \( F_r \) and \( F_i \) are cumulative distribution functions of \( \mu_r \) and \( \mu_i \), respectively. By the first step, \( F_r \) and \( F_i \) are \( \mathcal{B}(\mathbb{R}^m) \)-measurable. Clearly, \( F(x) = \mu(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) = \mu_r(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) + i\mu_i(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \preceq x \}) = F_r(x) + iF_i(x) \), \( \forall x \in \Omega \). By Propositions 11.38, 11.39, and 7.23, \( F \) is \( \mathcal{B}(\mathbb{R}^m) \)-measurable. This completes the second step.
In the last step, we consider the general case where $\mathcal{Y}$ is a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable. Fix any $y_0 \in \mathcal{Y}^*$. Define $\mu_{y_0} : \mathcal{B}_B(\mathcal{P}(\Omega)) \to \mathbb{K}$ by $\mu_{y_0}(E) = \langle \langle y_0, \mu(E) \rangle \rangle$, $\forall E \in \mathcal{B}_B(\mathcal{P}(\Omega))$. By Proposition 11.136, $\mu_{y_0}$ is a finite $\mathbb{K}$-valued measure on $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)))$ and $\mathcal{P} \circ \mu_{y_0} \leq \| \mu \| \mathcal{P} \circ \mu$. Define $F_{y_0} : \Omega \to \mathbb{K}$ by $F_{y_0}(x) = \langle \langle y_0, F(x) \rangle \rangle$, $\forall x \in \Omega$. Clearly, $F_{y_0}(x) = \langle \langle y_0, F(x) \rangle \rangle = \langle \langle y_0, \mu\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x \} \rangle \rangle = \mu_{y_0}(\{ \bar{x} \in \mathcal{P}(\Omega) \mid \bar{x} \leq x \})$, $\forall x \in \Omega$. Then, $F_{y_0}$ is a cumulative distribution function of $\mu_{y_0}$. By the first and the second steps, we have $F_{y_0}$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable. By Proposition 11.170, $F$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable. This completes the last step, and hence, the proof of the proposition.

**Proposition 12.55** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable, $\mu$ be a $\sigma$-finite $\mathcal{Y}$-valued measure on $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)))$ such that $\mathcal{P} \circ \mu(x_{1,2}) < \infty$, $\forall x_{1,2} \in \Omega$ with $x_{1,2} \leq x_{1,2}$, and $F : \Omega \to \mathcal{Y}$ be the cumulative distribution function of $\mu$ with origin $x_0 \in \Omega$ when $\Omega \neq \emptyset$. Then, $F$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable.

**Proof** The result is trivial if $\Omega = \emptyset$. Consider the case $\Omega \neq \emptyset$. Since $\Omega$ is a rectangle, then, by Proposition 12.52, $\mathcal{P}(\Omega)$ is also a rectangle and $\mathcal{O}$ is a region. Then, $\exists (x_i)_{i=1}^\infty$, $(\bar{x}_i)_{i=1}^\infty \subseteq \Omega$ such that $x_i \leq x_0 \leq \bar{x}_i$, $\forall i \in \mathbb{N}$, $\bigcap_{i=1}^\infty \mathcal{O}_{x_i, \bar{x}_i} = \Omega$, and $\mathcal{P}(\Omega) = \bigcup_{i=1}^\infty \mathcal{O}_{x_i, \bar{x}_i}$.

Fix any $i \in \mathbb{N}$, let $E_i := \mathcal{O}_{x_i, \bar{x}_i} \cap \mathcal{B}_E(\mathbb{R}^m)$ and $E_i := r_{\bar{x}_i, \bar{x}_i} \in \mathcal{B}_B(\mathbb{R}^m)$. Note that $\mathcal{P} \circ \mu(E_i) < \infty$, by the assumption on $\mu$. Let $(E_i, \mathcal{B}_E, \mu_{E_i})$ be the finite $\mathcal{Y}$-valued measure subspace of $(\mathcal{P}(\Omega), \mathcal{B}_B(\mathcal{P}(\Omega)), \mu)$, $\mathcal{O}_{E_i}$ be the subset topology on $E_i \subseteq \mathcal{O}$, and $\mathcal{E}_i := (E_i, \mathcal{O}_{E_i})$. Define $\tilde{F} : \mathcal{E}_i \to \mathcal{Y}$ by $\tilde{F}(x) = \mu_{E_i}(\{ \bar{x} \in E_i \mid \bar{x} \leq x \})$, $\forall x \in E_i$. By Proposition 1.51, $\tilde{F}$ is a cumulative distribution function of $\mu_{E_i}$, and is of locally bounded variation. By Proposition 12.54, $\mathcal{P}$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable. $\forall x \in \mathcal{E}_i$, let $\check{x} := x \land x_0 \in E_i$ and $\bar{x} := x \lor x_0 \in \mathcal{E}_i$. Then, $\Delta_F(\check{x}, \bar{x}) = \mu(\check{x}, \bar{x}) = \mu_{E_i}(\check{x}, \bar{x}) = \Delta_F(\check{x}, \bar{x})$, where the first equality follows from the fact that $F$ is a cumulative distribution function of $\mu$; the second equality follows from Proposition 11.115; and the third equality follows from the fact that $\tilde{F}$ is a cumulative distribution function of $\mu_{E_i}$. If $\exists k \in J := \{1, \ldots, m\}$ such that $\pi_{k_0}(x) = \pi_{k_0}(x_0)$, then $F(x) = \varnothing = \Delta_F(\check{x}, \bar{x})$. Otherwise, we have $(-1)^{n(x)}F(x) = \Delta_F(\check{x}, \bar{x})$, where $n(x) = \text{card}\{ k \in J \mid \pi_k(x) < \pi_k(x_0) \}$. Hence, we have $F(x) = (-1)^{n(x)}F(x) = (-1)^{n(x)}\Delta_F(\check{x}, \bar{x}) = (-1)^{n(x)}\sum_{J \subseteq J}(\check{J})\tilde{F}(\check{x}_J)$, where $\check{x}_J \in E_i$ is defined by $\pi_k(\check{x}_J) = \{ \pi_k(\check{x}) \mid k \in J \}$, $\forall k \in J$, $\forall k \in J \setminus \check{J}$, $\forall k \in J \setminus \check{J}$. $\forall j \in \{0, \ldots, 2^m - 1\}$, let $j_1 \ldots j_m$ be the representation of $j$ in base 2, $J_j := \{ k \in J \mid j_k = 0 \}$, $n_j := \text{card}(J_j)$. Define $O_j := \{ x \in \mathcal{E}_i$
Example 12.56 Let \( m \in \mathbb{N} \). Let \( \mathbb{R} \) be the second countable \( \sigma \)-finite real Banach measure space as defined in Example 11.28. \( \mathbb{R}^m \) be the second countable \( \sigma \)-finite product topological measure space according to Proposition 12.32. Since \( \mathbb{R}^m \) is a real Banach space, then \( \mathbb{R}^m : = ( ( \mathbb{R}^m , \mathbb{R}^m ) , \mu_{BM} ) \) is a second countable \( \sigma \)-finite real Banach measure space. \( \mu_{BM} \) is the \( m \)-dimensional Borel measure. By Carathéodory Extension Theorem 11.19, \( \mu_{BM} \) induces the \( m \)-dimensional Lebesgue outer measure \( \mu_{L,mo} : \mathbb{R}^m 2 \to [0, \infty] \subset \mathbb{R} \), which is defined by \( \mu_{L,mo}(E) = \inf \{ \lambda(A) \mid \lambda \in \mathcal{B}(\mathbb{R}^m ) , E \subseteq \bigcup_{i=1}^{\infty} A , \sum_{i=1}^{\infty} \mu_{BM}(A_{i}) \leq \lambda(E) \} \). \( \mathcal{B}_{L,m} := \{ E \subseteq \mathbb{R}^m \mid E \text{ is measurable with respect to } \mu_{L,mo} \} \) is the collection of \( m \)-dimensional Lebesgue measurable sets, and \( \mu_{L,m} := \mu_{L,mo}|_{\mathcal{B}_{L,m}} \) is the \( m \)-dimensional Lebesgue measure. By Carathéodory Extension Theorem 11.19, \( ( \mathbb{R}^m , \mathcal{B}(\mathbb{R}^m ) , \mu_{LM} ) \) is the completion of the \( m \)-dimensional Borel measure space \( ( \mathbb{R}^m , \mathcal{B}(\mathbb{R}^m ) , \mu_{BM} ) \), and is said to be the \( m \)-dimensional Lebesgue measure space. When \( m = 1 \), \( \mu_{BM} , \mu_{L,mo} , \mu_{L,m} , \) and \( \mathcal{B}_{L,m} \) are also denoted by \( \mu_B \), \( \mu_{LM} \), \( \mu_L \), and \( \mathcal{B}_L \), respectively.

\[ \forall E \subseteq \mathbb{R}^m \]
Since $\mathcal{O}_m \subseteq \mathcal{B}(\mathbb{R}^m)$, then $\mu_{L(m)}(E) \leq \hat{\mu}_{o}(E)$. On the other hand, if $\mu_{L(m)}(E) = \infty$, then $\infty = \mu_{L(m)}(E) \leq \hat{\mu}_{o}(E) \leq \infty$ and $\mu_{L(m)}(E) = \hat{\mu}_{o}(E)$.

If $\mu_{L(m)}(E) < \infty$, $\forall \in (0, \infty) \subset \mathbb{R}$, $\exists (A_i)_{i=1}^\infty \subseteq \mathcal{B}(\mathbb{R}^m)$ such that $E \subseteq \bigcup_{i=1}^\infty A_i$ and $\mu_{L(m)}(E) \leq \sum_{i=1}^\infty \mu_{Bm}(A_i) < \mu_{L(m)}(E) + \epsilon / 2$, $\forall i \in \mathbb{N}$, since $\mathbb{R}^m$ is a topological measure space, then $\exists O_i \in \mathcal{O}_m$ with $A_i \subseteq O_i$ such that $\mu_{Bm}(O_i \setminus A_i) < 2^{-i-1}\epsilon$. Then, $E \subseteq \bigcup_{i=1}^\infty O_i$ and $\hat{\mu}_{o}(E) \leq \sum_{i=1}^\infty \mu_{Bm}(O_i) = \sum_{i=1}^\infty (\mu_{Bm}(A_i) + \mu_{Bm}(O_i \setminus A_i)) < \mu_{L(m)}(E) + \epsilon / 2 + \epsilon / 2 = \mu_{L(m)}(E) + \epsilon$. By the arbitrariness of $\epsilon$, we have $\hat{\mu}_{o}(E) \leq \mu_{L(m)}(E)$. Hence, $\mu_{L(m)}(E) = \hat{\mu}_{o}(E)$.

By the translational invariance of $\mathbb{R}$, we have $\mathbb{R}^m$ is translational invariant. Then, by the above, $\mu_{L(m)}$ and $\mu_{Lm}$ are translational invariant.

\begin{definition}[Lebesgue-Stieltjes Integral] Let $m \in \mathbb{N}$, $\Omega \in \mathcal{B}(\mathbb{R}^m)$ be a region with the subset topology, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $F : \Omega \to \mathcal{Y}$ be of locally bounded variation, $(\mathfrak{P}(\Omega), \mathcal{B}(\mathfrak{P}(\Omega)), \mu)$ be the unique $\sigma$-finite $\mathcal{Y}$-valued measure space that admits $F$ as a cumulative distribution function as prescribed in Theorem 12.50. $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \in \mathcal{B}(\mathfrak{P}(\Omega))$, and $g : U \to \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}(\mathfrak{P}(\Omega))$-measurable (or $\mathcal{B}_{\mathbb{R}^m}$-measurable). We will denote the integral $\int_U g(x) \, d\mu(x)$ by $\int_U g(x) \, dF(x)$ or $\int_U g(x) \, dF(x)$, whenever the first integral makes sense.
\end{definition}

12.5 Absolute and Lipschitz Continuity

\begin{definition} Let $I \subseteq \mathbb{R}$ be an interval and $\mathcal{Y}$ be a normed linear space. A function $f : I \to \mathcal{Y}$ is said to be absolutely continuous at $x_0 \in I$ if $\forall \alpha \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta(\epsilon) \in (0, a] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - \alpha \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \cdots \leq x_n \leq \bar{x}_n \leq x_0 + \alpha$ with $x_i, \bar{x}_i \in I$, $i = 1, \ldots, n$, and $\sum_{i=1}^n \mu_{I}(x_1, \bar{x}_i) = \sum_{i=1}^n (\bar{x}_i - x_1) < \delta(\epsilon)$, we have $\sum_{i=1}^n \| \Delta f(x_i, \bar{x}_i) \| = \sum_{i=1}^n \| f(x_i) - f(x_1) \| < \epsilon$. We will say that $f$ is absolutely continuous if it is absolutely continuous at any $x_0 \in I$. We will say $f$ is absolutely continuous on $E \subseteq I$ if it is absolutely continuous at any $x_0 \in E$.

Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, and $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$. $F : \Omega \to \mathcal{Y}$ is said to be absolutely continuous at $x_0 \in \Omega$ if the following conditions holds.

(i) $\forall \alpha > 1$, $\forall i_0 \in \mathcal{I} := \{1, \ldots, m\}$, let $D_{i_0} := \{ x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega \ni \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{i_0}(\bar{x}) = x \}$, where $\pi_{i_0} : \mathbb{R}^m \to \mathbb{R}^{m-1}$ is defined by $\pi_{i_0}(x) = (\pi_1(x), \ldots, \pi_{i_0-1}(x), \pi_{i_0+1}(x), \ldots, \pi_m(x))$, $\forall x \in \mathbb{R}^m$, and $F_{i_0} : D_{i_0} \to \mathcal{Y}$ be defined by $F_{i_0}(x) = F(\bar{x})$, $\forall \bar{x} \in \Omega$ with $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{i_0}(\bar{x}) = x$. The function $F_{i_0}$ is absolutely continuous at $\pi_{i_0}(x_0)$.\)
(ii) \( \exists a \in (0, \infty) \subset \mathbb{R}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\varepsilon) \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall \left(\bar{x}_i\right)_{i=1}^{n} \subseteq \bigcup_{a=1}^{m} T_{x-a1_m, x+a1_m} \cap \Omega \) with \( \bar{x}_i \equiv \bar{x}_i, \forall i \in \{1, \ldots, n\}, (r_{x, i} \bar{x}_i)_{i=1}^{n} \) being pairwise disjoint, and \( \sum_{i=1}^{n} \mu_{B_0}(r_{x, i} \bar{x}_i) < \delta(\varepsilon), \) we have \( \sum_{i=1}^{n} \| \Delta F(r_{x, i} \bar{x}_i) \| < \varepsilon. \)

\( F \) is said to be absolutely continuous if it is absolutely continuous at any \( x_0 \in \Omega. \) \( F \) is said to be absolutely continuous on \( E \subseteq \Omega \) if it is absolutely continuous at any \( x_0 \in E. \)

In the above definition, the definition for general \( \mathbb{R}^m \) case subsumes the definition for \( \mathbb{R} \) case. We explicitly specify the \( \mathbb{R} \) case for its direct implications.

**Proposition 12.59** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \subset B_0 \) be a rectangle with subset topology \( \mathcal{O}, \mathcal{Y} \) be a normed linear space over \( \mathbb{K}, \) and \( F : \Omega \to \mathcal{Y} \) be absolutely continuous at \( x_0 \in \Omega. \) Then, \( F \) is continuous at \( x_0. \)

**Proof** We will prove the result using mathematical induction on \( m. \)

1° \( m = 1. \) By Definition 12.58, \( F \) is continuous at \( x_0. \) This case is proved.

2° Assume that the result holds for \( m \leq k - 1 \in \mathbb{N}. \)

3° Consider the case when \( m = k \in \{2, 3, \ldots\}. \) By (i) of Definition 12.58, \( \exists \varepsilon \in (0, \infty) \subset \mathbb{R}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\varepsilon) \in (0, \infty) \subset \mathbb{R}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \Delta F_0 \in \mathcal{B}_0 \) such that \( \| F_0(x) \| \leq \| F_0(\varepsilon) \| < \varepsilon, \forall x \in \mathbb{B}_0(\varepsilon) \cap \Omega. \)

By (ii) of Definition 12.58, \( \exists a \in (0, \infty) \subset \mathbb{R}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\varepsilon) \in (0, \infty) \subset \mathbb{R}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \Delta F_0 \in \mathcal{B}_0 \) such that \( \| F_0(x) \| \leq \| F_0(\varepsilon) \| < \varepsilon, \forall x \in \mathbb{B}_0(\varepsilon) \cap \Omega. \)

Take \( \delta(\varepsilon) := \min\{a, \sqrt[2]{\delta(2^{-m}\varepsilon)}, \delta(2^{-m}\varepsilon), \ldots, \delta_m(2^{-m}\varepsilon)\} \in (0, \infty) \subset \mathbb{R}. \forall x \in \mathbb{B}_0(x_0, \delta(\varepsilon)) \cap \Omega, \) we have \( \| F(x) - F(x_0) \| \leq \| \Delta F(x_0) \| + \sum_{j \in J} \| F(x_j) - F(x_0) \|, \) where \( x := x_0 \) and \( x := x \) is defined by \( \pi(x) = \{\pi(x) \mid \forall i \} \), \( \forall i \in \mathbb{J} \) \( \mathbb{J} = \mathbb{J} \setminus J, \forall i \in \mathbb{J} \setminus J. \)

Clearly, \( x \in \mathbb{B}_0(x_0, \delta(\varepsilon)) \cap \Omega. \) Since \( \mu_{B_0}(x) \leq \| x - x_0 \| < \delta(2^{-m}\varepsilon), \) then \( \| \Delta F(x_0) \| < \delta(2^{-m}\varepsilon). \) \( \forall J \subset \mathbb{J}, \) fix any \( x_0 \in J \setminus J, \) we have \( \| F(x) - F(x_0) \| < \| \Delta F(x_0) \| + \sum_{j \in J} \| F(x_j) - F(x_0) \|. \)

Clearly, \( \pi_0(x) \in \mathbb{D}_0 \) and \( \| F(x) - F(x_0) \| < \| \Delta F(x_0) \| + \sum_{j \in J} \| F(x_j) - F(x_0) \| = \| x - x_0 \| < \| x_0 \| + \| x - x_0 \| < \delta(2^{-m}\varepsilon). \) Then, \( \| F(x) - F(x_0) \| < \varepsilon. \) This shows that \( F \) is continuous at \( x_0. \)
This completes the induction process. This completes the proof of the proposition.

Proposition 12.60 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $F : \Omega \to \mathcal{Y}$, and $x_0 \in \Omega$. Then, $F$ is absolutely continuous at $x_0$ if, and only if,

(a) $\forall i \in \{0, \ldots, m-1\}$, $\forall J \subset \bar{J} := \{1, \ldots, m\}$ with $\text{card}(J) = i$, let $\pi_J : \mathbb{R}^m \to \mathbb{R}^i$ be such that $\pi_J(\bar{x}) = (\pi_i(\bar{x}))_{i \in J}$, $\forall \bar{x} \in \mathbb{R}^m$, $\pi_{\bar{J} \setminus J} : \mathbb{R}^m \to \mathbb{R}^{m-i}$ be such that $\pi_{\bar{J} \setminus J}(\bar{x}) = (\pi_i(\bar{x}))_{i \in \bar{J} \setminus J}$, $\forall \bar{x} \in \mathbb{R}^m$, $M : \mathbb{R}^i \times \mathbb{R}^{m-i} \to \mathbb{R}^m$ be such that $M(x_a, x_b) = \bar{x}$ with $x_a = \pi_J(\bar{x})$ and $x_b = \pi_{\bar{J} \setminus J}(\bar{x})$, $\forall x_a \in \mathbb{R}^i$ and $\forall x_b \in \mathbb{R}^{m-i}$, $x_f := \pi_J(x_0)$, $\bar{x}_f := \pi_{\bar{J} \setminus J}(x_0)$, $\bar{\Omega}_f := \pi_{\bar{J} \setminus J}(\Omega)$, and $F_J : \bar{\Omega}_f \to \mathcal{Y}$ be defined by $F_J(\bar{x}) = F(M(x_f, \bar{x}), \bar{x} \in \bar{\Omega}_f, \exists a_J \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_J(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n \subseteq \bar{\Omega}_f$ being pairwise disjoint, and $\sum_{j=1}^n \mu_{B(m-1)}(r_{\bar{x}_j, \bar{x}_j}) < \delta_J(\epsilon)$, we have $\sum_{j=1}^n \|\Delta_{F_J}(r_{\bar{x}_j, \bar{x}_j})\| < \epsilon$.

Proof We will prove the result using mathematical induction on $m$.

1° The result is clearly a restatement of Definition 12.58. Hence, the case is proved.

2° Assume that the result holds for $m \leq k-1 \in \mathbb{N}$.

3° Consider the case when $m = k \in \{2, 3, \ldots\}$.

"Sufficiency" We need to show that $F$ is absolutely continuous at $x_0$ assuming (a). Let $i = 0$, (a) implies (ii) of Definition 12.58. $\forall i_0 \in \bar{J}$, let $\pi_{i_0}$, $\pi_{i_0}$, $D_{i_0}$, and $F_{i_0}$ be as defined in (i) of Definition 12.58. By (a) holds for any $i \in \{1, \ldots, m-1\}$ and any $J \subset \bar{J}$ with $\text{card}(J) = i$ such that $i_0 \in J$, we have (a) holds for $F_{i_0}$ at $\pi_{i_0}(x_0)$. By the inductive assumption, $F_{i_0}$ is absolutely continuous at $\pi_{i_0}(x_0)$. Hence, (i) of Definition 12.58 holds for $F$ at $x_0$. Then, $F$ is absolutely continuous at $x_0$.

"Necessity" We need to show that (a) holds assuming that $F$ is absolutely continuous at $x_0$. By (ii) of Definition 12.58, (a) holds when $i = 0$. By (i) of Definition 12.58, $\forall i_0 \in \bar{J}$, let $\pi_{i_0}$, $\pi_{i_0}$, $D_{i_0}$, and $F_{i_0}$ be as defined in (i) of Definition 12.58, we have $F_{i_0}$ is absolutely continuous at $\pi_{i_0}(x_0)$. By the inductive assumption, (a) holds for $F_{i_0}$. Then, (a) holds for $F$, $\forall i \in \{1, \ldots, m-1\}$, $\forall J \subset \bar{J}$ with $\text{card}(J) = i$ and $i_0 \in J$. Hence, (a) holds for $F$. This completes the proof for the necessity.

This completes the induction process.

This completes the proof of the proposition.

Proposition 12.61 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $f : \Omega \to \mathcal{Y}$, $g : \Omega \to \mathcal{Z}$, $x_0 \in \Omega$, and $h : \Omega \to \mathcal{Y} \times \mathcal{Z}$ be defined by $h(x) = (f(x), g(x))$, $\forall x \in \Omega$. Then, $h$ is
We will prove the result using mathematical induction on \( n \).

1° **Proof.**

"Sufficiency" By \( f \) being absolutely continuous at \( x_0 \), \( \exists \varepsilon > 0 \) such that for all \( \delta > 0 \), there is an \( \varepsilon' > 0 \) such that for all \( \varepsilon > 0 \), \( \delta = \varepsilon' \), and \( \varepsilon/2 \), we have

\[
\sum_{i=1}^{n} \| f(x_i) - f(x_{i-1}) \| < \varepsilon.
\]

By \( g \) being absolutely continuous at \( x_0 \), \( \exists \varepsilon > 0 \) such that for all \( \delta > 0 \), there is an \( \varepsilon' > 0 \) such that for all \( \varepsilon > 0 \), \( \delta = \varepsilon' \), and \( \varepsilon/2 \), we have

\[
\sum_{i=1}^{n} \| g(x_i) - g(x_{i-1}) \| < \varepsilon.
\]

Let \( a := \min \{ \varepsilon, \varepsilon' \} \in (0, \infty) \subset \mathbb{R} \), and \( \varepsilon'' := \min \{ \varepsilon, \varepsilon' \} \in (0, \infty) \subset \mathbb{R} \). Then, \( \sum_{i=1}^{n} \| f(x_i) - f(x_{i-1}) \| < \varepsilon'' \), and \( \sum_{i=1}^{n} \| g(x_i) - g(x_{i-1}) \| < \varepsilon'' \). Hence, \( f \) and \( g \) are absolutely continuous at \( x_0 \).

2° Assume that \( f \) and \( g \) are absolutely continuous at \( x_0 \). Let \( \varepsilon > 0 \), \( \delta = \varepsilon \), and \( \varepsilon/2 \). Then, \( \sum_{i=1}^{n} \| f(x_i) - f(x_{i-1}) \| < \varepsilon \), and \( \sum_{i=1}^{n} \| g(x_i) - g(x_{i-1}) \| < \varepsilon \). Hence, \( f \) and \( g \) are absolutely continuous at \( x_0 \).

3° Consider the case \( m = k \in \{ 2, 3, \ldots \} \).

"Sufficiency" Assume that \( f \) and \( g \) are absolutely continuous at \( x_0 \). Let \( \varepsilon > 0 \), \( \delta = \varepsilon \), and \( \varepsilon/2 \). Then, \( \sum_{i=1}^{n} \| f(x_i) - f(x_{i-1}) \| < \varepsilon \), and \( \sum_{i=1}^{n} \| g(x_i) - g(x_{i-1}) \| < \varepsilon \). Hence, \( f \) and \( g \) are absolutely continuous at \( x_0 \).
12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY
609

\[ \sum_{i=1}^{n} \| (\Delta f(\bar{x}_i, \dot{x}_i), \Delta g(\bar{x}_i, \dot{x}_i)) \| = \sum_{i=1}^{n} (\| \Delta f(\bar{x}_i, \dot{x}_i) \|^2 + \| \Delta g(\bar{x}_i, \dot{x}_i) \|)^{1/2} \leq \sum_{i=1}^{n} \| \Delta f(\bar{x}_i, \dot{x}_i) \| + \sum_{i=1}^{n} \| \Delta g(\bar{x}_i, \dot{x}_i) \| < \epsilon. \] Hence, \( h \) satisfies (ii) of Definition 12.58. Then, \( h \) is absolutely continuous at \( x_0 \).

“Necessity” Let \( h \) be absolutely continuous at \( x_0 \). \( \forall i_0 \in J, \) let \( D_{i_0}, \pi_{i_0}, f_{i_0}, g_{i_0}, \) and \( h_{i_0} \) be defined as in Definition 12.58. By (i) of Definition 12.58, \( h_{i_0} \) is absolutely continuous at \( \pi_{i_0}(x_0) \). By the inductive assumption, \( f_{i_0} \) and \( g_{i_0} \) are absolutely continuous at \( x_0 \). Hence, \( f \) and \( g \) satisfy (i) of Definition 12.58.

By (ii) of Definition 12.58, \( \exists a \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall (\bar{x}_i)_{i=1}^{n}, (\dot{x}_i)_{i=1}^{n} \subset \mathbb{R} \) such that \( \bar{x}_i \neq \dot{x}_i, \forall i \in \{1, \ldots, n\}, (\bar{x}_i, \dot{x}_i)_{i=1}^{n} \) being pairwise disjoint, and \( \sum_{i=1}^{n} \| \Delta f(\bar{x}_i, \dot{x}_i) \| < \delta(\epsilon) \), we have \( \epsilon > \sum_{i=1}^{n} \| \Delta h(\bar{x}_i, \dot{x}_i) \| \geq \sum_{i=1}^{n} \| \Delta f(\bar{x}_i, \dot{x}_i) \| \). Hence, \( f \) satisfies (ii) of Definition 12.58. Then, \( f \) is absolutely continuous at \( x_0 \). By symmetry, \( g \) is absolutely continuous at \( x_0 \).

This completes the induction process. This completes the proof of the proposition.

\[ \Box \]

Proposition 12.62 Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) be a rectangle with subset topology \( \mathcal{O} \), \( \mathcal{O} \) be a normed linear space over \( K \), \( F : \Omega \to \mathcal{O} \), and \( \Omega_1, \Omega_2 \subseteq \Omega \) be rectangles in \( \mathbb{R}^m \) such that \( \Omega_1 \cup \Omega_2 = \Omega \). Then, the following statements hold.

(i) If \( F \) is absolutely continuous on \( E \subseteq \Omega \), then \( F|_{\Omega_1} \) is absolutely continuous on \( E \cap \Omega_1 \).

(ii) If \( F|_{\Omega_1} \) and \( F|_{\Omega_2} \) are absolutely continuous and \( \Omega_1 \) and \( \Omega_2 \) are both relatively open or both relatively closed (with respect to the subset topology \( \mathcal{O} \)), then \( F \) is absolutely continuous.

Proof

(i) \( \forall x_0 \in E \cap \Omega_1 \), since \( F \) is absolutely continuous at \( x_0 \), then \( F|_{\Omega_1} \) is absolutely continuous at \( x_0 \). By the arbitrariness of \( x_0 \), \( F|_{\Omega_1} \) is absolutely continuous on \( E \cap \Omega_1 \).

(ii) We will distinguish two exhaustive cases: Case 1: \( \Omega_1 \) and \( \Omega_2 \) are relatively open in \( \Omega \); Case 2: \( \Omega_1 \) and \( \Omega_2 \) are relatively closed in \( \Omega \).

Case 1: \( \Omega_1 \) and \( \Omega_2 \) are relatively open in \( \Omega \). Fix any \( x_0 \in \Omega \). We will prove the statement that “if \( F|_{\Omega_1} \) is absolutely continuous at \( x_0 \) when \( x_0 \in \Omega_1 \), and \( \Omega_1 \) and \( \Omega_2 \) are relatively open in \( \Omega \), then \( F \) is absolutely continuous at \( x_0 \)” by mathematical induction on \( m \).

1° \( m = 1 \). Without loss of generality, assume \( x_0 \in \Omega_1 \). Since \( \Omega_1 \) is open relative to \( \Omega \), then \( \exists a \in (0, \infty) \subset \mathbb{R} \) such that \( (x_0 - a, x_0 + a) \cap \Omega_1 = (x_0 - a, x_0 + a) \cap \Omega \). By \( F|_{\Omega_1} \) being absolutely continuous at \( x_0 \), we can easily show that that \( F \) is absolutely continuous at \( x_0 \). This completes the first step in the induction process.

2° Assume that the result holds for \( m \leq k - 1 \in \mathbb{N} \).

3° Consider the case when \( m = k \in \{2, 3, \ldots\} \). Without loss of generality, assume \( x_0 \in \Omega_1 \). \( \forall i_0 \in J := \{1, \ldots, m\} \), let \( D_{i_0} \) and \( F_{i_0} : D_{i_0} \to \mathbb{Y} \) be as
defined in Definition 12.58. Let $D_{i_0,j} := \{ x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega \exists \cdot \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) and \pi_{i_0}^0(\bar{x}) = x \}$ and $F_{i_0,j} : D_{i_0,j} \to Y$ be defined by $F_{i_0,j}(x) = F|_{\Omega_j}(\bar{x})$, $\forall x \in D_{i_0,j}$, where $\bar{x} \in \Omega_j$ is such that $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{i_0}^0(\bar{x}) = x$, $j = 1, 2$. Fix any $j \in \{1, 2\}$. Since $\Omega_j \subseteq \Omega$ is a relatively open rectangle, then $D_{i_0,j} \subseteq D_{i_0}$ is a relatively open rectangle. By $F|_{\Omega_j}$ being absolutely continuous at $x_0$ when $x_0 \in \Omega_j$, we have $F_{i_0,j}$ is absolutely continuous at $\pi_{i_0}^0(x_0)$ when $\pi_{i_0}^0(x_0) \in D_{i_0,j}$. It is easy to see that $F_{i_0,j} = F|_{D_{i_0,j}}$ and $D_{i_0,1} \cup D_{i_0,2} = D_{i_0}$. By the inductive assumption, $F_{i_0}$ is absolutely continuous at $\pi_{i_0}^0(x_0) \in D_{i_0}$. Hence, $F$ satisfies (i) of Definition 12.58.

Since $\Omega_1$ is relatively open in $\Omega$, then $\exists a \in (0, \infty) \subseteq \mathbb{R}$ such that $\cap_{x_0 = a} \Omega_1 \cap \Omega = \cap_{x_0 = a} \Omega_1 \cap \Omega$. By $F|_{\Omega_1}$ being absolutely continuous at $x_0$ and (ii) of Definition 12.58, $\exists \alpha \in (0, a) \subseteq \mathbb{R}, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$. Fix any $j \in \{1, 2\}$. Then, by Proposition 4.10, $\exists a \in (0, \infty) \subseteq \mathbb{R}$ such that $(x_0 - a, x_0 + a) \cap \Omega \subseteq \Omega \cap \Omega \subseteq \Omega$. $F|_{\Omega_1}$ is absolutely continuous at $x_0$ implies that $F$ is absolutely continuous at $x_0$. This subcase is proved. Case 2b: $x_0 \in \Omega_2$. Then, $x_0 \in \Omega_1 \cap \Omega_2$. By $F|_{\Omega_1}$ being absolutely continuous at $x_0$, $\exists a \in (0, \infty) \subseteq \mathbb{R}, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta(\epsilon) \in (0, a] \subseteq \mathbb{R}, \forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \cdots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_j, \bar{x}_j \in \Omega_i$, $j = 1, \ldots, n$, and $\sum_{j=1}^{n} (\bar{x}_j - x_j) < \delta(\epsilon)$, we have $\sum_{j=1}^{n} \| F(\bar{x}_j) - F(x_j) \| < \epsilon$, $i = 1, 2$. Take $a := \min\{a_1, a_2\} \subseteq \mathbb{R}, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_1(\epsilon/2), \delta_2(\epsilon/2)\}$, $\delta_1(\epsilon/2) \subseteq (0, a] \subseteq \mathbb{R}, \forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \cdots \leq x_n \leq \bar{x}_n \leq x_0 + a$ with $x_j, \bar{x}_j \in \Omega_j$, $j = 1, \ldots, n$, and $\sum_{j=1}^{n} (\bar{x}_j - \bar{x}_j) < \delta(\epsilon)$. Without loss of generality, assume $\exists j_0 \in \{1, \ldots, n - 1\}$ such that $\bar{x}_{j_0} \leq x_0 \leq \bar{x}_{j_0+1}$. (In case $j_0$ does not exist, add a pair of $x_0$'s into the collection of $x_j$'s and $\bar{x}_j$’s.) Then, $x_1 \in \Omega = \Omega_1 \cup \Omega_2$. Without loss of generality, assume $x_1 \in \Omega_1$. Then $x_1, \bar{x}_1, \ldots, x_{j_0}, \bar{x}_{j_0} \in \Omega_1$, since $\Omega_1$ is an rectangle (interval, since $m = 1$), and $\sum_{j=1}^{j_0} (\bar{x}_j - x_j) < \delta(\epsilon) \leq \delta(\epsilon/2)$. Then, $\sum_{j=1}^{j_0} \| F(\bar{x}_j) - F(x_j) \| < \epsilon/2$. By a similar argument, we have
\[ \sum_{j=j_0+1}^{n} \| F(\bar{x}_j) - F(x_j) \| < \epsilon/2. \] Then, \[ \sum_{j=1}^{n} \| F(\bar{x}_j) - F(x_j) \| < \epsilon. \] Hence, \( F \) is absolutely continuous at \( x_0 \). This subcase is proved.

Hence, in both subcases, \( F \) is absolutely continuous at \( x_0 \). This completes the proof for the first step of the induction process.

2° Assume that the result holds for \( m \leq k - 1 \in \mathbb{N} \).

3° Consider the case when \( m = k \in \{2, 3, \ldots \} \). Without loss of generality, assume \( x_0 \in \Omega_1 \). \( \forall \alpha \in J \), let \( D_{i_0} \) and \( F_{i_0} : D_{i_0} \to \mathbb{Y} \) be defined as in Definition 12.58. Let \( D_{i_0,j} := \{ x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega_j \exists \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{i_0}(\bar{x}) = x \} \) and \( F_{i_0,j} : D_{i_0,j} \to \mathbb{Y} \) be defined by \( F_{i_0,j}(x) = F_{i_0}(\bar{x}) \), \( \forall \bar{x} \in D_{i_0,j} \), where \( \bar{x} \in \Omega_j \) is such that \( \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \) and \( \pi_{i_0}(\bar{x}) = x \), \( j = 1, 2 \). Fix any \( j \in \{1, 2\} \). Since \( \Omega_j \subseteq \Omega \) is a relatively closed rectangle, then \( D_{i_0,j} \subseteq D_{i_0} \) is a relatively closed rectangle. By \( F|_{\Omega_1} \) being absolutely continuous at \( x_0 \) when \( x_0 \in \Omega_j \), we have \( F_{i_0,j} \) is absolutely continuous at \( \pi_{i_0}(x_0) \) when \( \pi_{i_0}(x_0) \in D_{i_0,j} \). It is easy to see that \( F_{i_0,j} = F_{i_0}|_{D_{i_0,j}} \) and \( D_{i_0,1} \cup D_{i_0,2} = D_{i_0} \). By the inductive assumption, \( F_{i_0} \) is absolutely continuous at \( \pi_{i_0}(x_0) \in D_{i_0} \). Hence, \( F \) satisfies (i) of Definition 12.58.

We will further distinguish two exhaustive and mutually exclusives subcases: Case 2α: \( x_0 \notin \Omega_2 \); Case 2β: \( x_0 \in \Omega_2 \).

Case 2α: \( x_0 \notin \Omega_2 \). Then, \( x_0 \in \Omega \setminus \Omega_2 \). Since \( \Omega_2 \) is relatively closed in \( \Omega \), by Proposition 4.10, \( \exists a \in (0, \infty) \subset \mathbb{R} \) such that \( r_{x_0-a_1,m,x_0+a_1} \cap \Omega \subseteq \Omega \setminus \Omega_2 \subseteq \Omega_1 \). By \( F|_{\Omega_1} \) being absolutely continuous at \( x_0 \), \( \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall (\bar{x}_i)_{i=1}^{n} \subseteq \mathbb{R}_{x_0-a_1,m,x_0+a_1} \cap \Omega \) with \( \bar{x}_i \subseteq \bar{x}_j \), \( \bar{x}_i \subseteq \bar{x}_j \), \( \forall i \in \{1, \ldots, n\} \), \( (\bar{x}_i, \bar{x}_j)_{i,j=1}^{n} \) being pairwise disjoint, and \( \sum_{i=1}^{n} \mu_{\mathbb{R}}(r_{\bar{x}_i, \bar{x}_j}) < \delta(\epsilon) \), we have \( \sum_{i=1}^{n} \| \Delta F(\bar{x}_i, \bar{x}_j) \| < \epsilon \). Hence, \( F \) satisfies (ii) of Definition 12.58. Then, \( F \) is absolutely continuous at \( x_0 \). This subcase is proved.

Case 2β: \( x_0 \in \Omega_2 \). Then, \( x_0 \in \Omega_1 \cap \Omega_2 \). Fix any \( i \in \{1, 2\} \). By \( F|_{\Omega_i} \) being absolutely continuous at \( x_0 \) and (ii) of Definition 12.58, \( \exists \delta\epsilon_i \in (0, \infty) \subset \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta\epsilon_i \in (0, \infty) \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall (\bar{x}_j)_{j=1}^{n} \subseteq \mathbb{R}_{x_0-a_1,m,x_0+a_1} \cap \Omega \) with \( \bar{x}_j \subseteq \bar{x}_j \), \( \forall j \in \{1, \ldots, n\} \), \( (\bar{x}_j, \bar{x}_j)_{j=1}^{n} \) being pairwise disjoint, and \( \sum_{j=1}^{n} \mu_{\mathbb{R}}(r_{\bar{x}_j, \bar{x}_j}) < \delta\epsilon_i \), we have \( \sum_{j=1}^{n} \| \Delta F(\bar{x}_j, \bar{x}_j) \| < \epsilon \).

Take \( a := \min\{a_1, a_2\} \in (0, \infty) \subset \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), let \( \delta(\epsilon) := \min\{\delta\epsilon(\epsilon/2), \delta(\epsilon/2)\} \in (0, \infty) \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall (\bar{x}_j)_{j=1}^{n} \subseteq \mathbb{R}_{x_0-a_1,m,x_0+a_1} \cap \Omega \) with \( \bar{x}_j \subseteq \bar{x}_j \), \( \forall j \in \{1, \ldots, n\} \), \( (\bar{x}_j, \bar{x}_j)_{j=1}^{n} \) being pairwise disjoint, and \( \sum_{j=1}^{n} \mu_{\mathbb{R}}(r_{\bar{x}_j, \bar{x}_j}) < \delta(\epsilon) \), \( \forall j \in \{1, \ldots, n\} \), since \( \Omega = \Omega_1 \cup \Omega_2 \) and \( \Omega_1 \) and \( \Omega_2 \) are relatively closed rectangles in the rectangle \( \Omega \), then \( r_{\bar{x}_j, \bar{x}_j} = \bigcup_{j=1}^{n} r_{\bar{x}_j, \bar{x}_j} \cap \Omega_k \), where \( \bar{x}_j,k \in \mathbb{R}_{x_0-a_1,m,x_0+a_1} \cap \Omega_k \), \( \bar{x}_j,k \subseteq \bar{x}_j,k \), \( k = 1, 2 \), and the two sets in the union are disjoint. To see the above, note that \( r_{\bar{x}_j, \bar{x}_j} \cap \Omega_k \) is a rectangle, \( k = 1, 2 \). If \( r_{\bar{x}_j, \bar{x}_j} \subseteq \Omega_1 \), for some \( l \in \{1, 2\} \), then let \( \bar{x}_j,l = \bar{x}_j,k \), \( \bar{x}_j,l = \bar{x}_j,k \), \( \bar{x}_j,3-l = x_0 \) and \( \bar{x}_j,3-l = x_0 \). The desired result holds. Consider the remaining po-
sibility, we must have $\bar{x}_j \in \Omega_1 \setminus \Omega_{3-l}$ and $\tilde{x}_j \in \Omega_{3-l} \setminus \Omega_1$, for some $l \in \{1, 2\}$. Without loss of generality, assume $l = 1$. $\Omega_1$ and $\Omega_2$ are $m$-dimensional rectangles and none contains the other, their union is a $m$-dimensional rectangle $\Omega$, in which both $\Omega_1$ and $\Omega_2$ are relatively closed.

The only possibility for this to happen is that $\pi_i(\Omega_1) = \pi_i(\Omega_2) = \pi_i(\Omega), \forall i \in \bar{J} \setminus \{i_0\}$, for some $i_0 \in \bar{J}$, and $\pi_{i_0}(\Omega_1)$ and $\pi_{i_0}(\Omega_2)$ are overlapping relatively closed intervals in $\pi_{i_0}(\Omega)$. Let $\tilde{x}_{j,1} = \tilde{x}_j \in \Omega_1 \setminus \Omega_2$. Since $\bar{r}_{\tilde{x}_{j,1},\tilde{x}_j} \subseteq \Omega$, then $\bar{r}_{\tilde{x}_{j,1},\tilde{x}_j} \cap \Omega_1 \subseteq \Omega$ is a relatively closed rectangle with respect to $\bar{r}_{\tilde{x}_{j,1},\tilde{x}_j}$. Then, $\emptyset \neq \bar{r}_{\tilde{x}_{j,1},\tilde{x}_j} \cap \Omega_1 =: \bar{r}_{\tilde{x}_{j,1},\tilde{x}_j,1} \subseteq \bar{r}_{\tilde{x}_{j,1},\tilde{x}_j} \subseteq \Omega$. This implies that $\tilde{x}_{j,1}, \tilde{x}_{j,1,1} \in \bar{r}_{x_0-a_1m,x_0+a_1m} \cap \Omega_1$ and $\tilde{x}_{j,1} \leq \tilde{x}_{j,1,1}$. Since $\tilde{x}_j \in \Omega_2 \setminus \Omega_1$, then $\emptyset \neq \bar{r}_{x_j,\tilde{x}_j,1} \cap \Omega_2 =: \bar{r}_{x_j,\tilde{x}_j,1} \subseteq \bar{r}_{x_j,\tilde{x}_j} \subseteq \Omega$, where $\tilde{x}_{j,1} \in \bar{r}_{x_0-a_1m,x_0+a_1m} \cap \Omega_2$ and $\tilde{x}_j \leq \tilde{x}_{j,1,1}$. Clearly, $\tilde{x}_{j,1,1} \in \Omega_1 \cap \Omega_2$ since $\Omega_2$ is relatively closed in $\Omega$ and $\tilde{x}_j \in \Omega_1 \cap \Omega_2$ since $\Omega_1$ is relatively closed in $\Omega$. Then, $\tilde{x}_j \in \bar{r}_{x_j,\tilde{x}_j} \cap \Omega_1 = \bar{r}_{x_j,\tilde{x}_j,1} \subseteq \bar{r}_{x_j,\tilde{x}_j} \subseteq \bar{r}_{x_j,\tilde{x}_j,1}$ and $\tilde{x}_j \leq \tilde{x}_{j,1,1}$. This implies that $\bar{r}_{x_j,\tilde{x}_j,1} \subseteq \Omega_1 \setminus \Omega_2$. Let $\tilde{x}_{j,2} = \tilde{x}_j$ and $\tilde{x}_{j,2} \in \bar{r}_{x_j,\tilde{x}_j,1}$ be such that $\pi_l(\tilde{x}_{j,2}) = \{ \pi_l(\tilde{x}_{j,2}) \mid \pi_l(\tilde{x}_{j,2}) = \pi_l(\tilde{x}_j), \forall l \in \bar{J} \}$. Then, $\tilde{x}_{j,2} \leq \tilde{x}_{j,2}$. Clearly, $r_{x_j,\tilde{x}_j,1} \cap r_{\tilde{x}_{j,2},\tilde{x}_j,2} = \emptyset$ and $r_{\tilde{x}_{j,2},\tilde{x}_j,2} \subseteq \bar{r}_{x_0-a_1m,x_0+a_1m}$. Since $r_{x_j,\tilde{x}_j,1} \cap r_{\tilde{x}_{j,2},\tilde{x}_j,2} = \emptyset$, then $r_{x_j,\tilde{x}_j,1} \cap r_{\tilde{x}_{j,2},\tilde{x}_j,2} = \emptyset$. These $\tilde{x}_{j,1}, \tilde{x}_{j,2}, \tilde{x}_{j,1,1}$, and $\tilde{x}_{j,2}$ satisfy the conditions above.

Then, $\sum_{j=1}^n \mu_B(r_{\tilde{x}_j,\tilde{x}_j}) = \sum_{j=1}^n \sum_{k=1}^m \mu_B(r_{\tilde{x}_j,\tilde{x}_j,k}) < \delta(\epsilon)$ and $\sum_{j=1}^n \mu_B(r_{\tilde{x}_j,\tilde{x}_j}) < \delta(\epsilon/2), k = 1, 2$. This implies that $\sum_{j=1}^n \mu_B(r_{\tilde{x}_j,\tilde{x}_j}) < \delta(\epsilon/2), k = 1, 2$. This implies that $\sum_{j=1}^n \left\| \Delta F(r_{\tilde{x}_j,\tilde{x}_j}) \right\| = \sum_{j=1}^n \left\| \Delta F(r_{\tilde{x}_j,\tilde{x}_j}) \right\| < \epsilon$, where the first equality follows from Proposition 12.46. Hence, $F$ satisfies (ii) of Definition 12.58. Therefore, $F$ is absolutely continuous at $x_0$. This subcase is proved.

In both subcases, $F$ is absolutely continuous at $x_0$. This completes the induction process.

By the arbitrariness of $x_0$, $F$ is absolutely continuous. This completes the proof for Case 2.

In both cases, $F$ is absolutely continuous. This completes the proof of the proposition. \qed

**Definition 12.63** Let $X$ be a normed linear space over $K_1$, $Y$ be a normed linear space over $K_2$, $U \subseteq X$, $f : U \to Y$, and $x_0 \in U$; $f$ is said to be locally Lipschitz at $x_0$ (with Lipschitz constant $L_0 \in [0, \infty) \subseteq R$) if $\exists \delta \in (0, \infty) \subseteq R, \forall x_1, x_2 \in U \cap B_X(x_0, \delta)$, we have $\| f(x_1) - f(x_2) \| \leq L_0 \| x_1 - x_2 \|$. $f$ is said to be locally Lipschitz on $U$ if $f$ is locally Lipschitz at $x$, $\forall x \in U$. $f$ is said to be Lipschitz on $U$ (with Lipschitz constant $L \in [0, \infty) \subseteq R$) if $\forall x_1, x_2 \in U$, we have $\| f(x_1) - f(x_2) \| \leq L \| x_1 - x_2 \|$. Clearly, if $f$ is locally Lipschitz at $x_0$, then $f$ is continuous at $x_0$; if $f$ is Lipschitz on $U$, then it is uniformly continuous; if $f$ is Lipschitz on $U$, then it is locally Lipschitz on $U$. \qed
Definition 12.64 Let $\mathcal{X}$ be a normed linear space over $K_1$, $Y$ be a set, $Z$ be a normed linear space over $K_2$, $U \subseteq \mathcal{X}$, $f : U \times Y \to Z$, and $x_0 \in U$. $f$ is said to be locally Lipschitz at $x_0$ (with Lipschitz constant $L_0 \in [0, \infty) \subseteq \mathbb{R}$) uniformly over $Y$ if $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$, $\forall x_1, x_2 \in U \cap B_X(x_0, \delta)$, $\forall y \in Y$, we have $\| f(x_1, y) - f(x_2, y) \| \leq L_0 \| x_1 - x_2 \|$. $f$ is said to be locally Lipschitz on $U$ uniformly over $Y$ if it is locally Lipschitz at $x$ uniformly over $Y$, $\forall x \in U$. $f$ is said to be Lipschitz on $U$ (with Lipschitz constant $L \in [0, \infty) \subseteq \mathbb{R}$) uniformly over $Y$ if $\forall x_1, x_2 \in U$, $\forall y \in Y$, we have $\| f(x_1, y) - f(x_2, y) \| \leq L \| x_1 - x_2 \|$. Clearly, if $f$ is Lipschitz on $U$ uniformly over $Y$, then it is locally Lipschitz on $U$ uniformly over $Y$.

Proposition 12.65 Let $\mathcal{X}$ be a normed linear space over $K_1$, $Y$ be a set, $Z$ be a normed linear space over $K_2$, $U \subseteq \mathcal{X}$ be compact, $f : U \times Y \to Z$ be locally Lipschitz on $U$ uniformly over $Y$. Assume that $\forall x \in U$, $\sup_{y \in Y} \| f(x, y) \| < +\infty$. Then, $f$ is Lipschitz on $U$ uniformly over $Y$.

Proof \hspace{1cm} $\forall x \in U$, by the local Lipschitz continuity of $f$, $\exists L_x \in [0, \infty) \subseteq \mathbb{R}$, $\exists \delta_x \in (0, \infty) \subseteq \mathbb{R}$, $\forall x_1, x_2 \in O_x := U \cap B_X(x, \delta_x)$, $\forall y \in Y$, we have $\| f(x_1, y) - f(x_2, y) \| \leq L_x \| x_1 - x_2 \|$. Let $O_x := B_X(x, \delta_x/2)$. Then, $U \subseteq \bigcup_{x \in U} O_x$, which is an open covering of $U$. By the compactness of $U$, there exists a finite set $X_N \subseteq U$ such that $U \subseteq \bigcup_{x \in X_N} O_x$. Let $L_M := \max \{ \max_{x \in X_N} L_x, \frac{2}{\min_{x \in X_N} \delta_x} \max_{\hat{x}_1, \hat{x}_2 \in X_N} (L_{\hat{x}_1} \delta_{\hat{x}_1}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 + \sup_{y \in Y} \| f(\hat{x}_1, y) \| + \sup_{y \in Y} \| f(\hat{x}_2, y) \|), 0 \} \in [0, \infty) \subseteq \mathbb{R}$. $\forall x_1, x_2 \in U$, $\forall y \in Y$, $\exists \hat{x}_1 \in X_N$ such that $x_1 \in O_{\hat{x}_1}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_2 \in O_{\hat{x}_1}$; Case 2: $x_2 \notin O_{\hat{x}_1}$. Case 1: $x_2 \in O_{\hat{x}_1}$. Then, $\| f(x_1, y) - f(x_2, y) \| \leq L_{\hat{x}_1} \| x_1 - x_2 \| \leq L_M \| x_1 - x_2 \|$. Case 2: $x_2 \notin O_{\hat{x}_1}$. Then, $\| x_2 - \hat{x}_1 \| \leq \delta_{\hat{x}_1}$ and $\| x_2 - x_1 \| \geq \| x_2 - \hat{x}_1 \| - \| x_1 - \hat{x}_1 \| > \delta_{\hat{x}_1}/2 \geq \frac{\min_{x \in X_N} \delta_x}{2} > 0$. $\exists \hat{x}_2 \in X_N$ such that $x_2 \in O_{\hat{x}_2}$. Then, $\| f(x_1, y) - f(x_2, y) \| \leq \| f(x_1, y) - f(x_2, y) \| + \| f(\hat{x}_1, y) - f(\hat{x}_2, y) \| + \| f(\hat{x}_2, y) - f(x_2, y) \| \leq L_{\hat{x}_1} \| x_1 - \hat{x}_1 \| + \| f(\hat{x}_1, y) \| + \| f(\hat{x}_2, y) \| + L_{\hat{x}_2} \| x_2 - \hat{x}_2 \| \leq \sup_{y \in Y} \| f(\hat{x}_1, y) \| + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 \leq L_M \frac{\min_{x \in X_N} \delta_x}{2} \leq L_M \| x_1 - x_2 \|$. Hence, in both cases, we have $\| f(x_1, y) - f(x_2, y) \| \leq L_M \| x_1 - x_2 \|$. Then, $f$ is Lipschitz on $U$ uniformly over $Y$. This completes the proof of the proposition. □

Proposition 12.66 Let $\mathcal{X}$ be a normed linear space over $K$, $Y$ be a set, $Z$ be a normed linear space over $K$, $U \subseteq \mathcal{X}$ be convex, $f : U \times Y \to Z$ be partial differentiable with respect to $x$. Assume that $\exists L \in [0, \infty) \subseteq \mathbb{R}$ such that $\| \frac{\partial f}{\partial x}(x, y) \| \leq L$, $\forall (x, y) \in U \times Y$. Then, $f$ is Lipschitz on $U$ with Lipschitz constant $L$ uniformly over $Y$. 

Proof. \( \forall x_1, x_2 \in U, \forall y \in Y, \) by Mean Value Theorem 9.23, \( \exists t \in (0, 1) \subset \mathbb{R} \) such that \( \|f(x_1, y) - f(x_2, y)\| \leq \left\| \frac{\partial f}{\partial x}(x_1 + (1-t)x_2, y) \right\| \|x_1 - x_2\| \). Hence, \( f \) is Lipschitz on \( U \) with Lipschitz constant \( L \).

This completes the proof of the proposition.

\[ \Box \]

**Proposition 12.67** Let \( I \subseteq \mathbb{R} \) be an interval, \( Y \) be a normed linear space over \( \mathbb{K}_1 \), \( Z \) be a normed linear space over \( \mathbb{K}_2 \), \( U \subseteq Y \), \( f : I \rightarrow U \) be absolutely continuous at \( x_0 \in I \), \( g : U \rightarrow Z \) be locally Lipschitz at \( y_0 := f(x_0) \in U \). Then, \( g \circ f \) is absolutely continuous at \( x_0 \).

Proof. By \( g \) being locally Lipschitz at \( y_0 \), then \( \exists L_0 \in [0, \infty) \subseteq \mathbb{R} \), \( \exists \delta_y \in (0, \infty) \subseteq \mathbb{R} \), \( \forall y_1, y_2 \in U \cap B_y(y_0, \delta_y) \), we have \( \|g(y_1) - g(y_2)\| \leq L_0 \|y_1 - y_2\| \). Since \( f \) is absolutely continuous at \( x_0 \), then it is continuous at \( x_0 \) by Proposition 12.59. Then, \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \) such that \( \forall x \in B_R(x_0, \delta_0) \cap I \), we have \( f(x) \in U \cap B_y(y_0, \delta_y) \). By the absolute continuity of \( f \) at \( x_0 \), \( \exists \delta_f \in (0, \delta_0) \subseteq \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists \delta_f(\epsilon) \in (0, \delta_f) \in \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall x_0 - a \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq x_0 + a \) with \( x_i, x_i \in I \), \( i = 1, \ldots, n \), and \( \sum_{i=1}^{n} (x_i - x_i) < \delta_f(\epsilon) \), we have \( \sum_{i=1}^{n} \|f(x_i) - f(x_i)\| < \epsilon \).

For the function \( g \circ f \), take \( a = a_f \in (0, \infty) \subseteq \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), let \( \delta(\epsilon) = \delta_f(\epsilon/(1 + L_0)) \in (0, a) \subseteq \mathbb{R} \). \( \forall n \in \mathbb{N} \), \( \forall x_0 - a \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq x_0 + a \) with \( x_i, x_i \in I \), \( i = 1, \ldots, n \), and \( \sum_{i=1}^{n} (x_i - x_i) < \delta(\epsilon) \), we have \( \sum_{i=1}^{n} \|f(x_i) \cap B_y(y_0, \delta_y) \) and \( \sum_{i=1}^{n} \|g \circ f(x_i) - \|f(x_i) - f(x_i)\| \). Hence, \( g \circ f \) is absolutely continuous at \( x_0 \). This completes the proof of the proposition.

\[ \Box \]

**Proposition 12.68** Let \( I \subseteq \mathbb{R} \) be an interval, \( Y \) be a normed linear space over \( \mathbb{K}_1 \), \( Z \) be a normed linear space over \( \mathbb{K}_2 \), \( f : I \rightarrow Y \) and \( g : I \rightarrow Y \) be absolutely continuous at \( x_0 \in I \), \( A : I \rightarrow B(Y, Z) \) be absolutely continuous at \( x_0 \), and \( h : I \rightarrow Z \) be defined by \( h(x) = A(x)f(x) \), \( \forall x \in I \). Then, \( f + g \) and \( h \) are absolutely continuous at \( x_0 \).

Proof. By Propositions 12.61, 9.14, 12.67, and 12.66, we have that \( f + g \) is absolutely continuous at \( x_0 \).

By Propositions 12.61, 9.17, 12.67, and 12.66, we have that \( h \) is absolutely continuous at \( x_0 \).

This completes the proof of the proposition.

\[ \Box \]

**Proposition 12.69** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in B_B(\mathbb{R}^m) \) be a rectangle, \( Y \) be a normed linear space over \( \mathbb{K}_1 \), \( Z \) be a normed linear space over \( \mathbb{K}_2 \), \( F : \Omega \rightarrow Y \) be absolutely continuous at \( x_0 \in \Omega \), and \( A \in B(Y, Z) \). Then, \( h : \Omega \rightarrow Z \), defined by \( h(x) = Af(x) \), \( \forall x \in \Omega \), is absolutely continuous at \( x_0 \).

Proof. We will prove this result using mathematical induction on \( m \).

1° \( m = 1 \). Clearly, the result follows from Proposition 12.68.
2° Assume that the result holds for \( m \leq k \in \mathbb{N} \).

3° Consider the case when \( m = k + 1 \). We will prove that \( h \) is absolutely continuous at \( x_0 \) by Definition 12.58. Fix any \( i_0 \in J := \{1, \ldots, m\} \), let \( D_{i_0}, \pi_{i_0}, \) and \( F_{i_0} \) be defined as in (i) of Definition 12.58. Let \( h_{i_0} : D_{i_0} \to \mathbb{R} \) be defined by \( h_{i_0}(x) = h(x) = A F(x) = A F_{i_0}(x) \), \( \forall x \in D_{i_0} \subseteq \mathbb{R}^k \), \( \forall \bar{x} \in \Omega \) with \( \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \) and \( \pi_{i_0}(\bar{x}) = x \). By the assumption, \( F \) is absolutely continuous at \( x_0 \), then \( F_{i_0} \) is absolutely continuous at \( \pi_{i_0}(x_0) \). Thus, (i) of Definition 12.58 holds.

Since \( F \) is absolutely continuous at \( x_0 \), by (ii) of Definition 12.58, \( \exists 0 \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall (x_i)_{i=1}^n \subseteq \mathbb{R}, (r_i)_{i=1}^n \subseteq \mathbb{R}, (x_i = x) \subseteq \Omega \) with \( \bar{x} = x = x \), \( (r_i = r_0) \subseteq \mathbb{R} \), \( (x_i = x_i) \subseteq \mathbb{R} \), \( (r_i = r_i) \subseteq \mathbb{R} \), \( (x_i = x_i) \subseteq \mathbb{R} \), \( (r_i = r_i) \subseteq \mathbb{R} \). If \( \sum_{i=1}^n \| x_i - x \| \leq \delta(\epsilon) \), we have \( \| F(x_i) - F(x) \| \leq \epsilon \). It is easy to see that for \( \delta(\epsilon) \). Hence, (ii) of Definition 12.58 holds.

Thus, \( h \) is absolutely continuous at \( x_0 \).

This completes the induction process and therefore the proof of the proposition. \( \square \)

**Proposition 12.70** Let \( X \) be a normed linear space over \( \mathbb{K} \), \( Y \) be a set, \( Z \) be a normed linear space over \( \mathbb{K} \), \( U \subseteq X \) be endowed with subset topology \( O \), \( f : U \times Y \to Z \), \( g : U \to \mathbb{R} \) be continuous, \( (c_i)_{i \in \mathbb{Z}} \subseteq \mathbb{R} \) be a sequence of strictly increasing real numbers with \( \lim_{i \to -\infty} c_i = -\infty \) and \( \lim_{i \to \infty} c_i = \infty \), \( U_i := \{ x \in U \mid c_i \leq g(x) \leq c_{i+1} \} \), \( \forall i \in \mathbb{Z} \). Assume that

(i) \( U \) is locally convex according to Definition 7.124;

(ii) \( f|_{U_i \times Y} \) is locally Lipschitz on \( U_i \) uniformly over \( Y \), \( \forall i \in \mathbb{Z} \).

Then, \( f \) is locally Lipschitz on \( U \) uniformly over \( Y \).

**Proof** Fix any \( x_0 \in U = \bigcup_{i \in \mathbb{Z}} U_i \). Then, \( g(x_0) \in \mathbb{R} \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \exists i \in \mathbb{Z} \) such that \( c_i < g(x_0) < c_{i+1} \); Case 2: \( \exists i \in \mathbb{Z} \) such that \( g(x_0) = c_i \). Case 1: \( \exists i \in \mathbb{Z} \) such that \( c_i < g(x_0) < c_{i+1} \). Then, \( x_0 \in U_i \). Since \( g \) is continuous, then \( \exists \delta_i \in (0, \infty) \subseteq \mathbb{R} \) such that \( B_\delta(x_0, \delta_i) \cap U \subseteq U_i \), \( \forall x_0, x_2 \in B_\delta(x_0, \delta) \subseteq B_\delta(x_0, \delta_i) \subseteq U_i \), \( \forall y \in Y \), we have \( \| f(x_1, y) - f(x_2, y) \| \leq L_1 \| x_1 - x_2 \| \). Hence, \( f \) is locally Lipschitz at \( x_0 \) uniformly over \( Y \).

Case 2: \( \exists i \in \mathbb{Z} \) such that \( g(x_0) = c_i \). Then, \( x_0 \in U_{i-1} \cap U_i \). By (i) and the continuity of \( g \), \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \) such that \( U \cap B_\delta(x_0, \delta_0) \subseteq U_{i-1} \cap U_i \) is convex. By \( f|_{U_{i-1} \times Y} \) being locally Lipschitz on \( U_{i-1} \) uniformly over \( Y \), we have \( f|_{U_{i-1} \times Y} \) is locally Lipschitz at \( x_0 \) uniformly over \( Y \), \( j = i - 1, i \).
By Definition 12.64, \( \exists L_j \in [0, \infty) \subset \mathbb{R}, \ \exists \delta_j \in (0, \delta_{x_0}) \subset \mathbb{R}, \ \forall x_1, x_2 \in B_{Y Y}(x_0, \delta_j) \cap U_j, \ \forall y \in Y, \) we have \( \| f(x_1, y) - f(x_2, y) \| \leq L_j \| x_1 - x_2 \|, \) \( j = i - 1, i. \) Take \( L := L_{i-1} \lor L_i \in [0, \infty) \subset \mathbb{R} \) and \( \delta := \delta_1 \land \delta_2 \in (0, \delta_{x_0}) \subset \mathbb{R}. \) \( \forall x_1, x_2 \in B_Y(x_0, \delta) \cap U =: D, \ \forall y \in Y, \) by Proposition 6.40, \( D \) is convex. If \( \exists j \in \{i-1, i\} \) such that \( x_1, x_2 \in U_j, \) then \( \| f(x_1, y) - f(x_2, y) \| \leq L_j \| x_1 - x_2 \| \leq L \| x_1 - x_2 \|. \) On the other hand, without loss of generality, assume that \( x_1 \in U_{i-1} \) and \( x_2 \in U_i. \) Consider the line segment connecting \( x_1 \) and \( x_2. \)

Claim 12.70.1 \( \exists t_0 \in [0, 1] \subset \mathbb{R} \) such that \( x_{t_0} := t_0 x_1 + (1 - t_0) x_2 \in U_{i-1} \cap U_i \cap D. \)

Proof of claim: Since \( D \) is convex, then \( x_t := tx_1 + (1 - t)x_2 \in D \subseteq U_{i-1} \cup U_i, \ \forall t \in [0, 1] \subset \mathbb{R}. \) Define \( t_0 := \sup \{ t \in [0, 1] \subset \mathbb{R} \ | \ x_t \in U_i \} =: \sup A \)

Clearly, \( 0 \in A \) and \( 0 \leq t_0 \leq 1. \) Suppose \( t_0 \notin A. \) This leads to \( t_0 \in (0, 1) \subset \mathbb{R}. \) Then, \( \exists (t_n)_{n=1}^\infty \subseteq (0, t_0) \cap A \) such that \( \lim_{n \to \infty} t_n = t_0. \ \forall n \in \mathbb{N}, \) by \( t_n \in A, \) we have \( g(x_{t_n}) \geq c_i. \) Then, \( g(x_{t_0}) = \lim_{n \to \infty} g(x_{t_n}) \geq c_i, x_{t_0} \in U_i, \) and \( t_0 \in A. \) This is a contradiction. Hence, \( t_0 \in A \) and \( x_{t_0} \in U_2. \) If \( t_0 = 1, \) then \( x_{t_0} = x_1 \in U_{i-1} \) and \( x_{t_0} \in U_{i-1} \cap U_i. \) Hence, the claim holds. On the other hand, if \( t_0 < 1, \) then \( \forall t \in (t_0, 1) \subset \mathbb{R}, t \notin A \) and \( x_t \in U_{i-1}. \) This implies that \( g(x_t) \leq c_i, \forall t \in (t_0, 1) \subset \mathbb{R}. \) Then, \( g(x_{t_0}) = \lim_{t \to t_0^+} g(x_t) \leq c_i \) and \( x_{t_0} \in U_{i-1}. \) Hence, \( x_{t_0} \in U_{i-1} \cap U_i. \) The claim holds as well. This completes the proof of the claim.

By Claim 12.70.1, \( \| f(x_1, y) - f(x_2, y) \| \leq \| f(x_1, y) - f(x_{t_0}, y) \| + \| f(x_{t_0}, y) - f(x_2, y) \| \leq L_{i-1} \| x_1 - x_{t_0} \| + L_i \| x_2 - x_{t_0} \| = L_{i-1} (1 - t_0) \| x_1 - x_2 \| + L_i \| x_2 - x_{t_0} \| \leq L \| x_1 - x_2 \|. \)

In both cases, we have that \( \| f(x_1, y) - f(x_2, y) \| \leq L \| x_1 - x_2 \|. \) Then, \( f \) is locally Lipschitz at \( x_0 \) uniformly over \( Y. \)

In both cases, we have that \( f \) is locally Lipschitz at \( x_0 \) uniformly over \( Y. \) By the arbitrariness of \( x_0, f \) is locally Lipschitz on \( U \) uniformly over \( Y. \) This completes the proof of the proposition.

Definition 12.71 Let \( a, b \in \mathbb{R}, I := [a, b] \subset \mathbb{R} \) be the semi-open interval with \( a \) and \( b \) as end points, \( I := ([a, b], \mathcal{B}, \mu) \) be the finite metric measure subspace of \( \mathbb{R}, \) \( Y \) be a normed linear space, and \( f : I \to Y \) be \( \mathcal{B} \)-measurable. We will write

\[
\int_a^b f(x) \, dx := \int_a^b f \, d\mu_B := \begin{cases} \int_a^b f \, d\mu_B & \text{if } b \geq a \\ -\int_b^a f \, d\mu_B & \text{if } b < a \end{cases}
\]

whenever the right-hand-side makes sense.

By Lemma 11.73, the interval \( I \) may be \([a, b], \mathcal{B}, \mu\), or \([a, b), \mathcal{B}, \mu\), or \((a, b), \mathcal{B}, \mu\), which will not change the value of the integral.
Fact 12.72 Let \( a, b, c \in \mathbb{R} \), \( I := [a, b] \) \( \subset \mathbb{R} \), \( I := ((I, |\cdot|), \mathcal{B}, \mu) \) be the finite metric measure subspace of \( \mathbb{R} \), \( \mathcal{Y} \) be a separable Banach space, and \( f : I \to \mathcal{Y} \) be absolutely integrable over \( I \). Then, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx = \int_a^b f(x) \, dx = \int_a^b f(x) \, dx \in \mathcal{Y} \).

**Proof** We will distinguish six exhaustive and mutually exclusive cases:

1. Case 1: \( a \leq c \leq b \); Case 2: \( a \leq b \leq c \); Case 3: \( b \leq a \leq c \); Case 4: \( b \leq c \leq a \); Case 5: \( c \leq a \leq b \); Case 6: \( b \leq c \leq a \). 

Let \( J_1 := (a, c) \subset \mathbb{R} \), \( J_2 := (c, b) \subset \mathbb{R} \). By Proposition 11.92, \( \int_a^b f(x) \, dx = \int_a^b f(x) \, dx + \int_a^c f(x) \, dx \).

Then, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx - \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_a^c f(x) \, dx \).

Case 3: \( b \leq a \leq c \).

Then, \( \int_b^c f(x) \, dx = \int_a^b f(x) \, dx - \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_a^c f(x) \, dx \).

Case 4: \( b \leq c \leq a \).

Then, \( \int_b^a f(x) \, dx = \int_b^c f(x) \, dx + \int_b^a f(x) \, dx \).

Case 6: \( c \leq b \leq a \).

Then, \( \int_c^b f(x) \, dx = \int_c^a f(x) \, dx + \int_c^b f(x) \, dx \).

This completes the proof of the fact. \( \square \)

Proposition 12.73 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) be a rectangle with the subset topology \( \mathcal{O} \), \( ((P(\Omega), |\cdot|), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathcal{Y} \) be a normed linear space over \( \mathbb{K} \), and \( F : \Omega \to \mathcal{Y} \) be absolutely continuous. Then, the following statements hold.

(i) \( F \) is of locally bounded variation.

(ii) If, in addition, \( \mathcal{Y} \) is a Banach space over \( \mathbb{K} \), then there exists a unique \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space \( (P(\Omega), \mathcal{B}, \nu) \) such that \( \nu \) admits \( F \) as a cumulative distribution function. Furthermore, \( \mathcal{P} \circ \nu(r_{x_1, x_2}) = T_F(r_{x_1, x_2}) \), \( \forall x_1, x_2 \in \Omega \) with \( x_1 \equiv x_2 \), and \( \mathcal{P} \circ \nu(P(\Omega)) = T_F \) and \( \mathcal{P} \circ \nu \ll \mu \).

**Proof** (i) By Proposition 12.52, \( \Omega \) is a region and \( P(\Omega) \in \mathcal{B}_B(\mathbb{R}^m) \) is a rectangle. By Proposition 12.59, \( F \) is continuous. Then, (i) of Definition 12.41 holds.
∀x_1, x_2 ∈ Ω with x_1 ≤ x_2, consider the closed rectangle r_{x_1, x_2} ⊆ Ω. ∀x_0 ∈ r_{x_1, x_2}, F is absolutely continuous at x_0. By (ii) of Definition 12.58, ∃a_{x_0} ∈ (0, ∞) ⊂ R, ∀ε ∈ (0, ∞) ⊂ R, ∃δ_{x_0}(ε) ∈ (0, ∞) ⊂ R, ∀n ∈ N, ∀i \in N, (\tilde{x}_i)_{i=1}^n ⊆ r_{x_0-a_{x_0}1_m, x_0+a_{x_0}1_m} \cap Ω with \tilde{x}_i ≤ x_0, ∀i \in \{1, \ldots, n\}, being pairwise disjoint, and \sum_{i=1}^n\mu(B_m(\tilde{x}_i, 1)) < δ_{x_0}(ε), we have \sum_{i=1}^n\|\Delta F(\tilde{x}_i, 1)\| < ε. Then, \tilde{r}_{x_1, x_2} ⊆ \bigcup_{x \in \tilde{r}_{x_1, x_2}} r_{x-a_{x}1_m, x+a_{x}1_m}. By the compactness of \tilde{r}_{x_1, x_2}, \exists a finite set N ⊆ \tilde{r}_{x_1, x_2} such that \tilde{r}_{x_1, x_2} ⊆ \bigcup_{x \in N} r_{x-a_{x}1_m, x+a_{x}1_m}. Let a_0 := 1 \in R, a := \max_{x \in N} a_x ∈ (0, ∞) ⊂ R, and δ := \min_{x \in N} δ_x(ε) ∈ (0, ∞) ⊂ R. ∀n ∈ N, \forall (\tilde{x}_i)_{i=1}^n ⊆ \tilde{r}_{x_1, x_2} with \tilde{x}_i ≤ x_0, ∀i \in \{1, \ldots, n\}, \bigcup_{i=1}^n r_{\tilde{x}_i, \tilde{x}_i} = r_{x_1, x_2}, and sets in the union being pairwise disjoint, we have, by Proposition 12.46,

$$\sum_{i=1}^n\|\Delta F(\tilde{x}_i, 1)\| = \sum_{i=1}^n\|\sum_{x \in N}\sum_{j=1}^{n_x}\Delta F(\tilde{x}_i, x, j, 1)\| ≤ \sum_{x \in N}\sum_{i=1}^n\sum_{j=1}^{n_x}\|\Delta F(\tilde{x}_i, x, j, 1)\|$$

where n_x ∈ Z, ∀x ∈ N, (\tilde{x}_i, x, j)_{i \in \{1, \ldots, n_x\}, x \in N, j \in \{1, \ldots, n_x\}, \tilde{x}_i \subseteq \tilde{r}_{x_1, x_2}, ∀x ∈ N, ∀j \in \{1, \ldots, n_x\}, \bigcup_{x \in N} \bigcup_{j=1}^{n_x} r_{\tilde{x}_i, x, j, 1} = r_{x-a_{x}1_m, x+a_{x}1_m}, ∀x \in N, \forall i \in \{1, \ldots, n\}. It is easy to see that n_x ≤ 1 + \lfloor 2^{n_x}/\delta_x(ε) \rfloor. Then, \sum_{i=1}^n\|\Delta F(\tilde{x}_i, x, j, 1)\| ≤ \sum_{x \in N} \sum_{i=1}^n\|\Delta F(\tilde{x}_i, x, j, 1)\| ≤ \sum_{x \in N} \sum_{i=1}^n\|\Delta F(\tilde{x}_i, x, j, 1)\| < \sum_{x \in N} \sum_{i=1}^n\|\Delta F(\tilde{x}_i, x, j, 1)\| < \text{card}(N)(1 + \lfloor 2^{n_x}/\delta_x(ε) \rfloor). Hence, F satisfies (ii) of Definition 12.41.

∀x_1, x_2 ∈ Ω with x_1 < x_2, define \tilde{G} : \tilde{r}_{x_1, x_2} → (0, ∞) ⊂ R by \tilde{G}(x) = T_F(\tilde{x}_i, x), ∀x ∈ \tilde{r}_{x_1, x_2}. We will show that \tilde{G} is continuous on the right. Fix any x_0 ∈ \tilde{r}_{x_1, x_2}. Consider the sets S_J := \{x ∈ \tilde{r}_{x_1, x_2} | π_j(x) = π_j(x_0), ∀i \in J, ∃J ⊆ \{1, \ldots, m\}, J ≠ ∅. Fix any J ⊆ \{1, \ldots, m\} with J ≠ ∅. Clearly, S_J is compact. ∀x ∈ S_J, by (ii) of Definition 12.58, ∃a_x ∈ (0, ∞) ⊂ R, ∀ε ∈ (0, ∞) ⊂ R, ∃δ_x(ε) ∈ (0, ∞) ⊂ R, ∀n ∈ N, ∀i \in N, (\tilde{x}_i)_{i=1}^n \subseteq r_{x-a_{x}1_m, x+a_{x}1_m} \cap \tilde{r}_{x_1, x_2} with \tilde{x}_i ≤ x_0, ∀i \in \{1, \ldots, n\}, (\tilde{x}_i, \tilde{x}_i)_{i=1}^n being pairwise disjoint, and \sum_{i=1}^n\mu(B_m(\tilde{x}_i, 1)) < δ_x(ε), we have \sum_{i=1}^n\|\Delta F(\tilde{x}_i, 1)\| < ε. Then, S_J ⊆ \bigcup_{x ∈ S_J} r_{x-a_{x}1_m, x+a_{x}1_m}. By the compactness of S_J, \exists a finite set N_J ⊆ S_J such that S_J ⊆ \bigcup_{x ∈ N_J} r_{x-a_{x}1_m, x+a_{x}1_m}. Let N := \bigcup_{J ⊆ \tilde{J}, J ≠ ∅} N_J. Let a := \max_{x \in N} a_x ∈ (0, ∞) ⊂ R and a := \min_{x \in N} a_x ∈ (0, ∞) ⊂ R. ∀ε ∈ (0, ∞) ⊂ R, let δ := \min_{x \in N} δ_x(ε) ∈ (0, ∞) ⊂ R. Let δ be the minimum positive real root of \prod_{i=1}^m(π_i(x_0) - π_i(x_1) + δ) - \prod_{i=1}^m(π_i(x_0) - π_i(x_1)) = δ/2 and δ := δ ^ a ∈ (0, ∞) ⊂ R. ∀x ∈ B_{R_m}(x_0, a) \cap \tilde{r}_{x_1, x_2} with x_0 = \tilde{x}, we have |\tilde{G}(x) - G(x_0)| = |T_F(\tilde{x}_i, x) - T_F(\tilde{x}_i, x_0)| = \sum_{J ⊆ \tilde{J}, J ≠ ∅} |T_F(\tilde{x}_i, x) - T_F(\tilde{x}_i, x_0)|, where \tilde{x}_i \in \tilde{r}_{x_1, x_2} is such that π_i(\tilde{x}_i) = π_i(x_0), ∀i \in \tilde{J}, π_i(\tilde{x}_i) = π_i(x_1),
12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

619

∀i ∈ J¯ \ J; and x̂J ∈ rx0 ,x̄ is such that π i (x̂J ) = π i (x̄), ∀i ∈ J,
π i (x̂J ) = π i (x0 ), ∀i ∈ J¯ \ J; and the secondSequality follows from Proposition 12.47. ∀J ⊆ J¯ with J 6= ∅, rx̌J ,x̂J ⊆ x∈NJ r◦x−ax 1m ,x+ax 1m implies
S
that rx̌J ,x̂J = x∈NJ (r◦x−ax 1m ,x+ax 1m ∩ rx̌J ,x̂J ). Then, we have Ḡ(x̄) −
P
P
Ḡ(xP
¯ J6=∅ TF (rx̌J ,x̂J ∩ rx−ax 1m ,x+a
0) ≤
x∈N
J⊆J,
P x 1m ). Note that, ∀x ∈
N , J⊆J¯, J6=∅ µBm (rx̌J ,x̂J ∩ rx−ax 1m ,x+ax 1m ) ≤ J⊆J,
¯ J6=∅ µBm (rx̌J ,x̂J ) ≤
Qm
Qm
(π
(x
)
−
π
(x
))
≤ δ/2 < δ. Then,
(π
(x
)
−
π
(x
)
+
δ̄)
−
i
0
i
1
i
0
i
1
i=1
Pi=1
)
<
ǫ/(1
+
card(N
)). Hence, Ḡ(x̄)−
∩r
T
(r
¯
x−a
1
,x+a
1
F
x̌
,x̂
x m
x m
J
J
J⊆J, J6=∅P
Ḡ(x0 ) < x∈N ǫ/(1 + card(N )) < ǫ. This shows that Ḡ is continuous on
the right at x0 . By the arbitrariness of x0 , we have Ḡ is continuous on
the right. Hence, (iii) of Definition 12.41 holds. Therefore, F is of locally
bounded variation. This completes the proof of (i).
(ii) By Theorem 12.50, there exists a unique σ-finite Y-valued measure
space (P(Ω), B, ν) such that F is a cumulative distribution function of ν,
⋖
P◦ν(rx1 ,x2 ) = TF ( rx1 ,x2 ), ∀x1 , x2 ∈ Ω with x1 = x2 , and P◦ν(P(Ω)) = TF .
All we need to show is P ◦ ν ≪ µ. First, we consider the special
⋖
case Ω = rx1 ,x2 ⊆ IRm , where x1 , x2 ∈ IRm with x1 = x2 . ∀x ∈ rx1 ,x2 ,
F being absolutely continuous at x implies that ∃λx ∈ (0, ∞) ⊂ IR,
∀ǫ ∈ (0, ∞) ⊂ IR, ∃δx (ǫ) ∈ (0, λx ] ⊂ IR, ∀n ∈ IN, ∀ ( x̂i )ni=1 , ( x̄i )ni=1 ⊆
⋖
( rx̂i ,x̄i )ni=1 being pairΩ ∩ rx−λx 1m ,x+λx 1P
m with x̂i = x̄i , ∀i = 1, . . . , n,P
n
n
wise disjoint, and
i=1 k ∆F (rx̂i ,x̄i ) k < ǫ.
S i=1 µ(r◦x̂i ,x̄i ) < δx (ǫ), we have
Then, rx1 ,x2 ⊆ x∈rx ,x rx−λx 1m ,x+λx 1m ⊂ IRm . By the compactness of
1 2
a finite set XN := {x̃1 , . . . , x̃p } ⊆ rx1 ,x2 , where p ∈ IN,
rx1 ,x2 , there exists S
such that rx1 ,x2 ⊆ pi=1 r◦x̃i −λx̃ 1m ,x̃i +λx̃ 1m ⊂ IRm .
i

i

ǫ
∀E ∈ B with µ(E) = 0, ∀ǫ ∈ (0, ∞) ⊂ IR, let δ(ǫ) := mini∈{1,...,p} δx̃i ( 2p
)
∈ (0, ∞) ⊂ IR. By Proposition 11.23, ∃O ∈ OIRm with E ⊆ O such that
µBm (O\E) < δ(ǫ). Then, µBm (O) = µBm (O\E)+µBm (E) = µBm (O\E) <
δ(ǫ). By Proposition 12.40, O is a region with P(O) = O. By Definition 12.38, there exists N ⊆ IN and ∃ ( x̂i,1 )i∈N , ( x̂i,2 )i∈N ⊆ O such that
S
⋖
rx̂i,1 ,x̂i,2 and the sets in the union
x̂i,1 = x̂i,2 , ∀i ∈ IN, and O = i∈N P
are pairwise disjoint. Then,
µ
(O)
=
i∈N µBm (rx̂i,1 ,x̂i,2 ) < δ(ǫ) and P ◦
P Bm
ν(E) ≤ P◦ν(O∩rx1 ,x2 ) = i∈N P◦ν(rx̂i,1 ,x̂i,2 ∩rx1 ,x2 ). By Definition 12.41,
for any fixed i ∈ N , define rx̃i,1 ,x̃i,2 := rx̂i,1 ,x̂i,2 ∩ rx1 ,x2 with x̃i,j ∈ rx1 ,x2 ,
S
⋖
j = 1, 2, and x̃i,1 = x̃i,2 , we have O ∩ rx1 ,x2 = i∈N rx̃i,1 ,x̃i,2 . By possibly
splitting rectangles into smaller disjoint pieces in the collection, without
loss of generality, assume that ∀i ∈ N , ∃j ∈ {1, . . . , p} such that rx̃i,1 ,x̃i,2
 ⊆
rx̃j −λx̃j 1m ,x̃j +λx̃j 1m . Then, ∀i ∈ N , P ◦ ν(rx̃i,1 ,x̃i,2 ) = TF rx̃i,1 ,x̃i,2 =
Pn
sup n∈Z+ , x̌1,1 ,x̌1,2 ,...,x̌n,1 ,x̌n,2 ∈rx̃ ,x̃ , x̌j,1 ⋖
j=1 ∆F (rx̌j,1 ,x̌j,2 ) .
= x̌j,2 , ∀j∈{1,...,n}
i,1 i,2
Sn
r
=rx̃
,r
∩rx̌
=∅, ∀1≤k<j≤n
j=1 x̌j,1 ,x̌j,2
j,1 ,x̌j,2
i,1 ,x̃i,2 x̌k,1 ,x̌k,2
ni
ni

Then, there exists ni ∈ IN, ( x̌i,j,1 )j=1 , ( x̌i,j,2 )j=1 ⊆ rx̃i,1 ,x̃i,2 such that x̌i,j,1
Sni
⋖
= x̌i,j,2 , ∀j ∈ {1, . . . , ni }, j=1
rx̌i,j,1 ,x̌i,j,2 = rx̃i,1 ,x̃i,2 and the sets in the
Pni
union are pairwise disjoint, and P◦ν(rx̃i,1 ,x̃i,2 ) < j=1
∆F (rx̌i,j,1 ,x̌i,j,2 ) +


Lemma 12.74. Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space and $\nu$ be a measure on $(X, \mathcal{B})$. The following statements hold.

(i) If $\nu \ll \mu$ and $\nu$ is finite, then for any measurable set $E \subseteq X$, we have $\nu(E) < \epsilon$ for some $\epsilon > 0$.

(ii) If $\nu \ll \mu$ and $\nu$ is finite, then for any measurable set $E \subseteq X$, we have $\nu(E) = 0$.

Proof. (i) We will prove the result by an argument of contradiction. Suppose the result does not hold. Then, there exists a set $E \subseteq X$ with $\mu(E) = 0$ and $\nu(E) > 0$. Let $E_n := \bigcap_{j=n}^{\infty} E_j$, $\forall n \in \mathbb{N}$. Then, $\bigcup_{n=1}^{\infty} E_n \supseteq X$. By Proposition 11.5, we have $\nu \ll \mu$ and $\mu(\bigcup_{n=1}^{\infty} E_n) = \mu(X) > \epsilon_0$. Hence, the result holds.

(ii) By the assumption, $\forall E \subseteq X$ with $\mu(E) = 0$, we have $\nu(E) = 0$. By the arbitrariness of $\epsilon$, we have $\nu(E) = 0$. By the arbitrariness of $\epsilon$, we have $\nu(E) = 0$. Thus, the result holds.
Proposition 12.75 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \subseteq \mathcal{B}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathbb{K}$, $\nu$ be a $\sigma$-finite $\mathcal{Y}$-valued measure on the measurable space $(\Omega, \mathcal{B}(\Omega), \nu)$, and $\mathcal{X} := (\Omega, \mathcal{B}(\Omega), \mu)$ be the metric measure subspace of $\mathbb{R}^m$. Assume that $\mathcal{P} \circ \nu \ll \mu$ and $\mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \equiv x_2$. By Proposition 12.52, we may define $F : \Omega \to \mathcal{Y}$ to be the cumulative distribution function of $\nu$ with origin $x_s$, where $x_s$ is an arbitrary fixed vector in $\Omega$ if $\Omega \neq \emptyset$. Then, $F$ is of locally bounded variation and is absolutely continuous.

In particular, if $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $M \subseteq \mathcal{Y}$ be a separable subspace, and $f : \mathcal{P}(\Omega) \to M$ be $\mathcal{B}$-measurable. We may define $\nu$ to be the $\mathcal{Y}$-valued measure with kernel $f$ over $\mathcal{X}$ as outlined in Proposition 11.116. Then, $f$ is the Radon-Nikodym derivative of $\nu$ with respect to $\mu$. Assume that $\forall x_1, x_2 \in \Omega$ with $x_1 \equiv x_2$, $f$ is absolutely integrable on $r_{x_1, x_2}$. Then, $\nu$ satisfies the assumption of the previous paragraph. Let $F : \Omega \to \mathcal{Y}$ be the cumulative distribution function of $\nu$ with origin at $x_s$ as defined in Proposition 12.52. Then, $F$ is absolutely continuous.

Proof We will prove that $F$ is absolutely continuous at $x_0$, $\forall x_0 \in \Omega$, by Proposition 12.60. Fix any $x_0 \in \Omega$, any $i \in \{0, \ldots, m-1\}$, and any $J \subseteq \mathbb{R}^i \cap \Omega_{x_0}$ with $\operatorname{card}(J) = i$. Let $\pi_j : \mathbb{R}^m \to \mathbb{R}^i$, $\pi_{J\setminus J} : \mathbb{R}^m \to \mathbb{R}^{m-i}$, $M : \mathbb{R}^i \times \mathbb{R}^{m-i} \to \mathbb{R}^m$, $x_f$, $\hat{x}_f$, $\Omega_f$, and $\tilde{F} : \hat{\Omega}_f \to \mathcal{Y}$ be as defined in (a) Proposition 12.60. Define $\tilde{x} := (\pi_j(x_s), \pi_{J\setminus J}(x_0)) \in \tilde{\Omega}_f$, $E := r_{x_f \cap \pi_j(x_s), \pi_{J\setminus J}(x_0)}$, and $E_n := x_{f-1}^{-1}/n, x_{f+1}/n \cap \tilde{\Omega}_f$, $\forall n \in \mathbb{N}$. Clearly, $V_n := M(E \times E_n) \subseteq \Omega$ and $\mu_{E_n}(P(V_n)) = \mu_{E_n}(V_n) = \mu_{E_n}(E)\mu_{E_n}(E_n) \leq (2/n)^{m-i} \mu_{E_n}(E) < \infty$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \to \infty} \mu_{E_n}(P(V_n)) = \mu_{E_n}(\bigcap_{n=1}^{\infty} P(V_n)) = 0$, by Proposition 11.5. Then, by $\mathcal{P} \circ \nu \ll \mu$, we have $\mathcal{P} \circ \nu(\bigcap_{n=1}^{\infty} P(V_n)) = 0$. Since $\Omega$ is a rectangle, then $\forall x_1, x_2 \in \Omega$ with $x_1 \equiv x_2$. Then, $\mathcal{P} \circ \nu(P(V_n)) \to \mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$. Let $\alpha := 1/(2n_0) \in (0, \infty) \subseteq \mathbb{R}$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, by Lemma 12.74, $\exists \delta(\epsilon) \in (0, \infty) \subseteq \mathbb{R}$, $\forall E \in \mathbb{B}(\mathbb{P}(V_n))$ with $\mu(E) = \mu_{E_n}(E) < \delta(\epsilon)$, we have $\mathcal{P} \circ \nu(E) < \epsilon$. Take $\tilde{\delta}(\epsilon) := \delta(\epsilon)/(1 + \mu_{E_n}(E)) \in (0, \infty) \subseteq \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\tilde{x}_j)_{j=1}^{n}, (\tilde{x}_j)_{j=1}^{n}$ being pairwise disjoint, and $\sum_{j=1}^{\infty} \mu_{E_n}(r_{\tilde{x}_j, \tilde{x}_j}) < \tilde{\delta}(\epsilon)$, we have $\sum_{j=1}^{\infty} \| \Delta_{E}(r_{\tilde{x}_j, \tilde{x}_j}) \| = \sum_{j=1}^{\infty} \| \Delta_{E}(E \times r_{\tilde{x}_j, \tilde{x}_j}) \| = \sum_{j=1}^{\infty} \| \mu_{E_n}(E \times r_{\tilde{x}_j, \tilde{x}_j}) \| = \sum_{j=1}^{\infty} \mathcal{P} \circ \nu(E \times r_{\tilde{x}_j, \tilde{x}_j}) \leq \sum_{j=1}^{\infty} \mathcal{P} \circ \nu(E \times r_{\tilde{x}_j, \tilde{x}_j}) \leq \mathcal{P}(E \times \bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j}) \subseteq \mathcal{P}(E \times \bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j})$. Since $\bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j} \subseteq E_{n_0}$ with $\mu_{E_n}(E \times r_{\tilde{x}_j, \tilde{x}_j}) < \tilde{\delta}(\epsilon)$, then $\mu_{E_n}(E \times \bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j}) = \mu_{E_n}(E \times \bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j}) < \delta(\epsilon)$ and $\mu_{E_n}(E \times \bigcup_{j=1}^{\infty} r_{\tilde{x}_j, \tilde{x}_j}) \subseteq \bigcup_{j=1}^{\infty} M(E \times r_{\tilde{x}_j, \tilde{x}_j}) = \bigcup_{j=1}^{\infty} \mathcal{P}(M(E \times r_{\tilde{x}_j, \tilde{x}_j})) \subseteq \mathcal{P}(M(E \times E_{n_0}))$...
$P(V_{n_0})$, where the second equality, the first set containment, the third equality, and the second set containment follow from Proposition 12.39. Hence, $\mathcal{P} \circ \nu(M(E \times (\bigcup_{j=1}^n r_{x_j}, \tilde{x}_j))) < \epsilon$. Therefore, $\sum_{j=1}^n \| \Delta F_j(r_{x_j}, \tilde{x}_j) \| < \epsilon$.

By Proposition 12.60, $F$ is absolutely continuous at $x_0$. By the arbitrariness of $x_0$, $F$ is absolutely continuous. By Proposition 12.73, $F$ is of locally bounded variation. This completes the proof of the first paragraph of the proposition.

By Proposition 11.116, $(\mathcal{P}(\Omega), \mathcal{B}(\mathcal{P}(\Omega)), \nu)$ is a $\sigma$-finite $\mathcal{Y}$-valued measure space and $\mathcal{P} \circ \nu = \bar{\nu}$, where $\bar{\nu}$ is a $\sigma$-finite measure on $(\mathcal{P}(\Omega), \mathcal{B}(\mathcal{P}(\Omega)))$ defined by $\bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mu \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \mathcal{B}(\mathcal{P}(\Omega))$. Hence, $\mathcal{P} \circ \nu \ll \mu$. By Definition 11.166, $f$ is the Radon-Nikodym derivative of $\nu$ with respect to $\mu$. $\forall x_1, x_2 \in \Omega$ with $x_1 \preceq x_2$, we have $r_{x_1, x_2} \subseteq \mathcal{P}(\Omega)$. By the assumption, we have $\mathcal{P} \circ \nu(r_{x_1, x_2}) < \infty$. Thus, $\nu$ satisfies all the assumption of the first paragraph of the proposition. This implies that $F$ is absolutely continuous. This completes the proof of the proposition. \qed

The preceding results states that if a function of a rectangle in $\mathbb{R}^m$ into a normed linear space, then, the function being absolutely continuous implies it is of locally bounded variation. If the normed linear space is a Banach space, then there exists a unique Banach space valued measure whose cumulative distribution function is the given function, and this measure is absolutely continuous with respect to the restriction of $\mu_{\mathcal{B}m}$ to the principal. This will further lead to the existence of a unique Radon-Nikodym derivative of the Banach space valued measure with respect to the restriction of $\mu_{\mathcal{B}m}$, if the Banach space is separable and reflexive whose dual is also separable. When, the range space is only a Banach space, then the Banach space valued measure may or may not admit a Radon-Nikodym derivative with respect to the restriction of $\mu_{\mathcal{B}m}$ to the principal.

Starting from a measurable function of the principal of a rectangle in $\mathbb{R}^m$ to a separable subspace of a Banach space, we may define a $\sigma$-finite Banach space valued measure on the principal. The cumulative distribution function of the Banach space valued measure with origin at some $x_0$ inside the rectangle exists and is absolutely continuous under the mild condition that the measurable function we started with is absolutely integrable on any semi-open rectangles in $\mathbb{R}^m$ with corners inside the rectangle. In this case, we may further conclude that $f$ is the Radon-Nikodym derivative of the Banach space valued measure with respect to the restriction of $\mu_{\mathcal{B}m}$ to the principal.

12.6 Fundamental Theorem of Calculus

Definition 12.76 Let $X := (X, \rho)$ be a metric space and $S \subseteq X$. The diameter of $S$ is defined to be $\text{dia}(S) = \max\{\sup_{x_1, x_2 \in S} \rho(x_1, x_2), 0\} \in \mathbb{R}$. If $S$ is compact and bounded, then $\text{dia}(S) < \infty$.
[0, \infty] \subset \mathbb{R}_e.

**Definition 12.77** Let \( m \in \mathbb{N}, \, E \subseteq \mathbb{R}^m \) and \( \mathcal{I} \subseteq \mathcal{B}(\mathbb{R}^m) \) be a collection of nondegenerate rectangles in \( \mathbb{R}^m \) (that is, \( \exists \) nondegenerate intervals \( I_1, \ldots, I_m \) such that \( U = \prod_{i=1}^{m} I_i, \, \forall U \in \mathcal{I} \)), i.e., \( \mu_{\mathcal{B}_m}(U) > 0, \, \forall U \in \mathcal{I} \). We will say that \( \mathcal{I} \) covers \( E \) in the sense of Vitali with index \( c \in (0, \frac{1}{\sqrt{m}}] \subset \mathbb{R} \) if \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \, \forall x \in E, \, \exists U \in \mathcal{I} \) such that \( x \in U, \, \text{dia}(U) < \varepsilon, \) and the shortest side of the rectangle \( U \) has length at least \( c \text{dia}(U) \).

**Lemma 12.78** (Vitali) Let \( m \in \mathbb{N}, \, \mathbb{R}^m \) be endowed with the usual positive cone, \( E \subseteq \mathbb{R}^m \) with \( \mu_{\mathcal{B}_m}(E) < +\infty \) and \( \mathcal{I} \subseteq \mathcal{B}(\mathbb{R}^m) \) be a collection of nondegenerate rectangles in \( \mathbb{R}^m \) that covers \( E \) in the sense of Vitali with index \( c \in (0, \frac{1}{\sqrt{m}}] \subseteq \mathbb{R} \). Then, \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), there exists a finite pairwise disjoint subcollection \( \{ U_1, \ldots, U_n \} \subseteq \mathcal{I} \), where \( n \in \mathbb{Z}_+ \), such that \( \mu_{\mathcal{B}_m}(E \setminus (\bigcup_{i=1}^{m} U_i)) < \varepsilon \).

**Proof** Let \( \hat{\mathcal{I}} := \{ U \subseteq \mathbb{R}^m \mid U \in \mathcal{I} \} \). Then, \( \hat{\mathcal{I}} \) is a collection of nondegenerate closed rectangles in \( \mathbb{R}^m \) that covers \( E \) in the sense of Vitali with index \( c \). We will first show that the result holds for \( \hat{\mathcal{I}} \). Fix any \( \varepsilon \in (0, \infty) \subset \mathbb{R} \). By Example 12.56, \( \mu_{\mathcal{B}_m}(E) = \inf_{(O), \subseteq \cup_{i=1}^{m} O, \sum_{i=1}^{m} \mu_{\mathcal{B}_m}(O_i) < +\infty} \). Then, \( \exists O \subseteq \mathcal{O}_{\mathbb{R}^m} \) such that \( E \subseteq O \) and \( \mu_{\mathcal{B}_m}(O) < +\infty \). By neglecting rectangles in \( \hat{\mathcal{I}} \), we may without loss of generality, assume that \( U \subseteq O \) and the shortest side of \( U \) has length at least \( c \text{dia}(U) \), \( \forall U \in \hat{\mathcal{I}} \).

\( \forall k \in \mathbb{Z}_+ \), assume that pairwise disjoint \( U_1, \ldots, U_k \in \hat{\mathcal{I}} \) has already been chosen. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( E \subseteq \bigcup_{i=1}^{k} U_i \); Case 2: \( E \setminus (\bigcup_{i=1}^{k} U_i) \neq \emptyset \). Case 1: \( E \subseteq \bigcup_{i=1}^{k} U_i \). Then, \( \mu_{\mathcal{B}_m}(E \setminus (\bigcup_{i=1}^{k} U_i)) = 0 \). The result holds for \( \hat{\mathcal{I}} \). Case 2: \( E \setminus (\bigcup_{i=1}^{k} U_i) \neq \emptyset \). Note that, for any closed rectangle \( U \in \hat{\mathcal{I}} \) with \( \text{dia}(U) =: \tilde{\rho} > 0 \), let \( x_0 \in U \) be the center of \( U \), such that \( \mathcal{T}_{x_0 - \frac{1}{4} \tilde{\rho}} \subseteq U \subseteq \mathcal{T}_{x_0} = \mathcal{T}_{\frac{1}{2} \tilde{\rho}} \), where \( \tilde{\rho} := \sqrt{1 - (m - 1)c^2} \tilde{\rho} \) is the upper bound for the longest side of the rectangle \( U \). This implies that \( c^m \tilde{\rho}^m \leq \mu_{\mathcal{B}_m}(U) \leq \tilde{\rho}^m = (1 - (m - 1)c^2)^{m/2} \tilde{\rho}^m \). Let \( l_k \) be the supremum of the diameters of rectangles in \( \hat{\mathcal{I}} \) that do not intersect \( U_1, \ldots, U_k \). Clearly, \( l_k \leq \frac{\sqrt{\mu_{\mathcal{B}_m}(E)}}{c} \). Then, \( \exists x_0 \in E \setminus (\bigcup_{i=1}^{k} U_i) \). Since \( \cup_{i=1}^{k} U_i \) is closed, then, by Proposition 4.10, \( \tilde{d} := \inf_{x \in \cup_{i=1}^{k} U_i} |x - x_0| > 0 \). Then, by the assumption, \( \exists \tilde{U} \in \hat{\mathcal{I}} \) such that \( x_0 \in \tilde{U} \) and \( \text{dia}(U) < \frac{\sqrt{\mu_{\mathcal{B}_m}(E)}}{c} \). Then, \( \tilde{U} \cap (\bigcup_{i=1}^{k} U_i) = 0 \). This shows that \( l_k \geq \text{dia}(\tilde{U}) > 0 \). Hence, \( \exists U_{k+1} \in \hat{\mathcal{I}} \) such that \( U_1, \ldots, U_{k+1} \in \hat{\mathcal{I}} \) are pairwise disjoint and \( \text{dia}(U_{k+1}) \geq l_k/2 \). Inductively, we either have the result holds for \( \hat{\mathcal{I}} \) or \( \exists (U_k)_{k=1}^{\infty} \subseteq \hat{\mathcal{I}} \) such that \( \text{dia}(U_{k+1}) \geq l_k/2, \forall k \in \mathbb{Z}_+ \) and the sequence is pairwise disjoint.

In the latter case, we have \( O \supseteq \bigcup_{k=1}^{\infty} U_k \) and \( \mu_{\mathcal{B}_m}(O) \geq \mu_{\mathcal{B}_m}(\bigcup_{k=1}^{\infty} U_k) = \sum_{k=1}^{\infty} \mu_{\mathcal{B}_m}(U_k) \geq \sum_{k=0}^{\infty} 2^{-m} c^m \tilde{\rho}^m \). Note that \( (l_k)_{k=0}^{\infty} \) is nonincreasing. Hence, \( \lim_{k \in \mathbb{N}} l_k = 0 \). Then, \( \exists n \in \mathbb{N} \) such that \( \sum_{i=n+1}^{\infty} \mu_{\mathcal{B}_m}(U_i) < \pi^{-m/2} \Gamma(\frac{m}{2} + 1)\epsilon \), where we have referred
to the special Gamma function. Let \( R := E \setminus \bigcup_{k=1}^{n} U_k \). Without loss of generality, assume \( \text{dia}(U_k) =: p_k \in (0, \infty) \subset \mathbb{R} \) and the center of \( U_k \) is \( x_k \), \( \forall k \in \mathbb{N} \). \( \forall \bar{x} \in R \), since \( \bigcup_{k=1}^{n} U_k \) is closed, then, by Proposition 4.10, \( d \leq \inf_{x \in \bigcup_{k=1}^{n} U_k} |x - \bar{x}| > 0 \). By the assumption, \( \exists U \in \bar{T} \) such that \( \bar{x} \in U \) and \( 0 < \text{dia}(U) < d \). Then, \( U \cap \bigcup_{k=1}^{n} U_k = \emptyset \).

By the fact that \( \lim_{k \to \infty} U_k = 0 \), we have \( \exists i_0 \in \{ n + 1, n + 2, \ldots \} \) such that \( U \cap U_{i_0} \neq \emptyset \) and \( U \cap U_k = \emptyset \), \( \forall k \in \{ 1, \ldots, i_0 - 1 \} \). Then, \( \text{dia}(U) \leq |i_0 - 1| \leq 2 \text{dia}(U_{i_0}) = 2p_{i_0} \). Let \( \bar{x} \in U \cap U_{i_0} \). Then, \( |\bar{x} - x_{i_0}| \leq |\bar{x} - \bar{x}| + |\bar{x} - x_{i_0}| \leq \text{dia}(U) + \frac{1}{2}p_{i_0} \leq \frac{3}{2}p_{i_0} \). Then, \( \bar{x} \in V_{i_0} := \bar{B}_{\mathbb{R}^m}(x_{i_0}, \frac{3}{2}p_{i_0}) \subset \mathbb{R}^m \).

Then, \( \bar{x} \in \bigcup_{k=n+1}^{\infty} V_k := \bigcup_{k=n+1}^{\infty} \bar{B}_{\mathbb{R}^m}(x_k, \frac{3}{2}p_k) \subset \mathbb{R}^m \). Hence, \( R \subseteq \bigcup_{k=n+1}^{\infty} V_k \). Then, \( \mu_{\text{lm}}(R) \leq \mu_{\text{lm}}(\bigcup_{k=n+1}^{\infty} V_k) \leq \sum_{k=n+1}^{\infty} \mu_{\text{lm}}(V_k) = \sum_{k=n+1}^{\infty} \frac{\pi^{m/2 - m} \Gamma^m}{(m+1)^{m+1}} \sum_{k=n+1}^{\infty} c^m p_k^m \leq \frac{\pi^{m/2 - m} \Gamma^m}{(m+1)^{m+1}} \sum_{k=n+1}^{\infty} c^m p_k^m \). 

This completes the proof of the lemma. 

**Definition 12.79** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( E \in \mathcal{O}_{\mathbb{R}^m} \) with \( \mu_{\text{lm}}(E) > 0 \), \( \mathcal{E} := ((E, |\cdot|), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \) as defined in Proposition 11.29, \( \forall \) be a separable normed linear space over \( K \), \( c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R} \), and \( f \in L_1(\mathcal{E}, \mathcal{Y}) \).

A point \( x \in E \) is said to be a rectangular Lebesgue point with regularity \( c \) of \( f \) if \( \limsup_{r \to 0^+} \sup_{x_1, x_2 \in \mathbb{R}^m \text{ with } \frac{1}{m}(x_1 + x_2) = x} \frac{1}{\mu(x_1, x_2) \cap E} \int_{x_1, x_2} \mathbb{E} \| f(y) - f(x) \| \, d\mu(y) = 0. \)

**Proposition 12.80** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( E \in \mathcal{O}_{\mathbb{R}^m} \) with \( \mu_{\text{lm}}(E) > 0 \), \( \mathcal{E} := ((E, |\cdot|), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \) as defined in Proposition 11.29, \( \forall \) be a separable normed linear space over \( K \), \( c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R} \), \( f \in L_1(\mathcal{E}, \mathcal{Y}) \), and \( A := \{ x \in E \mid x \text{ is a rectangular Lebesgue point of } f \text{ with regularity } c \} \).

Then, \( A \in \mathcal{B}_{\mathbb{R}^m} \) and \( \mu_{\text{lm}}(E \setminus A) = 0. \)

**Proof** Since \( \mathcal{E} \) is \( \sigma \)-finite, then \( \exists (E_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( \bigcup_{n=1}^{\infty} E_n = E \) and \( \mu(E_n) < +\infty, \forall n \in \mathbb{N} \). \( \forall r \in (0, \infty) \subset \mathbb{R} \), \( \forall x \in E \), define \( T_{r,c} : E \to [0, \infty) \subset \mathbb{R} \) and \( M_{r,c} : E \to [0, \infty) \subset \mathbb{R} \) by, \( \forall x \in E \),

\[
(T_{r,c}(f))(x) = \sup_{x_1, x_2 \in \mathbb{R}^m \text{ with } \frac{1}{m}(x_1 + x_2) = x} \frac{1}{\mu(x_1, x_2) \cap E}
\]
\[ \int_{x_1,x_2 \in \mathbb{R}^m} \| f(y) - f(x) \| \, d\mu(y) \]

\[ (M_{r,c}f)(x) = \sup_{(x_1,x_2) \subseteq \mathbb{R}^m} \frac{1}{\mu(x_1,x_2)} \int_{x_1,x_2 \cap E} P \circ f \, d\mu \]

\[ \forall x \in E, \text{ since } E \subseteq \mathcal{O}_{\mathbb{R}^m} \text{ then } \exists \delta \in (0, \infty) \subseteq \mathbb{R} \text{ such that } r_{x_1}^2 - \delta_{x_1,x_2} + \delta_{x_2} \subseteq E. \]

\[ \forall x_1,x_2 \subseteq \mathbb{R}^m \text{ with } x = \frac{1}{2}(x_1 + x_2), \text{ } |x_2 - x_1| = r > 0 \text{ and } \min(x_2 - x_1) \geq cr, \]

we have \( \mu(x_1,x_2 \cap E) = \mu_{x_1}(x_2 \cap E) \geq \mu_{B_{x_1}}(x_2 \cap E) \geq \mu_{B_{x_1}}(x_2 \cap r_{x_1}^2 - \delta_{x_1,x_2} + \delta_{x_2}) \geq \|(2\delta) \cap (cr)\)^m > 0. \]

Then, \( (M_{r,c}f)(x) \leq (2\delta) \cap (cr)^m \| f \|_1 < \infty \) and \( (T_{r,c}f)(x) \leq (M_{r,c}f)(x) + \| f(x) \| < \infty. \) Hence, \( T_{r,c}f \) and \( M_{r,c}f \) are well-defined. Define \( T_{r,c} : E \rightarrow [0, \infty) \subseteq \mathbb{R}_{+} \) and \( M_{r,c} : E \rightarrow [0, \infty) \subseteq \mathbb{R}_+ \) by

\[ (T_{r,c}f)(x) = \limsup_{r \rightarrow 0^+} (T_{r,c}f)(x) \quad (M_{r,c}f)(x) = \limsup_{r \rightarrow 0^+} (M_{r,c}f)(x); \quad \forall x \in E \]

Clearly, \( T_{r,c} \) and \( M_{r,c} \) are well-defined. Then, \( T_{r,c} \) and \( M_{r,c} \) are functions of \( L_1(\mathbb{R}, \mathcal{B}) \) to nonnegative real-valued functions of \( E \), and \( T_{r,c} \) and \( M_{r,c} \) are functions of \( L_1(\mathbb{E}, \mathcal{B}) \) to nonnegative extended real-valued functions of \( E \).

Then, \( A = \{ x \in E \mid (T_{r,c}f)(x) = 0 \} \) and \( E \setminus A = \{ x \in E \mid (T_{r,c}f)(x) > 0 \} = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n,k} := \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \{ x \in E_n \mid (T_{r,c}f)(x) > 1/k \}. \)

We will show that \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) = 0 \), \( \forall n \in \mathbb{N}, \forall k \in \mathbb{N} \), which implies that \( A_{n,k} \subseteq \mathcal{B}_{\mathbb{R}} \) and \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) = 0 \), \( \forall n \in \mathbb{N}, \forall k \in \mathbb{N} \). This further implies that \( E \setminus A \subseteq \mathcal{B}_{\mathbb{R}}, \mu_{\mathcal{M}_{\mathbb{R}}}(E \setminus A) = 0, \) and \( A \subseteq \mathcal{B}_{\mathbb{R}} \).

Fix any \( n \in \mathbb{N} \) and any \( k \in \mathbb{N} \). Suppose \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) > 0 \). By monotonicity of outer measures, we have \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) \leq \mu_{\mathcal{M}_{\mathbb{R}}}(E_n) = \mu_{\mathcal{M}_{\mathbb{R}}}(E_{n} - \epsilon_0) \leq +\infty. \) Let \( \epsilon_0 = \frac{\mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k})}{\delta_{x_2}} \in (0, +\infty) \subseteq \mathbb{R} \). By Propositions 11.82 and 4.11, there exists a continuous function \( g : E \rightarrow \mathbb{Y} \) such that \( g \in \mathcal{L}_1(\mathbb{E}, \mathcal{B}) \) and \( \| g - f \|_1 < \epsilon_0. \) Let \( h := f - g. \) Then, \( \| h \|_1 < \epsilon_0. \) Note that \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x) + (T_{r,c}h)(x), \forall r \in (0, \infty) \subseteq \mathbb{R} \), \( \forall x \in E \).

Then, by Proposition 3.85, \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x) + (T_{r,c}h)(x), \forall r \in (0, \infty) \subseteq \mathbb{R} \), \( \forall x \in E \). Since \( g \) is continuous, we have \( (T_{r,c}g)(x) = 0, \forall x \in E \). Then, \( (T_{r,c}h)(x) \leq (T_{r,c}h)(x), \forall x \in E. \)

Note that \( (T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \| h(x) \| ; \quad \forall r \in (0, \infty) \subseteq \mathbb{R}, \forall x \in E \)

Then, by Proposition 3.85, we have \( (T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \| h(x) \|, \forall x \in E \). Hence, \( (T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \| h(x) \|, \forall x \in E \).

Then, \( A_{n,k} = \{ x \in E_n \mid (M_{r,c}h)(x) > \frac{1}{2k} \} \subseteq \{ x \in E_n \mid \| h(x) \| > \frac{1}{2k} \} =: A_{n,k} \subseteq \mathcal{B}_{\mathbb{R}} \). By monotonicity of outer measures, we have \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) \leq \mu_{\mathcal{M}_{\mathbb{R}}}(E_n \setminus A_{n,k}) \leq \mu_{\mathcal{M}_{\mathbb{R}}}(\{ x \in E \mid \| h(x) \| > \frac{1}{2k} \}) =: \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) = \mu_{\mathcal{B}_{\mathbb{R}}}(A_{n,k}) = \int_{E_{\mathbb{R}}} \chi_{A_{n,k}} \, d\mu \leq \int_{E_{\mathbb{R}}} (2kP \circ h) \, d\mu = 2k \| h \|_1 < 2k\epsilon_0, \) where the second equality follows from the fact that \( A_{n,k} \subseteq \mathcal{B} \) since \( h \) is \( \mathcal{B} \)-measurable. By countable subadditivity of outer measures, we have \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) \leq \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) + \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) < \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) + 2k\epsilon_0. \) Hence, \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) > 3k\epsilon_0. \) Clearly, \( \mu_{\mathcal{M}_{\mathbb{R}}}(A_{n,k}) \leq \mu_{\mathcal{M}_{\mathbb{R}}}(E_n) < +\infty. \)
Define $\mathcal{I} := \{ r_{x_1, x_2} \subseteq \mathbb{R}^m \mid x = \frac{1}{7}(x_1 + x_2) \in E, |x_2 - x_1| =: r > 0, \min(x_2 - x_1) \geq c r, \int r_{x_1, x_2} \cap E \mathcal{P} \circ h \, d\mu > \frac{1}{3k} \mu(r_{x_1, x_2} \cap E) \}$. Then, $\mathcal{I}$ covers $\tilde{A}_{n,k}$ in the sense of Vitali with index $c$. By Vitali’s Lemma 12.78, there exists pairwise disjoint rectangles $(r_{x_{i,1}, x_{i,2}})^{m}_{i=1} \subseteq \mathcal{I}$ with $\tilde{m} \in \mathbb{Z}^+$ such that

$$\mu_{\text{Lmo}}(\tilde{A}_{n,k} \setminus (\bigcup_{i=1}^{m} r_{x_{i,1}, x_{i,2}})) < \epsilon_0.$$ 

Let $B = (\bigcup_{i=1}^{m} r_{x_{i,1}, x_{i,2}}) \cap E \in \mathcal{B}$. Then, $\mu(B) = \mu_{\text{Lmo}}(B) \geq \mu_{\text{Lmo}}(B \cap \tilde{A}_{n,k}) = \mu_{\text{Lmo}}(\tilde{A}_{n,k}) - \mu_{\text{Lmo}}(\tilde{A}_{n,k} \setminus B) > (3k - 1) \epsilon_0$, where the first inequality follows from the monotonicity of outer measures; and the second equality follows from the measurability of $B$. Then, $\| h \|_t = \int E \mathcal{P} \circ h \, d\mu \geq \int E \mathcal{P} \circ h \, d\mu = \sum_{i=1}^{m} \int r_{x_{i,1}, x_{i,2}} \cap E \mathcal{P} \circ h \, d\mu > \sum_{i=1}^{m} \frac{1}{3k} \mu(r_{x_{i,1}, x_{i,2}} \cap E) = \frac{\tilde{m} \mu(B)}{3k} \epsilon_0 \geq \epsilon_0 > \| h \|_t$. This is a contradiction. Therefore, we must have $\mu_{\text{Lmo}}(\tilde{A}_{n,k}) = 0$.

This completes the proof of the proposition. \(\Box\)

In the rest of the notes, we will denote the $i$th unit vector in $\mathbb{R}^m$ by $e_{m,i}$, $\forall m \in \mathbb{N}$, $\forall i \in \{1, \ldots, m\}$.

**Proposition 12.81** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\mathcal{Y}$ be a separable normed linear space over $\mathbb{K}$, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be an open set, and $F : \Omega \rightarrow \mathcal{Y}$ be continuous on the right (or continuous on the left) and $\mathcal{B}_B(\mathbb{R}^m)$-measurable. Then, the following statements hold.

(i) $\forall i \in \{1, \ldots, m\}$, let $D_i F : U_i \rightarrow \mathcal{B}(\mathbb{K}, \mathcal{Y})$ denote the partial derivative of $F$ with respect to its $i$th coordinate in its domain and $U_i := \text{dom}(D_i F)$. Then, $U_i \in \mathcal{B}_B(\mathbb{R}^m)$ and $D_i F$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable.

(ii) Let $F^{(1)} : U \rightarrow \mathcal{B}(\mathbb{K}^m, \mathcal{Y})$ be the Fréchet derivative of $F$, whose domain is $U := \text{dom}(F^{(1)}) \subseteq \bigcap_{i=1}^{m} U_i \subseteq \Omega$. Then, $U \in \mathcal{B}_B(\mathbb{R}^m)$ and $F^{(1)}$ is $\mathcal{B}_B(\mathbb{R}^m)$-measurable.

**Proof**

(i) Fix any $i \in \{1, \ldots, m\}$. Consider the case that $F$ is continuous on the right. Let $X := \mathbb{K}$, $D := \pi_i(\Omega) \in \mathcal{B}_B(\mathbb{R})$, which is an open set. $\forall s \in D$, clearly span$(\tilde{A}_D(s)) = \mathbb{K} = X$. Then, $D_i F : U_i \rightarrow \mathcal{B}(\mathbb{K}, \mathcal{Y})$, where $U_i := \text{dom}(D_i F) \subseteq \Omega$. Clearly, $\mathcal{B}(\mathbb{K}, \mathcal{Y})$ is isometrically isomorphic to $\mathcal{Y}$. Then, $D_i F : U_i \rightarrow \mathcal{Y}$. Clearly, $\forall x \in U_i$, $D_i F(x) = \lim_{h \rightarrow 0}(F(x + he_{m,i}) - F(x))/h \in \mathcal{Y}$.

**Claim 12.81.1** $U_i = \bigcap_{n=1}^{\infty} \bigcup_{x \in \Omega} \bigcup_{h_1 \in \mathbb{Q} \cup \{1/n\}} \bigcup_{h_2 \in \mathbb{Q}} U_{i,n,h_1,h_2} := U_i \in \mathcal{B}_B(\mathbb{R}^m)$, where $U_{i,n,h_1,h_2} := \{ x \in \Omega \mid \| (F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2 \| < 1/n \}$, $\forall n \in \mathbb{N}$, $\forall h_1, h_2 \in \mathbb{Q}$ with $h_1 \neq 0$ and $h_2 \neq 0$.

**Proof of claim:** $\forall x \in U_i$, $D_i F(x) = \lim_{h \rightarrow 0}(F(x + he_{m,i}) - F(x))/h \in \mathcal{Y}$. By Proposition 4.45, $\forall n \in \mathbb{N}$, $\exists \delta \in \mathbb{Q}$ with $\delta > 0$, such that $x + he_{m,i} \in \Omega$, $\forall h \in \mathcal{B}_R(0, \delta)$, and $\forall h_1, h_2 \in \mathcal{B}_R(0, \delta) \setminus \{0\}$, we have $\| (F(x + h_1 e_{m,i}) - F(x + h e_{m,i})) - (F(x + h_2 e_{m,i}) - F(x + h e_{m,i})) \| < 1/n$. Then, $\forall x \in U_i$, $D_i F(x)$ is measurable.
12.6. FUNDAMENTAL THEOREM OF CALCULUS

\[ F(x) / \bar{h}_1 - (F(x + h_2 e_{m,i}) - F(x)) / \bar{h}_2 \| < 1/n. \]

Then, \( \forall h_1, h_2 \in \mathbb{Q} \) with \( 0 < |h_1| < \delta \) and \( 0 < |h_2| < \delta \), we have \( \|(F(x + h_1 e_{m,i}) - F(x)) / h_1 - (F(x + h_2 e_{m,i}) - F(x)) / h_2\| < 1/n \). Thus, \( x \in U_{i,n,h_1,h_2} \). Hence, \( x \in \bar{U}_i \).

By the arbitrariness of \( x \), we have \( U_i \subseteq \bar{U}_i \).

On the other hand, \( \forall x \in \bar{U}_i \). \( \forall n \in \mathbb{N}, \exists \delta \in \mathbb{Q} \) with \( \delta > 0 \), such that, \( \forall h \in \mathcal{B}_h(0, \delta), x + h e_{m,i} \in \Omega \), and \( \forall h_1, h_2 \in \mathbb{Q} \) with \( 0 < |h_1| < \delta \) and \( 0 < |h_2| < \delta \), we have \( \|(F(x + h_1 e_{m,i}) - F(x)) / h_1 - (F(x + h_2 e_{m,i}) - F(x)) / h_2\| < 1/n \). \( \forall h_1, h_2 \in \mathcal{B}_h(0, \delta) \setminus \{0\}, \exists (h_{j,l})_{l=1}^\infty \subseteq (0 \cap \mathcal{B}_h(0, \delta)) \setminus \{0\} \) such that \( h_j < h_{j,l}, \forall l \in \mathbb{N}, \) and \( \lim_{l \to \infty} h_{j,l} = h_j, j = 1, 2, \forall l \in \mathbb{N} \), \( \|(F(x + h_{1,l} e_{m,i}) - F(x)) / h_{1,l} - (F(x + h_{2,l} e_{m,i}) - F(x)) / h_{2,l}\| < 1/n \).

Then,

\[
\begin{align*}
\| & (F(x + h_{1,l} e_{m,i}) - F(x)) / h_{1,l} - (F(x + h_{2,l} e_{m,i}) - F(x)) / h_{2,l} \\
& = \left\| \lim_{l \to \infty} \left( \frac{F(x + h_{1,l} e_{m,i}) - F(x)}{h_{1,l}} - \frac{F(x + h_{2,l} e_{m,i}) - F(x)}{h_{2,l}} \right) \right\| \\
& = \lim_{l \to \infty} \left\| \frac{F(x + h_{1,l} e_{m,i}) - F(x)}{h_{1,l}} - \frac{F(x + h_{2,l} e_{m,i}) - F(x)}{h_{2,l}} \right\| \\
& \leq 1/n
\end{align*}
\]

where the first equality follows from the fact that \( F \) is continuous on the right; and the second equality follows from Propositions 3.66, 3.67, and 7.23. By Proposition 4.45, \( \lim_{h \to 0} (F(x + h e_{m,i}) - F(x)) / h \in \mathcal{Y} \). Hence, \( x \in \text{dom}(D_1 F) = U_i \). By the arbitrariness of \( x \), we have \( U_i \subseteq \bar{U}_i \). Therefore, \( U_i = \bar{U}_i \).

By Propositions 7.21, 7.23, 11.38, and 11.39, \( U_{i,n,h_1,h_2} \in \mathcal{B}_h(\mathbb{R}^m) \), \( \forall n \in \mathbb{N}, \forall h_1, h_2 \in \mathbb{Q} \) with \( h_1 \neq 0 \) and \( h_2 \neq 0 \). Then, \( U_i = \bar{U}_i \in \mathcal{B}_h(\mathbb{R}^m) \).

\( \forall n \in \mathbb{N} \), define \( H_n : \Omega \to \mathcal{Y} \) by, \( \forall x \in \Omega \),

\[
H_n(x) = \left\{ \begin{array}{ll}
\frac{n}{\partial_y} (F(x + \frac{1}{n} e_{m,i}) - F(x)) & x \in O := \Omega \cap (\Omega - \frac{1}{n} e_{m,i}) \\
x \in \Omega \setminus O
\end{array} \right.
\]

By Propositions 7.23, 11.38, 11.39, and 11.41, \( H_n \) is \( \mathcal{B}_h(\mathbb{R}^m) \)-measurable. Then, \( H_n |_{U_i} \) is \( \mathcal{B}_h(\mathbb{R}^m) \)-measurable. Note that, \( \forall x \in U_i, D_1 F(x) = \lim_{n \to \infty} H_n(x) = \lim_{n \to \infty} H_n |_{U_i} (x) \). By Proposition 11.48, \( D_1 F \) is \( \mathcal{B}_h(\mathbb{R}^m) \)-measurable.

The case when \( F \) is continuous on the left can be proved similarly.

(ii) Consider the case that \( F \) is continuous on the right. By Proposition 9.9, \( \forall x \in U = \text{dom}(F^{(1)}) \), we must have \( x \in U_i, \forall i \in \{1, \ldots, m\} \). Then, \( U \subseteq \bigcap_{i=1}^m U_i \), \( \forall x \in \Omega \), since \( \Omega \) is open, then \( \text{span} \left( A_{\Omega}(x) \right) = \mathbb{K}^m \).

**Claim 12.81.2** \( U = \bigcap_{n=1}^\infty \bigcup_{\delta \in \mathcal{O}} \bigcap_{|h_1| < \delta} \bigcap_{|h_2| < \delta} \bigcap_{|h_m| < \delta} V_{n,h_1,\ldots,h_m} =: \bar{V} \in \mathcal{B}_h(\mathbb{R}^m) \),

where \( V_{n,h_1,\ldots,h_m} := \left\{ x \in \bigcap_{i=1}^m U_i \left| \| F(x + \sum_{i=1}^m h_i e_{m,i}) - F(x) - \sum_{i=1}^m h_i D_1 F(x) \| \leq \frac{1}{n} \sqrt{\sum_{i=1}^m h_i^2} \right. \right\} \), \( \forall n \in \mathbb{N}, \forall h_1, \ldots, h_m \in \mathbb{Q} \).
Proof of claim: \( \forall x \in U, \) we have \( x \in \bigcap_{i=1}^{m} U_i. \) By Definition 9.3 and Proposition 9.9, the candidate for \( F^{(1)}(x) \) is \( \left[ D_1 F(x) \ldots D_m F(x) \right] := L \in B(\mathbb{K}^m, \mathbb{Y}). \) By \( x \in \text{dom} \left( F^{(1)} \right), \) we have \( \forall n \in \mathbb{N}, \exists \delta \in \mathbb{Q} \) with \( \delta > 0, \) such that \( B_{\mathbb{R}^m} \left( x, \sqrt{m \delta} \right) \subseteq \Omega, \) and \( \forall \delta \in B_{\mathbb{R}^m} \left( 0, \sqrt{m \delta} \right), \) we have \( \| F(x + h) - F(x) - Lh \| \leq \frac{1}{n} \| h \|. \) Then, \( \forall h := (h_1, \ldots, h_m) \in \mathbb{Q}^m \) with \( |h_i| < \delta, \ i = 1, \ldots, m, \) we have \( \| F(x + h) - F(x) - Lh \| = \| F(x + \sum_{i=1}^{m} h_i e_{m, i}) - F(x) - \sum_{i=1}^{m} h_i D_i F(x) \| \leq \frac{1}{n} \| h \| = \frac{1}{n} \sqrt{\sum_{i=1}^{m} h_i^2}. \) Thus, \( x \in V_{n,h_1,\ldots,h_m}. \) Hence, \( x \in V. \) By the arbitrariness of \( x, \) we have \( U \subseteq V. \)

On the other hand, \( \forall x \in V. \) Then, \( \forall n \in \mathbb{N}, \exists \delta \in \mathbb{Q} \) with \( \delta > 0, \) such that \( B_{\mathbb{R}^m} (x, \delta) \subseteq \Omega \) and, \( \forall h := (h_1, \ldots, h_m) \in \mathbb{Q}^m \) with \( |h_i| < \delta, \ i = 1, \ldots, m, \) we have \( x \in V_{n,h_1,\ldots,h_m} \subseteq \bigcap_{i=1}^{m} U_i. \) \( \forall \bar{x} \in B_{\mathbb{K}^m} (x, \delta) \cap \Omega = B_{\mathbb{R}^m} (x, \delta), \) let \( (\bar{h}_1, \ldots, \bar{h}_m) := \bar{x} - x. \) Clearly, \( |\bar{h}_i| < \delta, \ i = 1, \ldots, m, \) then, \( \exists (h_{i,l})_{l=1}^{\infty} \subseteq \mathbb{Q} \) with \( |h_{i,l}| < \delta \) such that \( \bar{h}_i < h_{i,l}, \forall l \in \mathbb{N}, \) and \( \lim_{l \to \infty} h_{i,l} = h_i, \ i = 1, \ldots, m. \) By the assumption,

\[
\| F(\bar{x}) - F(x) - L(\bar{x} - x) \|
= \left\| \lim_{l \to \infty} \left( F(x + \sum_{i=1}^{m} h_{i,l} e_{m, i}) - F(x) - \sum_{i=1}^{m} h_{i,l} D_i F(x) \right) \right\|
= \lim_{l \to \infty} \left\| F(x + \sum_{i=1}^{m} h_{i,l} e_{m, i}) - F(x) - \sum_{i=1}^{m} h_{i,l} D_i F(x) \right\|
\leq \lim_{l \to \infty} \frac{1}{n} \left( \sum_{i=1}^{m} h_{i,l}^2 \right)^{1/2} = \frac{1}{n} \left( \sum_{i=1}^{m} h_i^2 \right)^{1/2} = \frac{1}{n} |\bar{x} - x|
\]

where the first equality follows from the fact that \( F \) is continuous on the right; the second equality follows from Propositions 3.66, 3.67, 7.21, and 7.23; and the first inequality follows from the fact that \( x \in V_{n,h_1,\ldots,h_m}. \) By Definition 9.3, \( F^{(1)}(x) = L. \) Hence, \( x \in \text{dom} \left( F^{(1)} \right) = U. \) By the arbitrariness of \( x, \) we have \( V \subseteq U. \) Therefore, \( U = V. \)

By Propositions 7.21, 7.23, 11.38, and 11.39, \( V_{n,h_1,\ldots,h_m} \in \mathcal{B}_B \left( \mathbb{R}^m \right), \) \( \forall n \in \mathbb{N}, \forall h_1, \ldots, h_m \in \mathbb{Q}. \) Then, \( U = V \in \mathcal{B}_B \left( \mathbb{R}^m \right). \) \( \square \)

Note that \( \forall x \in U, \) \( [\text{D}_1 F(x) \ldots \text{D}_m F(x)] = [ (\text{D}_1 F)_U(x) \ldots (\text{D}_m F)_U(x) ] \). By Propositions 11.38, 11.39, 11.41, and 11.139, \( F^{(1)} \) is \( \mathcal{B}_B \left( \mathbb{R}^m \right) \)-measurable. The case when \( F \) is continuous on the left can be proved similarly.

This completes the proof of the proposition. \( \square \)

**Theorem 12.82** Let \( I := [a,b] \subseteq \mathbb{R} \) with \( a, b \in \mathbb{R} \) and \( a < b, \) \( \mathbb{Y} \) be a separable Banach space over \( \mathbb{K}, \) \( f : I \to \mathbb{Y} \) be absolutely integrable over \( I \) with respect to \( \mu_\mathbb{Y}, \) and \( F : I \to \mathbb{Y} \) be defined by \( F(x) = \int_a^x f(t) \, dt, \forall x \in I. \) Assume that \( f \) is continuous at \( x_0 \in I. \) Then, \( F \) is Fréchet differentiable at \( x_0 \) and \( F^{(1)}(x_0) = f(x_0). \) (Note that, when \( \mathbb{K} = \mathbb{C}, \) \( I \) is viewed as a subset of \( \mathbb{C} \) in calculations of \( F^{(1)}). \)
12.6. FUNDAMENTAL THEOREM OF CALCULUS

Proof Clearly, $F$ is well-defined by Proposition 11.92. Note that \( \text{span}(A_I(x_0)) = K \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the continuity of $f$ at $x_0$, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \), \( \forall x \in I \cap B_R(x_0, \delta) \), we have \( \|f(x) - f(x_0)\| < \epsilon \). Then,

\[
\|F(x) - F(x_0) - f(x_0)(x - x_0)\| = \|\int_{x_0}^{x} f(t) \, dt - \int_{x_0}^{x} f(t) \, dt - \int_{x_0}^{x} f(x_0) \, dt\| = \|\int_{x_0}^{x} (f(t) - f(x_0)) \, dt\| = \int_{x_0}^{x} \|f(t) - f(x_0)\| \, dt \leq \epsilon \|f\|_{L^\infty(I)} \|x - x_0\|.
\]

where the second equality follows from Fact 12.72; the third equality and the first inequality follow from Proposition 11.92; and the second inequality follows from Proposition 11.78. Hence, $F'(x_0) = f(x_0)$. This completes the proof of the theorem.

\[ \square \]

Theorem 12.83 Let $I := [a,b] \subset \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a < b$, $I := ([I,|\cdot|],B,\mu)$ be the finite complete metric measure subspace of $\mathbb{R}$, $\mathcal{Y}$ be a Banach space over $K$, and $F : I \rightarrow \mathcal{Y}$ be $C_1$. (Note that, when $K = C$, $I$ is viewed as a subset of $C$ in calculations of $F'(1)$.) Then, $F(1) : I \rightarrow \mathcal{Y}$ is absolutely integrable over $\mathbb{I}$ and $F(b) - F(a) = \int_{a}^{b} F'(t) \, dt = \int_{a}^{b} F(1)(t) \, dt \, dB$.

Proof Since $F(1) : I \rightarrow \mathcal{Y}$ is continuous, then, by Proposition 11.37, $F(1)$ is $B$-measurable. Since $I$ is compact, then $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|F(1)(x)\| \leq M, \forall x \in I$. By Proposition 11.78, we have $F(1)$ is absolutely integrable over $I$. By Proposition 7.126, $F(1) : I \rightarrow N := \text{span}(F(1)(I)) \subset \mathcal{Y}$ and $N$ is a separable Banach subspace of $\mathcal{Y}$. Define $F_a : I \rightarrow N \subset \mathcal{Y}$ by $F_a(x) = \int_{a}^{x} F(1)(t) \, dt = \int_{a}^{x} F(1)(t) \, dB$, $\forall x \in I$. By Proposition 11.92, $F_a$ is well defined. By Theorem 12.82, $F_a$ is Fréchet differentiable and $F_{a}(1) = F(1)$. Then, define $g : I \rightarrow \mathcal{Y}$ by $g = F - F_a$. Then, by Proposition 9.15, $g$ is Fréchet differentiable and $g'(x) = \partial_y, \forall x \in I$. By Mean Value Theorem 9.23, $\|g(x) - g(a)\| \leq 0, \forall x \in I$. Hence, $g(x) = g(a) = F(a), \forall x \in I$. Then, $g(b) = F(b) = F(a) = F(b) - F_a(b) = F(b) - \int_{a}^{b} F(1)(t) \, dt$. Therefore, $F(b) - F(a) = \int_{a}^{b} F'(1)(t) \, dt$. This completes the proof of the theorem.

\[ \square \]

Proposition 12.84 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in B_{B}(\mathbb{R}^m)$ be a region, $X := ((P(\Omega),|\cdot|),B,\mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a finite-dimensional Banach space over $K$, $f : P(\Omega) \rightarrow \mathcal{Y}$ be $B_{B}(\mathbb{R}^m)$-measurable, $J := \{1, \ldots, m\}$, and $x_0 := (x_{0,1}, \ldots, x_{0,m}) \in \Omega$. Then, we may define the $\sigma$-finite $\mathcal{Y}$-valued measure $\nu$ with kernel $f$ over $\mathcal{X}$ according to Proposition 11.116. By Proposition 11.167, $f$ is the unique Radon-Nikodym derivative of $\nu$ with respect
to $\mu$. Assume that, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\bar{r}_{x_1, x_2} \subseteq \Omega$, we have $f$ is bounded over $\bar{r}_{x_1, x_2} \subseteq \Omega$. Let $F : \Omega \to \mathcal{Y}$ be a cumulative distribution function of $\nu$. Then, $\forall x := (x_1, \ldots, x_m) \in \Omega$ with $\bar{r}_{x, \hat{x}} \subseteq \Omega$, where $\hat{x} := x_0 \wedge x$ and $\hat{x} := x_0 \vee x$, we have

$$F(x) = \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \ldots, s_m) \, ds_m \cdots ds_1 - \sum_{J \subseteq \hat{J}, J \neq \emptyset} (-1)^{\text{card}(J)} F(x_J)$$

and the order of integration can be arbitrary, where $x_J \in \bar{r}_{x, \hat{x}} \subseteq \Omega$ is defined by $\pi_i(x_J) = \left\{ \begin{array}{l} \pi_i(x_0) = x_{0,i} \quad i \in J \\
\pi_i(x) = x_i \quad i \in \hat{J} \setminus J, \forall i \in \hat{J}. \end{array} \right.$

**Proof**  $\mathcal{Y}$ is separable and is $\sigma$-compact, since it is finite-dimensional. Fix any $x \in \Omega$ with $\bar{r}_{x, \hat{x}} \subseteq \Omega$. Then, $r_{x, \hat{x}} \subseteq \Omega$. By Definition 12.42 and Definition 11.166, we have $\Delta_F(r_{x, \hat{x}}) = \nu(r_{x, \hat{x}}) = \int_{r_{x, \hat{x}}} f \, d\mu = \int_{r_{x, \hat{x}}} f \, d\mu_{\mathcal{B}}^m$. By the assumption and Fubini’s Theorem 12.31, we have

$$\int_{r_{x, \hat{x}}} f \, d\mu_{\mathcal{B}}^m = \int_{x_{0,1} \wedge x_1}^{x_1} \cdots \int_{x_{0,m} \wedge x_m}^{x_m} f(s_1, \ldots, s_m) \, ds_m \cdots ds_1 = \Delta_F(r_{x, \hat{x}})$$

and the order of integration can be arbitrary, where $n(x) = \text{card}(\{ i \in \hat{J} \mid \pi_i(x_0) > \pi_i(x)\}) =: \text{card}(\hat{J})$. By Definition 12.41,

$$\Delta_F(r_{x, \hat{x}}) = \sum_{J \subseteq \hat{J}} (-1)^{\text{card}(J)} F(x_J)$$

$$= (-1)^{n(x)} \int_{x_{0,1}}^{x_1} \cdots \int_{x_{0,m}}^{x_m} f(s_1, \ldots, s_m) \, ds_m \cdots ds_1$$

where $\bar{x}_J \in \bar{r}_{x, \hat{x}} \subseteq \mathbb{R}^m$ is defined by, $\forall i \in \hat{J},$

$$\pi_i(\bar{x}_J) = \left\{ \begin{array}{l} \pi_i(\hat{x}) \quad \forall i \in \hat{J} \\
\pi_i(\hat{x}) \quad \forall i \in \hat{J} \setminus J \end{array} \right.$$

$$= \left\{ \begin{array}{l} \pi_i(x) \quad \forall i \in (\hat{J} \cap J) \cup ((\hat{J} \setminus J) \cap (\hat{J} \setminus J)) \\
\pi_i(x_0) \quad \text{otherwise} \end{array} \right.$$

$$= \left\{ \begin{array}{l} \pi_i(x) \quad \forall i \in \hat{J} \setminus (\hat{J} \Delta J) \\
\pi_i(x_0) \quad \forall i \in \hat{J} \Delta \hat{J} \end{array} \right. = \pi_i(x_{J \Delta J})$$

This leads to

$$\Delta_F(r_{x, \hat{x}}) = \sum_{J \subseteq \hat{J}} (-1)^{\text{card}(J)} F(x_{J \Delta J})$$
First consider the special case when \( \mu(\Omega) < \infty \) and \( f \) is absolutely integrable over \( \mathbb{K} \). Since \( \Omega \) is a nonempty open rectangle, then \( \Omega \) is bounded. Define \( \bar{f} : \pi_\Omega(\Omega) \to \mathbb{Y} \) by \( \bar{f}(s, x) = (-1)^{|J|}f(M_J(s, \pi_J(x))) \chi_{\pi_J(x)}(x) \chi_{\pi_{\bar{J}}(\bar{x})}(s) \), \( \forall (s, x) \in \pi_\Omega(\Omega) \times \pi_{\bar{J}}(\bar{x}) \). By Propositions 7.23, 11.39, and 11.38, \( \bar{f} \) is \( \mathcal{B}_\mathbb{Y}(\mathbb{R}^{\text{card}(J)+m}) \)-measurable. Note

\[
\sum_{J \subseteq \bar{J}} (-1)^{|J|} \int_{x_0, \ldots, x_{\text{card}(J)}} f(s_1, \ldots, s_{\text{card}(J)}) \, ds_{\text{card}(J)} = (-1)^{|J|} \int_{x_0, \ldots, x_{\text{card}(J)}} f(s_1, \ldots, s_{\text{card}(J)}) \, ds_{\text{card}(J)}
\]

Then, we can conclude

\[
\sum_{J \subseteq \bar{J}} (-1)^{|J|} F(x_J) = \int_{x_0, \ldots, x_{\text{card}(J)}} f(s_1, \ldots, s_{\text{card}(J)}) \, ds_{\text{card}(J)}
\]

The above equation is equivalent to (12.1). This completes the proof of the proposition.

**Proposition 12.85** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_\mathbb{Y}(\mathbb{K}) \) be an open rectangle, \( \mathbb{X} := ((\Omega, | \cdot |), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathcal{B} \) be a separable Banach space, \( f : \Omega \to \mathcal{B} \) be \( \mathcal{B} \)-measurable, \( x_0 \in \Omega \), and \( \bar{J} := \{1, \ldots, m\} \). Assume that \( f \) is absolutely integrable over \( [x_a, x_b] \) with respect to \( \mu \), \( \forall x_a, x_b \in \Omega \) with \( x_a \leq x_b \). Define \( \pi_J : \mathbb{R}^m \to \mathbb{R}^{\text{card}(J)} \) by \( \pi_J(x) = (\pi_i(x))_{i \in J} \), \( \forall x \in \mathbb{R}^m \), \( \forall J \subseteq \bar{J} \); \( (\pi_J(\Omega), \mathcal{B}_J, \mu_J) \) be the measure subspace of \( \mathbb{R}^{\text{card}(J)} \), \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \); \( M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{\text{m-card}(J)} \to \mathbb{R}^m \) by \( M_J(s, t) \in \mathbb{R}^m \) is such that \( \pi_J(M_J(s, t)) = s \) and \( \pi_{J \setminus \bar{J}}(M_J(s, t)) = t \), \( \forall s \in \mathbb{R}^{\text{card}(J)} \), \( \forall t \in \mathbb{R}^{\text{m-card}(J)} \), \( \forall J \subseteq \bar{J} \). Then, the following statements hold.

1. \( \forall \emptyset \neq J \subseteq \bar{J}, \exists U_J \in \mathcal{B} \) with \( \mu(\Omega \setminus U_J) = 0 \) such that \( f(M_J(\cdot, \pi_J(x))) : \pi_J(\Omega) \to \mathcal{B} \) is absolutely integrable over \( \pi_J(x) \) with respect to \( \mu_J \), \( \forall x \in U_J \); and \( F : \Omega \to \mathcal{B} \), defined by, \( \forall x \in \Omega \), \( F(x) = \left\{ \begin{array}{ll} (-1)^{|J|} \int_{\pi_J(x)} f(M_J(s, \pi_J(x))) \, ds_J & x \in U_J \\ \emptyset & x \in \Omega \setminus U_J \end{array} \right. \), where \( \hat{x} := x_0 \wedge x \), \( \hat{x} := x_0 \lor x \), and \( \bar{J} := \{ i \in \bar{J} | \pi_i(x_0) > \pi_i(x) \} \), is \( \mathcal{B} \)-measurable.

2. Furthermore, if, in addition, \( \mathcal{B} \) is finite dimensional then \( U_J = \{ x \in \Omega | f(M_J(\cdot, \pi_J(x))) : \pi_J(\Omega) \to \mathcal{B} \) is absolutely integrable over \( \pi_J(\hat{x}) \) with respect to \( \mu_J \} \).

**Proof** First consider the special case when \( \mu(\Omega) < \infty \) and \( f \) is absolutely integrable over \( \mathbb{K} \). Since \( \Omega \) is a nonempty open rectangle, then \( \Omega \) is bounded. Define \( \bar{f} : \pi_\Omega(\Omega) \to \mathcal{B} \) by \( \bar{f}(s, x) = (-1)^{|J|}f(M_J(s, \pi_J(x))) \chi_{\pi_J(x)}(x) \chi_{\pi_{\bar{J}}(\bar{x})}(s) \), \( \forall (s, x) \in \pi_\Omega(\Omega) \times \pi_{\bar{J}}(\bar{x}) \). By Propositions 7.23, 11.39, and 11.38, \( \bar{f} \) is \( \mathcal{B}_\mathcal{B}(\mathbb{R}^{\text{card}(J)+m}) \)-measurable. Note
Theorem 12.31, this choice of special case on \( E \). By the assumption of the proposition, we can apply the result of the

\[
\int_{\pi_j(\Omega) \times \Omega} \mathcal{P} \circ f(s, x) \, d(\mu_J \times \mu)(s, x)
\]

\[
\leq \int_{\pi_j(\Omega)} \int_{\Omega} \| f(M_J(s, \pi_{J,j}(x))) \| \, d\mu(x) \, d\mu_J(s)
\]

\[
= \int_{\pi_j(\Omega)} \int_{\pi_{j,j}(\Omega)} \int_{\pi_j(\Omega)} \| f(M_J(s, t)) \| \, d\mu_J(t) \, d\mu_J(s)
\]

\[
\leq \mu_J(\pi_j(\Omega)) \int_{\pi_j(\Omega)} \int_{\pi_{j,j}(\Omega)} \| f(x) \| \, d\mu_{J_{j,j}}(\pi_{J,j}(x)) \, d\mu_J(\pi_j(x))
\]

\[
= \mu_J(\pi_j(\Omega)) \int_{\Omega} \mathcal{P} \circ f \, d\mu < \infty
\]

where the first inequality and the equalities follow from Tonelli’s Theorem 12.29; and the second inequality follows from Proposition 11.83. Then, we have that \( \bar{f} \) is absolutely integrable over \( \pi_j(\Omega) \times \Omega \) with respect to \( \mu_J \times \mu \). By Fubini’s Theorem 12.30, \( \exists p_j : \Omega \to \mathcal{Y} \) and \( U_j \in \mathcal{B} \) such that \( \mu(\Omega \setminus U_j) = 0 \); \( f(x) : \pi_j(\Omega) \to \mathcal{Y} \) is absolutely integrable over \((\pi_j(\Omega), \mathcal{B}_j, \mu_j)\), \( \forall x \in U_j \); \( p_j : \Omega \to \mathcal{Y} \), defined by, \( \forall x \in \Omega \), \( p_j(x) = \left\{ \begin{array}{ll}
\int_{\pi_j(\Omega)} f(s, x) \, d\mu_j(s) & x \in U_j \\
\bar{f}(x) & x \in \Omega \setminus U_j
\end{array} \right. \) \( \text{is } \mathcal{B}\)-measurable; \( p_j \) is absolutely integrable over \( X \), and \( \int_{\pi_j(\Omega) \times \Omega} f(s, x) \, d(\mu_J \times \mu)(s, x) = \int_{\Omega} p_j(x) \, d\mu(x) \). Clearly, \( F(x) = p_j(x) \), \( \forall x \in \Omega \). Hence, \( F \) is \( \mathcal{B} \)-measurable. Then, \( f(M_J(\cdot, \pi_{J,j}(x))) : \pi_j(\Omega) \to \mathcal{Y} \) is absolutely integrable over \( \pi_j(\Omega) \), \( \mathcal{B}_j, \mu_j \). Then, \( U_j := \{ x \in \Omega \mid f(x) : \pi_j(\Omega) \to \mathcal{Y} \text{ is absolutely integrable over } (\pi_j(\Omega), \mathcal{B}_j, \mu_j) \} \). By Fubini’s Theorem 12.31, this choice of \( U_j \) is admissible for the definition of \( p_J \) and \( p_j \) is \( \mathcal{B} \)-measurable. This proves the result in this special case.

Now consider the general case. Since \( \Omega \) is an open rectangle and \( x_0 \in \Omega \), then there exists \( (\bar{x}_j)_{j=1}^\infty, (\bar{x})_{j=1}^\infty \subseteq \Omega \) such that \( \cdots < \bar{x}_j < \cdots < \bar{x}_1 < x_0 < \bar{x}_1 < \cdots < \bar{x}_j < \cdots \), and \( \Omega = \bigcup_{j=1}^\infty E_j := \bigcup_{j=1}^\infty E_{\bar{x}_j, \bar{x}_j} \). Fix any \( j \in \mathbb{N} \). By the assumption of the proposition, we can apply the result of the special case on \( E_j \). There exists \( U_{j,j} \in \mathcal{B} \) with \( \mu(E_j \setminus U_{j,j}) = 0 \) such that \( f(M_J(\cdot, \pi_{J,j}(x))) : \pi_j(\Omega) \to \mathcal{Y} \) is absolutely integrable over \( \pi_j(\bar{x}_j, \bar{x}) \) with respect to \( \mu_j \), \( \forall x \in U_{j,j} \); and \( F_j : E_j \to \mathcal{Y} \), defined by, \( \forall x \in E_j \),

\[
F_j(x) = \left\{ \begin{array}{ll}
(-1)^{\text{card}(J \cap j)} \int_{\pi_{j}(\bar{x}_j, \bar{x})} f(M_J(s, \pi_{j,j}(x))) \, d\mu_J(s) & x \in U_{j,j} \\
\bar{f}(x) & x \in E_j \setminus U_{j,j}
\end{array} \right.
\]
12.6. FUNDAMENTAL THEOREM OF CALCULUS

is $\mathcal{B}$-measurable. Let $U_j := \bigcup_{j=1}^{\infty} U_{j,j} \in \mathcal{B}$. Then, $0 \leq \mu(\Omega \setminus U_j) = \mu(\bigcup_{j=1}^{\infty} (E_j \setminus U_{j,j})) \leq \mu(\bigcup_{j=1}^{\infty} (E_j \setminus U_{j,j})) \leq 0$. Define $F$ according to the statement of the proposition, we observe that $F(x) = F_j(x), \forall x \in U_{j,j} \setminus (\bigcup_{j=1}^{\infty} U_{j,j})$, if $x \in U_j; F(x) = \vartheta y_j, \forall x \in \Omega \setminus U_j$. By Proposition 11.41, $F$ is $\mathcal{B}$-measurable. Clearly, $f(M_j(\cdot, \pi_j(\cdot))(x) : \pi_j(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_j(r_{x,\hat{x}}) \subseteq \pi_j(\Omega)$ with respect to $\mu_j, \forall x \in U_j$.

If, in addition, $\mathcal{Y}$ is finite dimensional, then, by the special case, $U_j, j = \{x \in E_j \mid f(M_j(\cdot, \pi_j(\cdot))(x)) : \pi_j(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_j(r_{x,\hat{x}})$ with respect to $\mu_j\}, \forall j \in \mathbb{N}$. Then, $U_j = \bigcup_{j=1}^{\infty} U_{j,j} = \{x \in \Omega \mid f(M_j(\cdot, \pi_j(\cdot))(x)) : \pi_j(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_j(r_{x,\hat{x}})$ with respect to $\mu_j\}$. This completes the proof of the proposition.

Theorem 12.86 (Fundamental Theorem of Calculus 1) Let $m \in \mathbb{N}, \mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}\mathcal{B} (\mathbb{R}^m)$ be an open rectangle, $X := ((\Omega, \cdot), \mathcal{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a separable Banach space over $\mathfrak{K}, f : \Omega \rightarrow \mathcal{Y}$ be $\mathcal{B}$-measurable, $\hat{J} := \{1, \ldots, m\}, \text{ and } x_0, \hat{x} \in \Omega \in \hat{\Omega}$. Assume that $f$ is absolutely integrable over $r_{x_a,\hat{x}}$ with respect to $\mu, \forall x_a, \hat{x} \in \Omega$ with $x_a \lessgtr \hat{x}$. Let $\nu$ be the $\sigma$-finite $\mathcal{Y}$-valued measure with kernel $f$ over $X$ according to Proposition 11.116. Define $F : \Omega \rightarrow \mathcal{Y}$ to be the cumulative distribution function of $\nu$ with origin at $x_0$ according to Proposition 12.52. $F$ is absolutely continuous by Proposition 12.75, and $f = \frac{d\nu}{d\mu}$ a.e. in $X$. Let $D_i F : U_i \rightarrow \mathcal{Y}, \forall i \in \hat{J}$, where $D_i$ is the derivative of a function with respect to the $i$th coordinate variable of $\Omega, \forall i \in \hat{J}$, define $\pi_i : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$ by $\pi_i(x) = (\pi_i(x))_j \in \hat{J}\setminus\{i\}$, $\forall x \in \mathbb{R}^m$; also define $M_i : \mathbb{R} \times \mathbb{R}^{m-1} \rightarrow \mathbb{R}^m$ by $\pi_i(M_i(x, y)) = x$ and $\pi_i(M_i(x, y)) = y, \forall x \in \mathbb{R}, \forall y \in \mathbb{R}^{m-1}$. Then, the following statements hold.

(i) $F(x) = (-1)^{\text{card}(J)} \int_{r_{x,\hat{x}}} f \, d\mu, \forall x \in \Omega$, where $\hat{x} := x_0 \wedge x, \check{x} := x_0 \vee x$, and $J := \{i \in \hat{J} \mid \pi_i(x_0) > \pi_i(x)\}$.

(ii) $U_i \in \mathcal{B}, \mu(\Omega \setminus U_i) = 0, D_i F$ is $\mathcal{B}$-measurable; $\exists \tilde{U}_i \in \mathcal{B}$ with $\mu(\Omega \setminus \tilde{U}_i) = 0, \forall x \in \tilde{U}_i, f(M_i(\pi_i(x), \cdot)) : \pi_i(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_i(r_{x,\hat{x}})$ with respect to $\mu_i$, and $\exists p_i : \Omega \rightarrow \mathcal{Y}$ defined by, $\forall x \in \Omega, p_i(x) = \begin{cases} (-1)^{\text{card}(J) \setminus \{i\}} \int_{\pi_i(r_{x,\hat{x}})} f(M_i(\pi_i(x), s)) \, d\mu_i(s) & x \in \tilde{U}_i \\ \vartheta y & x \in \Omega \setminus \tilde{U}_i \end{cases}$

when $m > 1, p_i(x) = f(x), \forall x \in \Omega, \tilde{U}_i = \emptyset$, when $m = 1$; where $\pi_i(\Omega), \mathcal{B}_i, \mu_i$ is the measure subspace of $\mathbb{R}^{m-1}$; such that $p_i$ is $\mathcal{B}$-measurable, and $D_i F(x) = p_i(x)$ a.e. $x \in X, \forall i \in \hat{J}$.

(iii) If, in addition, $\mathcal{Y}$ is finite dimensional, then $\tilde{U}_i = \{x \in \Omega \mid f(M_i(\pi_i(x), \cdot)) : \pi_i(\Omega) \rightarrow \mathcal{Y}$ is absolutely integrable over $\pi_i(r_{x,\hat{x}})$ with respect to $\mu_i\}$. 
Proof. By Proposition 12.75, $F$ is absolutely continuous. By Definition 12.41 and Definition 11.166, we have $\Delta_F(\bar{x}, \bar{y}) = \nu(\bar{x}, \bar{y}) = \int_{\bar{x}}^\infty f \, d\mu$.

Since $\nu$ is the cumulative distribution function of $\nu$ with origin at $x_0$, then $\Delta_F(\bar{x}, \bar{y}) = (-1)^{\text{card}(J)} F(x)$. Thus, we have $F(x) = (-1)^{\text{card}(J)} \int_{\bar{x}}^\infty f \, d\mu$, $\forall x \in \Omega$. This establishes (i).

(ii) By Proposition 12.81, $D_i F$ is $\mathcal{B}$-measurable, $\forall i \in J$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $m = 1$; Case 2: $m > 1$. Case 1: $m = 1$. Then, $i = 1$. $F(x) = (-1)^{\text{card}(J)} \int_{\bar{x}}^\infty f \, d\mu = \int_{\bar{x}}^\infty f \, d\mu_B$, $\forall x \in \Omega$. Since $\Omega$ is a nonempty open interval, then $\exists (E_j)_{j=1}^\infty := ((a_j, b_j))_{j=1}^\infty \subseteq \mathcal{O}_\mathbb{R}$ such that $\Omega = \bigcup_{j=1}^\infty E_j$ and $-\infty < \cdots < a_n < \cdots < a_2 < a_1 < x_0 < b_1 < b_2 < \cdots < b_n < \cdots < \infty$.

Fix any $j \in \mathbb{N}$. Let $E_j := (E_j, B_j, \mu_j)$ be the finite measure subspace of $(\Omega, \mathcal{B}, \mu)$. Then, $f \in L_1(E_j, \mu)$. Let $c \in (0, 1)$ be a constant. Define $A_j := \{ x \in E_j \mid x \text{ is a rectangular Lebesgue point of } f \mid_{E_j} \text{ with regularity } c \}$. By Proposition 12.80, $A_j \in \mathcal{B}_L$ and $\mu_\mathbb{L}(E_j \setminus A_j) = 0$. Fix any $x \in A_j$.

We have $\limsup_{r \to 0^+} \frac{1}{r} \sup_{\substack{\bar{x}_1, \bar{x}_2 \in \mathbb{R} \text{ with } \frac{1}{2}(\bar{x}_1 + \bar{x}_2) = x, \frac{1}{2}|\bar{x}_2 - \bar{x}_1| = r, \frac{1}{2}|\bar{x}_2 - \bar{x}_1| \geq c r}} \int_{\bar{x}_1}^{\bar{x}_2} \| f(x) \| \, d\mu(y) = 0$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists r_0 \in (0, \infty) \subseteq \mathbb{R}$, $\forall r \in (0, r_0) \subseteq \mathbb{R}$, we have $0 \leq \frac{1}{r} \sup_{\substack{\bar{x}_1, \bar{x}_2 \in \mathbb{R} \text{ with } \frac{1}{2}(\bar{x}_1 + \bar{x}_2) = x, \frac{1}{2}|\bar{x}_2 - \bar{x}_1| = r, \frac{1}{2}|\bar{x}_2 - \bar{x}_1| \geq c r}} \int_{\bar{x}_1}^{\bar{x}_2} \| f(x) \| \, d\mu(y) < \epsilon$. Let $\bar{x}_1, \bar{x}_2 \subseteq \mathbb{R}$ be such that $\frac{1}{2}(\bar{x}_1 + \bar{x}_2) = x, \bar{x}_2 - x < r_0/2$, and $\bar{x}_1, \bar{x}_2 \subseteq E_j$, $\forall \bar{x} \in \bar{x}_1, \bar{x}_2$ with $\bar{x} \neq x$, we have, where $h := \bar{x} - x$,

\[
0 \leq \left\| F(\bar{x}) - F(x) - f(x)(\bar{x} - x) \right\| \\
= \left\| \int_{x}^{x+h} f(y) \, d\mu(y) - \int_{\bar{x}}^{\bar{x}+h} f(x) \, d\mu(y) \right\| \\
= \left\| \int_{x}^{x+h} (f(y) - f(x)) \, d\mu(y) \right\| \\
= \left\| \int_{\bar{x} \wedge (x+h), x \vee (x+h)} (f(y) - f(x)) \, d\mu(y) \right\| \\
\leq \int_{x-|h|, x+|h|} \| f(y) - f(x) \| \, d\mu(y) \\
= \frac{1}{\mu_B(x-|h|, x+|h|)} \int_{x-|h|, x+|h|} \| f(y) - f(x) \| \, d\mu_B(y) \\
\leq 2 |\bar{x} - x|$

where the first equality follows from Fact 12.72 and Proposition 11.75; the second equality follows from Proposition 11.92; the third equality follows from Definition 12.71; the second inequality follows from Proposition 11.92; and the last inequality follows from the fact $x \in A_j$ and the choice of $r_0$. Then, $x \in \text{dom}(D_1 F) = U_1$ and $D_1 F(x) = f(x)$. By the arbitrariness
of \( x \), we have \( A_j \subseteq U_1 \). Clearly, \( \bigcup_{j=1}^{\infty} A_j \subseteq U_1 \subseteq \Omega \). This implies that
\[
0 \leq \mu(\Omega \setminus U_1) = \mu_B(\Omega \setminus U_1) = \mu_L(\Omega \setminus U_1) \leq \mu_L(\bigcup_{j=1}^{\infty} (E_j \setminus \{A_j\})) \leq \mu_L(\bigcup_{j=1}^{\infty} (E_j \setminus A_j)) = 0. \]
Hence, \( \mu(\Omega \setminus U_1) = 0 \). By \( D_1F \) being \( B \)-measurable and \( f \) being \( B \)-measurable, we have \( D_1F(x) = f(x) \) a.e. \( x \in \Omega \). This case is proved.

Case 2: \( m > 1 \). Without loss of generality, consider \( i = 1 \). Let \( X_1 := (\pi_1(\Omega), B, \bar{\mu}) \) be the measure subspace of \( \mathbb{R} \). By Proposition 12.85, \( \exists U_1 \in B \) with \( \mu(\Omega \setminus U_1) = 0 \), and \( p_1 \) as defined in the statement of the theorem, such that \( p_1 \) is \( B \)-measurable. Let \( E := \Omega \setminus U_1 \). By Lemma 12.27, \( \exists U_{1,1} \in B \) with \( \mu_1(\Omega \setminus U_{1,1}) = 0 \) such that \( \bar{\mu}(E_{\{\pi_1(x)\}1}) = 0, \forall \pi_1(x) \in U_{1,1} \), where
\[
E_{\{\pi_1(x)\}} := \{ s \in \pi_1(\Omega) \mid M_1(s, \pi_1(x)) \in E \}.
\]

By Fubini’s Theorem 12.30, we have \( F(x) = \int_{\pi_1(x)} p_1(M_1(s, \pi_1(x))) \, d\bar{\mu}(s), \forall x \in \Omega \) with \( \pi_1(x) \in U_{1,1} \).

Then, by Case 1, we have \( D_1F(x) = p_1(x) \) a.e. \( \pi_1(x) \in X_1 \), \( \forall \pi_1(x) \in U_{1,1} \). Since \( D_1F \) and \( p_1 \) are \( B \)-measurable and \( \bar{\mu} \) is separable, then, by Propositions 7.23, 11.39, and 11.38, \( D_1F - p_1 \) is \( B \)-measurable. Then, \( U_1 := \{ x \in \Omega \mid D_1F(x) = p_1(x) \} \in B \). By Lemma 12.27, \( \mu(\Omega \setminus U_1) = 0 \):
\[
\mu(E) = \int_{\pi_1(x)} \bar{\mu}(E_{\{\pi_1(x)\}}) \, d\mu_1(\pi_1(x)) = \int_{U_{1,1}} \bar{\mu}(E_{\{\pi_1(x)\}}) \, d\mu_1(\pi_1(x)) = \int_{U_{1,1}} 0 \, d\mu_1 = 0,
\]
where \( E_{\{\pi_1(x)\}} := \{ s \in \pi_1(\Omega) \mid M_1(s, \pi_1(x)) \in E \} \) and \( \bar{\mu}(E_{\{\pi_1(x)\}}) = 0 \) follows from \( D_1F(x) = p_1(x) \) a.e. \( \pi_1(x) \in X_1 \), \( \forall \pi_1(x) \in U_{1,1} \). Hence, \( D_1F = p_1 \) a.e. in \( X_1 \), and \( U_1 \in B \) with \( \mu(\Omega \setminus U_1) = 0 \). This case is also proved.

In both cases, (ii) holds. Hence, (ii) is true.

(iii) This follows immediately from (ii) and Proposition 12.85.

This completes the proof of the theorem. \( \square \)

Proposition 12.87 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}(\mathbb{R}^m) \) be region, \( Z \) be a normed linear space, \( G: \Omega \to Z \), \( \vec{J} := \{1, \ldots, m\} \), and \( x_0 \in \Omega \). \( \forall J \subseteq \vec{J} \), define \( \pi_J: \mathbb{R}^m \to \mathbb{R}^{\text{card}(J)} \) by \( \pi_J(x) = (\pi_J(x))_{\vec{J} \setminus J} \), \( \forall x \in \mathbb{R}^m \); define \( M_J: \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \to \mathbb{R}^m \) by \( M_J(s, t) \in \mathbb{R}^m \) is such that \( \pi_J(M_J(s, t)) = s \) and \( \pi_J, J(M_J(s, t)) = t \), \( \forall s \in \mathbb{R}^{\text{card}(J)}, \forall t \in \mathbb{R}^{m-\text{card}(J)} \); also define \( G_J: \pi_J(\Omega) \to \mathbb{Y} \) by \( G_J(s) = G(M_J(s, \pi_J(x_0))), \forall s \in \pi_J(\Omega) \). Then, we have, \( \forall x \in \Omega \),
\[
(-1)^{\text{card}(J)} \Delta_G(x_\hat{\pi}, \hat{x}) = G(x) - G(x_0) - \sum_{J \subset J, \vec{J} \neq J} (-1)^{\text{card}(J \cap J)} \Delta_G_J(\pi_J(x_\hat{\pi}, \hat{x})) \quad (12.2)
\]
where \( \hat{x} := x \land x_0, \hat{x} := x \lor x_0; \) and \( \vec{J} := \{ i \in \vec{J} \mid \pi_i(x_0) > \pi_i(x) \} \).

Proof We will prove the result using mathematical induction on \( m \).

1° \( m = 1 \). This is obvious.

2° Assume the claim holds for \( m \leq k \in \mathbb{N} \).
Consider the case when \( m = k + 1 \in \{2, 3, \ldots\} \). Fix any \( x \in \Omega \).
Define \( \bar{G} : \pi_j \cup \{ \pi_{k+1} \}(\Omega) \to \mathbb{Z} \) by
\[
\bar{G}(s) = G(M_j \cup \{ \pi_{k+1}(x) \}), \quad \forall s \in \pi_j \cup \{ \pi_{k+1} \}(\Omega).
\]
Define \( \bar{G}_j : \pi_j(\Omega) \to \mathbb{Z} \) in terms of \( G \), \( \forall J \subseteq \bar{J} \setminus \{ k + 1 \} \). Then, we have the following sequence of arguments.

\[
(-1)^{\text{card}(J)} \Delta G_{\{\bar{x}, \bar{x}\}} = (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \left( \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}} \right) - \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
- \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
\]

\[
= G(x) - G_{\{\pi_{k+1} \}}(x) - \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
- \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
\]

\[
= G(x) - G_{\{\pi_{k+1} \}}(x) - \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
- \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
\]

\[
= G(x) - G_{\{\pi_{k+1} \}}(x) - \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
- \sum_{J \subset J \cup \{ \pi_{k+1} \}, J \neq \emptyset} (-1)^{\text{card}(J \cup \{ \pi_{k+1} \})} \Delta G_{\{\pi_j \cup \{ \pi_{k+1} \} \}}(x_0) \\
\]
where the first equality follows from Definition 12.41; the second equality follows from the inductive assumption 2; the fourth equality follows from Definition 12.41. This establishes the claim in this step, and completes the induction process.

This completes the proof of the proposition. □

Theorem 12.88 (Fundamental Theorem of Calculus II) Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B} (\mathbb{R}^m) \) be an open rectangle, \( X := ((\Omega, 1 \cdot 1), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathcal{Y} \) be a separable reflexive Banach space over \( K \) with \( \mathcal{Y}^* \) being separable, \( F : \Omega \to \mathcal{Y} \) be absolutely continuous, \( \bar{J} := \{1, \ldots, m\} \), \( F^{(1)} : U \to \mathbb{B}(\mathbb{R}^m, \mathcal{Y}), D_i F : U_i \to \mathcal{Y}, \forall i \in \bar{J} \), where \( D_i \) is the derivative of a function with respect to the \( i \)-th coordinate variable of \( \Omega \), and \( x_0 \in \Omega \). \( \forall J \subseteq \bar{J} \), define \( \pi_J : \mathbb{R}^m \to \mathbb{R}^{\text{card}(J)} \) by \( \pi_J(x) = (\pi_i(x))_{i \in J} \), \( \forall x \in \mathbb{R}^m \); define \( M_J : \mathbb{R}^\text{card}(J) \times \mathbb{R}^{m-\text{card}(J)} \to \mathbb{R}^m \) by \( M_J(s, t) \in \mathbb{R}^m \) is such that \( \pi_J(M_J(s, t)) = s \) and \( \pi_{J \setminus J}(M_J(s, t)) = t \), \( \forall s \in \mathbb{R}^{\text{card}(J)}, \forall t \in \mathbb{R}^{m-\text{card}(J)} \); also define \( F_J : \Omega \to \mathcal{Y} \) by \( F_J(s) = F(M_J(s, \pi_{J \setminus J}(x_0))) \), \( \forall s \in \pi_J(\Omega) \); and let \( \mathcal{X}_J := (\pi_J(\Omega), \mathcal{B}, \mu_J) \) be the measure subspace of \( \mathbb{R}^{\text{card}(J)} \) when \( J \neq \emptyset \). Then, the following statements hold.

(i) There exists a set of functions \( f_{J, x_0} : \pi_J(\Omega) \to \mathcal{Y} \), which is \( \mathcal{B}_J \)-measurable, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), such that \( f_{J, x_0}'s \) are absolutely integrable on \( \pi_J(r_{x_0, x_0}) \) with respect to \( \mu_J \), \( \forall x_a, x_b \in \Omega \) with \( x_a \leq x_b \), and we have, \( \forall x \in \Omega \),

\[
F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(J \setminus J)} \int_{\pi_J(r_{\hat{x}, \hat{x}})} f_{J, x_0} \, d\mu_J + F(x_0) \quad (12.3)
\]

where \( \hat{x} := x \land x_0, \hat{x} := x \lor x_0 \); and \( \bar{J} := \{ i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x) \} \).

The set of functions \( \{ f_{J, x_0} \}_{J \subseteq \bar{J}, J \neq \emptyset} \) is unique in the sense that if \( \{ g_{J, x_0} \}_{J \subseteq \bar{J}, J \neq \emptyset} \) is another set of such functions satisfying (12.3), then \( f_{J, x_0} = g_{J, x_0} \) a.e. in \( \mathcal{X}_J, \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \).

(ii) If we expand absolutely continuous function \( F \) at \( \hat{x} \in \Omega \), instead of \( x_0 \), then we will have a set of functions \( f_{J, \hat{x}} : \pi_J(\Omega) \to \mathcal{Y} \), which is \( \mathcal{B}_J \)-measurable, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), satisfying the conclusion of (i) with \( x_0 \) replaced by \( \hat{x} \). The \( f_{J, \hat{x}}'s \) relate to \( f_{J, x_0}'s \) according to the following formula, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), \( \exists U_J \in \mathcal{B}_J \) with \( \mu_J(\pi_J(\Omega) \setminus U_J) = 0 \),

such that

\[
f_{J, \hat{x}}(s) = \sum_{J \subseteq \bar{J}} \sum_{J \subseteq \bar{J}} \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \setminus J)} \int_{\pi_J(r_{\hat{x}, \hat{x}})} f_{J, x_0}(\pi_J(M_J(s, \pi_{J \setminus J}(t)))) \, d\mu_{J \setminus J}(\pi_{J \setminus J}(t)) + f_{J, x_0}(s), \quad \forall s \in \bar{U}_J \quad (12.4)
\]

and \( f_{J, \hat{x}}(s) = \varnothing \), \( \forall s \in \pi_J(\Omega) \setminus \bar{U}_J \), where \( \hat{x} := x \land x_0, \hat{x} := x \lor x_0 \), and \( \bar{J} := \{ i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x) \} \); and all involved integrations are
such that the integrands are absolutely integrable over the specified integration domain with respect to the given measures, respectively.

(iii) \( U_i \in \mathcal{B}, \mu(\Omega \setminus U_i) = 0, D_i F \) is \( \mathcal{B} \)-measurable, and, \( \forall i \in \mathcal{B} \),
\[
D_i F(x) = f_{i(x)}(\pi_i(x)) \quad \text{a.e.} \quad \pi_i(x) \in \mathcal{X}_{\{i\}}, \quad \forall \pi_{J \setminus \{i\}}(x) \in \mathcal{X}_{J \setminus \{i\}},
\]
where \( x_i := M_{\{i\}}(\pi_i(x_0), \pi_{J \setminus \{i\}}(x)) \).

(iv) \( U \subseteq \bigcap_{i=1}^{\infty} U_i, U \in \mathcal{B}, F^{(1)} \) is \( \mathcal{B} \)-measurable, and, \( \forall x \in U \), we have \( F^{(1)}(x) = \left[ D_1 F(x) \cdots D_m F(x) \right] \). If, in addition, \( \forall J \subseteq J \) with \( J \neq \emptyset \), \( \forall x \in \Omega, \exists \delta(J, x) \in (0,\infty) \subset \mathbb{R}, \exists c(J, x) \in [0,\infty) \subset \mathbb{R} \) such that \( \text{ess sup}_{y \in \mathbb{B}_{\mathbb{R}^m}(x, \delta(J, x))} \| f_{J,x_0}(\pi_J(y)) \| \leq c(J, x) \) and \( \mathbb{B}_{\mathbb{R}^m}(x, \delta(J, x)) \subseteq \Omega \), then \( \mu(\Omega \setminus U) = 0 \).

(v) If, in addition, \( Y \) is finite dimensional, then \( \hat{U}_J = \left\{ s \in \pi_J(\Omega) \mid f_{J,x_0}(\pi_J(M_{J}(s, \pi_{J \setminus J}(t)))) \right\} \) is absolutely integrable over \( \pi_{J \setminus J}(\pi_{\hat{J}}(\hat{x})) \) with respect to \( \mu_{J \setminus J}(\pi_{J \setminus J}(t)), \forall \hat{J} \subseteq \hat{J} \), \( \forall J \subseteq \hat{J} \) with \( J \neq \emptyset \).

In the rest of this book, \( f_{J,x_0} \) will be called the stream of \( F \) in \( J \) coordinate subspace with respect to \( x_0 \).

**Proof** (i) By \( F = F_J \) being absolutely continuous and Definition 12.41, \( \forall J \subseteq \hat{J} \) with \( J \neq \emptyset \), \( F_J \) is absolutely continuous. By Proposition 12.73, there exists a unique \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu_J \) on \( (\pi_J(\Omega), \mathcal{B}_J) \) such that \( F_J \) is a cumulative distribution function of \( \nu_J \). Furthermore, \( \mathcal{P} \circ \nu_J \ll \mu_J \). Since \( \mathcal{Y} \) is a separable reflexive Banach space with \( \mathcal{Y}^* \) being separable, then, by Radon-Nikodym Theorem 11.171, there exists a unique \( f_{J,x_0} : \pi_J(\Omega) \to \mathcal{Y} \), which is \( \mathcal{B}_J \)-measurable, such that \( f_{J,x_0} = \frac{d\nu_J}{d\mu_J} \) a.e. in \( \mathcal{X}_J \). \( \forall x_a, x_b \in \Omega \) with \( x_a \leq x_b \), we have \( \mathcal{P} \circ \nu_J(\pi_J(x_a, x_b)) = \mathcal{P} \circ f_{J,x_0} \, d\mu_J = T_{F_J}(\pi_J(x_a, x_b)) < \infty \), where the first equality follows from Definition 11.166; the second equality follows from Proposition 12.73; and the inequality follows from the fact \( F_J \) is absolutely continuous and therefore of locally bounded variation which implies this according to Definition 12.41. Hence, \( f_{J,x_0} \) is absolutely integrable over \( \pi_J(x_a, x_b) \) with respect to \( \mu_J \). Then, \( \forall x \in \Omega \), we have \( \Delta_{F_J}(\pi_J(x_a, x_b)) = \nu_J(\pi_J(x_a, x_b)) = \int_{\pi_J(x_a, x_b)} f_{J,x_0}(s) \, d\mu_J(s) \). By Proposition 12.87, we have
\[
(-1)^{\text{card}(J)} \Delta_{F_J}(x_a, x_b) = F(x) - F(x_0) - \sum_{J \subseteq \hat{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap J)} \Delta_{F_J}(\pi_J(x_a, x_b))
\]
which implies that
\[
(-1)^{\text{card}(J)} \int_{\pi_J(x_a, x_b)} f_{J,x_0}(s) \, d\mu_J(s)
\]
This is because rearranging the above formula, we have
\[
F(x) = F(x_0) + \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, x_0}(s) \, d\mu_J(s)
\]

Reorganizing the above formula, we have
\[
F(x) = F(x_0) - \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, x_0}(s) \, d\mu_J(s)
\]

This establishes (12.3). The uniqueness of the set of functions \((f_{j, x_0})_{J \subseteq J, J \neq \emptyset}\) is clear from the uniqueness of \(F_j\)'s and the uniqueness of Radon-Nikodym derivatives. Hence, (i) holds.

(ii) By (i), there exists a set of functions \(f_{j, \hat{x}} : \pi J(\Omega) \to Y\), which is \(\mathcal{B}\)-measurable, \(\forall J \subseteq J\) with \(J \neq \emptyset\), such that \(f_{j, \hat{x}}\)'s are absolutely integrable on \(\pi J(x_0, x)\) with respect to \(\mu_J\), \(\forall x, x_0 \in \Omega\) with \(x_0 \leq x\), and we have
\[
F(x) = \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, \hat{x}} \, d\mu_J + F(\hat{x}),
\]
where \(\hat{x} := x \setminus J\), \(\bar{x} := x \setminus J\), and \(\hat{J} := \{ i \in J \mid \pi_i(\hat{x}) > \pi_i(x) \}\). Now, let \(x = \hat{x} + M_J(s, \{0_{\text{card}(J')}, \forall J \subseteq J\} \cup \{1\})\), \(\forall J \subseteq J\) with \(J \neq \emptyset\). By the preceding formula, we have
\[
F(x) = \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, \hat{x}} \, d\mu_J + F(\hat{x}).
\]
This is because \(\forall J \subseteq J\), the domain of integration will be an empty set and that leads to \(\partial_y\) for that integral. Now, substitute the formula for \(F(x)\) and \(F(\hat{x})\) using (12.3), we have
\[
\sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, \hat{x}} \, d\mu_J = \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, \hat{x}} \, d\mu_J
\]

Rearrange terms in the above, we obtain
\[
\sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, \hat{x}} \, d\mu_J = \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, x_0} \, d\mu_J - \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap J')} \int_{\pi J(x)} f_{j, x_0} \, d\mu_J
\]
\[
= \prod_{J \subseteq J, \hat{J} \neq \emptyset} \left( \int_{\pi J(x_0)} f_{j, x_0}(s_1, \ldots, s_{\text{card}(J)}) \, d\mu_B(s_{\text{card}(J)}) \cdots d\mu_B(s_1) \right)
\]
\[
- \prod_{J \subseteq J, \hat{J} \neq \emptyset} \left( \int_{\pi J(x_0)} f_{j, x_0}(s_1, \ldots, s_{\text{card}(J)}) \, d\mu_B(s_{\text{card}(J)}) \cdots d\mu_B(s_1) \right)
\]
\[
\begin{align*}
\int_{J} & \left( \prod_{j \in J} \left( \int_{\pi_j(x)} \sum_{s_{\text{card}(j)}} f_{J,x_0}(s_1, \ldots, s_{\text{card}(j)}) \right) \right) \\
& \cdot d\mu_B(s_{\text{card}(j)}) \cdots d\mu_B(s_1) - \sum_{J \subseteq J} \prod_{j \notin J} \left( \int_{\pi_j(x)} \right) \\
& f_{J,x_0}(s_1, \ldots, s_{\text{card}(j)}) d\mu_B(s_{\text{card}(j)}) \cdots d\mu_B(s_1) \\
& = \sum_{J \subseteq J} \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& = \sum_{J \subseteq J} \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& = \sum_{J \subseteq J} \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& \cdot \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& = \sum_{J \subseteq J} \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& \cdot \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& = \sum_{J \subseteq J} \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& \cdot \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
& \cdot \sum_{J \subseteq J} (-1)^{\text{card}(J \cap j)} + \text{card}(J \cap \hat{j}) \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \\
\end{align*}
\]

where \( \hat{x}_{1,x,x_0} := x_{1,1,x,x_0} \cap x_{2,1,x,x_0}; \hat{x}_{1,x,x_0} := x_{1,1,x,x_0} \cup x_{2,1,x,x_0}; \)
\( x_{1,j,x,x_0} := M_j(\pi_j(x), \pi_j(x_0)); x_{2,j,x,x_0} := M_j(\pi_j(x), \pi_j(x_0)); \)
the second to fifth equalities follow from intuition, we have not used rigorous notation in these equations in expressing the joint integrals as repeated integrals and the other way around, this should not cause concern since from the first equality to the fifth, all that we have done is breaking down \((1)^{\text{card}(J \cap j)} \int_{\pi_j(r_{\hat{x},x,x_0})} \cdot d\mu_j \)
into pieces along the intermediate point \( \hat{x} \); the sixth equality follows from the idea that the cor-
responding integral will be zero if its domain of integration is an empty set, which will happen since \( \pi_i(x) = \pi_i(x), \forall i \in J \setminus J; \) the seventh equality follows from the interchange of summing variables \( J \) and \( J; \) the ninth equality follows from Fubini’s Theorem 12.30, and \( \hat{U}_{j,J} \in B_j \) with \( \mu_j(\pi_j(\Omega) \setminus \hat{U}_{j,J}) = 0, \) and \( p_{j,J} : \pi_j(\Omega) \to Y, \forall J \subseteq J, \) defined by \( p_{j,J}(s) := \int_{\pi_j,J} f_{J,x_0}(\pi_j(M_j(s, \pi_j,J(t)))) d\mu_{j,J}(\pi_j,J(t)) \) \( \forall s \in \hat{U}_{j,J} \) \( s \in \pi_j(\Omega) \setminus \hat{U}_{j,J} \) is \( B_j \)-measurable; and the tenth equality follows from Proposition 11.92. Let \( \hat{U}_J := \bigcap_{J \subseteq J} \hat{U}_{j,J}, \forall J \subseteq J \) with \( \hat{J} \neq \emptyset. \) Clearly, \( \hat{U}_J \in B_j \) and \( \mu_j(\pi_j(\Omega) \setminus \hat{U}_J) = 0. \) The definition of \( f_i,x \)'s (12.4) satisfies the above equalities, \( \forall J \subseteq J \) with \( J \neq \emptyset. \) Then, by (i), this set of functions \( (f_i,x)_{J \subseteq J, J \neq \emptyset} \) is the unique one we seek.

(iii) Fix any \( i \in J. \) Without loss of generality, let \( i = 1. \) By Proposition 12.81, \( U_1 \in B \) and \( D_1 F \) is \( B \)-measurable. By Fundamental Theorem of Calculus 12.86, every term on the right-hand-side of (12.3) is differentiable with respect to the \( i \)th coordinate variable of \( \Omega \) almost everywhere in \( X. \) \( \forall J \subseteq J \) with \( 1 \in J, \exists U_{i,J} \in B \) with \( \mu(\Omega \setminus U_{i,J}) = 0, \) \( D_1 \left( (-1)^{\text{card}(J \cap J)} \int_{\pi_j,J \cap (r,x)} f_{J,x_0} d\mu_J \right) : U_{i,J} \to Y \) is \( B \)-measurable, \( \exists \pi_{i,J} \in B \) with \( \mu(\Omega \setminus \pi_{i,J}) = 0, \) and \( \exists p_{i,J} : \Omega \to Y \) defined by \( p_{i,J} = \left\{ \begin{array}{ll} (-1)^{\text{card}(J \cap J)} \int_{\pi_j,J \cap (r,x)} f_{J,x_0} d\mu_J \\

\vartheta_y \\

, \pi_j,J \cap (r,x) \in B_j \end{array} \right. \) \( x \in \pi_{i,J}, \forall x \in \Omega \setminus \pi_{i,J} \) \( \pi_{i,J} \cap (r,x) \in B_j \) \( x \in \Omega, \) when \( J \supseteq \{1\}, p_{i,J} = f_{J,x_0}(\pi_j(x)), \forall x \in \Omega, \) \( \pi_{i,J} = \Omega, \) when \( J = \{1\}, \) such that \( p_{i,J} \) is \( B \)-measurable and \( D_1 \left( (-1)^{\text{card}(J \cap J)} \int_{\pi_j,J \cap (r,x)} f_{J,x_0} d\mu_J \right) = p_{i,J}(x) \) \( a.e., x \in X. \) \( \forall J \subseteq J \) with \( 1 \notin J, \) we have \( D_1 \left( (-1)^{\text{card}(J \cap J)} \int_{\pi_j,J \cap (r,x)} f_{J,x_0} d\mu_J \right) = \vartheta_y, \forall x \in \Omega. \) In this case, we have \( \hat{U}_{i,J} = \Omega. \) Let \( \hat{U}_J := \bigcap_{J \subseteq J, J \neq \emptyset} \hat{U}_{i,J} \). Then, \( \hat{U}_J \in B \) and \( \mu(\Omega \setminus \hat{U}_J) = 0. \) Hence, \( D_1 F(x) = \sum_{J \subseteq J, J \neq \emptyset} p_{i,J}(x), \forall x \in \hat{U}_1, \) and \( D_1 F(x) = \sum_{J \subseteq J, J \neq \emptyset} p_{i,J}(x) \) \( a.e., x \in X. \) This implies that \( \mu(\Omega \setminus U_1) = 0. \) Now, expand the function \( F \) at \( x_1 \) (which is as defined in the statement of the theorem.) By (12.3), we have \( F(x) = F(x_1) + \int_{\pi_1,J \cap (r,x)} f_{J,x_0} d\mu_J, \forall x = x_1 + s \in \pi_{J \setminus 1} \) \( \in \Omega \) with \( s \in \mathbb{R}. \) By Fundamental Theorem of Calculus I, Theorem 12.86, we have \( D_1 F(x) = f_{1,J \setminus 1}(\pi_{1,J \setminus 1}(x)) \) \( a.e., \pi_{1,J \setminus 1}(x) \in X_{\pi_{1,J \setminus 1}(x)}, \forall \pi_{1,J \setminus 1}(x) \in X_{\pi_{1,J \setminus 1}(x)}. \) Thus, (iii) holds.

(iv) By Proposition 12.81, \( U \subseteq \bigcap_{i=1}^m U_i, \) \( F^{(i)} \) is \( B \)-measurable, and \( \forall x \in U, \) we have \( F^{(i)}(x) = \left[ D_1 F(x) \cdots D_m F(x) \right]. \) All we need to show is that \( \mu(\Omega \setminus U) = 0 \) under the additional assumption that \( f_{i,J \cap J} \)'s are locally bounded.

\( \forall i \in J, \forall J \subseteq J \) with \( i \in J, \) by the proof of (iii), \( \exists \hat{U}_{i,J} \in B \) with \( \mu(\Omega \setminus \hat{U}_{i,J}) = 0, \) such that \( D_i H_J(x) := D_i \left( (-1)^{\text{card}(J \cap J)} \int_{\pi_j,J \cap (r,x)} f_{J,x_0} d\mu_J \right) (x) \)
$= p_{J,i}(x)$, $\forall x \in \bar{U}_{J,i}$. Then, let $\bar{U}_{J,i} := U_{J,i} \cap \bar{U}_{J,i} \cap \bar{U}_{J,i} \in B$, we have $\mu(B \setminus \bar{U}_{J,i}) = 0$ and $D_i H_J(x) = (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{J,i}(\pi_J(M_{i}(\pi_i(x), \pi_{J \setminus \{i\}}(s)))) d\mu_{J \setminus \{i\}}(\pi_J(\pi_{J \setminus \{i\}}(s))) =: H_{J,i}(x), \forall x \in \bar{U}_{J,i}$.

Let $\bar{U} := \left( \bigcap_{i=1}^{m} \bigcap_{J \subseteq \bar{J}} \bar{U}_{J,i} \right) \cap \left( \bigcap_{J \subseteq \bar{J}, J \neq \emptyset} \bar{U}_{J,i} \right) \in B$. Clearly, $\mu(B \setminus \bar{U}) = 0$. Fix any $\bar{x} \in \bar{U}$. Then, by (ii), there exists a unique set of functions $(f_{J,i})_{J \subseteq \bar{J}, J \neq \emptyset}$ such that $F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{J,i}(\pi_J(M_{i}(\pi_i(x), \pi_{J \setminus \{i\}}(s)))) d\mu_{J \setminus \{i\}}(\pi_J(\pi_{J \setminus \{i\}}(s))) + \bar{H}_J(x), \forall x \in \bar{U}$. Furthermore, $f_{J,i}$'s are related to $\bar{f}_{J,i}$'s by (12.4). Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By $f_{J,i}$'s being locally bounded on their domain of definition and (12.4), we can conclude that $f_{J,i}$'s are also locally bounded on their domain of definition. $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, $\forall x \in B$, $\exists \bar{c}(J, x) \in [0, \infty) \subset \mathbb{R}$ such that $\text{ess sup}_{y \in B_{\bar{U}}(x, \bar{c}(J,x))} \| f_{J,i}(\pi_J(y)) \| \leq \bar{c}(J, x)$ and $B_{\bar{U}}(x, \bar{c}(J, x)) \subseteq B_{\bar{U}}$.

Let $\delta_0 := \min_{J \subseteq \bar{J}, J \neq \emptyset} \delta(J, x) \in (0, \infty) \subset \mathbb{R}$ and $\bar{c}_0 := \max_{J \subseteq \bar{J}, J \neq \emptyset} \bar{c}(J, x) \in [0, \infty) \subset \mathbb{R}$. Then, $\text{ess sup}_{y \in B_{\bar{U}}(x, \delta_0)} \| f_{J,i}(\pi_J(y)) \| \leq \bar{c}_0$ and $B_{\bar{U}}(x, \delta_0) \subseteq \Omega$. $\forall i \in \bar{J}$, we have

$$D_i F(x) = f_{i}(x) \cdot \pi_i(x) + \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{J,i}(\pi_J(M_{i}(\pi_i(x), \pi_{J \setminus \{i\}}(s)))) \cdot d\mu_{J \setminus \{i\}}(\pi_J(\pi_{J \setminus \{i\}}(s)))$$

where the first equality follows from the preceding discussion and the second equality follows from (12.4), $\forall i \in \bar{J}, \forall J \subseteq \bar{J}$ with $i \in J$, $\exists h_i(x) \in (0, \infty) \subset \mathbb{R}$ such that $\| H_J(x + h_i e_{m_i}) - H_J(x) - h_i \| \leq m^{-1} 2^{-m} | h_i |$. Let $\delta_1 := \min_{J \subseteq \bar{J}} \min_{J \subseteq \bar{J}} \delta_J(x) \wedge \delta_0 \wedge \frac{\epsilon}{2(1 + \epsilon)} \wedge \frac{1}{2(1 + \epsilon)} \wedge 1 \in (0, \infty) \subset \mathbb{R}$. This leads to, $\forall x \in B_{\bar{U}}(x, \delta_1)$,

$$\| F(x) - F(x) - \sum_{i=1}^{m} D_i F(x) \pi_i(x - x) \|$$

$$= \| \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{J,i} \cdot d\mu_J - \sum_{i=1}^{m} f_{i}(x) \cdot \pi_i(x)$$

$$- \bar{x} \| = \| \sum_{J \subseteq \bar{J}, \text{card}(J) \geq 2} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{J,i} \cdot d\mu_J$$

$$+ \sum_{i=1}^{m} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{i}(x) \cdot d\mu_{i}(t)$$

$$- \sum_{i=1}^{m} (-1)^{\text{card}(\bar{J} \cap \{i\})} \int_{\bar{J} \setminus \{i\}} f_{i}(x) \cdot \pi_i(\bar{x}) \cdot d\mu_{i}(t) \|$$
where the first equality follows from the above discussion; the first inequality follows from Proposition 11.92; the second inequality follows from Proposition 11.92; and the third inequality follows from choice of \( \delta_1 \) and the preceding discussion. Then, \( \bar{x} \in \text{dom}(F(\bar{\cdot})) \). By the arbitrariness of \( \bar{x} \), we have \( \bar{U} \subseteq \text{dom}(F(\bar{\cdot})) = U \). Hence, \( \mu(\Omega \setminus U) = 0 \).

(v) This is a direct consequence of Fubini’s Theorem 12.31 and the proof of (ii).

This completes the proof of the theorem. \( \square \)

**Theorem 12.89 (Integration by Parts)** Let \( I \subseteq \mathbb{R} \) be a nonempty open or closed interval, \( \mathbb{I} := ((P(I), \cdot, \cdot), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \), \( Y \) be a separable reflexive Banach space over \( K \), \( Z \) be a separable reflexive Banach space over \( K \) with \( Z^* \) being separable, and \( W := (Y, Z) \) be separable, \( a : P(I) \to W \) be \( \mathcal{B} \)-measurable and absolutely integrable over \( \mathbb{I} \), \( \nu_A \) the \( W \)-valued measure with kernel \( a \) over \( \mathbb{I} \), and \( A : I \to W \) be a cumulative distribution function for \( \nu_A \), \( f : P(I) \to Y \) be \( \mathcal{B} \)-measurable and absolutely integrable over \( \mathbb{I} \), \( \nu_F \) the \( Y \)-valued measure with kernel \( f \) over \( \mathbb{I} \), and \( F : I \to Y \) be a cumulative distribution function for \( \nu_F \). Then, \( H : I \to Z \) defined by \( H(x) = A(x)F(x), \forall x \in I \), is absolutely continuous and is a cumulative distribution function of \( \nu_H \), which is the \( Z \)-valued measure with kernel \( h : P(I) \to Z \) over \( \mathbb{I} \), where \( h \) is defined by \( h(x) = A(x)f(x) + a(x)F(x), \forall x \in P(I) \), and \( Af, aF, \) and \( h \) are absolutely integrable over \( \mathbb{I} \). As a consequence, \( \forall b, c \in I \) with \( b \leq c \), we have

\[
\int_{b}^{c} A(x)f(x) \, dx + \int_{b}^{c} a(x)F(x) \, dx = A(c)F(c) - A(b)F(b) \tag{12.5}
\]
By Proposition 7.66, \( W \) is a separable Banach space. Fix any \( x_0 \in I \). Let \( F_0 : I \to Y \) be the cumulative distribution function of \( \nu_F \) with origin \( x_0 \), and \( A_0 : I \to W \) be the cumulative distribution function of \( \nu_A \) with origin \( x_0 \). By Proposition 12.75, \( F_0 \) and \( A_0 \) are absolutely continuous and therefore continuous. Since \( f \) and \( a \) are absolutely integrable over \( I \), then \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \|F_0(x)\|_Y \leq M \) and \( \|A_0(x)\|_W \leq M \), \( \forall x \in I \). Hence, \( F_0 \) and \( A_0 \) are measurable.

By Propositions 7.23, 7.65, 11.38, and 11.39, \( \exists \) continuous, \( \therefore \) fundamental Theorem of Calculus I 11.86, \( x \in \mathbb{R} \). Then, \( F_0 \) and \( A_0 \) are absolutely integrable over \( I \), and \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \|F_0(x)\|_Y \leq M \) and \( \|A_0(x)\|_W \leq M \), \( \forall x \in I \). By Proposition 11.37, \( F \) and \( A \) are \( B \)-measurable. By Propositions 7.23, 7.65, 11.38, and 11.39, \( Af, aF \) and \( h \) are \( B \)-measurable. Then, \( Af, aF \) are absolutely integrable over \( I \). Hence, \( h \) is absolutely integrable over \( I \) by Proposition 11.83. By Proposition 12.68, \( H \) is absolutely continuous.

We will first consider the case that \( I \) is an open interval. By Fundamental Theorem of Calculus I 11.86, \( F^{(1)} = D_1F : U_F \to Y \) and \( A^{(1)} = D_1A : U_A \to W \) are such that \( U_F, U_A \in \mathcal{B}, \mu(I \setminus U_F) = 0 = \mu(I \setminus U_A) \), \( F^{(1)} \) and \( A^{(1)} \) are \( B \)-measurable, \( F^{(1)} = f \) a.e. in \( I \), and \( A^{(1)} = a \) a.e. in \( I \).

Then, \( \exists \mu_U, \mu_A \in \mathcal{B} \) such that \( \mu(I \setminus U_F) = 0 = \mu(I \setminus U_A) \) and \( F^{(1)}(x) = f(x), \forall x \in U_F \), and \( A^{(1)}(x) = a(x), \forall x \in U_A \).

By Propositions 9.17 and 9.19 and the Chain Rule, Theorem 9.18, we have \( H^{(1)}(x) = A(x)f(x) + a(x)F(x) = h(x) \). By Fundamental Theorem of Calculus II 12.88, \( H^{(1)} = D_1H : U_H \to \mathbb{R} \) is such that \( U_H \in \mathcal{B}, \mu(I \setminus U_H) = 0 \), and \( H^{(1)} \) is \( B \)-measurable. Then, \( E := (I \setminus U_H) \cup \{ x \in U_H \mid H^{(1)}(x) \neq h(x) \} \in \mathcal{B} \) and \( E \subseteq (I \setminus U_F) \cup (I \setminus U_A) \).

Hence, \( \mu(E) = 0 \) and \( h = H^{(1)} \) a.e. in \( I \). By Fundamental Theorem of Calculus II 12.88, \( H(x) = \int_{x_0}^x h \cdot \mu \), \( \forall x \in I \), \( h \) is \( B \)-measurable and absolutely integrable over \( r_{x,a,x_b} \), \( \forall x_a, x_b \in I \) with \( x_a \leq x_b \), and \( H^{(1)}(x) = h(x) \) a.e. in \( I \). Thus, by Lemma 11.44, \( h = h \) a.e. in \( I \). Then, by Proposition 11.92, we have and \( H(x) = H(x_0) = \int_{x_0}^x h(s) \cdot ds, \forall x \in I \).

By Proposition 11.92 and Fact 12.72, \( \forall b, c \in I \) with \( b \leq c \), we have \( A(c)F(c) - A(b)F(b) = H(c) - H(x_0) - (H(b) - H(x_0)) = \int_{x_0}^c h(x) \cdot dx - \int_{x_0}^b h(x) \cdot dx = \int_{c}^b (A(x)f(x) + a(x)F(x)) \cdot dx + \int_{b}^c a(x)F(x) \cdot dx \).

Define \( \mu_H \) to be the \( \mathbb{Z} \)-valued measure with kernel \( h \) over \( I \). By the preceding paragraph, we have \( H \) is the cumulative distribution function of \( \mu_H \). This completes the proof for this case.

Next, we consider the case when \( I \) is a closed interval. We will define \( \tilde{I} \) to be an open interval such that \( I \subseteq \tilde{I} \). Define \( \tilde{f} : \tilde{I} \to Y \) by \( f(x) = \begin{cases} f(x) & x \in P(I) \\ \varnothing & x \in \tilde{I} \setminus P(I) \end{cases} \). Similarly define \( \tilde{a} \) in terms of \( a \). Define \( \tilde{F} : \tilde{I} \to Y \) by \( \tilde{F}(x) = \begin{cases} F(x) & x \in I \\ F(\min_{s \in I} s) & x < \min_{s \in I} s \\ F(\max_{s \in I} s) & x > \max_{s \in I} s \end{cases} \). Similarly define \( \tilde{A} \) in terms of \( A \). Then, the result can be applied for \( \tilde{I} := ((\tilde{I}, \cdot, |), \tilde{B}, \tilde{\mu}) \) for the functions.
\( j, F, a, A, \) to yield \( \hat{H} : \hat{I} \rightarrow \mathbb{Z} \) and \( \hat{h} : \hat{I} = P(\hat{I}) \rightarrow \mathbb{Z}. \) Restrict these functions to \( I \) to yield the desired result.

This completes the proof of the theorem. \( \square \)

**Lemma 12.90 (Vitali)** Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( E \subseteq \mathbb{R}^m \) with \( \mu_{\text{loc}}(E) < +\infty, \mathcal{I} \subseteq \mathcal{B}_{\mathbb{R}}(\mathbb{R}^m) \) be a collection of nondegenerate closed rectangles in \( \mathbb{R}^m \) that covers \( E \) in the sense of Vitali with index \( c \in (0, \frac{1}{\sqrt{m}}] \subseteq \mathbb{R}, \) and \( \tilde{I} := \{ V \subseteq \mathbb{R}^m \mid V \) is a nondegenerate rectangle in \( \mathbb{R}^m \) with center \( x \in \mathbb{R}^m, \) such that the rectangle \((\frac{1}{2}(V-x)+x) \in \mathcal{I} \). Assume that \( \tilde{I} \) covers \( E \) in the sense of Vitali with index \( c. \) Then, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \) there exists a countable subcollection \((V_i)_{i \in N} \subseteq \tilde{I} \) with \( N \subseteq \mathbb{N}, \) (each \( V_i \) corresponds to \( U_i \in \mathcal{I} \) as outlined in the definition of \( \tilde{I} \)) and \( \exists n \in \mathbb{Z}_+ \) such that \( \{1, \ldots, n\} \subseteq N, \)

\[
\{ V_1, \ldots, V_n \} \subseteq \tilde{I} \text{ is pairwise disjoint with } \mu_{\text{loc}}(E \setminus \bigcup_{i=1}^n V_i) < \epsilon, \quad \text{and } E \subseteq \bigcup_{i=1}^n V_i \cup \bigcup_{i \in N, i > n} U_i \text{ with } \sum_{i \in N, i > n} \mu_{\text{BM}}(U_i) \leq \epsilon.
\]

**Proof** Fix any \( \epsilon \in (0, \infty) \subseteq \mathbb{R}. \) By Example 12.56, \( \mu_{\text{loc}}(E) = \inf_{(O_i)_{i \in \mathbb{N}, E \subseteq \bigcup_{i=1}^\infty O_i} \sum_{i=1}^\infty \mu_{\text{BM}}(O_i) < +\infty. \) Then, \( \exists \tilde{O} \subseteq \mathcal{O}_{\mathbb{R}^m} \) such that \( E \subseteq \tilde{O} \) and \( \mu_{\text{BM}}(\tilde{O}) < +\infty. \) By neglecting rectangles in \( \tilde{I}, \) we may without loss of generality assume that \( V \subseteq \bigcup_{i=1}^n V_i \) and the shortest side of \( V \) has length at least \( c \cdot \text{dia}(V), \forall V \in \tilde{I}. \forall k \in \mathbb{Z}_+, \) assume that pairwise disjoint \( V_1, \ldots, V_k \in \tilde{I} \) has already been chosen. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( E \subseteq \bigcup_{i=1}^k V_i; \) Case 2: \( E \setminus \bigcup_{i=1}^k V_i \neq \emptyset. \) Case 1: \( E \subseteq \bigcup_{i=1}^k V_i. \) Then, we let \( n = k \) and \( N = \{1, \ldots, n\}. \) Clearly, \( \mu_{\text{loc}}(E \setminus \bigcup_{i=1}^n V_i) = \mu_{\text{loc}}(\emptyset) = 0 < \epsilon, \quad E \subseteq \bigcup_{i=1}^n V_i \cup \bigcup_{i \in N, i > n} U_i \) with \( \sum_{i \in N, i > n} \mu_{\text{BM}}(U_i) = 0 < \epsilon. \) The result holds. Case 2: \( E \setminus \bigcup_{i=1}^k V_i \neq \emptyset. \) Note that, for any closed rectangle \( V \in \tilde{I} \) with \( \text{dia}(V) = \tilde{p} > 0, \) let \( x_0 \in V \) be the center of \( V, \) such that \( \tilde{p} = \sqrt{1 - (m-1)c^2} \tilde{p}, \) where \( \tilde{p} = 1 - (m-1)c^2 \tilde{p} \) is the upper bound for the longest side of the rectangle \( V. \) This implies that \( c^m \tilde{p}^m \leq \mu_{\text{BM}}(V) \leq \tilde{p}^m = (1 - (m-1)c^2)^{\frac{m}{2}}. \) Let \( l_k \) be the supremum of the diameters of rectangles in \( \tilde{I} \) that do not intersect \( V_1, \ldots, V_k. \) Clearly, \( l_k \leq \frac{\sqrt{\mu_{\text{BM}}(O)}}{\epsilon} < +\infty. \) Then, \( \exists x_0 \in E \setminus \bigcup_{i=1}^k V_i. \) Since \( \bigcup_{i=1}^k V_i \) is closed, then, by Proposition 4.10, \( d := \inf_{x \in \bigcup_{i=1}^k V_i} |x - x_0| > 0. \) Then, by the assumption, \( \exists V \in \tilde{I} \) such that \( x_0 \in V \) and \( \text{dia}(V) < d \) and the shortest side of the rectangle \( V \) has length at least \( c \cdot \text{dia}(V). \) Then, \( V \cap (\bigcup_{i=1}^k V_i) = \emptyset. \) This shows that \( l_k \geq \text{dia}(V) > 0. \) Hence, \( \exists V_{k+1} \in \tilde{I} \) such that \( V_1, \ldots, V_{k+1} \in \tilde{I} \) are pairwise disjoint and \( \text{dia}(V_{k+1}) \geq l_k/2. \) Inductively, we either have the result holds or \( \exists (V_k)_{k=1}^\infty \subseteq \tilde{I} \) such that \( \text{dia}(V_{k+1}) \geq l_k/2, \forall k \in \mathbb{Z}_+. \) The sets in the sequence are pairwise disjoint.

In the latter case, we have \( O \supseteq \bigcup_{k=1}^\infty V_k \) and \( +\infty > \mu_{\text{BM}}(O) \geq \mu_{\text{BM}}(\bigcup_{k=1}^\infty V_k) = \sum_{k=1}^\infty \mu_{\text{BM}}(V_k) \geq \sum_{k=0}^\infty 2^{-m} \epsilon m^k. \) Then, we set \( N = \mathbb{N}. \) Note that \( (l_k)_{k=0}^\infty \) is nonincreasing. Hence, \( \lim_{k \to \infty} l_k = 0. \) Then, \( \exists n \in \mathbb{N} \)
such that \( \sum_{i=n+1}^{\infty} \mu_B m(V_i) < 5^{-m} e^m \epsilon \). Let \( R := E \setminus \left( \bigcup_{k=1}^{n} V_k \right) \). Without loss of generality, assume \( \text{diam}(V_k) =: p_k \in (0, \infty) \subset \mathbb{R} \) and the center of \( V_k \) is \( x_k, \forall k \in \mathbb{N} \). \( \forall \delta \in R \), since \( \bigcup_{k=1}^{n} V_k \) is closed, then, by Proposition 4.10, 
\[ d := \inf_{x \in \bigcup_{k=1}^{n} V_k} |x - \bar{x}| > 0. \]
By the assumption, \( \exists \bar{V} \in \mathcal{I} \) such that \( \bar{x} \in \bar{V} \) and \( 0 < \text{diam}(\bar{V}) < d \) and the shortest side of the rectangle \( \bar{V} \) has length at least \( c \text{diam}(V) \). Then, \( V \cap (\bigcup_{k=1}^{n} V_k) = \emptyset \). By the fact that \( \lim_{k \to \infty} l_k = 0 \), we have \( 3l_0 \in \{ n + 1, n + 2, \ldots \} \) such that \( V \cap V_{l_0} \neq \emptyset \) and \( V \cap V_{l_k} = \emptyset, \forall k \in \{ 1, \ldots, l_0 - 1 \} \). Then, \( \text{diam}(V) \leq l_{i_0 - 1} \leq 2 \text{diam}(V_{i_0}) = 2p_{i_0} \). Let \( \hat{x} \in V \cap V_{i_0} \). Then, \( |\hat{x} - x_{i_0}| \leq |\hat{x} - \bar{x}| + |\bar{x} - x_{i_0}| \leq \text{diam}(V) + \frac{1}{2} p_{i_0} \leq \frac{5}{2} p_{i_0} \). Then, \( \hat{x} \in \mathbb{R}^{x_{i_0} - \frac{5}{2} p_{i_0} 1_m, x_{i_0} + \frac{5}{2} p_{i_0} 1_m} \subseteq \mathbb{R}^{1_m, \frac{5}{2} p_{i_0} 1_m + x_{i_0}} \subseteq \mathbb{R}^{1_m, 5 p_{i_0}} \). Hence, \( R \subseteq \bigcup_{k=n+1}^{\infty} U_k \). Then, \( \mu_{X,m}(R) \leq \mu_{X,m}(\bigcup_{k=n+1}^{\infty} U) \leq \sum_{i=n+1}^{\infty} \mu_B m(U_k) = \sum_{i=n+1}^{\infty} 5^m c^{-m} \mu_B m(U_k) < \epsilon \). Note that \( E \subseteq (\bigcup_{k=1}^{n} V_k) \cup (\bigcup_{k=n+1}^{\infty} U) \) and \( \sum_{i \in N, i > n} \mu_B m(U_i) < \epsilon \). Hence, the result holds.

This completes the proof of the lemma. \( \square \)

**Theorem 12.91 (Change of Variable)** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \subset \mathbb{R}^m \) be an open subset, \( \Omega \subset \mathbb{R}^m \) be an open subset, \( F : \Omega \to \Omega \) be a homeomorphism, \( F_1 = F_{\text{inv}} : \Omega \to \Omega \), \( \mathbb{I} := \langle (\Omega, | \cdot |), B, \mu \rangle \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathbb{X} := \langle (\Omega, | \cdot |), B, \bar{\mu} \rangle \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), and \( F_{i*} \bar{\mu} \) be the induced measure on \( (\Omega, B) \) and \( F_{i*} \mu \) be the induced measure on \( (\Omega, B) \) as defined in Proposition 12.9. Then, \( \mathbb{X} \) and \( \langle (\Omega, | \cdot |), B, F_{i*} \mu \rangle \) is homeomorphically isosmeasurable under \( F \) and \( \mathbb{X} \) and \( \langle (\Omega, | \cdot |), B, F_{i*} \mu \rangle \) is homeomorphically isosmeasurable under \( F_1 \). Assume that \( F^{(1)} : U \to \mathbb{R}_m, \mathbb{R}^m \) with \( \mu(\Omega \setminus U) = 0 \) satisfies, \( \forall x_1, x_2 \in \Omega \) with \( x_1 \equiv x_2 \) and \( \mathbb{I}_{x_1, x_2} \subset \Omega \), we have \( \sup_{x_1, x_2, \gamma \in \Omega} \left| \det(F^{(1)}(x)) \right| \leq M_{x_1, x_2} \in (0, \infty) \subset \mathbb{R}, F^{(1)}_i : U \to \mathbb{R}_m, \mathbb{R}^m \) with \( \bar{\mu}(\Omega \setminus U) = 0 \) satisfies, \( \forall \hat{x}_1, \hat{x}_2 \in \Omega \) with \( \hat{x}_1 \equiv \hat{x}_2 \) and \( \mathbb{I}_{\hat{x}_1, \hat{x}_2} \subset \Omega \), we have \( \sup_{\hat{x}_1, \hat{x}_2, \gamma \in \Omega} \left| \det(F^{(1)}(\hat{x})) \right| \leq M_{\hat{x}_1, \hat{x}_2} \in (0, \infty) \subset \mathbb{R} \), and \( \tilde{\Omega} := \{ \hat{x} \in \Omega \ | \ det(F^{(1)}(\hat{x})) \neq 0 \} \) is such that \( \bar{\mu}(\Omega \setminus \tilde{\Omega}) = 0 = \mu(\Omega \setminus \tilde{\Omega}) \). Then, \( F_{i*} \bar{\mu} \ll \bar{\mu} \) and \( \frac{\text{d}F_{i*} \bar{\mu}}{\text{d}\bar{\mu}} = g_1 \) a.e. in \( \mathbb{X} \), \( F_{i*} \mu \ll \mu \) and \( \frac{\text{d}F_{i*} \mu}{\text{d}\mu} = g_1 \) a.e. in \( \mathbb{X} \), where \( g : \Omega \to (0, \infty) \subset \mathbb{R} \) is given by \( g(\hat{x}) = \begin{cases} \left| \det(F^{(1)}(\hat{x})) \right| & \hat{x} \in \tilde{\Omega} \\ 1 & \hat{x} \in \Omega \setminus \tilde{\Omega} \end{cases} \). \( \forall \hat{x} \in \Omega \). Furthermore, let \( \mathcal{Y} \) be a separable Banach space, \( \forall f : \Omega \to \mathcal{Y} \) that is absolutely integrable over \( \mathbb{X} \), we have 
\[ \int_{\Omega} f \, d\mu = \int_{\tilde{\Omega}} f(F_i(\hat{x})) g(\hat{x}) \, d\bar{\mu}(\hat{x}) \]  
(12.6a)
and the right-hand-side integrand is absolutely integrable over \( \mathbb{X} \), \( \forall f : \Omega \to \mathcal{Y} \).
That is absolutely integrable over \( \mathbb{X} \), we have

\[
\int_{\Omega} \bar{f} \, d\bar{\mu} = \int_{\Omega} \bar{f}(F(x))g_1(x) \, d\mu(x)
\]

and the right-hand-side integrand is absolutely integrable over \( \mathbb{X} \).

**Proof**

Fix \( c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R} \). The result is trivial if \( \bar{\Omega} = \emptyset \). Consider the case \( \bar{\Omega} \neq \emptyset \). Then, \( \Omega = F_i(\bar{\Omega}) \neq \emptyset, \mu(\Omega) > 0, \) and \( \bar{\mu}(\bar{\Omega}) > 0 \). First, we need the following intermediate result.

**Claim 12.91.1** \( \forall \bar{x} \in \bar{U}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in \mathbb{B}_R = (0_m, \delta) \) with \( h > 0_m \) and \( \min h \geq c \, |\bar{h}| \), we have \( \mathbb{T}_{x-h, \bar{x}+h} \subseteq \Omega \) and \( F_i \mu(\mathbb{T}_{x-h, \bar{x}+h}) \leq (1 + \varepsilon)(g(\bar{x}))^{-1} \mu(\mathbb{T}_{x-h, \bar{x}+h}) \).

Symmetrically, \( \forall x \in \bar{U} = F_i(\bar{U}), \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in \mathbb{B}_R = (0_m, \delta) \) with \( h > 0_m \) and \( \min h \geq c \, |\bar{h}| \), we have \( \mathbb{T}_{x-h, x+h} \subseteq \Omega, \) and \( F_i \mu(\mathbb{T}_{x-h, x+h}) \leq (1 + \varepsilon)(g(F(x)))^{-1} \mu(\mathbb{T}_{x-h, x+h}) \).

**Proof of claim:** \( \forall \bar{x} \in \bar{U}, \) we have \( F_i^{(1)}(\bar{x}) \) exists and \( \det(F_i^{(1)}(\bar{x})) \neq 0 \). This implies that \( F_i^{(1)}(F_i(\bar{x})) = (F_i^{(1)}(\bar{x}))^{-1} \) and \( \det(F_i^{(1)}(F_i(\bar{x}))) = (\det(F_i^{(1)}(\bar{x})))^{-1} \). By Definition 9.3, \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \varepsilon_1 \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in \mathbb{B}_R = (0_m, \delta_1) \), we have \( \bar{x} + \bar{h} \in \Omega \) and

\[
|F_i(\bar{x} + \bar{h}) - F_i(\bar{x}) - F_i^{(1)}(\bar{x})\bar{h}| \leq \varepsilon \, |\bar{h}|
\]

Then, \( \forall h \in \mathbb{B}_R = (0_m, \delta) \) with \( h > 0_m \) and \( \min h \geq c \, |\bar{h}| \), we have \( F_i(\mathbb{T}_{x-h, \bar{x}+h}) \subseteq F_i(\bar{x}) + F_i^{(1)}(\bar{x})(\mathbb{T}_{x-h, h}) + \mathbb{B}_R = (0_m, \varepsilon \, |\bar{h}|) \). This implies that \( F_i \mu(\mathbb{T}_{x-h, \bar{x}+h}) = \mu(F_i(\mathbb{T}_{x-h, \bar{x}+h}) \leq \mu_B(\mathbb{T}_{x-h, h}) \subseteq F_i(\mathbb{T}_{x-h, h}) + \mathbb{B}_R = (0_m, \varepsilon \, |\bar{h}|) \). Clearly, \( \mu_B(F_i^{(1)}(\bar{x})(\mathbb{T}_{x-h, h})) = g(\bar{x}) \mu_B(\mathbb{T}_{x-h, h}) \), which can be proved by decomposing the matrix \( F_i^{(1)}(\bar{x}) \) into basic row operations, and the equality holds for each row operation. Then, for sufficiently small \( \varepsilon > 0 \), we have \( F_i \mu(\mathbb{T}_{x-h, \bar{x}+h}) \leq (1 + \varepsilon)(g(\bar{x}))^{-1} \mu_B(\mathbb{T}_{x-h, h}) \), since \( F_i^{(1)}(\bar{x})(\mathbb{T}_{x-h, h}) \) is a closed parallelogram with all of its sides has length of at least some constant proportion of \( |\bar{h}| \). This proves the first paragraph of the claim.

\( \forall x \in \bar{U}, \) we have \( F(x) \in F(\bar{U}) = \bar{U} \). Then, \( F_i^{(1)}(x) = (F_i^{(1)}(F_i(x)))^{-1} \) and \( |\det(F_i^{(1)}(x))| = |(\det(F_i^{(1)}(F_i(x))))^{-1} = (g(F(x)))^{-1} \). By symmetry, \( \forall x \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall h \in \mathbb{B}_R = (0_m, \delta) \) with \( h > 0_m \) and \( \min h \geq c \, |h| \), we have \( \mathbb{T}_{x-h, x+h} \subseteq \Omega, \) and \( \mu(F(\mathbb{T}_{x-h, x+h})) = F_i \mu(\mathbb{T}_{x-h, x+h}) \leq (1 + \varepsilon)(g(F(x)))^{-1} \mu(\mathbb{T}_{x-h, x+h}) = (1 + \varepsilon)(g(F(x)))^{-1} \). This proves the second paragraph of the claim.
This completes the proof of the claim. 

We first show that $F_i \mu \ll \bar{\mu}$ under the condition that $\tilde{\Omega}$ is a nonempty bounded open set and $\left| \det(F_i^{(1)}(\bar{x})) \right| \leq M \in (0, \infty) \subset \mathbb{R}$, $\forall \bar{x} \in \tilde{U}$. Then, $0 < \bar{\mu}(\tilde{\Omega}) < \infty$. We need the following intermediate result.

Claim 12.91.2 \( \forall \tilde{O} \in \mathcal{B} \) with \( \tilde{O} \) being open and \( \bar{\mu}(\tilde{O}) > 0 \), we have \( \bar{\mu}(\tilde{O}) \geq \mu(\tilde{O})/M := \mu(F_i(\tilde{O}))/M. \)

Proof of claim: Fix any \( \tilde{O} \in \mathcal{B} \) with \( \tilde{O} \) being open and \( \bar{\mu}(\tilde{O}) > 0 \). Then, \( \tilde{O} \in \mathcal{B} \) and is open. \( \forall \bar{x} \in \tilde{O} := \tilde{O} \cap \tilde{U} \) by Claim 12.91.1, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\bar{x}, \epsilon) \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in B_{\mathbb{R}^m}(0_m, \delta(\bar{x}, \epsilon)) \) with \( \bar{h} > 0_m \) and \( \min \bar{h} \geq c/|\bar{h}| \), we have \( F_i \mu(\mathfrak{r}_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}) \leq (1 + \epsilon)g(\bar{x})\bar{\mu}(\mathfrak{r}_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}). \) The collection of such \( \mathfrak{r}_{\bar{x} - \bar{h}, \bar{x} + \bar{h}} \) is denoted \( \tilde{I} \). We will restrict \( \tilde{I} \) to consist of all of the above mentioned \( \mathfrak{r}_{\bar{x} - \bar{h}, \bar{x} + \bar{h}} \)'s which are subsets of \( \tilde{O} \). Clearly, \( \tilde{I} \) covers \( \tilde{O} \) in the sense of Vitali with index \( c. \) Define \( \bar{\tilde{I}} := \{ V \subseteq \mathbb{R}^m \mid V \text{ is a nondegenerate rectangle in } \mathbb{R}^m \text{ with center } x \in \mathbb{R}^m \text{ such that } (\frac{1}{2}(V - x) + x) \in \tilde{I} \}. \)

Clearly, \( \bar{\tilde{I}} \subseteq \tilde{I}. \) It is easy to see that \( \bar{\tilde{I}} \) covers \( \tilde{O} \) in the sense of Vitali with index \( c. \) By Vitali's Lemma 12.90, \( \exists N \subseteq \mathbb{N}, \exists (V_i)_{i \in N} \subseteq \bar{\tilde{I}}, \forall n \in \mathbb{Z}^+ \) such that \( \{1, \ldots, n\} \subseteq N, \{V_1, \ldots, V_n\} \) is pairwise disjoint, \( \bar{\mu}(\tilde{O} \setminus (\bigcup_{i=1}^n V_i)) < \frac{\epsilon}{4} \bar{\mu}(\tilde{O}), \tilde{O} \subseteq (\bigcup_{i=1}^n V_i)\cup(\bigcup_{i \in N, i > n} U_i), \) and \( \sum_{i \in N, i > n} \mu(U_i) \leq \frac{\epsilon}{4} \bar{\mu}(\tilde{O}), \) where \( V_i = \mathfrak{r}_{\bar{x}_i - \bar{h}_i, \bar{x}_i + \bar{h}_i} \in \bar{\tilde{I}} \subseteq \tilde{I} \) and \( U_i = \frac{1}{2}(V_i - \bar{x}_i) + \bar{x}_i = \mathfrak{r}_{\bar{x}_i - \frac{1}{2}h_i, \bar{x}_i + \frac{1}{2}h_i} \in \tilde{I}, \forall i \in N. \) Then, we have the following line of arguments.

\[
\mu(\tilde{O}) = \mu(\tilde{O} \cup (\tilde{O} \setminus \tilde{O})) = \mu(\tilde{O}) + \mu(\tilde{O} \setminus \tilde{O})
\]

\[
= \mu(F_i(\tilde{O}) \cup (\bigcup_{i \in N, i > n} U_i))
\]

\[
\leq \sum_{i=1}^n \mu(F_i(V_i)) + \sum_{i \in N, i > n} \mu(F_i(U_i)) = \mu(F_i(\tilde{O})) + \sum_{i \in N, i > n} F_i \mu(U_i)
\]

\[
\leq \sum_{i=1}^n (1 + \epsilon)g(\bar{x}_i)\bar{\mu}(\mathfrak{r}_{\bar{x}_i - \bar{h}_i, \bar{x}_i + \bar{h}_i}) + \sum_{i \in N, i > n} (1 + \epsilon)g(\bar{x}_i)\bar{\mu}(\mathfrak{r}_{\bar{x}_i - \frac{1}{2}h_i, \bar{x}_i + \frac{1}{2}h_i})
\]

\[
\leq (1 + \epsilon)\tilde{M} \sum_{i=1}^n \bar{\mu}(\mathfrak{r}_{\bar{x}_i - \bar{h}_i, \bar{x}_i + \bar{h}_i}) + (1 + \epsilon)\tilde{M} \sum_{i \in N, i > n} \bar{\mu}(\mathfrak{r}_{\bar{x}_i - \frac{1}{2}h_i, \bar{x}_i + \frac{1}{2}h_i})
\]

\[
= (1 + \epsilon)\tilde{M} \sum_{i=1}^n \bar{\mu}(V_i) + (1 + \epsilon)\tilde{M} \sum_{i \in N, i > n} \bar{\mu}(U_i)
\]

\[
\leq (1 + \epsilon)\tilde{M} \bar{\mu}(\bigcup_{i=1}^n V_i) + (1 + \epsilon)\tilde{M} \bar{\mu}(\tilde{O})
\]

\[
\leq (1 + \epsilon)\tilde{M} \bar{\mu}(\tilde{O}) + (1 + \epsilon)\tilde{M} \frac{\epsilon}{4} \bar{\mu}(\tilde{O}) = (1 + \epsilon)\frac{4 + \epsilon}{4} \tilde{M} \bar{\mu}(\tilde{O})
\]
where \( \bar{O} := O \cap \tilde{U} = F_i(\bar{O} \cap \tilde{U}) = F_i(\tilde{O}) \); the third equality follows from the fact that \( O \setminus \bar{O} \subseteq \Omega \setminus \tilde{U} \); the first inequality follows from monotonicity of measure; the second inequality follows from the subadditivity of measure; the fourth equality follows from Proposition 12.9; the third inequality follows from the choice of \( \mathcal{I} \) and \( \tilde{Z} \); the fifth inequality follows from the monotonicity of \( \mu \); and from the preceding discussion; and the last inequality follows from the monotonicity of measure. Since \( \epsilon > 0 \) is arbitrary, then we have \( \mu(\bar{O}) \leq M\tilde{\mu}(\bar{O}) \).

This completes the proof of the claim. \( \square \)

By the arbitrariness of \( \bar{O} \), we have \( F_i(\bar{O}) = 0 \). By the arbitrariness of \( E \), we have \( F_i(\bar{E}) = 0 \). By the arbitrariness of \( \bar{E} \), we have \( F_i(\mu) = 0 \). By the arbitrariness of \( \mu \), we have \( F_i(\mu) \leq \tilde{\mu} \). This proves that \( F_i(\mu) \leq \tilde{\mu} \).

By symmetry, \( F_i(\mu) \leq \mu \).

By Radon-Nikodym Theorem 11.69, \( \frac{dF_i(\mu)}{d\mu} := \hat{g} \) exists and is unique almost everywhere on \( \tilde{X} \), and \( \frac{dF_i(\mu)}{d\mu} := \hat{g} \) exists and is unique almost everywhere on \( \tilde{X} \), where \( \hat{g} : \Omega \to [0, \infty) \subset \mathbb{R} \) is \( \mathcal{B} \)-measurable (or simply \( \mathcal{B}_0(\mathbb{R}^m) \)-measurable); and \( \hat{g} : \Omega \to [0, \infty) \subset \mathbb{R} \) is \( \mathcal{B} \)-measurable (or simply \( \mathcal{B}_0(\mathbb{R}^m) \)-measurable). By Definitions 12.7 and 12.8 and Proposition 11.168, it is clear that \( \hat{g}(x) = \frac{1}{g(F_i(x))} \) a.e. \( x \in \tilde{X} \) and \( \hat{g}(x) = \frac{1}{g(F_i(x))} \) a.e. \( \tilde{x} \in \tilde{X} \). Then, by changing \( \hat{g} \) on a set of measure zero in \( \mu \) and changing \( \hat{g} \) on a set of measure zero in \( \mu \), we have \( \hat{g}(x) = \frac{1}{g(F_i(x))}, \) \( \forall x \in \tilde{X} \).

By Propositions 11.92, 12.11, and 11.168, we have (12.6) holds with \( g \) substituted by \( \hat{g} \) and \( q_1 \) substituted by \( \hat{g} \). We will only need to show that \( \hat{g} = g \) a.e. in \( \tilde{X} \). Then, the theorem is proved.

First, consider the special case when \( \Omega \) is bounded. Then, \( \chi_{\tilde{X}, \Omega} \) is absolutely integrable over \( \tilde{X} \). By the preceding paragraph, we have \( \infty > \mu(\Omega) = \int_{\Omega} \chi_{\tilde{X}, \Omega} \, d\mu = \int_{\Omega} \chi_{\tilde{X}, \Omega}(F_i(x)) \hat{g}(x) \, d\tilde{\mu}(x) = \int_{\Omega} \chi_{\tilde{X}, \Omega}(x) \hat{g}(x) \, d\tilde{\mu}(x) \). Then, we have \( \int_{\Omega} \hat{g} \, d\tilde{\mu} < \infty \). This implies that \( \hat{g} \in L_1(\tilde{\mu}, \mathbb{R}) \). Let \( A := \{ \tilde{x} \in \tilde{X} \mid x \mapsto \hat{g}(x) \} \) be a rectangular Lebesgue point with regularity \( c \) of \( \hat{g} \). By Proposition 12.80, \( \hat{A} \in \mathcal{B}_0(\mathbb{R}^m \setminus \tilde{A}) = 0 \). Define the set \( \tilde{B} := \{ \tilde{x} \in \tilde{X} \mid \hat{g}(\tilde{X}) \leq g(\tilde{X}) \} \). By the measurability of \( g \) and \( \hat{g} \) and Propositions 7.23, 11.38, 11.39, and 11.35, we have \( \tilde{B} \in \mathcal{B}_m \). We need the following intermediate result.

Claim 12.91.3 \( \tilde{B} \supseteq \tilde{A} \cap \tilde{U} \).

12.6. FUNDAMENTAL THEOREM OF CALCULUS

649
CHAPTER 12. DIFFERENTIATION AND INTEGRATION

Proof of claim: Fix any \( \bar{x} \in \bar{A} \cap \bar{U} \). Fix any \( \epsilon \in (0, \infty) \subset \mathbb{R} \).

By Claim 12.91.1, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \forall \bar{h} \in B_{\mathbb{R}^m}(0_m, \delta) \) with \( \bar{h} > 0_m \) and \( \min \bar{h} \geq c |\bar{h}| \), we have \( \tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}} \subseteq \bar{\Omega} \) and \( F_{\bar{x}} \mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}) \leq (1 + \epsilon) g(\bar{x}) \mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}) \). By Definition 12.79, \( \exists r_0 \in (0, \infty) \subset \mathbb{R} \) such that \( \forall r \in (0, r_0) \), we have \( B_{\mathbb{R}^m}(\bar{x}, r) \subseteq \bar{\Omega} \) and

\[
\frac{1}{\mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}})} \int_{\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}} |\hat{g}(y) - \hat{g}(\bar{x})| \, d\hat{\mu}(y) < \epsilon
\]

Take \( \bar{h} \in \mathbb{R}^m \) with \( |\bar{h}| < r_0 \wedge \delta, \bar{h} > 0_m \), and \( \min \bar{h} \geq c |\bar{h}| \), we have

\[
\hat{g}(\bar{x}) - \epsilon \leq \hat{g}(\bar{x}) - \frac{1}{\mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}})} \int_{\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}} |\hat{g}(y) - \hat{g}(\bar{x})| \, d\hat{\mu}(y) \\
\leq \hat{g}(\bar{x}) - \frac{1}{\mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}})} \int_{\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}} \hat{g}(y) \, d\hat{\mu}(y) \\
\leq \frac{1}{\mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}})} \int_{\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}} \hat{g}(y) \, d\hat{\mu}(y) = \frac{1}{\mu(\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}})} \int_{\tau_{\bar{x} - \bar{h}, \bar{x} + \bar{h}}} 1 \, d\mu \\
\leq (1 + \epsilon) g(\bar{x})
\]

where the first inequality follows from the fact \( |\bar{h}| < r_0 \); the second inequality follows from Proposition 11.92; the third inequality follows from basic algebra; the first equality follows from (12.6a) and the second to last paragraph preceding the claim; the second equality follows from Proposition 11.75; the third equality follows from Proposition 12.9; the fourth inequality follows from the monotonicity of measure; and the last inequality follows from \( |\bar{h}| < \delta \).

By the arbitrariness of \( \epsilon \), we have \( \hat{g}(\bar{x}) \leq g(\bar{x}) \) and \( \bar{x} \in \bar{B} \). By the arbitrariness of \( \bar{x} \), we have \( \bar{A} \cap \bar{U} \subseteq \bar{B} \).

This completes the proof of the claim. \( \Box \)

By Claim 12.91.3, we have \( 0 \leq \bar{\mu}(\bar{\Omega} \setminus \bar{B}) \leq \mu_{Lm}(\Omega \setminus (\bar{A} \cap \bar{U})) \leq \mu_{Lm}(\Omega \setminus \bar{A}) + \bar{\mu}(\bar{\Omega} \setminus \bar{U}) = 0 \). Thus, \( \bar{\mu}(\bar{\Omega} \setminus \bar{B}) = 0 \).

Define the set \( C := \{ x \in \Omega \mid \hat{g}(x) \leq g_1(x) \} \). By symmetry, we have \( C \in \mathcal{B} \) and \( \mu(\Omega \setminus C) = 0 \). On the set \( \bar{G} := \bar{B} \cap F(C) \), we have \( \hat{g}(\bar{x}) \leq g(\bar{x}) \) and \( \hat{g}(F_1(\bar{x})) \leq g_1(F_1(\bar{x})) \). This implies that \( \hat{g}(\bar{x}) \leq g(\bar{x}) \leq g_1(\bar{x}) \leq \frac{1}{g(\bar{x})} \hat{g}(\bar{x}), \forall \bar{x} \in \bar{G} \). Note that \( 0 \leq \bar{\mu}(\bar{G}) \leq \bar{\mu}(\bar{\Omega} \setminus \bar{B}) + \bar{\mu}(\bar{G} \setminus F(C)) = \bar{\mu}(\bar{G} \setminus F(C)) = \mu(\Omega \setminus C) \). By \( \mu(\Omega \setminus C) = 0 \) and \( F_1 \mu \leq \mu \), we have \( F_1 \bar{\mu}(\Omega \setminus C) = 0 \). Then, we have \( \bar{\mu}(\Omega \setminus C) = 0 \). Hence, \( \hat{g} = g \) a.e. in \( \bar{X} \). This proves the special case.

Finally, consider the general case. Since \( \Omega \) is an nonempty open set and \( \mathbb{R}^m \) is second countable, then \( \exists (\bar{x}_i)_{i=1}^\infty, (\bar{x}_i)_{i=1}^\infty \subseteq \bar{\Omega} \) such that \( \bar{x}_i \neq \bar{x}_i \) and \( \tau_{\bar{x}_i, \bar{x}_i} \subseteq \bar{\Omega}, \forall i \in \mathbb{N} \), and \( \bar{\Omega} = \bigcup_{i=1}^\infty \tau_{\bar{x}_i, \bar{x}_i}, \forall i \in \mathbb{N} \), let
\( \bar{X}_i := (\langle r_{\bar{x},\bar{x}}, |\cdot| \rangle, \bar{B}_i, \bar{\mu}_i) \) be the finite metric measure subspace of \( \bar{X} \).
By the special case, we have \( \bar{g} \mid_{\bar{r}_{\bar{x},\bar{x}}} = g|_{r_{\bar{x},\bar{x}}} \) a.e. in \( \bar{X}_i \). Let \( \bar{D} := \{ \bar{x} \in \bar{\Omega} \mid \bar{g}(\bar{x}) \neq g(\bar{x}) \} \). By Propositions 7.23, 11.38, 11.39, and 11.35, \( \bar{D} \in \bar{B} \).
Then, we have \( 0 \leq \bar{\mu}(\bar{D}) = \bar{\mu}(\bigcup_{i=1}^{\infty} (\bar{D} \cap r_{\bar{x},\bar{x}})) \leq \sum_{i=1}^{\infty} \bar{\mu}(\bar{D} \cap r_{\bar{x},\bar{x}}) = 0 \).
Hence, \( \bar{g} = g \) a.e. in \( \bar{X} \).
This completes the proof of the theorem. \( \square \)

### 12.7 Representation of \((C_{K}(\Omega, Y))^*\)

**Theorem 12.92 (Riesz Representation Theorem)** Let \( X \) be a normed linear space over \( K \), \( Y \) be Banach space over \( K \), \( \Omega \subseteq X \) be compact with the subset topology \( O \), and \( k \in \mathbb{N} \). Assume that, \( \forall x \in \Omega, \exists \delta_x \in (0, \infty) \subset \mathbb{R} \) such that \( \Omega \cap B_X(x, \delta_x) = x \) is a conic segment and \( \text{span}(A_\Omega(x)) = X \).
Let \( \bar{Z} := C_{\bar{K}}(\Omega, Y) \) be the Banach space over \( \bar{K} \) as prescribed in Example 9.70.
\( \forall f \in \bar{Z}^*, \exists \mu := (\mu_0, \ldots, \mu_k) \in \prod_{i=0}^{k} (C(\Omega, B_i(X, Y)))^* \) such that \( f(z) = \sum_{i=0}^{k} \langle \mu_i, z^{(i)} \rangle \), \( \forall z \in \bar{Z} \), \( \| \mu \| = \left( \sum_{i=0}^{k} \| \mu_i \|^2 \right)^{1/2} = \| f \|_{\bar{Z}} \).
Furthermore, \( \forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^{k} (C(\Omega, B_i(X, Y)))^* \), define \( h : \bar{Z} \to \bar{K} \) by \( h(z) = \sum_{i=0}^{k} \langle \nu_i, z^{(i)} \rangle \), \( \forall z \in \bar{Z} \). Then, \( h \in \bar{Z}^* \) and \( \| h \|_{\bar{Z}} \leq \| \nu \|_{\bar{Z}} \).

**Proof** Let \( W := \prod_{i=0}^{k} C(\Omega, B_i(X, Y)) =: \prod_{i=0}^{k} \bar{W}_i \). By Proposition 7.66 and Example 7.32, \( \bar{W}_i \) is a Banach space over \( \bar{K} \), \( \forall i \in \{0, \ldots, k\} \).
Then, by Propositions 7.22 and 4.31, \( W \) is a Banach space over \( \bar{K} \), \( \forall i \in \{0, \ldots, k\} \), we have \( z^{(i)} \in \bar{W}_i \). Then, we may define a mapping \( T : \bar{Z} \to \bar{W} \) by \( T(z) = (z^{(0)}, \ldots, z^{(k)}) \), \( \forall z \in \bar{Z} \). Clearly, \( T \) is linear and norm-preserving. Hence, \( T \in B(\bar{Z}, \bar{W}) \) and is injective.

**Claim 12.92.1** \( R(T) \subseteq \bar{W} \) is closed.

**Proof of claim:** Clearly, we have \( R(T) \subseteq \bar{R}(T) \subseteq \bar{W} \). \( \forall w \in \bar{R}(T), \exists (z_n)_{n=1}^{\infty} \subseteq \bar{Z} \) such that \( w = \lim_{n \to \infty} T(z_n) \) in \( \bar{W} \). Then, \( (T(z_n))_{n=1}^{\infty} \subseteq \bar{W} \) is a Cauchy sequence. Since \( T \) is norm-preserving, then \( (z_n)_{n=1}^{\infty} \subseteq \bar{Z} \) is a Cauchy sequence. By the completeness of \( \bar{Z} \), \( \exists z_0 \in \bar{Z} \) such that \( \lim_{n \to \infty} z_n = z_0 \) in \( \bar{Z} \).
By the continuity of \( T \), we have \( w = \lim_{n \to \infty} T(z_n) = T(z_0) \in \bar{W} \).
Hence, \( w \in R(T) \).
By the arbitrariness of \( w \), we have \( R(T) \subseteq R(T) \).
Hence, \( R(T) \) is closed. This completes the proof of the claim. \( \square \)

By Claim 12.92.1 and Proposition 4.39, \( R(T) \subseteq \bar{W} \) is a Banach space over \( \bar{K} \). By Open Mapping Theorem 7.103, \( T_{\text{inv}} : R(T) \to \bar{Z} \) and \( T_{\text{inv}} \in B(R(T), \bar{Z}) \).
\( \forall f \in \bar{Z}^* \), define \( g : R(T) \to \bar{K} \) by \( g(w) = f(T_{\text{inv}}(w)) \), \( \forall w \in R(T) \).
Clearly, \( g \) is a bounded linear functional on \( R(T) \). By Hahn-Banach Extension Theorem 7.83, \( \exists \mu \in W^* \) such that \( \mu(w) = g(w), \forall w \in R(T) \), and \( \| \mu \|_{W^*} = \| g \|_{R(T)} \leq \| f \|_{\bar{Z}} \).
These last inequality follows from the fact that \( T \) is norm-preserving and so is \( T_{\text{inv}} \).
By Example 7.76, $\mathcal{W}^* = \prod_{i=0}^k \mathcal{W}_i^* = \prod_{i=0}^k (C(\Omega, B_i(\mathcal{X}, \mathcal{Y})))^*$. Then, $\mu = (\mu_0, \ldots, \mu_k) \in \mathcal{W}^*$ and $f(z) = g(T(z)) = \langle \langle \mu, T(z) \rangle \rangle = \sum_{i=0}^k \langle \langle \mu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$.

$\forall z : (\nu_0, \ldots, \nu_k) \in \mathcal{W}^*$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \langle \langle \nu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h(z) = \langle \langle \nu, T(z) \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|_{\mathcal{W}^*} \|T\|_{B(\mathcal{Z}, \mathcal{W})} \leq \|\nu\|_{\mathcal{W}^*}$, where the last inequality follows from the fact that $T$ is norm-preserving.

Then, $\|f\|_{\mathcal{Z}^*} \leq \|\mu\|_{\mathcal{W}^*} \leq \|f\|_{\mathcal{Z}^*}$. Hence, $\|f\|_{\mathcal{Z}^*} = \|\mu\|_{\mathcal{W}^*}$. This completes the proof of the theorem.

\[ \blacksquare \]

**Theorem 12.93 (Riesz Representation Theorem)** Let $m \in \mathbb{Z}_+$, $\mathcal{X} := \mathbb{K}^m$ be the Euclidean space, $I_1, \ldots, I_m \subset \mathbb{R}$ be nondegenerate compact intervals, $\Omega := \prod_{i=1}^m I_i \subset \mathcal{X}$ with the subset topology $\mathcal{O}$, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $k \in \mathbb{N}$, and $\mathcal{Z} := C_0(\Omega, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as prescribed in Example 9.70. $\forall f \in \mathcal{Z}^*$, $\exists \mu := (\mu_0, \ldots, \mu_k) \in \prod_{i=0}^k M_{f_1}(\Omega, (B_i(\mathcal{X}, \mathcal{Y})))^* = \prod_{i=0}^k M_f(\Omega, B_{i}(\mathcal{X}, \mathcal{Y}))(B_i(\mathcal{X}, \mathcal{Y})^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\mu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left(\sum_{i=0}^k (P \circ \mu_i(\Omega))^2\right)^{1/2} = \|f\|_{\mathcal{Z}^*}$. Furthermore, $\forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^k M_{f_1}(\Omega, (B_i(\mathcal{X}, \mathcal{Y})))^*$, define $h : \mathcal{Z} \rightarrow \mathbb{K}$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\nu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|$.

**Proof** Let $\mathcal{W} := \prod_{i=0}^k C(\Omega, B_i(\mathcal{X}, \mathcal{Y})) = \prod_{i=0}^k W_i$. By Proposition 7.66 and Example 9.70, $\mathcal{W}_i$ is a Banach space over $\mathbb{K}$, $\forall i \in \{0, \ldots, k\}$. It is easy to show that $\Omega \subseteq \mathcal{X}$ satisfies the assumption of Theorem 12.92. By Riesz Representation Theorem 11.204, we have $\mathcal{W}^* = \prod_{i=0}^k C(\Omega, B_i(\mathcal{X}, \mathcal{Y}))^* = \prod_{i=0}^k M_f(\Omega, B_{i}(\mathcal{X}, \mathcal{Y}))(B_i(\mathcal{X}, \mathcal{Y})^*)$. Then, the result follows immediately from Theorem 12.92. This completes the proof of the theorem.

\[ \blacksquare \]

**Theorem 12.94 (Riesz Representation Theorem)** Let $\mathcal{X}$ and $\mathcal{Y}$ be finite-dimensional Banach spaces over $\mathbb{K}$, $\Omega \subseteq \mathcal{X}$ be compact with the subset topology $\mathcal{O}$, and $k \in \mathbb{N}$. Assume that, $\forall x \in \Omega$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap B_{\chi}(x, \delta_x) = x$ is a conic segment and span$(A_{\Omega}(x)) = \mathcal{X}$.

Let $\mathcal{Z} := C_0(\Omega, \mathcal{Y})$ be the Banach space over $\mathbb{K}$ as prescribed in Example 9.70. $\forall f \in \mathcal{Z}^*$, $\exists \mu := (\mu_0, \ldots, \mu_k) \in \prod_{i=0}^k M_{f_1}(\Omega, (B_i(\mathcal{X}, \mathcal{Y})))^* = \prod_{i=0}^k M_f(\Omega, B_{i}(\mathcal{X}, \mathcal{Y}))(B_i(\mathcal{X}, \mathcal{Y})^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\mu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left(\sum_{i=0}^k (P \circ \mu_i(\Omega))^2\right)^{1/2} = \|f\|_{\mathcal{Z}^*}$. Furthermore, $\forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^k M_{f_1}(\Omega, (B_i(\mathcal{X}, \mathcal{Y})))^*$, define $h : \mathcal{Z} \rightarrow \mathbb{R}$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle \langle d\nu_i, z(i) \rangle \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathcal{Z}^*$ and $\|h\|_{\mathcal{Z}^*} \leq \|\nu\|$.

**Proof** Let $\mathcal{W} := \prod_{i=0}^k C(\Omega, B_i(\mathcal{X}, \mathcal{Y})) = \prod_{i=0}^k W_i$. The result follows immediately from Theorem 12.92 if we can prove that $\mathcal{W}^* = \prod_{i=0}^k M_{f_1}(\Omega, (B_i(\mathcal{X}, \mathcal{Y})))^* = \prod_{i=0}^k M_f(\Omega, B_{i}(\mathcal{X}, \mathcal{Y}))(B_i(\mathcal{X}, \mathcal{Y})^*)$. The first
equality follows from Example 7.76 and Riesz Representation Theorem 11.205. Since $\Omega \subseteq \mathcal{X}$ is compact then it is locally compact. Since $\mathcal{X}$ is a finite dimensional Banach space, then it is separable by Proposition 11.190. By Propositions 4.37 and 4.38, $\Omega$ is separable metric space. By Theorem 11.198, we have $\mathcal{M}_{f_i}(\Omega, (B_{i} (X, \mathcal{Y} ) )) = \mathcal{M}_{f}(\Omega, B_{0} (\Omega), (B_{i} (X, \mathcal{Y} ))^{*})$, $\forall i \in \{0, \ldots, k\}$. This completes the proof of the theorem.

\[\square\]

**Example 12.95** Let $\mathbb{R}^2$ be endowed with the usual positive cone, $\Omega := \{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 > 0, \angle (x_1 + ix_2) \neq 0 \}$, $\widehat{\Omega} := \{ (\rho, \theta) \in \mathbb{R}^2 \mid \rho > 0, \theta \in (0, 2\pi) \}$, and $\Gamma : \Omega \rightarrow \widehat{\Omega}$ be defined by $\Gamma(x_1, x_2) := (\sqrt{x_1^2 + x_2^2}, \angle (x_1 + ix_2))$, and $F_i : \Omega \rightarrow \Omega$ be defined by $F_i(\rho, \theta) = \Gamma(\rho \cos(\theta), \rho \sin(\theta))$. It is easy to check that $F_i(\rho, \theta) = \rho(\cos(\theta), \sin(\theta))$ and $\det(F_i(\rho, \theta)) = \rho > 0, \forall (\rho, \theta) \in \Omega$. Hence, $\tilde{\Omega} = \Omega$ is defined as in Change of Variable Theorem 12.16. Then, all assumption of that theorem are satisfied. Let $\mathcal{Y}$ be a separable Banach space, $\forall f : \Omega \rightarrow \mathcal{Y}$ that is absolutely integrable, we have $\int_{\Omega} f(x_1, x_2) \, d\mu_{\mathcal{B}}(x_1, x_2) = \int_{\widehat{\Omega}} f(\Gamma(\rho, \theta)) \rho \, d\mu_{\mathcal{B}}(\rho, \theta)$. Furthermore, $\forall f : \mathbb{R}^2 \rightarrow \mathcal{Y}$ that is absolutely integrable over $\mathbb{R}^2$, we have $\int_{\mathbb{R}^2} f(x_1, x_2) \, d\mu_{\mathcal{B}}(x_1, x_2) = \int_{\widehat{\Omega}} f(\Gamma(\rho, \theta)) \rho \, d\mu_{\mathcal{B}}(\rho, \theta)$, where the first equality follows from the fact that $\mu_{\mathcal{B}}(\mathbb{R}^2 \setminus \Omega) = 0$ and Proposition 11.92.

\[\diamondsuit\]

**Example 12.96** Let $m \in \mathbb{N}$ with $m \geq 3$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega := \{ (x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_i^2 + x_j^2 > 0, \angle (x_i + ix_j) \neq 0 \}$, $\widehat{\Omega} := \{ (\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) \in \mathbb{R}^m \mid \rho > 0, \theta_i \in (0, \pi), i = 1, \ldots, m-2, \varphi \in (0, 2\pi) \}$, and $\Gamma : \Omega \rightarrow \widehat{\Omega}$ be defined by $\Gamma(x_1, \ldots, x_m) := (\sqrt{\sum_{i=1}^{m-1} x_i^2}, \arccos(\frac{x_1}{\sqrt{\sum_{i=1}^{m-1} x_i^2}}), \ldots, \arccos(\frac{x_m}{\sqrt{\sum_{i=1}^{m-1} x_i^2}}), \angle (x_1 + ix_{m-2}), F_i : \Omega \rightarrow \Omega$ be defined by $F_i(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) = \Gamma(\rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \cos(\varphi), \rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \sin(\varphi), \rho \sin(\theta_1) \cdots \sin(\theta_{m-3}) \cos(\theta_m), \rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \cos(\theta_2), \rho \cos(\theta_1))$. It is easy to check that $F(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) = \rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \cos(\varphi), \rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \sin(\varphi), \rho \sin(\theta_1) \cdots \sin(\theta_{m-3}) \cos(\theta_m), \rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \cos(\theta_2), \rho \cos(\theta_1))$. Hence, $\tilde{\Omega} = \Omega$ as defined in Change of Variable Theorem 12.91. Then, all assumption of that theorem are satisfied. Let $\mathcal{Y}$ be a separable Banach space, $\forall f : \Omega \rightarrow \mathcal{Y}$ that is absolutely integrable, we have $\int_{\Omega} f(x_1, \ldots, x_m) \, d\mu_{\mathcal{B}}(x_1, \ldots, x_m) = \int_{\widehat{\Omega}} f(\Gamma(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi)) \rho \, d\mu_{\mathcal{B}}(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi)$. Furthermore, $\forall f : \mathbb{R}^m \rightarrow \mathcal{Y}$ that is absolutely integrable over $\mathbb{R}^m$, we have $\int_{\mathbb{R}^m} f(x_1, \ldots, x_m) \, d\mu_{\mathcal{B}}(x_1, \ldots, x_m) = \int_{\widehat{\Omega}} f(\Gamma(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi)) \rho \, d\mu_{\mathcal{B}}(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi)$, where the first equality follows from the fact that $\mu_{\mathcal{B}}(\mathbb{R}^m \setminus \Omega) = 0$ and Proposi-
tion 11.92.

**Example 12.97**  Let \( \mathbb{R}^3 \) be endowed with the usual positive cone, \( \Omega := \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 > 0, \angle(x_1 + ix_2, x_3) \neq 0 \} \), \( \Omega := \{ (\rho, \varphi, z) \in \mathbb{R}^3 \mid \rho > 0, \varphi \in (0, 2\pi) \} \), and \( F : \Omega \to \Omega \) be defined by \( F(x_1, x_2, x_3) := (\sqrt{x_1^2 + x_2^2}, \angle(x_1 + ix_2, x_3)) \), and \( F_i : \Omega \to \Omega \) be defined by \( F_i(\rho, \varphi, z) = F_{\text{inv}}(\rho, \varphi, z) = (\rho \cos(\varphi), \rho \sin(\varphi), z) \). It is easy to check that \( F_i^{(1)} : \Omega \to \mathbb{R}^3 \) and \( F_i^{(1)}(\rho, \varphi, z) \), and \( \det(F_i^{(1)}(\rho, \varphi, z)) \) is absolutely integrable, we have

\[
\int_{\Omega} \rho \, d\mu(\rho, \varphi, z) = \int_{\Omega} \rho \, d\mu(\rho, \varphi, z), \quad \forall (\rho, \varphi, z) \in \bar{\Omega}.
\]

Hence, \( \bar{\Omega} = \Omega \) as defined in Change of Variable Theorem 12.91. Then, all assumption of that theorem are satisfied. Let \( \mathcal{F} \) be a separable Banach space, \( \forall f : \Omega \to \mathcal{F} \) is absolutely integrable, we have

\[
\int_{\Omega} \rho \, d\mu(\rho, \varphi, z) = \int_{\Omega} \rho \, d\mu(\rho, \varphi, z).
\]

Furthermore, \( \forall f : \mathbb{R}^3 \to \mathcal{F} \) is absolutely integrable over \( \mathbb{R}^3 \), we have

\[
\int_{\mathbb{R}^3} \rho \, d\mu(\rho, \varphi, z) = \int_{\mathbb{R}^3} \rho \, d\mu(\rho, \varphi, z),
\]

where the first equality follows from the fact that \( \mu_{\text{inv}}(\mathbb{R}^3 \setminus \Omega) = 0 \) and Proposition 11.92.

**Proposition 12.98**  Let \( I \subseteq \mathbb{R} \) be an open interval, \( \tilde{I} \subseteq \mathbb{R} \) be an open interval, \( F : I \to \tilde{I} \) be a homeomorphism, \( F_i : F_{\text{inv}} : \tilde{I} \to I \), and \( F \) and \( F_i \) are absolutely continuous, \( \Gamma := ((I, \sigma), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \) and \( \bar{\Gamma} := ((\tilde{I}, \sigma), \mathcal{B}, \tilde{\mu}) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \), and \( \mathcal{F} \) be a separable Banach space. Assume that \( F_i^{(1)} : \tilde{I} \to \mathbb{R} \) satisfies: \( \sup_{x \in \mathbb{R}} \left| F_i^{(1)}(x) \right| < \infty \), \( \forall x_1, x_2 \in \tilde{I} \) with \( x_1 \leq x_2 \)

and \( \tilde{U} := \{ x \in \tilde{U} \mid F_i^{(1)}(x) \neq 0 \} \) is such that \( \tilde{\mu}(\tilde{U} \setminus \tilde{U}) = 0 = \mu(I \setminus F_i(\tilde{U})) \).

Then, \( \forall f : I \to \mathbb{F} \) with \( f \) being absolutely integrable over \( I \), we have

\[
\int_{F(i)} f \, d\mu = \int_{\tilde{U}} f(F_i(x))g(x) \, d\mu(x)
\]

and \( \forall \tilde{f} : \tilde{I} \to \mathcal{F} \) with \( \tilde{f} \) being absolutely integrable over \( \tilde{I} \), we have

\[
\int_{\tilde{I}} \tilde{f} \, d\tilde{\mu} = \int_{F_i(\tilde{I})} \tilde{f}(F(x))g_1(x) \, d\mu(x)
\]

where \( g : \tilde{I} \to (0, \infty) \subset \mathbb{R} \) and \( g_1 : I \to (0, \infty) \subset \mathbb{R} \) are given by \( g(x) = \begin{cases} F_i^{(1)}(x) \mid_{x \in \tilde{U}} \end{cases} \), \( g_1(x) = \frac{1}{F_i^{(1)}(x)} \), \( \forall x \in I \).

**Proof**  By Fundamental Theorems of Calculus, Theorems 12.88 and 12.86, and Proposition 12.81, we have \( F^{(1)} = DF = D_i F \) is \( \mathcal{B} \)-measurable and \( DF = f \, a.e. \) in \( \tilde{I} \), where \( f : I \to \mathbb{R} \) satisfies \( F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) \, dx \).
\[ \int_{x_1}^{x_2} f(s) \, ds, \forall x_1, x_2 \in I \text{ with } x_1 \leq x_2. \] Similar result holds for \( F_i \). Then, the proposition is a direct consequence of Theorem 12.91. \( \square \)

**Theorem 12.99 (Riemann-Lebesgue)** Let \( Y \) be a separable Banach space over \( \mathbb{R} \), \( f \in L_1(\mathbb{R}, Y) \). Then,

\[
\lim_{\|g\| \to \infty} \int \! f(t) \sin(p t) \, d\mu_B(t) = \vartheta_Y
\] (12.7)

**Proof** Since \( f \in \tilde{L}_1(\mathbb{R}, Y) \), then \( f \) is absolutely integrable over \( \mathbb{R} \). Then, \( \tilde{f}_p : \mathbb{R} \to \tilde{Y} \) defined by \( \tilde{f}_p(t) := f(t) \sin(pt), \forall t \in \mathbb{R} \), is absolutely integrable over \( \mathbb{R} \), \( \forall p \in \mathbb{R} \), by Lebesgue Dominated Convergence Theorem 11.91.

First, we consider the special case when \( f = y \chi_{I, \mathbb{R}} \) with \( I \) being an interval and \( y \in \tilde{Y} \) with \( y \neq \vartheta_Y \). Then, \( I \) must be a finite interval, since \( f \in L_1(\mathbb{R}, Y) \). The case when \( I = \emptyset \) is trivial. Let \( \mathcal{T} := [a, b] \) with \( a, b \in \mathbb{R} \) and \( a \leq b \). Then, \( \lim_{\|g\| \to \infty} \int \! f(t) \sin(p t) \, d\mu_B(t) = \lim_{\|g\| \to \infty} \int \! y \sin(pt) \, d\mu_B(t) = \lim_{\|g\| \to \infty} y \frac{1}{p} (\cos(pa) - \cos(pb)) \mu_B(t) = \vartheta_Y \), where the first equality follows from \( I \) being an interval and \( \mathcal{T} = [a, b] \); the second equality follows from Proposition 11.92 and Theorem 12.83. This special case is proved.

Next, we consider the general case when \( f \in \tilde{L}_1(\mathbb{R}, Y) \). By Proposition 11.66, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^{\infty} = \{ \varphi_n : \mathbb{R} \to \tilde{Y}, \forall n \in \mathbb{N} \} \) such that \( \lim_{n \to \infty} \| \varphi_n(x) - f(x) \|_{\mathbb{L}_1(\mathbb{R}, Y)} = 0 \) for \( \varphi_n(x) = \sum_{i=1}^{n} y_i \chi_{E_i, \mathbb{R}} \) where \( n \in \mathbb{N} \), \( y_1, \ldots, y_n \in \tilde{Y} \) are distinct and none equals \( \vartheta_Y \), \( E_1, \ldots, E_n \in B_{\mathbb{R}}(\mathbb{R}) \) are nonempty, pairwise disjoint and of finite measure. By Proposition 11.7, \( \exists a, b \in \mathbb{R} \) with \( a < b \) such that \( \varphi_n(a, b) \) and \( \varphi_n(b, a) \) are nonempty. Let \( f_1 := \varphi_n \chi_{[a, b], \mathbb{R}} \). Then, \( \int \! f \sin(pt) \, d\mu_B < \frac{\varepsilon}{2} \). By Proposition 11.182, \( \exists \) a continuous function \( g : [a, b] \to \tilde{Y} \) such that \( g \in \tilde{L}_1([a, b], Y) \) and \( \| g - f_1 \|_{\mathbb{L}_1([a, b], Y)} < \frac{\varepsilon}{4} \). By Proposition 5.39, \( g \) is uniformly continuous. \( \exists \) \( m, n \in \mathbb{N} \) such that \( |x_1 - x_2| \leq \frac{b-a}{m} \), we have \( \| g(x_1) - g(x_2) \| < \frac{\varepsilon}{3(b-a)} \). Construct a step function \( h : \mathbb{R} \to \tilde{Y} \) by \( h(x) = \begin{cases} 0 & x \in \mathbb{R} \setminus [a, b] \\ g(a + \frac{x-a}{m}) & x \in [a, b] \end{cases} \). Then, we have

\[
\int_{[a, b]} |h|_{[a, b]} - g \, d\mu_B < \frac{\varepsilon}{2}, \quad \text{Hence,} \quad \int_{[a, b]} |f - h| \, d\mu_B \leq \int_{[a, b]} |f - f_1| \, d\mu_B + \int_{[a, b]} |h| \, d\mu_B < \frac{\varepsilon}{2} + \int_{[a, b]} |f_1| \, d\mu_B + g \, d\mu_B \leq \frac{\varepsilon}{2} + \int_{[a, b]} |f_1| \, d\mu_B < \varepsilon. \] Then, \( 0 \leq \liminf_{\|g\| \to \infty} \| \int \! f_p \, d\mu_B \| \leq \limsup_{\|g\| \to \infty} \| \int \! f \sin(pt) \, d\mu_B(t) \| + \int \! f_1 \sin(pt) \, d\mu_B(t) \| < \varepsilon. \]
Proposition 12.100 Let \( J := [a,b] \subset \mathbb{R} \) be a compact interval \((a \leq b)\), \( g : J \to \mathbb{R} \) be of bounded variation (as defined in Definition 12.41) and is nondecreasing; and \( f : J \to \mathbb{R} \) be \( \mathcal{B}(\mathbb{R}) \)-measurable. Assume that \( f \) is bounded, that is, \( \exists M \geq 0 \) such that \(|f(x)| \leq M, \forall x \in J\). If \( f \) is Riemann-Stieltjes integrable with respect to \( g \) over \( J \) (as defined in Definition 29.2 of Bartle (1976)) and let the Riemann-Stieltjes integral be denoted by \( I \), then \( I = \int_{[a,b]} f \, dg \), where the right-hand-side is the Lebesgue-Stieltjes integral defined in Definition 12.57.

Proof By Proposition 12.39, \( P(J) = r_{a,b} \). By Proposition 12.49, there exists a unique finite measure \( \mu \) on the measurable space \((P(J), \mathcal{B}(P(J)))\) such that \( g \) is a cumulative distribution function of \( \mu \), \( \mu([x_1,x_2]) = g(x_2) - g(x_1), \forall a \leq x_1 \leq x_2 \leq b \), and \( \mu(P(J)) = \mu([a,b]) = g(b) - g(a) = T_\gamma < \infty \). Since \( f \) is bounded, by Bounded Convergence Theorem 11.77, \( f \) is absolutely integrable over \( P(J) \) with respect to \( \mu \). Then, by Proposition 11.92, \( I := \int_{[a,b]} f \, dg \in \mathbb{R} \).

By Riemann Criterion for Integrability 30.1 of Bartle (1976), \( \forall n \in \mathbb{N} \), there exists a partition \( P_n \) of \( J \) such that, \( \forall \) refinement \( P := (x_0, \ldots, x_{k_n}) \) of \( P_n \), we have

\[
\left| \sum_{j=1}^{k_n} M_j(g(x_j) - g(x_{j-1})) - I \right| < 2^{-n}; \left| \sum_{j=1}^{k_n} m_j(g(x_j) - g(x_{j-1})) - I \right| < 2^{-n}
\]

where \( M_j := \sup \{ f(x) \in \mathbb{R} \mid x \in [x_{j-1}, x_j] \} \) and \( m_j := \inf \{ f(x) \in \mathbb{R} \mid x \in [x_{j-1}, x_j] \} \). Without loss of generality, we may assume that \( P_{n+1} \) is a refinement of \( P_n \) (otherwise, define \( P_{n+1} \) to be the partition that contains all of the points in \( P_n \) and \( P_{n+1} \)). Then,

\[
I = \lim_{n \to \infty} k_n \int_{[a,b]} \sum_{j=1}^{k_n} M_j g(x_j) - g(x_{j-1}) \, d\mu = \lim_{n \to \infty} \int_{[a,b]} \sum_{j=1}^{k_n} M_j 1_{[x_{j-1}, x_j]} \, d\mu
\]

\[
= \lim_{n \to \infty} \int_{[a,b]} \sum_{j=1}^{k_n} m_j 1_{[x_{j-1}, x_j]} \, d\mu = \lim_{n \to \infty} \int_{[a,b]} h_n \, d\mu
\]
where \( h_n : r_{a,b} \to \mathbb{R} \) and \( \bar{h}_n : r_{a,b} \to \mathbb{R}, \forall n \in \mathbb{N}; \) and the second and fourth equalities follow from Proposition 11.75. Clearly, \(-M \leq h_n(x) \leq h_{n+1}(x) \leq f(x) \leq \bar{h}_{n+1}(x) \leq \bar{h}_n(x) \leq M, \forall x \in r_{a,b}, \forall n \in \mathbb{N}. \) By Proposition 11.48, we have \( \lim_{n \to \infty} h_n = h : r_{a,b} \to \mathbb{R} \) and \( \lim_{n \to \infty} \bar{h}_n = \bar{h} : r_{a,b} \to \mathbb{R} \) are \( \mathcal{B}((r_{a,b}), \mathbb{R}) \)-measurable. By Lebesgue Dominated Convergence Theorem 11.91, we have \( h \) and \( \bar{h} \) are absolutely integrable over \((r_{a,b}, \mathcal{B}((r_{a,b}), \mathbb{R}), \mu)\) and \( \int_{r_{a,b}} h \, d\mu = \lim_{n \to \infty} \int_{r_{a,b}} h_n \, d\mu = I = \lim_{n \to \infty} \int_{r_{a,b}} \bar{h}_n \, d\mu = \int_{r_{a,b}} \bar{h} \, d\mu. \) Note that \( h(x) \leq f(x) \leq \bar{h}(x), \forall x \in r_{a,b} \) and \( f \) is \( \mathcal{B}((\mathbb{R}), \mathbb{R}) \)-measurable by assumption, then, by Proposition 11.92, \( I = \int_{r_{a,b}} h \, d\mu \leq \int_{r_{a,b}} f \, d\mu = \bar{I} \leq \int_{r_{a,b}} \bar{h} \, d\mu = I. \)

This completes the proof of the proposition. \( \square \)

The above can be generalized further. Let \( g : J \to \mathbb{R} \) be of bounded variation, but not necessarily nondecreasing and everything else as assumed in the preceding proposition. Then, by Theorem 12.50, there exists a unique finite \( \mathbb{R} \)-valued measure \( \mu \) on the measurable space \((P(J), \mathcal{B}((P(J))))\) such that \( g \) is a cumulative distribution function of \( \mu, \mathcal{P} \circ \mu(r_1, r_2) = T_g(r_1, r_2), \forall r_1 \leq x_1 \leq x_2 \leq b, \) and \( \mathcal{P} \circ \mu(P(J)) = \mathcal{P} \circ \mu(r_{a,b}) = T_g < \infty. \) By Jordan Decomposition Theorem 11.162, there exists a unique pair of mutually singular finite measures \( \mu_+ \) and \( \mu_- \) on \((r_{a,b}, \mathcal{B}((r_{a,b})), \mathcal{P} \circ \mu(r_{a,b}))\), such that \( \mu = \mu_+ - \mu_- \) and \( \mathcal{P} \circ \mu = \mu_+ + \mu_- \). Let \( G \) be a cumulative distribution function for \( \mathcal{P} \circ \mu \) as delineated in Proposition 12.51 and \( G \) is of bounded variation. We need to assume that \( f \) is Riemann-Stieltjes integrable with respect to \( g \) and \( G \) over \( J \), respectively, and denote these integrals as \( I \) and \( I_G \), respectively. It is easy to see that \( \frac{1}{2}(g+G) \) and \( \frac{1}{2}(G-g) \) are cumulative distribution functions for \( \mu_+ \) and \( \mu_- \), respectively. Then, they are nondecreasing since \( \mu_+ \) and \( \mu_- \) are measures. Furthermore, they are continuous on the right since \( g \) and \( G \) are of bounded variation. Hence, \( \frac{1}{2}(g+G) \) and \( \frac{1}{2}(G-g) \) are of bounded variation. Thus, by Theorem 29.5 of Bartle (1976) and Proposition 12.100, we have \( \frac{1}{2}(I + I_G) = \int_{r_{a,b}} f \, d\mu_+ \) and \( \frac{1}{2}(I_G - I) = \int_{r_{a,b}} f \, d\mu_- \). This leads to \( I = \frac{1}{2}(I + I_G) - \frac{1}{2}(I_G - I) = \int_{r_{a,b}} f \, d\mu_+ - \int_{r_{a,b}} f \, d\mu_- = \int_{r_{a,b}} f \, d\mu = \bar{I}, \) where the last equality follows from Propositions 11.145 and 11.146.

**Theorem 12.101 (Integration by Parts)** Let \( J := [a,b] \subseteq \mathbb{R} \) be a compact interval, \( \mathbb{I} := (P(J), \mathcal{B}, \mu) \) be the finite metric measure subspace of \( \mathbb{R}, f : J \to R \) be of bounded variation and a cumulative distribution function of the finite \( \mathbb{R} \)-valued measure space \((P(J), \mathcal{B}, \nu_f)\), \( F : J \to \mathbb{R} \) be the cumulative distribution function for the finite measure space \((P(J), \mathcal{B}, \mathcal{P} \circ \nu_f)\) as delineated in Proposition 12.51 and is of bounded variation, \( g : J \to R \) be of bounded variation and a cumulative distribution function of the finite \( \mathbb{R} \)-valued measure space \((P(J), \mathcal{B}, \nu_g)\), \( G : J \to \mathbb{R} \) be the cumulative distribution function for the finite measure space \((P(J), \mathcal{B}, \mathcal{P} \circ \nu_g)\) as delineated in Proposition 12.51 and is of bounded variation. Assume that \( f \) is Riemann-Stieltjes integrable with respect to \( g \) and \( G \) over \( J \), respectively, and denote the Riemann-Stieltjes integral of \( f \) with respect to \( g \) over \( J \) as \( I_1 \); and \( g \) is
Riemann-Stieltjes integrable with respect to \( f \) and \( F \) over \( J \), respectively, and denote the Riemann-Stieltjes integral of \( g \) with respect to \( f \) over \( J \) as \( I_2 \). Then, we have

\[
\int_{\tau_a, b} f(x) \, dg(x) + \int_{\tau_a, b} g(x) \, df(x) = I_1 + I_2 = f(b)g(b) - f(a)g(a) \quad (12.8)
\]

**Proof** By Proposition 12.54, \( f - f(a) \) and \( g - g(a) \) are \( \mathcal{B}_B(\mathbb{R}) \)-measurable. Hence, \( f \) and \( g \) are \( \mathcal{B}_B(\mathbb{R}) \)-measurable, by Propositions 11.38, 7.23, and 11.39. Since \( f \) and \( g \) are of bounded variation, then \( f \) and \( g \) are bounded. The first equality in the theorem statement follows from Proposition 12.100 and the discussion immediately after it. The second equality in the theorem statement follows from Integration by Parts 29.7 of Bartle (1976).

This completes the proof of the theorem. \( \Box \)

**Proposition 12.102** Let \( I \subseteq \mathbb{R} \) be an open interval, \( \mathcal{X} := ((I, \cdot), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \), \( \mathcal{Y} \) be separable reflexive Banach space over \( \mathbb{K} \) with \( \mathcal{Y}^* \) being separable, \( (I, \mathcal{B}, \nu) \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space, \( F : I \to \mathcal{Y} \) be a cumulative distribution function of \( \nu \), \( f : I \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( F \) is absolutely continuous, and \( DF(x) = f(x) \) a.e. \( x \in \mathcal{X} \). Then, \( \frac{d\nu}{df} = f \) a.e. in \( \mathcal{X} \).

**Proof** By Fundamental Theorem of Calculus II, Theorem 12.88, \( F(x) = F(a) + \int_a^x f(t) \, dt \). By Proposition 11.116, we may define \( \check{\nu} \) to be \( \sigma \)-finite \( \mathcal{Y} \) valued measure with kernel \( f \) over \( \mathcal{X} \). Then, by Definition 11.166, \( f \) is the Radon-Nikodym derivative of \( \check{\nu} \) with respect to \( \mu \). By Definition 12.42, \( F \) is a cumulative distribution function of \((I, \mathcal{B}, \check{\nu})\). By Theorem 12.50 and Proposition 12.73, we have \( \nu = \check{\nu} \). Then, we have \( \check{f} = \frac{d\nu}{df} \) a.e. in \( \mathcal{X} \). \( \Box \)

**Lemma 12.103** Let \( \Omega \) be a set, \( \mathcal{S} \) be a \( \pi \)-system on \( \Omega \), \( \mathcal{B} \) be the \( \sigma \)-algebra on \( \Omega \) generated by \( \mathcal{S} \), \( \mathcal{Y} \) be a normed linear space, and \( \mu_1 \) and \( \mu_2 \) be finite \( \mathcal{Y} \)-valued measures on \((\Omega, \mathcal{B})\). Assume that \( \mu_1(E) = \mu_2(E) \), \( \forall E \in \mathcal{S} \). Then, \( \mu_1(E) = \mu_2(E) \), \( \forall E \in \mathcal{B} \).

**Proof** Define \( \mathcal{D} := \{ E \in \mathcal{B} \mid \mu_1(E) = \mu_2(E) \} \). Clearly, \( \mathcal{S} \subseteq \mathcal{D} \subseteq \mathcal{B} \).

We will show that \( \mathcal{D} \) is a monotone class on \( \Omega \). Then, by Monotone Class Lemma 12.19, we have \( \mathcal{D} = \mathcal{B} \). Hence, the result holds.

Clearly, (i) \( \emptyset, \Omega \in \mathcal{S} \subseteq \mathcal{D} \). (ii) \( \forall E_1, E_2 \in \mathcal{D} \) with \( E_1 \subseteq E_2 \), we have \( \mu_1(E_2 \setminus E_1) = \mu_2(E_2) - \mu_1(E_1) = \mu_2(E_2) - \mu_2(E_1) = \mu_2(E_2 \setminus E_1) \), where the first and last equalities follows from Definitions 11.108 and 11.109; and the second equality follows from the fact \( E_1, E_2 \in \mathcal{D} \). This implies that \( E_2 \setminus E_1 \in \mathcal{D} \). (iii) \( \forall (E_i)_{i=1}^\infty \subseteq \mathcal{D} \) with \( E_i \subseteq E_{i+1}, \forall i \in \mathbb{N} \), we have \( \bigcup_{i=1}^\infty E_i = \bigcup_{i=1}^\infty A_i \), where \( A_1 := E_1, A_2 := E_2 \setminus E_1, A_i := E_i \setminus E_{i-1}, \forall i \in \{3, 4, \ldots \} \). Clearly, \( A_i \)'s are pairwise disjoint. By (ii), we have \( A_i \in \mathcal{D}, \forall i \in \)
Hence, we have shown that \( S \) is some measure between \( \mu \). Then, \( \bar{g} \) is a homeomorphism.

On the other hand, \( g \) is not a \( S \)-homeomorphism. Thus, we may define \( \hat{\mu} := \mu |_S \) such that \( \hat{\mu} |_S = \mu |_S \).

Clearly, \( \hat{\mu} \) is countably additive on \( S \). Since \( \mu_{B_n} \) is countably additive, let \( \mathcal{A} \subseteq S_{n-2} \) be the algebra generated by \( S \). By Proposition 2.8, \( \forall A \in \mathcal{A} \), \( \exists n, m \in \mathbb{N}, \forall i_1, \ldots, i_{2n} \in [1, \ldots, m], \exists F_{i_1, \ldots, i_{2n}} \subseteq X \) with \( F_{i_1, \ldots, i_{2n}} \in \mathcal{O} \) or \( S_{n-1} \setminus F_{i_1, \ldots, i_{2n}} \in \mathcal{O} \) such that \( A = \bigcup_{i_1=1}^{m} \bigcap_{i_2=1}^{m} \cdots \bigcap_{i_{2n}=1}^{m} F_{i_1, \ldots, i_{2n}} \).

Thus, we may define \( \hat{\mu} : \mathcal{A} \rightarrow [0, \infty) \subset \mathbb{R} \) by \( \hat{\mu}(A) := n \mu_{B_n}(V_A), \forall A \in \mathcal{A} \).

Clearly, \( \hat{\mu} \) is an extension of \( \mu \) to \( \mathcal{A} \), and it is a finite measure on the algebra \( \mathcal{A} \). By Carathéodory Extension Theorem 11.19, \( \hat{\mu} \) admits unique extension to a finite measure \( \hat{\mu} \) on the \( \sigma \)-algebra \( \mathcal{B} \). Then, \( \hat{\mu} : \mathcal{B} \rightarrow [0, \infty) \subset \mathbb{R} \) and \( \hat{\mu}|_{\mathcal{O}} = \mu |_{\mathcal{O}} \).

Note that \( (S_{n-1}, | \cdot |) \) is second countable by Propositions 4.4 and 4.38. By Proposition 5.40, it is compact. Thus, \( (S_{n-1}, | \cdot |) \) is a compact (therefore locally compact) separable metric space and \( \mu \) is a finite measure on \( \mathcal{B} = \mathcal{B}_\mathbb{R}(S_{n-1}, \mathcal{O}) \). By Theorem 11.198, \( S_{n-1} := (S_{n-1}, | \cdot |), \mathcal{B}, \hat{\mu} \) is a finite compact separable metric measure space.

Example 12.105 Let \( \mathbb{I} := ((r_{-\pi, \pi}, | \cdot |), \mathcal{B}, \mu) \) be the finite metric measure subspace of \( \mathbb{R} \), and \( S_1 := ((S_1, | \cdot |), \mathcal{B}, \mu) \) be the finite metric measure space as defined in Example 12.104. Define a mapping \( g : r_{-\pi, \pi} \rightarrow S_1 \) by \( g(x) = (\cos(x), \sin(x)) \in S_1, \forall x \in r_{-\pi, \pi} \). It is easy to see that \( g \) is bijective and \( g \) is continuous.

But, \( g_{\text{inv}} \) is not continuous and therefore \( g \) is not a homeomorphism. \( \forall O \in \mathcal{O}_{S_1} \), we have \( g_{\text{inv}}(O) \in \mathcal{O}_1 \) since \( g \) is continuous.

On the other hand, \( \forall O \in \mathcal{O}_1 \), we will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \pi \in O \); and Case 2: \( \pi \notin O \). Case 1: \( \pi \in O \).

Then, \( \overline{O} := O \setminus \{ \pi \} \in \mathcal{O}_1, \overline{O} \in \mathcal{O}_\mathbb{R} \), and \( g(\overline{O}) \in \mathcal{O}_{S_1} \). Then, \( g(\overline{O}) = g(O) \cup \{ g(\pi) \} \in \mathcal{B} \). Case 2: \( \pi \notin O \).

Then, \( O \in \mathcal{O}_\mathbb{R} \) and \( g(O) \in \mathcal{O}_{S_1} \subseteq \mathcal{B} \).

Hence, we have shown that \( g(O) \in \mathcal{B}, \forall O \in \mathcal{O}_1 \). By Proposition 11.34, we have \( g(B) \in \mathcal{B}, \forall B \in \mathcal{B} \) and \( g_{\text{inv}}(B) \in \mathcal{B}, \forall B \in \mathcal{B} \). Hence, \( g_{\text{inv}} \) is an isomorphism between \( S_1 \) and \( (r_{-\pi, \pi}, \mathcal{B}, g_{\text{inv}}) \).
It is easy to see that \( \mu(I_1) = \hat{\mu}(g(I_1)) = g_{\text{inv}}(I_1) \), \( \forall I_1 \subseteq I \). By Lemma 12.103, \( \mu = g_{\text{inv}} \). Hence, \( g \) is an isomasure between \( I \) and \( S_1 \). \( \forall p \in [1, \infty) \subseteq \mathbb{R} \), let \( X \) be a separable Banach space over \( K \). By Example 11.179, we have \( L_p(I, X) = L_p(S_1, X) \) in the sense that \( \forall z \in L_p(I, X) \), we have \( z \circ g_{\text{inv}} \in L_p(S_1, X) \) and \( \forall z \in L_p(S_1, X) \), we have \( z \circ g \in L_p(I, X) \).

\[ \diamond \]

**12.8 Sobolev Spaces**

In this section, we will be concerned with absolutely continuous \( \mathcal{Y} \)-valued function on an open rectangle \( \Omega \) of \( \mathbb{R}^m \), with \( m \in \mathbb{N} \), where \( \mathcal{Y} \) is a separable Banach space over \( K \) with \( \mathcal{Y}^* \) being separable. We restrict ourself to consider such functions that further has a finite \( L_p(X, \mathcal{Y}) \) pseudo-norm, where \( p \in [1, \infty) \subseteq \mathbb{R} \). We will be using the notations of the Fundamental Theorem of Calculus II, Theorem 12.88, without repeating their definitions. For any such function \( f \), there exists a set of functions \( f_{J,x_0} : \pi_J(\Omega) \to \mathcal{Y} \), which are \( \mathcal{B}_J \)-measurable and absolutely integrable over any bounded rectangles of \( \mathcal{X}_J \), \( \forall J \subseteq J := \{1, \ldots, m\} \) with \( J \neq \emptyset \), and any fixed \( x_0 \in \Omega \). These are the stream functions of \( f \). They form a unique set of functions, except on a set of measure zero in \( \mathcal{X}_J \) for each of the stream functions. They together satisfy (12.3). Theorem 12.88 clearly states that these functions determines the partial derivative functions \( D_i f \), \( \forall i \in \mathcal{J} \). We will further restrict ourself to absolutely continuous functions \( f \) such that \( f_{J,x_0} \) is in \( L_p(X, \mathcal{Y}) \), \( \forall J \subseteq J \) with \( J \neq \emptyset \). It should be clear that the \( (f + g)_{J,x_0} = f_{J,x_0} + g_{J,x_0} \) a.e. in \( \mathcal{X}_J \), \( \forall J \subseteq J \) with \( J \neq \emptyset \), \( \forall f, g \) under consideration. This shows that such functions form a vector space over \( K \). We will define a norm on this space as

\[
\| f \|_{z,x_0} := (\int_\Omega P_p \circ f(x) \, d\mu(x) + \sum_{J \subseteq \mathcal{J}} \int_{\pi_J(\Omega)} P_p \circ f_{J,x_0}(s) \, d\mu_J(s))^{1/p}.
\]

The value of the norm does depends on the point \( x_0 \). But the functions included in \( Z \) are independent of the choice of \( x_0 \) when \( \mathcal{Y} = \mathbb{R}^n \) with \( n \in \mathbb{N} \). We will show that this defines a Banach space over \( K \), which will be called the 1st order Sobolev space.

**Example 12.106** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( p \in [1, \infty) \subseteq \mathbb{R} \), \( \Omega \in \mathcal{B}_\mathcal{Y} (\mathbb{R}^m) \) be a nonempty open rectangle, \( X := (\| \cdot \|, \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), and \( \mathcal{Y} \) be a separable Banach space over \( K \) with \( \mathcal{Y}^* \) being separable. Fix any \( x_0 \in \Omega \). Let \( Z := \{ f \in L_p(X, \mathcal{Y}) \mid f \text{ is absolutely continuous}, \text{ and the set of stream functions of } f \text{ with respect to } x_0 \text{ are in } L_p(X, \mathcal{Y}), \forall J \subseteq J \text{ with } J \neq \emptyset, \text{ respectively} \} \), where \( J := \{1, \ldots, m\} \), \( \mathcal{X}_J := (\pi_J(\Omega), \mathcal{B}_J, \mu_J) \) is as defined in Theorem 12.88. On this set \( Z \), we will define the norm \( \| \cdot \|_{z,x_0} : \)
12.8. SOBOLEV SPACES

\[
Z \rightarrow [0, \infty) \subset \mathbb{R} \text{ by }
\]

\[
\|f\|_{Z,x_0} := \left( \int_{\Omega} P_p \circ f(x) \, d\mu(x) + \sum_{J \in J} \int_{\pi_J(\Omega)} P_p \circ f_{J,x_0}(s) \, d\mu_J(s) \right)^{1/p} \tag{12.9}
\]

We will show that \((Z, \|\cdot\|_{Z,x_0}, \mathbb{K})\) forms a Banach space over \(\mathbb{K}\), which will be called the 1st order Sobolev space of \(X\) to \(Y\) under \(p\)-norm with respect to \(x_0 \in \mathbb{X}\) and denoted by \(W_{p,1,x_0}(X,Y)\).

\[\forall f_1 \in Z, \|\cdot\|_{Z,x_0} : Z \rightarrow [0, \infty) \subset \mathbb{R}. \text{ If } \|f_1\|_{Z,x_0} = 0, \text{ then } f_1 = 0. \text{ Hence, } f_1 + f_2 \in Z. \forall \alpha \in \mathbb{K}, \text{ by Proposition 12.69, } \alpha f_1 \text{ is absolutely continuous. It is easy to see that } (\alpha f_1)_{J,x_0} = \alpha f_{1,J,x_0} \text{ a.e. in } \mathbb{X}_J. \text{ Thus, we have } (\alpha f_1)_{J,x_0} \in \bar{L}_p(\mathbb{X}_J,Y), \forall J \subseteq J \text{ with } J \neq \emptyset. \text{ Hence, } \alpha f_1 \in Z. \text{ This paragraph then proves that } (Z, \mathbb{K}) \text{ is a subspace of } \bar{L}_p(\mathbb{X},Y) \text{ and hence a vector space over } \mathbb{K}.
\]

Clearly \(\|\cdot\|_{Z,x_0} : Z \rightarrow [0, \infty) \subset \mathbb{R}. \text{ If } \|f_1\|_{Z,x_0} = 0, \text{ we have } f_1(x) = \vartheta_Y \text{ a.e. } x \in \mathbb{X} \text{ and } f_{1,J,x_0}(x) = \vartheta_Y \text{ a.e. } x \in \mathbb{X}_J. \text{ Since } f_1 \text{ is absolutely continuous, then } f_1(x) = \vartheta_Y, \forall x \in \mathbb{X}. \text{ Thus, } f_1 = \vartheta_Z.
\]

Note that, \(\forall \alpha \in \mathbb{K},\)

\[
\|\alpha f_1\|_{Z,x_0} = \left( \int_{\Omega} P_p \circ (\alpha f_1)(x) \, d\mu(x) + \sum_{J \in J} \int_{\pi_J(\Omega)} P_p \circ (\alpha f_{1,J,x_0})(s) \, d\mu_J(s) \right)^{1/p} = \|\alpha\|_p \|f_1\|_{Z,x_0}
\]

To prove the triangle inequality, we will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\|f_1\|_{Z,x_0} \|f_2\|_{Z,x_0} = 0\); Case 2: \(\|f_1\|_{Z,x_0} \|f_2\|_{Z,x_0} > 0\). Without loss of generality, assume \(\|f_2\|_{Z,x_0} = 0\). Then, \(f_2 = \vartheta_Z \) and \(f_2(x) = \vartheta_Y, \forall x \in \Omega\). Then, we have \(\|f_1 + f_2\|_{Z,x_0} = \|f_1\|_{Z,x_0} = \|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}\). This case is proved.

Case 2: \(\|f_1\|_{Z,x_0} \|f_2\|_{Z,x_0} > 0\). Then, \(\|f_1\|_{Z,x_0} > 0\) and \(\|f_2\|_{Z,x_0} > 0\). Let \(\lambda := \|f_1\|_{Z,x_0}/\|f_2\|_{Z,x_0} \in (0, 1) \subset \mathbb{R} \). Then, \(\forall x \in \mathbb{X},\) we have

\[
\frac{\|f_1(x)\|_Y^p}{\|f_1\|_{Z,x_0} + \|f_2(x)\|_Y^p} \leq \left( \frac{\|f_1(x)\|_Y + \|f_2(x)\|_Y}{\|f_1\|_{Z,x_0} + \|f_2(x)\|_Y} \right)^p = \left( \frac{\|f_1(x)\|_Y}{\|f_1\|_{Z,x_0}} + (1 - \lambda) \frac{\|f_2(x)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p \leq \lambda \left( \frac{\|f_1(x)\|_Y}{\|f_1\|_{Z,x_0}} \right)^p + (1 - \lambda) \left( \frac{\|f_2(x)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p \leq \lambda \left( \frac{\|f_1(x)\|_Y}{\|f_1\|_{Z,x_0}} \right)^p + (1 - \lambda) \left( \frac{\|f_2(x)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p,
\]

where the first inequality follows from Definition 7.1; and the second inequality follows from the convexity of the function \(p^p, \forall l \in [0, \infty) \subset \mathbb{R} \). By Proposition 11.83, we have

\[
\int_0^1 \frac{P_p \circ f_1(x) \, d\mu(x)}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}} \leq \lambda \int_0^1 \frac{P_p \circ f_1(x) \, d\mu(x)}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}} + (1 - \lambda) \int_0^1 \frac{P_p \circ f_2(x) \, d\mu(x)}{\|f_2\|_{Z,x_0} + \|f_2\|_{Z,x_0}}.
\]
Fix any $J \subseteq \bar{J}$ with $J \neq \emptyset$, we have, $\forall s \in \pi_J(\Omega)$,
\[
\frac{\|f_{1,J,x_0}(s) + f_{2,J,x_0}(s)\|_p^p}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}}^p \leq \left( \frac{\|f_{1,J,x_0}(s)\|_Y + \|f_{2,J,x_0}(s)\|_Y}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}} \right)^p
\]
\[
= \left( \lambda \frac{\|f_{1,J,x_0}(s)\|_Y}{\|f_1\|_{Z,x_0}} + (1 - \lambda) \frac{\|f_{2,J,x_0}(s)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p \leq \lambda \left( \frac{\|f_{1,J,x_0}(s)\|_Y}{\|f_1\|_{Z,x_0}} \right)^p + (1 - \lambda) \left( \frac{\|f_{2,J,x_0}(s)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p
\]
where the first inequality follows from Definition 7.1; and the second inequality follows from the convexity of the function $\| \cdot \|_p^p$, $\forall t \in [0, \infty) \subset \mathbb{R}$. By Proposition 11.83, we have
\[
\frac{\lambda^{\|f_{J,x_0}(s)\|_p \circ f_{J,x_0} \circ f_{J,x_0}(s) \mu_j(s)}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}}^p} \leq \lambda^{\|f_{J,x_0}(s)\|_p \circ f_{J,x_0} \circ f_{J,x_0}(s) \mu_j(s)}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}}^p
\]
Thus, we have
\[
\frac{\|f_1 + f_2\|_{Z,x_0}^p}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}}^p \leq \frac{1}{\|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}}^p \left( \int_\Omega P_p \circ (f_1 + f_2) \mu(x) \right)
\]
\[
+ \sum_{J \subseteq \bar{J}, J \neq \emptyset} \int_{\pi_J(\Omega)} P_p \circ (f_{1,J,x_0} + f_{2,J,x_0}) \mu_J(s)
\]
\[
\leq \lambda \frac{\|f_1\|_{Z,x_0}}{\|f_1\|_{Z,x_0}}^p \left( \int_\Omega P_p \circ f_1 \mu(x) \right) + (1 - \lambda) \frac{\|f_2\|_{Z,x_0}}{\|f_2\|_{Z,x_0}}^p \left( \int_\Omega P_p \circ f_2 \mu(x) \right)
\]
\[
+ \sum_{J \subseteq \bar{J}, J \neq \emptyset} \left( \frac{\|f_{1,J,x_0}(s)\|_Y}{\|f_1\|_{Z,x_0}} \right)^p \left( \int_\Omega P_p \circ f_{1,J,x_0} \mu_j(s) \right)
\]
\[
+ (1 - \lambda) \left( \frac{\|f_{2,J,x_0}(s)\|_Y}{\|f_2\|_{Z,x_0}} \right)^p \left( \int_\Omega P_p \circ f_{2,J,x_0} \mu_j(s) \right)
\]
\[
= \frac{\|f_1\|_{Z,x_0}^p}{\|f_2\|_{Z,x_0}^p} \left( \int_\Omega P_p \circ f_1 \mu(x) + \sum_{J \subseteq \bar{J}, J \neq \emptyset} \int_{\pi_J(\Omega)} P_p \circ f_{1,J,x_0} \mu_j(s) \right)
\]
\[
+ \frac{1}{\|f_2\|_{Z,x_0}^p} \left( \int_\Omega P_p \circ f_2 \mu(x) + \sum_{J \subseteq \bar{J}, J \neq \emptyset} \int_{\pi_J(\Omega)} P_p \circ f_{2,J,x_0} \mu_j(s) \right)
\]
\[
= 1
\]
Hence, we have $\|f_1 + f_2\|_{Z,x_0} \leq \|f_1\|_{Z,x_0} + \|f_2\|_{Z,x_0}$. This case is proved.

Based on both of the two cases, we have the triangle inequality for $\| \cdot \|_Z$. So far, we have shown that $(Z, \| \cdot \|_{Z,x_0}, \mathcal{K})$ is a normed linear space.
Finally, we will show that the space is complete. Fix any Cauchy sequence \((f_l)_{l=1}^{\infty} \subseteq \mathcal{Z}\). By the definition of the norm \(\| \cdot \|_{\mathcal{Z}, x_0}\), we have \((f_l)_{l=1}^{\infty} \subseteq \mathcal{L}_p(\mathcal{X}, \mathcal{Y})\) is a Cauchy sequence, and \((f_{l,J,x_0})_{l=1}^{\infty} \subseteq \mathcal{L}_p(\mathcal{X}_J, \mathcal{Y})\) is a Cauchy sequence, \(\forall J \subseteq \bar{J}\) with \(J \neq \emptyset\). By the completeness of \(\mathcal{L}_p\) as proved in Example 11.179, we have \(\exists F \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y})\) and \(\exists F_{l,J,x_0} \in \mathcal{L}_p(\mathcal{X}_J, \mathcal{Y})\), such that \(\lim_{l \in \mathbb{N}} f_l = F\) in \(\mathcal{L}_p(\mathcal{X}, \mathcal{Y})\) and \(\lim_{l \in \mathbb{N}} f_{l,J,x_0} = F_{l,J,x_0}\) in \(\mathcal{L}_p(\mathcal{X}_J, \mathcal{Y})\), \(\forall J \subseteq \bar{J}\) with \(J \neq \emptyset\). This implies that \(\lim_{l \in \mathbb{N}} f_{l,J,x_0} = F\) a.e. in \(\mathcal{X}_J\) and \(\lim_{l \in \mathbb{N}} f_l = F\) a.e. in \(\mathcal{X}\). We need to construct a function \(F \in \mathcal{Z}\) such that its stream functions with respect to \(x_0\) are \(F_{l,J,x_0}\)'s. It should be noted that \(F_{l,J,x_0}\)'s are specified exactly except on a set of measure zero, which are sufficient to serve as stream functions. Note that \(\forall l \in \mathbb{N}, \forall x \in \Omega,\) we have

\[
 f_l(x) = \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \cap J')} \int_{\pi_{J}(r_{x}, x)} f_{l,J,x_0} d\mu_J + f_l(x_0)
\]

where \(\hat{x} := x \wedge x_0, \hat{x} := x \vee x_0, \hat{J} := \{ i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x) \}\). (This is exactly (12.3)). As \(l \to \infty\), we have

\[
 \lim_{l \in \mathbb{N}} (f_l(x) - f_l(x_0)) = \lim_{l \in \mathbb{N}} \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \cap J')} \int_{\pi_{J}(r_{x}, x)} f_{l,J,x_0} d\mu_J
\]

\[
 = \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \cap J')} \int_{\pi_{J}(r_{x}, x)} F_{l,J,x_0} d\mu_J
\]

where the last equality follows from Lebesgue Dominated Convergence Theorem 11.91, the fact that \(F_{l,J,x_0} \in \mathcal{L}_p(\mathcal{X}_J, \mathcal{Y})\) and the domain of integration are of finite measure. This formula determines the function \(F\) up to \(F(x_0)\). Since \(\lim_{l \in \mathbb{N}} f_l = F\) a.e. in \(\mathcal{X}\), then \(\exists x_1 \in \Omega,\) (possibly different from \(x_0\)) such that \(\lim_{l \in \mathbb{N}} f_l(x_1) = F(x_1)\). Then, we may define \(F(x_0) := \hat{F}(x_1) - \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \cap J')} \int_{\pi_{J}(r_{x}, x_1)} F_{l,J,x_0} d\mu_J\), where \(\hat{x}_1 := x_1 \wedge x_0, \hat{x}_1 := x_1 \vee x_0, \hat{J}_1 := \{ i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x_1) \}\). Therefore, we may define \(F : \Omega \to \mathcal{Y}\) by, \(\forall x \in \Omega,\)

\[
 F(x) = F(x_0) + \sum_{J \subseteq \bar{J}} (-1)^{\text{card}(J \cap J')} \int_{\pi_{J}(r_{x}, x_0)} F_{l,J,x_0} d\mu_J
\]

By Fundamental Theorem of Calculus I, Theorem 12.86, each term in the above sum is a absolutely continuous function on \(\pi_{J}(\Omega)\) and therefore absolutely continuous on \(\Omega\) by Definition 12.58. By Propositions 12.69 and 12.61, \(F\) is absolutely continuous on \(\Omega\). It is easy to see that \(F = \hat{F}\) a.e. in \(\mathcal{X}\), and \(F_{l,J,x_0}\)'s are the stream functions of \(F\) with respect to \(x_0\) by Fundamental Theorem of Calculus II, Theorem 12.88. Hence, \(F \in \mathcal{Z}\). Then, it is straightforward to check that \(\lim_{l \in \mathbb{N}} \| f_l - F \|_{\mathcal{Z}, x_0} = 0\) and \(\lim_{l \in \mathbb{N}} f_l = F\) in \(W_{p,1,x_0}(\mathcal{X}, \mathcal{Y})\).
Hence, any Cauchy sequence in $Z$ converges to an element in $Z$. This proves that $W_{p,1,x_0}(X,Y)$ is complete, and therefore, it is a Banach space over $K$. \hfill \Box

We note that the above definition of Sobolev space agrees with that in Zeidler (1995) when $m = 1$. In this case, the norm does not depend on the choice of $x_0 \in X$.

**Example 12.107** We will now show that $\forall f \in W_{p,1,x_0}(X,Y)$ implies that $f \in W_{p,1,x_0}(X,Y)$, $\forall \bar{x} \in \bar{X}$, under the assumptions of Example 12.106.

Fix any $f \in W_{p,1,x_0}(X,Y)$, we have that $f \in L_p(\mathcal{X},Y)$ and it is absolutely continuous, and its stream functions $f_{J,x_0} \in L_p(\mathcal{X},\mathbb{R})$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$. Fix any $\bar{x} \in \bar{X}$. The stream functions of $f$ with respect to $\bar{x}$ are given by (12.4): $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$.

$$f_{J,\bar{x}}(s) = \begin{cases} \sum_{J \subsetneq J \subsetneq \bar{J}} (-1)^{|J \cap \bar{J} \setminus J|} f_{J,x_0}(\pi_J(s), \pi_J(t))) \mu_{J,x_0}(\pi_J(s), \pi_J(t)) + f_{J,x_0}(s) & s \in \bar{U}_J \\ \theta_y & s \in \pi_J(\Omega) \setminus \bar{U}_J \end{cases}$$

where $\bar{J} := \{ i \in \bar{J} \mid \pi_i(x_0) > \pi_i(\bar{x}) \}$, $\bar{x} := x \wedge x_0$, and $\bar{\bar{x}} := \bar{x} \vee x_0$; and $\bar{U}_J \in B_J$ with $\mu_J(\pi_J(\Omega) \setminus \bar{U}_J) = 0$. All we need to do is to show that $f_{J,\bar{x}} \in L_p(X,Y)$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$, under the assumption that $f_{J,x_0} \in L_p(X,Y)$, $\forall J \subseteq \bar{J}$ with $J \neq \emptyset$.

Fix any $J \subseteq \bar{J}$ with $J \neq \emptyset$. Then, we have

$$\|f_{J,\bar{x}}\|_p \leq \|f_{J,x_0}\|_p + \sum_{J \subsetneq J \subsetneq \bar{J}} \left( \int_{\bar{U}_J} \left( \int_{\pi_J(s)}^{\pi_J(t)} f_{J,x_0}(\pi_J(s), \pi_J(t))) \mu_{J,x_0}(\pi_J(s), \pi_J(t)))^{p} \right) \frac{1}{p} \right)$$

$$= \|f_{J,x_0}\|_p + \sum_{J \subsetneq J \subsetneq \bar{J}} \left( \int_{\bar{U}_J} \left( \int_{\pi_J(s)}^{\pi_J(t)} f_{J,x_0}(\pi_J(s), \pi_J(t))) \mu_{J,x_0}(\pi_J(s), \pi_J(t)))^{p} \right) \frac{1}{p} \right)$$

$$\leq \|f_{J,x_0}\|_p + \sum_{J \subsetneq J \subsetneq \bar{J}} \left( \int_{\bar{U}_J} \left( \int_{\pi_J(s)}^{\pi_J(t)} f_{J,x_0}(\pi_J(s), \pi_J(t))) \mu_{J,x_0}(\pi_J(s), \pi_J(t)))^{p} \right) \frac{1}{p} \right)$$
12.8. SOBOLEV SPACES

\[ \left( \int_{J,x} \left( \frac{d\mu_J(s)}{|x|} \right)^{1/p} \right)^p \leq \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\tau_{J'}} f_{J'}(r_{\delta J'}) \right)^{p/q} \int_{\pi_{J'}(r_{\delta J'})} f_{J'}(r_{\delta J'}) \right) \]

\[ \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\pi_{J'}(\Omega)} f_{J'}(r_{\delta J'}) \right) \]

\[ \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\pi_{J'}(\Omega)} f_{J'}(r_{\delta J'}) \right) \]

\[ \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\pi_{J'}(\Omega)} f_{J'}(r_{\delta J'}) \right) \]

\[ \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\pi_{J'}(\Omega)} f_{J'}(r_{\delta J'}) \right) \]

\[ \| f_{J,x_0} \|_p + \sum_{J' \subseteq J} \left( \int_{U_{J'}} \left( \frac{d\mu_{J'}}{|x|} \right)^{p/q} \int_{\pi_{J'}(\Omega)} f_{J'}(r_{\delta J'}) \right) \]

where the first inequality follows from Minkowski’s Inequality, Theorem 11.174; the first equality follows from Proposition 7.85 and \( y_* (s) \in Y^* \) with \( \| y_* (s) \|_{Y^*} \leq 1, \forall s \in U_J \); the second equality follows from Proposition 11.92; the second inequality follows from Hölder’s Inequality, Theorem 11.178; the third inequality follows from Proposition 11.83; the fourth inequality follows from Fubini’s Theorem 12.30; and the last inequality follows from our assumption. This shows that \( f_{J,x} \in L_p(X_J, Y) \). Hence, the desired conclusion follows.

\[ \diamond \]

**Proposition 12.108** Let \( m \in \mathbb{N}, \Omega \in B_B(\mathbb{R}^m) \) be a nonempty open rectangle, \( X := ((\Omega, |\cdot|, \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( J := \{1, \ldots, m\} \), and \( \bar{y} \) be a separable Banach space over \( K \) with \( Y^* \) being separable, \( x_0 \in \Omega, \mathcal{Z} := W_{p,1,x_0}(X, Y), f \in \mathcal{Z}, D_i f : U_i \to \bar{y} \) be the partial derivative of \( f \) along the \( i \)th coordinate direction, \( i \in \bar{J}, f_{J,x_0} : \pi_J(\Omega) \to \bar{y} \) be the stream functions of \( f \) with respect to \( x_0, \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \). Then, we have the following results.

(i) \( \forall i \in \bar{J}, U_i \in B \) with \( \mu(\Omega \setminus U_i) = 0 \).

(ii) \( \forall t \in X_{\{1\}}, \forall s_1, s_2 \in X_{\{1\}}, \) we have

\[ D_i f(M_{\{1\}}(s,t)) = f_{\{1\},M_{\{1\}}(\pi_{\{1\}(x_0),t)}(s) \ a.e. \ s \in X_{\{1\}} \]

\[ f(M_{\{1\}}(s_1,t)) - f(M_{\{1\}}(s_2,t)) = \int_{s_2}^{s_1} f_{\{1\},M_{\{1\}}(\pi_{\{1\}(x_0),t)}(s) \ ds, \ and \]

\[ f_{\{1\},M_{\{1\}}(\pi_{\{1\}(x_0),t)} \in L_p(X_{\{1\}}, \bar{y}). \]
Proof The results are direct consequences of (iii) of Fundamental Theorem of Calculus II, Theorem 12.88, and its proof and Example 12.107.
Chapter 13

Hilbert Spaces

Hilbert spaces are vector spaces equipped with inner products. They possess a wealth of structural properties generalizing basic geometrical insights. The concepts of orthonormal basis, Fourier series, and least-square minimization all have natural setting in Hilbert spaces.

13.1 Fundamental Notions

Definition 13.1 A pre-Hilbert space $\mathcal{X}$ over $\mathbb{K}$ is a vector space $\mathcal{X}$ over $\mathbb{K}$, together with an inner product $\langle \cdot , \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{K}$ satisfying, $\forall x, y, z \in \mathcal{X}, \forall \lambda \in \mathbb{K}$,

(i) $\langle x, y \rangle = \langle y, x \rangle$;

(ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;

(iii) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$;

(iv) $\langle x, x \rangle \in [0, \infty) \subset \mathbb{R}$ and $\langle x, x \rangle = 0$ if, and only if, $x = \varnothing_X$.

Proposition 13.2 Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$. Define $\| \cdot \| : \mathcal{X} \to [0, \infty) \subset \mathbb{R}$ by $\| x \| = \sqrt{\langle x, x \rangle}$, $\forall x \in \mathcal{X}$. Then, $\forall x, y, z \in \mathcal{X}, \forall \alpha, \beta \in \mathbb{K}$, we have

(i) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;

(ii) $\langle z, \alpha x + \beta y \rangle = \overline{\alpha} \langle z, x \rangle + \overline{\beta} \langle z, y \rangle$;

(iii) $\langle \varnothing_X, x \rangle = 0 = \langle x, \varnothing_X \rangle$;

(iv) $| \langle x, y \rangle | \leq \| x \| \| y \|$, where equality holds if, and only if, $y = \varnothing_X$ or $x = \lambda y$ for some $\lambda \in \mathbb{K}$; (Cauchy-Schwarz Inequality)

(v) $\| x + y \| \leq \| x \| + \| y \|$, where equality holds if, and only if, $y = \varnothing_X$ or $x = \lambda y$, for some $\lambda \in [0, \infty) \subset \mathbb{R}$.
(vi) $\| \cdot \|$ is a norm of $X$. (Induced norm)

**Proof**

(i) $\langle \alpha x + \beta y, z \rangle = \langle \alpha x, z \rangle + \langle \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$, where the first equality follows from (ii) of Definition 13.1; and the second equality follows from (iii) of Definition 13.1.

(ii) $\langle z, \alpha x + \beta y \rangle = \langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle = \overline{\alpha} \langle z, x \rangle + \overline{\beta} \langle z, y \rangle$, where the first equality follows from (i) of Definition 13.1; the second equality follows from (ii); and the fourth equality follows from (i) of Definition 13.1.

(iii) $\langle \vartheta x, y \rangle = \langle y - y, x \rangle = \langle y, x \rangle - \langle y, x \rangle = 0$, where the second equality follows from (i). Then, $\langle x, \vartheta x \rangle = \langle \vartheta x, x \rangle = 0$.

(iv) We will distinguish two exhaustive and mutually exclusives cases:

Case 1: $y = \vartheta X$; Case 2: $y \neq \vartheta X$. Case 1: $y = \vartheta X$. Then, $|\langle x, y \rangle| = 0 = \|x\| \cdot 0 = \|x\| \|y\|$. This case is proved.

Case 2: $y \neq \vartheta X$. By (iv) of Definition 13.1, we have $\langle y, y \rangle > 0$.

$\forall \lambda \in \mathbb{R}$, we have $0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda^2 \langle y, y \rangle$, where the inequality follows from (iv) of Definition 13.1; the first equality follows from (i); and the second equality follows from (ii). Take $\lambda := \langle x, y \rangle / \langle y, y \rangle$. Then, $0 \leq \langle x, x \rangle - \lambda \langle x, y \rangle - 2 \lambda \langle y, y \rangle + \lambda^2 \langle y, y \rangle = \langle x, x \rangle - \|x\|^2 / \langle y, y \rangle$, where the equality follows from (i) of Definition 13.1. Hence, $|\langle x, y \rangle| \leq \|x\| \|y\|$. Equality holds if, and only if, $x - \lambda y = \vartheta X$. This case is also proved. Hence, (iv) holds.

(v) Note that $\|x+y\|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, y \rangle + \langle y, x \rangle = \|x\|^2 + 2 \Re (\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 \leq (\|x\| + \|y\|)^2$, where the first equality follows from (i) and (ii); the second equality follows from (i) of Definition 13.1; the second inequality follows from (iv). Hence, $\|x+y\| \leq \|x\| + \|y\|$. Equality holds if, and only if, $\Re (\langle x, y \rangle) = |\langle x, y \rangle| = \|x\| \|y\|$. By (iv), have equality holds if, and only if, $y = \vartheta X$ or $x = \lambda y$ for some $\lambda \in [0, \infty) \subset \mathbb{R}$.

(vi) Clearly, $\|x\| \in [0, \infty) \subset \mathbb{R}$. $\|x\| = 0 \iff \langle x, x \rangle = 0 \iff x = \vartheta X$. Hence, (i) of Definition 7.1 holds. By (v), (ii) of Definition 7.1 holds. Note also that $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha^2 \langle x, x \rangle} = |\alpha| \|x\|$, where the second equality follows from (i) and (ii); and the third equality follows from the fact that $\alpha \overline{\alpha} = |\alpha|^2$. Hence, (iii) of Definition 7.1 holds. Therefore, $\| \cdot \|$ defines a norm of $X$.

**Lemma 13.3** (Parallelogram Law) *Let $X$ be a pre-Hilbert space. $\forall x, y \in X$, we have $\|x + y\|^2 + \|x - y\|^2 = 2 \|x\|^2 + 2 \|y\|^2$.*

**Proof**

This is straightforward, and is therefore omitted. □

Pre-Hilbert space is a special kind of normed linear space. The concept of convergence, continuity, topology, and completeness, etc. apply in pre-Hilbert spaces.
Proposition 13.4 Let $X$ be a pre-Hilbert space. Then, the inner product is a continuous function of the product space $X \times X$.

Proof We will show that the inner product is continuous at any $(x_0, y_0) \in X \times X$. \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), let $M := \max\{\|x_0\|, \|y_0\|\} + 1 \in (0, \infty) \subset \mathbb{R}$ and $\delta := \min\{\epsilon, 1\}/M \in (0, 1] \subset \mathbb{R}$. Fix any $(x, y) \in B_{X \times X}((x_0, y_0), \delta)$, we have $|\langle x, y \rangle - \langle x_0, y_0 \rangle| = |\langle x-x_0, y \rangle + \langle x_0, y-y_0 \rangle| \leq |\langle x-x_0, y \rangle| + |\langle x_0, y-y_0 \rangle| \leq \|x-x_0\| \|y\| + \|x_0\| \|y-y_0\| \leq \|x-x_0\| (\|y\| + \|y-y_0\|) + M \|y-y_0\| \leq M (\|x-x_0\| + \|y-y_0\|) < M \delta \leq \epsilon$, where the second inequality follows from Proposition 13.2. Hence, the inner product is continuous at $(x_0, y_0)$. By the arbitrariness of $(x_0, y_0)$, we have the inner product is continuous. This completes the proof of the proposition. \(\square\)

Definition 13.5 A complete pre-Hilbert space is called a Hilbert space.

Definition 13.6 Let $X$ be a pre-Hilbert space, $x, y \in X$, and $S \subseteq X$. The vectors $x$ and $y$ are said to be orthogonal if $\langle x, y \rangle = 0$. We will then write $x \perp y$. The vector $x$ is said to be orthogonal to the set $S$ if $x \perp s$, for all $s \in S$. We will then write $x \perp S$.

Theorem 13.7 (Pythagorean) Let $X$ be a pre-Hilbert space over $\mathbb{K}$ and $x, y \in X$. If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. Furthermore, if $\mathbb{K} = \mathbb{R}$ and $\|x+y\|^2 = \|x\|^2 + \|y\|^2$, then $x \perp y$.

Proof Let $x \perp y$, then $\langle x, y \rangle = 0$. This implies that $\|x+y\|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2$, where the second equality follows from Proposition 13.2; and the third equality follows from Definition 13.1.

Let $\mathbb{K} = \mathbb{R}$ and $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. Then, we have $0 = \langle x, y \rangle + \langle y, x \rangle = 2 \langle x, y \rangle$, where the first equality follows from Proposition 13.2; and the second equality follows from Definition 13.1 and $\mathbb{K} = \mathbb{R}$. Hence, $x \perp y$. This completes the proof of the theorem. \(\square\)

Example 13.8 Let $n \in \mathbb{Z}_+$ and $X = \mathbb{K}^n$. Define the inner product by $\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i} \in \mathbb{K}$, $\forall x := (\xi_1, \ldots, \xi_n) \in \mathbb{K}^n$ and $\forall y := (\eta_1, \ldots, \eta_n) \in \mathbb{K}^n$. It is easy to check that the pair $X := (X, \langle \cdot, \cdot \rangle)$ is a pre-Hilbert space over $\mathbb{K}$. The induced norm is clearly the Euclidean norm. By Example 7.29, $X$ is complete. Therefore, $X$ is a Hilbert space, which is to be denoted simply by $\mathbb{K}^n$. \(\diamond\)

Example 13.9 Let $X$ be a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle$, $i = 1, 2$. Let $\mathcal{Z} := X_1 \times X_2$. Define inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}} : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{K}$ by $\langle (x_1, x_2), (y_1, y_2) \rangle_{\mathcal{Z}} = \langle x_1, y_1 \rangle_1 + \langle x_2, y_2 \rangle_2$, $\forall (x_1, x_2), (y_1, y_2) \in \mathcal{Z}$. It is easy to show that $\mathcal{Z}$ with this inner product is a pre-Hilbert space over $\mathbb{K}$. It is also easy to see that the induced norm of $\mathcal{Z}$ is equal to the norm of $\mathcal{Z}$ as defined in Proposition 7.22. By Proposition 4.31, $\mathcal{Z}$ is complete. Therefore, $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$. \(\diamond\)
The above example can be easily generalized to $\prod_{i=1}^{n} X_i$ case, where $n \in \mathbb{Z}_+$ and $X_i$ is a Hilbert space over $\mathbf{K}$, $\forall i \in \{1, \ldots, n\}$.

**Example 13.10** Let $X$ be a Hilbert space over $\mathbf{K}$ with inner product $\langle \cdot, \cdot \rangle_X$ and $Z := l_2(X)$. We will show that $Z$ is a Hilbert space over $\mathbf{K}$ with inner product $\langle \cdot, \cdot \rangle_Z$ defined by $\langle x, y \rangle_Z = \sum_{i=1}^{\infty} \langle \xi_i, \eta_i \rangle$, $\forall x := (\xi_1, \xi_2, \ldots) \in Z$ and $\forall y := (\eta_1, \eta_2, \ldots) \in Z$. To see that $x, y \in Z$, we note that $x \in Z$ implies that $\sum_{i=1}^{\infty} \| \xi_i \|^2 < \infty$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. Then, $\exists n_1 \in \mathbb{N}$ such that $\sum_{i=n_1}^{\infty} \| \xi_i \|^2 < \epsilon$. By $y \in Z$, $\exists n_2 \in \mathbb{N}$ such that $\sum_{i=n_2}^{\infty} \| \eta_i \|^2 < \epsilon$. Let $n_0 := \max\{n_1, n_2\} \in \mathbb{N}$. $\forall n, m \in \mathbb{N}$ with $n_0 \leq n \leq m$, we have $|\sum_{i=n}^{m} \langle \xi_i, \eta_i \rangle| \leq (\sum_{i=n}^{m} \| \xi_i \|^2)^{1/2} (\sum_{i=n}^{m} \| \eta_i \|^2)^{1/2} \leq (\sum_{i=n_0}^{\infty} \| \xi_i \|^2)^{1/2} (\sum_{i=n_0}^{\infty} \| \eta_i \|^2)^{1/2} < \epsilon$, where the first inequality follows from Hölder’s Inequality 7.8. Hence, $\left( \sum_{i=n_0}^{\infty} \langle \xi_i, \eta_i \rangle \right)_{i=n_0}^{\infty} \subseteq \mathbf{K}$ is a Cauchy sequence, which converges since $\mathbf{K}$ is complete. Hence, $\langle x, y \rangle_Z \in \mathbf{K}$. It is easy to show that (i) – (iv) of Definition 13.1 are satisfied. Hence, $Z$ is a pre-Hilbert space. It is easy to see that the induced norm is equal to the norm of $Z$ as defined in Example 7.10. By Example 7.33, we have $Z$ is complete and therefore a Hilbert space.

**Example 13.11** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $\mathcal{Y}$ be a separable Hilbert space over $\mathbf{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$ and $Z := L_2(\mathcal{X}, \mathcal{Y})$. We will show that $Z$ is a Hilbert space over $\mathbf{K}$ with inner product $\langle \cdot, \cdot \rangle_Z$ defined by $\left\langle [z_1], [z_2] \right\rangle_Z = \int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} \, d\mu(x)$, $\forall z_1, z_2 \in L_2(\mathcal{X}, \mathcal{Y})$.

First, we show that $\langle \cdot, \cdot \rangle_Z$ is well-defined. $\forall z_1, z_2 \in L_2(\mathcal{X}, \mathcal{Y})$, we have $z_1$ and $z_2$ are $\mathcal{B}$-measurable. By Propositions 11.39, 11.38, and 13.4, the function $h : \mathcal{X} \to \mathbf{K}$ defined by $h(x) = \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}}$, $\forall x \in \mathcal{X}$, is $\mathcal{B}$-measurable. $\mathcal{P} \circ h(x) = |h(x)| \leq \| z_1(x) \||z_2(x)\| = \mathcal{P} \circ z_1(x) \mathcal{P} \circ z_2(x)$, $\forall x \in \mathcal{X}$, where the inequality follows from Proposition 13.2. Then, $\int_{\mathcal{X}} \mathcal{P} \circ h \, d\mu \leq \int_{\mathcal{X}} (\mathcal{P} \circ z_1 \mathcal{P} \circ z_2) \, d\mu \leq (\int_{\mathcal{X}} \mathcal{P} \circ z_1 \, d\mu)^{1/2} (\int_{\mathcal{X}} \mathcal{P} \circ z_2 \, d\mu)^{1/2} = \| z_1 \|_2 \| z_2 \|_2 < \infty$, where the first inequality follows from Definition 11.79; and the second inequality follows from Hölder’s Inequality 11.178. Hence, $h$ is absolute integrable over $\mathcal{X}$. By Proposition 11.92, $h$ is integrable over $\mathcal{X}$ and $\int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} \, d\mu(x) \in \mathbf{K}$. $\forall \tilde{z}_1, \tilde{z}_2 \in L_2(\mathcal{X}, \mathcal{Y})$ with $[z_1] = [\tilde{z}_1]$ and $[z_2] = [\tilde{z}_2]$, by Example 11.173, we have $z_1 = \tilde{z}_1$ a.e. in $\mathcal{X}$ and $z_2 = \tilde{z}_2$ a.e. in $\mathcal{X}$. Define $\tilde{h} : \mathcal{X} \to \mathbf{K}$ by $\tilde{h}(x) = \langle \tilde{z}_1(x), \tilde{z}_2(x) \rangle_{\mathcal{Y}}$, $\forall x \in \mathcal{X}$. By Lemmas 11.45 and 11.46, $\tilde{h} = h$ a.e. in $\mathcal{X}$. By Proposition 11.92, $\int_{\mathcal{X}} \tilde{h} \, d\mu = \int_{\mathcal{X}} h \, d\mu$. Hence, $\langle [z_1], [z_2] \rangle_Z \in \mathbf{K}$ is well-defined.

Next, we show that $\langle \cdot, \cdot \rangle_Z$ satisfies (i) – (iv) of Definition 13.1. Fix any $z_1, z_2, z_3 \in L_2(\mathcal{X}, \mathcal{Y})$ and any $\lambda \in \mathbf{K}$. (i) $\langle [z_1], [z_2] \rangle_Z = \int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_{\mathcal{Y}} \, d\mu(x) = \int_{\mathcal{X}} \langle z_2(x), z_1(x) \rangle_{\mathcal{Y}} \, d\mu(x) = \int_{\mathcal{X}} \langle z_2(x), z_3(x) \rangle_{\mathcal{Y}} \, d\mu(x) = \int_{\mathcal{X}} \langle z_3(x), z_2(x) \rangle_{\mathcal{Y}} \, d\mu(x) = \int_{\mathcal{X}} \langle z_3(x), z_3(x) \rangle_{\mathcal{Y}} \, d\mu(x) + \int_{\mathcal{X}} \langle z_3(x), z_1(x) \rangle_{\mathcal{Y}} \, d\mu(x) + \int_{\mathcal{X}} \langle z_2(x), z_2(x) \rangle_{\mathcal{Y}} \, d\mu(x)$.
13.2. PROJECTION THEOREMS

The basic concept of the Projection Theorem is illustrated in the Figure 13.1.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{projection.png}
\caption{Projection onto a subspace.}
\end{figure}

**Theorem 13.12** Let $X$ be a pre-Hilbert space, $M \subseteq X$ be a subspace, and $x \in X$. Consider the problem $\min_{m \in M} \| x - m \|$. If there exists $m_0 \in M$ such that $\| x - m_0 \| \leq \| x - m \|$, $\forall m \in M$, then $m_0$ is the unique vector...
in $M$ that minimizes $\|x - m\|$. A necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x-m_0) \perp M$.

**Proof** We first show that if $m_0 \in M$ is a minimizing vector then $(x-m_0) \perp M$ by an argument of contradiction. Suppose that $(x-m_0) \not\perp M$. Then, $\exists m \in M$ such that $(x-m_0, m) = \delta \neq 0$. Without loss of generality, assume $\|m\| = 1$. Define $m_1 := m_0 + \delta m \in M$. Then, by Proposition 13.12, 

$\|x-m_1\|^2 = (x-m_0 - \delta m, x-m_0 - \delta m) = \|x-m_0\|^2 - (x-m_0, \delta m) - (\delta m, x-m_0) + \|\delta m\|^2 = \|x-m_0\|^2 - \delta^2 + |\delta|^2 \|m\|^2 = \|x-m_0\|^2 - |\delta|^2 < \|x-m_0\|^2$. This contradicts the assumption that $m_0$ is a minimizing vector. Hence, $(x-m_0) \perp M$.

Next, we show that if $(x-m_0) \perp M$, where $m_0 \in M$, then $m_0$ is the unique minimizing vector. For all $m \in M$, by Pythagorean Theorem 13.7, we have $\|x-m\|^2 = \|x-m_0 - (m-m_0)\|^2 = \|x-m_0\|^2 + \|m-m_0\|^2$. Then, $\|x-m\| > \|x-m_0\|$ if $m \neq m_0$. This completes the proof of the theorem. 

**Theorem 13.13 (The Classical Projection Theorem)** Let $X$ be a Hilbert space and $M \subseteq X$ be a closed subspace. For all $x \in X$, there exists a unique vector $m_0 \in M$ such that $\|x-m_0\| = \min_{m \in M} \|x-m\|$. Furthermore, a necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x-m_0) \perp M$.

**Proof** The uniqueness and orthogonality are immediate consequences of Theorem 13.12. We are only required to establish the existence of $m_0$. Let $\delta := \inf_{m \in M} \|x-m\| \in [0, \infty) \subseteq \mathbb{R}$. Then, $\exists (m_n)_{n=1}^\infty \subseteq M$ such that $\lim_{n \in \mathbb{N}} \|x-m_n\| = \delta$. For all $\epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have $\delta \leq \|x-m_n\| < \sqrt{\delta^2 + \epsilon^2/4}$. By Parallelogram Law 13.3, we have, \forall i, j \in \mathbb{N} with $n_0 \leq i$ and $n_0 \leq j$, $\|m_i - x + x - m_j\|^2 + \|x-m_i + x - m_j\|^2 = 2 \|x-m_j\|^2 + 2 \|x-m_i\|^2$. Then, $\|m_i - m_j\|^2 = 2 \|x-m_j\|^2 + 2 \|x-m_i\|^2 - 4 \|x-(m_i + m_j)/2\|^2 < 4(\delta^2 + \epsilon^2/4) - 4\delta^2 = \epsilon^2$ and $\|m_i - m_j\| < \epsilon$. This implies that $(m_n)_{n=1}^\infty$ is a Cauchy sequence, which must converge to $m_0 \in M$, since $M$ is a complete subspace by Proposition 4.39. By Proposition 7.21, we have $\|x-m_0\| = \lim_{n \in \mathbb{N}} \|x-m_n\| = \delta = \inf_{m \in M} \|x-m\|$. This completes the proof of the theorem.

In the above proof, we observe that the key to the existence of the minimizing vector is that $M$ is complete. Hence, we have the following modified version of the projection theorem.

**Theorem 13.14** Let $X$ be a pre-Hilbert space and $M \subseteq X$ be a complete subspace. For all $x \in X$, there exists a unique vector $m_0 \in M$ such that $\|x-m_0\| = \min_{m \in M} \|x-m\|$. Furthermore, a necessary and sufficient condition for $m_0 \in M$ being the unique minimizing vector is $(x-m_0) \perp M$.

**Proof** This is straightforward by the proof of the Projection Theorem 13.13, and is therefore omitted.
13.3 Dual of Hilbert Spaces

**Theorem 13.15 (Riesz-Fréchet)** Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$. Then, the following statements hold.

1. $\forall f \in \mathcal{X}^*$, there exists a unique $y_0 \in \mathcal{X}$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in \mathcal{X}$, and $\|f\|_{\mathcal{X}^*} = \|y_0\|_{\mathcal{X}}$. Therefore, we may define a mapping $\Phi : \mathcal{X}^* \rightarrow \mathcal{X}$ by $\Phi(f) = y_0$.

2. $\forall y \in \mathcal{X}$, define $g : \mathcal{X} \rightarrow \mathbb{K}$ by $g(x) = \langle x, y \rangle$, $\forall x \in \mathcal{X}$, then $g \in \mathcal{X}^*$.

3. The mapping $\Phi$ is bijective, uniformly continuous, norm preserving, and conjugate linear (that is $\Phi(\alpha f_1 + \beta f_2) = \overline{\alpha} \Phi(f_1) + \overline{\beta} \Phi(f_2)$, $\forall f_1, f_2 \in \mathcal{X}^*$, $\forall \alpha, \beta \in \mathbb{K}$).

4. If $\mathbb{K} = \mathbb{R}$, then $\Phi$ is an isometrical isomorphism between $\mathcal{X}^*$ and $\mathcal{X}$.

5. If $\mathbb{K} = \mathbb{C}$, let $\phi : \mathcal{X} \rightarrow \mathcal{X}^{**}$ be the natural mapping as defined in Remark 7.88, then $\phi$ is surjective and $\mathcal{X}$ is reflexive.

6. If $\mathbb{K} = \mathbb{C}$, then $\mathcal{X}^*$ with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}^*}$, defined by $\langle f, g \rangle_{\mathcal{X}^*} := \langle \Phi(g), \Phi(f) \rangle$, $\forall f, g \in \mathcal{X}$, is a Hilbert space, and it is reflexive.

Henceforth, we will denote $\Phi_{inv}(x) =: x^*$, $\forall x \in \mathcal{X}$. Furthermore, the following statement hold.

7. When $\mathbb{K} = \mathbb{C}$, let $\Phi_* : \mathcal{X}^{**} \rightarrow \mathcal{X}^*$ be the mapping of $\Phi$ if $\mathcal{X}$ is replaced by $\mathcal{X}^*$. Then, $\Phi_* = \Phi_{inv}$. This leads to the identity $(x^*)^* = x$, $\forall x \in \mathcal{X}$.

**Proof**

We will first show that, $\forall y \in \mathcal{X}$, the functional $g$ defined in the statement (ii) of the theorem is a bounded linear functional with $\|g\|_{\mathcal{X}^*} = \|y\|_{\mathcal{X}}$. By Definition 13.1, $g$ is a linear functional. By Proposition 13.2, we have $|g(x)| = |\langle x, y \rangle| \leq \|x\|_{\mathcal{X}} \|y\|_{\mathcal{X}}$. Hence, $g \in \mathcal{X}^*$ and $\|g\|_{\mathcal{X}^*} \leq \|y\|_{\mathcal{X}}$. Note that $g(y) = |\langle y, y \rangle| = \|y\|^2_{\mathcal{X}}$. Hence, $\|g\|_{\mathcal{X}^*} \geq \|y\|_{\mathcal{X}}$. Hence, $\|g\|_{\mathcal{X}^*} = \|y\|_{\mathcal{X}}$.

Next, we will show that, for any $f \in \mathcal{X}^*$, there exists a $y_0 \in \mathcal{X}$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in \mathcal{X}$. Let $M := \mathcal{N}(f)$. We will distinguish two exhaustive and mutually exclusives cases: Case 1: $M = \mathcal{X}$; Case 2: $M \subset \mathcal{X}$.

Case 1: $M = \mathcal{X}$. Take $y = \vartheta_{\mathcal{X}}$. Then, $f(x) = 0 = \langle x, y \rangle$, $\forall x \in \mathcal{X}$. This case is proved.

Case 2: $M \subset \mathcal{X}$. Then, $\exists x_0 \in \mathcal{X} \setminus M$. Since $f \in \mathcal{X}^*$, then $M$ is a closed subspace of $\mathcal{X}$ by Proposition 7.68. By Projection Theorem 13.13, $\exists m_0 \in M$ such that $(x_0 - m_0) \perp M$. Let $z_0 := x_0 - m_0 \in \mathcal{X} \setminus M$. Then, $f(z_0) \neq 0$. Without loss of generality, we may assume that $f(z_0) = 1$. $\forall x \in \mathcal{X}$, by the linearity of $f$, we have $f(x - f(x)z_0) = f(x) - f(x)f(z_0) = 0$ and $x - f(x)z_0 \in M$. Since $z_0 \perp M$, we have $0 = \langle x - f(x)z_0, z_0 \rangle = \langle x, z_0 \rangle - f(x)\|z_0\|^2$. Thus, $f(x) = \frac{\langle x, z_0 \rangle}{\|z_0\|^2}$. This completes the proof.
Then, \( f(x) = \langle x, z_0 \rangle / \|z_0\|^2 = \langle x, z_0/\|z_0\|^2 \rangle \). Let \( y_0 := z_0/\|z_0\|^2 \in \mathcal{X} \), we have \( f(x) = \langle x, y_0 \rangle \), \( \forall x \in \mathcal{X} \). This case is also proved.

Next, we will show that, for any \( f \in \mathcal{X}^* \), the vector \( y_0 \in \mathcal{X} \) is unique. Let \( y \in \mathcal{X} \) be another vector such that \( f(x) = \langle x, y \rangle \), \( \forall x \in \mathcal{X} \). Then, \( 0 = f(x) - f(x) = \langle x, y_0 \rangle - \langle x, y \rangle = \langle x, y_0 - y \rangle \), \( \forall x \in \mathcal{X} \). Then, \( 0 = \langle y_0 - y, y_0 - y \rangle = \|y_0 - y\|^2 \). Thus, \( y = y_0 \). Hence, \( y_0 \) is unique.

Based on the above, we have (i) and (ii) holds.

(iii) Note that, \( \forall f_1, f_2 \in \mathcal{X}^* \), \( \forall \alpha, \beta \in \mathbb{K} \), \( (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha \langle x, \Phi(f_1) \rangle + \beta \langle x, \Phi(f_2) \rangle = \langle x, \alpha \Phi(f_1) + \beta \Phi(f_2) \rangle \), \( \forall x \in \mathcal{X} \), where the second equality follows from (i); and the third equality follows from Proposition 13.2. Then, by (i), we have \( \Phi(\alpha f_1 + \beta f_2) = \alpha \Phi(f_1) + \beta \Phi(f_2) \). By (i), \( \|\Phi(f)\|_{\mathcal{X}} = \|f\|_{\mathcal{X}^*}, \forall f \in \mathcal{X}^* \). Therefore, \( \Phi \) is norm preserving. By (ii), \( \Phi \) is surjective. \( \forall f_1, f_2 \in \mathcal{X}^* \), \( \Phi(f_1) = \Phi(f_2) \) implies that \( 0 = \|\Phi(f_1) - \Phi(f_2)\|_{\mathcal{X}} = \|f_1 - f_2\|_{\mathcal{X}^*} \), which further implies that \( f_1 = f_2 \). Hence, \( \Phi \) is injective. Hence, \( \Phi \) is bijective. Clearly, \( \Phi \) is uniformly continuous, which follows from the fact that it is norm preserving and conjugate linear.

(iv) Let \( \mathcal{K} = \mathbb{R} \). By (iii), \( \Phi \) is linear. Then, by Definitions 6.28 and 7.24, \( \Phi \) is an isometrical isomorphism.

(v) Let \( \mathcal{K} = \mathbb{C} \). By Remark 7.88, \( \forall x_* \in \mathcal{X}^* \), \( \forall x \in \mathcal{X} \), we have \( \langle x, x_* \rangle = \langle \Phi(x), x_* \rangle \). Then, \( \forall y \in \mathcal{X} \), we have \( \langle x, y \rangle = \langle \Phi(x), y \rangle = \langle \Phi(x_*), \Phi(y) \rangle \). We will show that \( \Phi \) is surjective, which then implies that \( \Phi \) is isometrical isomorphism between \( \mathcal{X} \) and \( \mathcal{X}^{**} \), and \( \mathcal{X} \) is reflexive. \( \forall x_* \in \mathcal{X}^{**} \), \( \langle x, x_* \rangle = \langle \Phi(x), \Phi(x) \rangle \). Define \( G: \mathcal{X} \rightarrow \mathcal{K} \) by \( G(y) = \langle x, \Phi(x) \rangle, \forall y \in \mathcal{X} \). By (iii), it is easy to show that \( G \in \mathcal{X}^* \). By (i), \( G(y) = \langle y, \Phi(G) \rangle, \forall y \in \mathcal{X} \). Then, \( \langle x, x_* \rangle = \langle \Phi(x_*), \Phi(x) \rangle = \langle \Phi(G), \Phi(x) \rangle = \langle \Phi(G), \Phi(x) \rangle = \langle \phi(\Phi(G)), x \rangle, \forall x_* \in \mathcal{X}^* \). Thus, we have \( x_* = \phi(\Phi(G)) \in \mathcal{R}(\phi) \). By the arbitrariness of \( x_* \), we have \( \phi \) is surjective.

(vi) We will show that \( \langle \cdot, \cdot \rangle_{\mathcal{X}^*} \) is an inner product on \( \mathcal{X}^* \). \( \forall f, g, h \in \mathcal{X}^* \), \( \forall \lambda \in \mathbb{K} \), (a) \( \langle g, f \rangle_{\mathcal{X}^*} = \langle \Phi(f), \Phi(g) \rangle = \langle \Phi(g), \Phi(f) \rangle = \langle f, g \rangle_{\mathcal{X}^*} \), where the second equality follows from (i) of Definition 13.1. (b) \( \langle f + g, h \rangle_{\mathcal{X}^*} = \langle \Phi(h), \Phi(f + g) \rangle = \langle \Phi(h), \Phi(f) + \Phi(g) \rangle = \langle \Phi(h), \Phi(f) \rangle + \langle \Phi(h), \Phi(g) \rangle = \langle f, h \rangle_{\mathcal{X}^*} + \langle g, h \rangle_{\mathcal{X}^*} \), where the second equality follows from (iii); and the third equality follows from Definition 13.1. (c) \( \langle \lambda f, g \rangle_{\mathcal{X}^*} = \Phi(g), \Phi(\lambda f) \rangle = \langle \Phi(g), \lambda \Phi(f) \rangle = \lambda \langle \Phi(g), \Phi(f) \rangle = \lambda \langle f, g \rangle_{\mathcal{X}^*} \), where the second equality follows from (iii); and the third equality follows from Definition 13.1. (d) \( \langle f, f \rangle_{\mathcal{X}^*} = \|\Phi(f)\|^2 = \|f\|^2_{\mathcal{X}^*}, \in [0, \infty) \subset \mathbb{R} \), where the second equality follows from Definition 13.1; and the third equality follows from (iii). Clearly, \( \langle f, f \rangle_{\mathcal{X}^*} = 0 \) if, and only if, \( \|f\|_{\mathcal{X}^*} = 0 \) if, and only if, \( f = 0 \). Hence, \( \mathcal{X}^* \) with the inner product \( \langle \cdot, \cdot \rangle_{\mathcal{X}^*} \) form a pre-Hilbert space. The induced norm is equal to the \( \|\cdot\|_{\mathcal{X}^*} \). By Proposition 7.72, \( \mathcal{X}^* \) is complete and hence a Hilbert space. By Proposition 7.90, \( \mathcal{X}^* \) is reflexive.
(vii) Let $K = \mathbb{C}$. \forall x \in X$. Let $f = \Phi_*(x) = \Phi_*(\phi(x))$. Then, $f(x) = \langle x, \Phi(f) \rangle = \langle \phi(x))(f) = \langle f, \Phi_*(\phi(x)) \rangle \_X = \langle f, \Phi_*(x) \rangle \_X = \langle f, f \rangle \_X = \|f\|^2_X$, where the first equality follows from (i); the second equality follows from Remark 7.88; the third equality follows from (i) as applied to $X^*$; and the fourth equality follows from (v). By (iii), $\|f\|_X = \|x\|_X$ and $\|f\|_{X^*} = \|\Phi(f)\|_X$. This implies that $\langle x, \Phi(f) \rangle = \|x\|_X \|\Phi(f)\|_X$. By (iv) of Proposition 13.2, $\Phi(f) = \vartheta_X$ or $x = \lambda \Phi(f)$, for some $\lambda \in \mathbb{C}$.

If $\Phi(f) = \vartheta_X$, by (iii), $f = \vartheta_{X^*}$. Thus, by (iii), $x = \vartheta_X = \Phi(f)$ since $f = \Phi_*(x)$.

If $\Phi(f) \neq \vartheta_X$, we must have $x = \lambda \Phi(f)$, for some $\lambda \in \mathbb{C}$. Then, $\|f\|^2_{X^*} = \langle x, \Phi(f) \rangle = \lambda \langle \Phi(f), \Phi(f) \rangle = \lambda \|\Phi(f)\|^2_X = \lambda \|f\|^2_{X^*}$. Thus, $\lambda = 1$. This implies that $x = \Phi(f)$.

In both cases, we have $x = \Phi(f) = \Phi(\Phi_*(x))$. By the arbitrariness of $x$, we have $\Phi \circ \Phi_* = \text{id}_X$. By (iii) and Proposition 2.4, we have $\Phi_* = \Phi_{\text{inv}}$.

This completes the proof of the theorem. \hfill \Box

**Definition 13.16** Let $X$ be a Hilbert space over $K$, $S \subseteq X$. By Definition 7.95, $S^\perp \subseteq X^*$ is the set $S^\perp = \{ x_* \in X^* \mid \langle x_* , x \rangle = 0 , \forall x \in S \}$. We will abuse the notation and denote $S^\perp := \{ y \in X \mid \langle x , y \rangle = \langle y^* , x \rangle = 0 , \forall x \in S \}$, which is said to be the orthogonal complement of $S$.

**Proposition 13.17** Let $X$ be a Hilbert space over $K$ and $S, T \subseteq X$. Then,

(i) $S^\perp \subseteq X$ is a closed subspace;

(ii) if $S \subseteq T$, then $T^\perp \subseteq S^\perp$;

(iii) $S^{\perp \perp} = S^\perp$;

(iv) $S^{\perp \perp} = \text{span}(S)$.

**Proof** (i) are (ii) follow immediately from Proposition 7.98.

(iv) By Riesz-Fréchet Theorem 13.15, $X$ is reflexive. Then, $X = X^{**}$. By Definition 7.95, $\text{^1} S \subseteq X^*$ is well-defined and $\text{^1} S = S^\perp$. Then, by Proposition 7.98, we have $S^{\perp \perp} = \text{^1} (S^\perp) = \text{span}(S)$.

(iii) By (iv), we have $S \subseteq \text{span}(S) = S^{\perp \perp}$. By (ii), we have $S^{\perp \perp} \supseteq S^{\perp \perp \perp}$. Again by (iv), we have $S^{\perp} \subseteq \text{span}(S^{\perp}) = S^{\perp \perp \perp}$. Therefore, $S^{\perp} = S^{\perp \perp \perp}$. This completes the proof of the proposition. \hfill \Box

**Definition 13.18** Let $X$ be a vector space over the field $F$ and $M, N \subseteq X$ are subspaces. We say that $X$ is the direct sum of $M$ and $N$ if $\forall x \in X, \exists m \in M$ and $\exists n \in N$, such that $x = m + n$. In this case, we will write $X = M \oplus N$.

**Theorem 13.19** Let $M$ be a closed subspace of a Hilbert space $X$. Then, $X = M \oplus M^\perp$. 

Proof. \(\forall x \in \mathcal{X}\), by Projection Theorem 13.13, \(\exists! m_0 \in M\) such that 
\[ n_0 = (x - m_0) \perp M. \]
Hence, \(x = m_0 + n_0\) with \(m_0 \in M\) and \(n_0 \in M^\perp\).
Clearly, \(m_0\) and \(n_0\) is unique since \(m_0\) is unique. Hence, \(\mathcal{X} = M \oplus M^\perp\).
This completes the proof of the theorem. \(\square\)

Proposition 13.20 Let \(\mathcal{X}\) be a real Hilbert space, \(f : \mathcal{X} \times \mathcal{X} \to \mathbb{R}\) be
given by \(f(x,y) = \langle x, y \rangle\), \(\forall (x,y) \in \mathcal{X} \times \mathcal{X}\). Then, \(f\) is analytic with
arbitrarily large analytic radius, \(f^{(1)}(x,y)(h_x,h_y) = \langle h_x, y \rangle + \langle x, h_y \rangle\),
\(f^{(2)}(x,y)(h_x,h_y)(h_{x_1},h_{y_1}) = \langle h_{x_2}, h_{y_2} \rangle + \langle h_{x_1}, h_{y_1} \rangle\), and \(f^{(i+2)}(x,y) =
\vartheta_{\mathcal{B}_{\mathcal{X}}}(\mathcal{X} \times \mathcal{X}, \mathbb{R})\), \(\forall (x,y) \in \mathcal{X} \times \mathcal{X}\), \(\forall i \in \mathbb{N}\), \(\forall (h_x, h_y) \in \mathcal{X} \times \mathcal{X}\), \(\forall (h_{x_1}, h_{y_1}) \in \mathcal{X} \times \mathcal{X}\).

Proof. This follows immediately from Propositions 9.41 and 9.27 and
Riesz-Fréchet Theorem 13.15. \(\square\)

13.4 Hermitian Adjoints

Definition 13.21 Let \(\mathcal{X}\) and \(\mathcal{Y}\) be Hilbert spaces over \(\mathbb{K}\), \(A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})\), and
\(\Phi_\mathcal{X} : \mathcal{X}^* \to \mathcal{X}\) and \(\Phi_\mathcal{Y} : \mathcal{Y}^* \to \mathcal{Y}\) as defined in Riesz-Fréchet Theorem 13.15.
Define the Hermitian adjoint of \(A\) by \(A^* := \Phi_\mathcal{X} \circ A^\prime \circ \Phi_\mathcal{Y}^{\text{inv}} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})\). \(A\)
is said to be Hermitian if \(A^* = A\).

Proposition 13.22 Let \(\mathcal{X}, \mathcal{Y}, \) and \(\mathcal{Z}\) be Hilbert spaces over \(\mathbb{K}\), \(A, B \in \mathcal{B}(\mathcal{X}, \mathcal{Y})\), \(C \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})\), \(x \in \mathcal{X}\), \(y \in \mathcal{Y}\), and \(\lambda \in \mathbb{K}\). Then,

(i) \(\text{id}_\mathcal{X}^* = \text{id}_\mathcal{X}\);
(ii) \((A + B)^* = A^* + B^*\);
(iii) \((\lambda A)^* = \bar{\lambda} A^*\);
(iv) \((CA)^* = A^* C^*\);
(v) if \(A\) is bijective, then \((A^*)^{-1} = (A^{-1})^* =: A^{-*}\);
(vi) \(\langle Ax, y \rangle_\mathcal{Y} = \langle x, A^* y \rangle_\mathcal{X}\);
(vii) \(\langle y, Ax \rangle_\mathcal{Y} = \langle A^* y, x \rangle_\mathcal{X}\);
(viii) \((A^*)^* = A\);
(ix) \(\|A^*\|_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} = \|A\|_{\mathcal{B}(\mathcal{X}, \mathcal{Y})}\);
(x) \((\mathcal{R}(A))^\perp = \mathcal{N}(A^*)\) and \((\mathcal{R}(A^*))^\perp = \mathcal{N}(A)\);
(xi) if \(\mathcal{R}(A) \subseteq \mathcal{X}\) is closed, then \(\mathcal{R}(A) = (\mathcal{N}(A^*))^\perp\) and \(\mathcal{R}(A^*) =
(\mathcal{N}(A))^\perp\).
13.4. HERMITIAN ADJOINTS

\[(xii)\] Let \((X, \mathcal{B})\) be a measurable space and \(\mu\) be a \(\mathcal{B}(\mathcal{Y}, \mathbb{Z})\)-valued measure on \((X, \mathcal{B})\). Define \(\mu^*\) to be a mapping from \(\mathcal{B}\) to \(\mathcal{B}(\mathcal{Z}, \mathcal{Y})\) by: \(\mu^*(E) = (\mu(E))^*\), if \(E \in \text{dom}(\mu)\); and \(\mu^*(E)\) is undefined, if \(E \in \mathcal{B}\setminus\text{dom}(\mu)\).

Then, \(\mu^*\) is a \(\mathcal{B}(\mathcal{Z}, \mathcal{Y})\)-valued measure on \((X, \mathcal{B})\) with \(P \circ \mu^* = P \circ \mu\).

**Proof**

Let \(\Phi_X : X^* \to X\), \(\Phi_Y : Y^* \to Y\), and \(\Phi_Z : Z^* \to Z\) be as defined in Riesz-Fréchet Theorem 13.15.

(i) \(\text{id}_X^* = \Phi_X \circ \text{id}_X \circ \Phi_{X inv} = \Phi_X \circ \text{id}_X \circ \Phi_{X inv} = \text{id}_X\), where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(ii) \((A + B)^* = \Phi_X \circ (A + B)' \circ \Phi_{Y inv} = \Phi_X \circ (A' + B') \circ \Phi_{Y inv} = \Phi_X \circ A' \circ \Phi_{Y inv} + \Phi_X \circ B' \circ \Phi_{Y inv} = A^* + B^*\), where the first equality follows from Definition 13.21; the second equality follows from Proposition 7.110; and the third equality follows from Riesz-Fréchet Theorem 13.15.

(iii) \((\lambda A)^* = \Phi_X \circ (\lambda A)' \circ \Phi_{Y inv} = \Phi_X \circ (\lambda A') \circ \Phi_{Y inv} = \lambda \Phi_X \circ A' \circ \Phi_{Y inv} = \lambda A^*\), where the first equality follows from Definition 13.21; the second equality follows from Proposition 7.110; and the third equality follows from Riesz-Fréchet Theorem 13.15.

(iv) \((CA)^* = \Phi_X \circ (CA)' \circ \Phi_{Z inv} = \Phi_X \circ (A'C') \circ \Phi_{Z inv} = \Phi_X \circ A' \circ \Phi_{Y inv} \circ \Phi_Y \circ C' \circ \Phi_{Z inv} = \Phi_X \circ A' \circ \Phi_{Y inv} = \Phi_X \circ A' \circ \Phi_{y inv} = A^* \circ C^*\), where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(v) By Open Mapping Theorem 7.103, \(A^{-1} \in \mathcal{B}(\mathcal{Y}, X)\). Then, we have \((A^{-1})^* = \Phi_Y \circ (A^{-1})' \circ \Phi_{X inv} = \Phi_Y \circ (A')^{-1} \circ \Phi_{X inv} = (\Phi_X \circ A' \circ \Phi_{y inv})^{-1} = (A^*)^{-1}\), where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(vi) \(\langle A x, y \rangle_y = \langle \langle y^*, A x \rangle_y \rangle_y = \langle \langle A'y^*, x \rangle \rangle_X = \langle x, \Phi_X(A'y^*) \rangle_X = \langle x, A^*y \rangle_X\), where the first equality follows from Riesz-Fréchet Theorem 13.15; the second equality follows from Proposition 7.109; the third and fourth equalities follow from Riesz-Fréchet Theorem 13.15; and the last equality follows from Definition 13.21.

(vii) \(\langle y, A x \rangle_y = \langle A x, y \rangle_y = \langle x, A^*y \rangle_X = \langle A^*y, x \rangle_X\), where the first equality follows from Definition 13.1; the second equality follows from (vi); and the third equality follows from Definition 13.1.

(viii) \(\langle y, A x \rangle_y = \langle A^*y, x \rangle_X = \langle y, (A^*)^*x \rangle_y\), where the first equality follows from (vii); and the second equality follows from (vii). By the arbitrariness of \(x\) and \(y\), we have \(A = (A^*)^*\).

(ix) \(\|A^*\|_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} = \sup_{y \in \mathcal{Y}, \|y\|_\mathcal{Y} \leq 1} \|A^*y\|_\mathcal{X} = \sup_{y \in \mathcal{Y}, \|y\|_\mathcal{Y} \leq 1} \|\Phi_X \circ A' \circ \Phi_{y inv}(y)\|_\mathcal{X} = \|A\|_{\mathcal{B}(\mathcal{Y}, \mathcal{X})} = \|A\|_{\mathcal{B}(\mathcal{Y}, \mathcal{X})}\), where the first equality follows from Proposition 7.63; the second equality follows from Definition 13.21; the third equality follows from the fact that \(\Phi\) is bijective and norm preserving; and the fourth equality follows from Proposition 7.109.

(x) By Proposition 7.112 and Definition 13.16, we have \((\mathcal{R}(A))^\perp = \Phi_Y(\mathcal{N}(A')) = \mathcal{N}(A^*)\). Then, \((\mathcal{R}(A^*))^\perp = \mathcal{N}(\mathcal{N}(A^*)^*) = \mathcal{N}(A), \) where the first equality follows from the previous sentence; and the second equality
follows from (viii).

(xi) By Proposition 7.114 and Definition 13.16, we have \( \mathcal{R}(A) = \mathcal{N}(A^*) \). Furthermore, \( \mathcal{R}(A^*) = \mathcal{N}(A) \), where the first equality follows from the previous sentence; the second equality follows from (viii).

We will prove the result using Definition 11.108. Let \( \nu := \mathcal{P} \circ \mu \). Clearly, \( \nu \) is a measure on \((X, \mathcal{B})\). (a) \( \nu(\emptyset) = 0 \) and \( \mu(\emptyset) = \emptyset \). Then, \( \mu^*(\emptyset) = \emptyset = \emptyset \), where the last equality follows from (ii).

(b) \( \forall E \in \mathcal{B} \) with \( \nu(E) = \infty \), then \( \mu(E) \) is undefined and \( E \in \mathcal{B} \setminus \underline{\text{dom}}(\mu) \). Then, \( \mu^*(E) \) is undefined.

(c) \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty \), then \( \mu(E) \in B(\mathbb{R}, \mathbb{Z}) \) and \( E \in \underline{\text{dom}}(\mu) \). Then, \( \mu^*(E) = (\mu(E))^* \in B(\mathbb{R}, \mathbb{Z}) \). A pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^\infty E_i \), we have \( \sum_{i=1}^\infty \|\mu(E_i)\|_{B(\mathbb{R}, \mathbb{Z})} < \infty \) and \( \mu(E) = \sum_{i=1}^\infty \mu(E_i) \). Then, \( \sum_{i=1}^\infty \|\mu(E_i)\|_{B(\mathbb{R}, \mathbb{Z})} = \sum_{i=1}^\infty \|\mu(E_i)\|_{B(\mathbb{R}, \mathbb{Z})} = \sum_{i=1}^\infty \|\mu(E_i)\|_{B(\mathbb{R}, \mathbb{Z})} \), where the second equality follows from (ix). Furthermore, \( \sum_{i=1}^\infty \mu^*(E_i) = \sum_{i=1}^\infty (\mu(E_i))^* = (\sum_{i=1}^\infty (\mu(E_i))^*) = (\mu(E))^* = \mu^*(E) \), where the second equality follows from Propositions 3.66, 7.109, and 7.110, Riesz-Fréchet Theorem 13.15, and Definition 13.21.

(d) \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty \), we have

\[
\nu(E) = \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\|_{B(\mathbb{R}, \mathbb{Z})} = \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n (\mu(E_i))^* = \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu^*(E_i)\|_{B(\mathbb{R}, \mathbb{Z})}
\]

where the first equality follows from Definition 11.108; and the second equality follows from (ix).

This shows that \( \mu^* \) is a \( B(\mathbb{R}, \mathbb{Z}) \)-valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \mu^* = \mathcal{P} \circ \mu \).

This completes the proof of the proposition. \( \square \)

13.5 Approximation in Hilbert Spaces

Definition 13.23 Let \( X \) be a pre-Hilbert space and \( S \subseteq X \). We will say that \( S \) is an orthogonal set if \( x \perp y, \forall x, y \in S \) with \( x \neq y \). The set \( S \) is said to be orthonormal if \( S \) is an orthogonal set and \( \|x\| = 1, \forall x \in S \).

Proposition 13.24 Let \( X \) be a pre-Hilbert space over \( \mathbb{K} \) and \( S \subseteq X \) be an orthogonal set. Assume that \( \emptyset_X \notin S \). Then, \( S \) is a linearly independent set.
13.5. APPROXIMATION IN HILBERT SPACES

Proof Suppose $S$ is not a linearly independent set. Then, $\exists x_0 \in S$ such that $x_0 \notin \text{span}(S \setminus \{x_0\})$. Then, $\exists n \in \mathbb{Z}_+$, $\exists x_1, \ldots, x_n \in S \setminus \{x_0\}$ with $x_i \neq x_j, \forall 1 \leq i < j \leq n$, $\exists \alpha_1, \ldots, \alpha_n \in \mathbb{K}$, such that $x_0 = \sum_{i=1}^{n} \alpha_i x_i$. Then, $0 = \|x_0 - \sum_{i=1}^{n} \alpha_i x_i\|^2 = \langle x_0, x_0 \rangle + \sum_{i=1}^{n} \langle \alpha_i x_i, \alpha_i x_i \rangle = \|x_0\|^2 + \sum_{i=1}^{n} |\alpha_i|^2 \|x_i\|^2 > 0$, where the second equality follows from the fact that $S$ is an orthogonal set; and the inequality follows from the fact that $x_0 \neq \varnothing x$. This is a contradiction. Hence, $S$ is a linearly independent set. This completes the proof of the proposition.

In Hilbert space, orthonormal sets are greatly favored over other linearly independent sets, and it is reassuring to know that they exist and can be easily constructed. The next theorem presents the Gram-Schmidt orthogonalization procedure.

Theorem 13.25 (Gram-Schmidt Procedure) Let $n \in \mathbb{N} \cup \{\infty\}$, $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$, and $(x_i)_{i=1}^{n} \subseteq \mathcal{X}$ be a linearly independent set and vectors in the sequence are distinct. Then, there is an orthonormal sequence $(e_i)_{i=1}^{n} \subseteq \mathcal{X}$ such that $\text{span} \{(x_1, \ldots, x_k) \} = \text{span} \{(e_1, \ldots, e_k)\}$, $\forall k \in \mathbb{N}$ with $k \leq n$, and $\text{span} \{(x_i)_{i=1}^{n}\} = \text{span} \{(e_i)_{i=1}^{n}\}.$

Proof We construct $(e_i)_{i=1}^{n}$ recursively as follows.

1° For $i = 1$, clearly $x_1 \neq \varnothing x$ since $(x_i)_{i=1}^{n}$ is a linearly independent set. Let $e_1 := x_1/\|x_1\| \in \mathcal{X}$. Then, $\text{span} \{(e_1)\} = \text{span} \{(x_1)\}$ and the sequence $(e_i)_{i=1}^{1}$ is orthonormal. This completes the first step of the construction.

2° Assume that we have constructed $(e_i)_{i=1}^{k}$, where $k \in \mathbb{N}$ and $k < n$, such that the sequence is orthonormal and $\text{span} \{(e_1, \ldots, e_k)\} = \text{span} \{(x_1, \ldots, x_k)\}.$

3° Consider $k + 1$. Let $z_{k+1} := x_{k+1} - \sum_{i=1}^{k} \langle x_{k+1}, e_i \rangle e_i \in \mathcal{X}$. It is easy to show that $\langle z_{k+1}, e_i \rangle = 0, \forall i \in \{1, \ldots, k\}$. Then, $z_{k+1} \perp (e_i)_{i=1}^{k}$. By the assumption and Lemma 6.49, the vectors $x_1, \ldots, x_{k+1}$ are linear independent. Then, $z_{k+1} \neq \varnothing x$. Define $e_{k+1} := z_{k+1}/\|z_{k+1}\| \in \mathcal{X}$. The inductive assumption and the construction imply that $(e_i)_{i=1}^{k+1}$ is orthonormal and $\text{span} \{(e_1, \ldots, e_{k+1})\} = \text{span} \{(x_1, \ldots, x_{k+1})\}$.

This completes the induction process. Then, there exists an orthonormal $(e_i)_{i=1}^{n} \subseteq \mathcal{X}$ such that $\text{span} \{(e_1, \ldots, e_k)\} = \text{span} \{(x_1, \ldots, x_k)\}$, $\forall k \in \mathbb{N}$ with $k \leq n$.

If $n \in \mathbb{N}$, the theorem is proved. Consider the case when $n = \infty$. We need to show that $\text{span} \{(e_i)_{i=1}^{\infty}\} = \text{span} \{(x_i)_{i=1}^{\infty}\}$. $\forall x \in \text{span} \{(x_i)_{i=1}^{\infty}\}$, $\exists m \in \mathbb{N}, \exists \alpha_1, \ldots, \alpha_m \in \mathbb{K}$ such that $x = \sum_{i=1}^{m} \alpha_i x_i$. Then, $x \in \text{span} \{(x_1, \ldots, x_m)\} = \text{span} \{(e_1, \ldots, e_m)\} \subseteq \text{span} \{(e_i)_{i=1}^{m}\}$. Hence, we have $\text{span} \{(e_i)_{i=1}^{\infty}\} \subseteq \text{span} \{(x_i)_{i=1}^{\infty}\}$. By symmetry, we have $\text{span} \{(x_i)_{i=1}^{\infty}\} \subseteq \text{span} \{(e_i)_{i=1}^{\infty}\}$. Hence, $\text{span} \{(e_i)_{i=1}^{\infty}\} = \text{span} \{(x_i)_{i=1}^{\infty}\}$. This completes the proof of the theorem. □
Definition 13.26  Let \( X \) be a pre-Hilbert space over \( \mathbb{K} \), \( n \in \mathbb{N} \), \( y_1, \ldots, y_n \in X \). The Gram matrix of \( y_1, \ldots, y_n \) is
\[
\text{Gram}(y_1, \ldots, y_n) := \begin{bmatrix}
\langle y_1, y_1 \rangle & \langle y_1, y_2 \rangle & \cdots & \langle y_1, y_n \rangle \\
\langle y_2, y_1 \rangle & \langle y_2, y_2 \rangle & \cdots & \langle y_2, y_n \rangle \\
\vdots & \vdots & \ddots & \vdots \\
\langle y_n, y_1 \rangle & \langle y_n, y_2 \rangle & \cdots & \langle y_n, y_n \rangle 
\end{bmatrix}
\]

The Gram determinant of \( y_1, \ldots, y_n \) is \( \text{gram}(y_1, \ldots, y_n) := \det(\text{Gram}(y_1, \ldots, y_n)) \).

Proposition 13.27  Let \( n \in \mathbb{N} \), \( X \) be a pre-Hilbert space over \( \mathbb{K} \), \( y_1, \ldots, y_n \in X \). Then, \( \text{gram}(y_1, \ldots, y_n) \neq 0 \) if, and only if, \( y_1, \ldots, y_n \) are linearly independent.

Proof  We will prove the equivalent statement that \( \text{gram}(y_1, \ldots, y_n) = 0 \) if, and only if, \( y_1, \ldots, y_n \) are linearly dependent. “Sufficiency” Let \( y_1, \ldots, y_n \) be linearly dependent. Then, \( \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K} \), which are not all zeros, such that \( \sum_{i=1}^n \alpha_i y_i = \vartheta \). Then,
\[
0_{1 \times n} = \begin{bmatrix}
\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle & \cdots & \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle 
\end{bmatrix} = \begin{bmatrix}
\alpha_1 & \cdots & \alpha_n 
\end{bmatrix} \text{Gram}(y_1, \ldots, y_n)
\]
This shows that the rank(\( \text{Gram}(y_1, \ldots, y_n) \)) < \( n \). Therefore, \( \text{gram}(y_1, \ldots, y_n) = 0 \).

“Necessity” \( \text{gram}(y_1, \ldots, y_n) = 0 \) implies that rank(\( \text{Gram}(y_1, \ldots, y_n) \)) < \( n \). Then, the row vectors of the Gram matrix is linearly dependent. This implies that \( \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K} \), which are not all zeros, such that
\[
0_{1 \times n} = \begin{bmatrix}
\alpha_1 & \cdots & \alpha_n 
\end{bmatrix} \text{Gram}(y_1, \ldots, y_n)
\]
\[
= \begin{bmatrix}
\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle & \cdots & \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle 
\end{bmatrix}
\]
Then,
\[
0 = \begin{bmatrix}
\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle & \cdots & \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle 
\end{bmatrix} \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n 
\end{bmatrix}
\]
\[
= \sum_{j=1}^n \alpha_j \sum_{i=1}^n \alpha_i y_i y_j = \left\langle \sum_{i=1}^n \alpha_i y_i, \sum_{j=1}^n \alpha_j y_j \right\rangle
\]
\[
= \left\| \sum_{i=1}^n \alpha_i y_i \right\|^2
\]
Hence, \( \sum_{i=1}^n \alpha_i y_i = \vartheta \) and \( y_1, \ldots, y_n \) are linearly dependent.

This completes the proof of the proposition. \( \square \)
13.5. APPROXIMATION IN HILBERT SPACES

Theorem 13.28 Let \( X \) be a pre-Hilbert space over \( \mathbb{K} \), \( n \in \mathbb{N} \), \( y_1, \ldots, y_n \in X \), \( M = \text{span} \{ \{y_1, \ldots, y_n\} \} \), \( x \in X \), and \( \delta := \min_{m \in M} \|x-m\| \). Then, the following statements hold.

(i) There exists a unique \( m_0 \in M \) such that \( \delta = \|x-m_0\| \) and \( m_0 = \sum_{i=1}^{n} \alpha_i y_i \), where \( \alpha_1, \ldots, \alpha_n \in \mathbb{K} \) are solutions to the following normal equation

\[
\begin{bmatrix}
\langle y_1, y_1 \rangle & \cdots & \langle y_1, x \rangle \\
\vdots & \ddots & \vdots \\
\langle y_n, y_1 \rangle & \cdots & \langle y_n, x \rangle \\
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n \\
\end{bmatrix}
= \begin{bmatrix}
\langle x, y_1 \rangle \\
\vdots \\
\langle x, y_n \rangle \\
\end{bmatrix}
\]

(ii) If the vectors \( y_1, \ldots, y_n \) are linearly independent, then the Gram matrix \( \text{Gram}(y_1, \ldots, y_n) \) is invertible and \( \delta^2 = \frac{\text{gram}(y_1, \ldots, y_n, x)}{\text{gram}(y_1, \ldots, y_n)} \).

Proof

(i) Clearly, \( M \) is a finite-dimensional subspace of \( X \). By Theorem 7.36, \( M \) is complete. By Theorem 13.14, \( \exists m_0 \in M \) such that \( \delta = \|x-m_0\| \) and \( x-m_0 \perp M \). Let \( m_0 = \sum_{i=1}^{n} \alpha_i y_i \) with \( \alpha_1, \ldots, \alpha_n \in \mathbb{K} \).

Then, \( x - \sum_{i=1}^{n} \alpha_i y_i, y_j \rangle = 0 \), \( j = 1, \ldots, n \). By Proposition 13.2, we have \( \sum_{i=1}^{n} \alpha_i \langle y_i, y_j \rangle = \langle x, y_j \rangle \), \( j = 1, \ldots, n \). Thus, (13.1) holds.

(ii) If vectors \( y_1, \ldots, y_n \) are linearly independent, then the Gram determinant \( \text{gram}(y_1, \ldots, y_n) \neq 0 \), by Proposition 13.27, which further implies that the Gram matrix \( \text{Gram}(y_1, \ldots, y_n) \) is invertible. Note that

\[
\delta^2 = \|x-m_0\|^2 = \langle x-m_0, x-m_0 \rangle = \langle x-m_0, x \rangle = \langle x, x \rangle - \sum_{i=1}^{n} \alpha_i \langle y_i, x \rangle,
\]

where the third equality follows from the fact that \( x-m_0 \perp m_0 \in M \). Equivalently, we have \( \sum_{i=1}^{n} \alpha_i \langle y_i, x \rangle + \delta^2 = \langle x, x \rangle \). Combining this equation with the normal equation, we have the following matrix equation.

\[
\begin{bmatrix}
\langle y_1, y_1 \rangle & \cdots & \langle y_1, x \rangle \\
\vdots & \ddots & \vdots \\
\langle y_n, y_1 \rangle & \cdots & \langle y_n, x \rangle \\
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n \\
\end{bmatrix}
= \begin{bmatrix}
\langle x, y_1 \rangle \\
\vdots \\
\langle x, y_n \rangle \\
\end{bmatrix}
\]

By Cramer’s Rule, we have \( \delta^2 = \frac{\det(G_1)}{\det(G)} \), where

\[
G = \begin{bmatrix}
\langle y_1, y_1 \rangle & \cdots & \langle y_1, y_1 \rangle & 0 \\
\vdots & \ddots & \vdots & \vdots \\
\langle y_n, y_n \rangle & \cdots & \langle y_n, y_n \rangle & 0 \\
\langle y_1, x \rangle & \cdots & \langle y_n, x \rangle & 1 \\
\end{bmatrix} \quad \text{and} \quad G_1 = (\text{Gram}(y_1, \ldots, y_n))'.
\]

By Laplace Expansion Theorem, we have \( \det(G) = \det((\text{Gram}(y_1, \ldots, y_n))') = \det(\text{Gram}(y_1, \ldots, y_n)) = \text{gram}(y_1, \ldots, y_n) \).

Clearly, \( \det(G_1) = \text{gram}(y_1, \ldots, y_n, x) \). Hence, \( \delta^2 = \frac{\text{gram}(y_1, \ldots, y_n, x)}{\text{gram}(y_1, \ldots, y_n)} \).

This completes the proof of the theorem. \( \Box \)
In Gram Theorem 13.28, when \( y_1, \ldots, y_n \) are orthonormal, we denote \( y_i \) by \( e_i, \ i = 1, \ldots, n \). Then, \( \alpha_i = \langle x, e_i \rangle, i = 1, \ldots, n, \ m_0 = \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \) and \( \delta = \sqrt{\langle x, x \rangle - \sum_{i=1}^{n} |\langle x, e_i \rangle|^2} \).

**Proposition 13.29** Let \( X \) be Hilbert space over \( K \), \( (e_i)_{i=1}^{\infty} \subseteq X \) be orthonormal, and \( (\xi_i)_{i=1}^{\infty} \subseteq K \). Then, \( x = \sum_{i=1}^{\infty} \xi_i e_i \in X \) if, and only if, \( \sum_{i=1}^{\infty} |\xi_i|^2 < \infty \). In this case, \( \xi_i = \langle x, e_i \rangle, \ \forall i \in \mathbb{N} \).

**Proof** Define the partial sum \( s_i := \sum_{j=1}^{i} \xi_j e_j \in X, \ \forall i \in \mathbb{N} \).

“Sufficiency” Let \( \sum_{i=1}^{\infty} |\xi_i|^2 < \infty \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \ \exists i_0 \in \mathbb{N} \) such that \( \sum_{i=n_0+1}^{\infty} |\xi_i|^2 < \epsilon^2 \). \( \forall n, m \in \mathbb{N} \) with \( n_0 \leq n \leq m \), \( \|s_n - s_m\|^2 = \sum_{i=n_0+1}^{m} |\xi_i|^2 < \epsilon^2 \), where the second equality follows from the assumption that \( (e_i)_{i=1}^{\infty} \) is orthonormal. Hence, \( \|s_n - s_m\| < \epsilon \) and \( (s_i)_{i=1}^{\infty} \subseteq X \) is a Cauchy sequence. By the completeness of \( X \), \( \sum_{i=1}^{\infty} \xi_i e_i = \lim_{n\to\infty} s_n =: x \in X \).

“Necessity” Let \( x = \sum_{i=1}^{\infty} \xi_i e_i \in X \). Then, \( x = \lim_{n \in \mathbb{N}} s_n \) and \( \langle s_n \rangle_{n=1}^{\infty} \subseteq X \) is a Cauchy sequence. \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \ \exists i_0 \in \mathbb{N}, \ \forall n, m \in \mathbb{N} \) with \( n_0 \leq n \leq m \), we have \( \|s_n - s_m\| < \sqrt{\epsilon} \). Then, \( \sum_{i=n_0+1}^{m} |\xi_i|^2 = \sum_{i=n_0+1}^{m} |\xi_i|^2 = \sum_{i=n_0+1}^{m} |\xi_i|^2 < \epsilon^2 \). Hence, \( \sum_{i=1}^{\infty} |\xi_i|^2 < \infty \).

When \( x = \sum_{i=1}^{\infty} \xi_i e_i \in X \), by Propositions 13.4 and 3.66, we have \( \langle x, e_i \rangle = \lim_{n \in \mathbb{N}} \langle s_n, e_i \rangle = \xi_i, \ \forall i \in \mathbb{N} \). This completes the proof of the proposition.

**Proposition 13.30** (Bessel’s Inequality) Let \( X \) be a pre-Hilbert space over \( K \), \( x \in X \), and \( (e_i)_{i=1}^{\infty} \subseteq X \) be orthonormal. Then, \( \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 \).

**Proof** Let \( \alpha_i := \langle x, e_i \rangle \in K, \ \forall i \in \mathbb{N} \). \( \forall n \in \mathbb{N}, \ 0 \leq \|x - \sum_{i=1}^{n} \alpha_i e_i\|^2 = \|x\|^2 - \sum_{i=1}^{n} \alpha_i \langle e_i, x \rangle - \sum_{i=1}^{n} \overline{\alpha_i} \langle x, e_i \rangle + \sum_{i=1}^{n} |\alpha_i|^2 = \|x\|^2 - \sum_{i=1}^{n} |\alpha_i|^2 \). Hence, \( \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 \).

**Proposition 13.31** Let \( X \) be a Hilbert space over \( K \), \( x \in X \), and \( (e_i)_{i=1}^{\infty} \subseteq X \) be orthonormal. Then, \( m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in \text{span}(\{e_i\}_{i=1}^{\infty}) =: M, \langle x - m_0 \rangle \perp M, \text{ and } \|x - m_0\| = \min_{m \in M} \|x - m\| \).

**Proof** By Bessel’s Inequality 13.30, \( \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 < \infty \). By Proposition 13.29, \( m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in X \). Since \( \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in M, \forall n \in \mathbb{N}, \text{ then } m_0 \in M \) by Proposition 4.13. By Propositions 13.4 and 3.66, \( \langle x - m_0, e_j \rangle = \lim_{n \in \mathbb{N}} \langle x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle \|e_j\|^2 = 0, \forall j \in \mathbb{N} \). Hence, \( x - m_0 \in \text{span}(\{e_i\}_{i=1}^{\infty}) \perp = (\{e_i\}_{i=1}^{\infty}) \perp = (\langle e_i \rangle_{i=1}^{\infty}) \) = \( M^\perp \), where the first and the second equalities follow from Proposition 13.17. By Projection Theorem 13.13, \( \|x - m_0\| = \min_{m \in M} \|x - m\| \). This completes the proof of the proposition.

□
Based on Proposition 13.31, it is apparent that if $M = X$, then, $\forall x \in X$ can be expanded as an infinite series of $e_i$'s with coefficients equal to the Fourier coefficients $\langle x, e_i \rangle$. This motivates the following definition and proposition.

**Definition 13.32** Let $X$ be a Hilbert space, $n \in \mathbb{Z} \cup \{\infty\}$, and $(e_i)_{i=1}^n \subseteq X$ be orthonormal. The sequence $(e_i)_{i=1}^n$ is said to be complete if $\text{span}((e_i)_{i=1}^n) = X$.

**Proposition 13.33** Let $X$ be a Hilbert space, $n \in \mathbb{Z} \cup \{\infty\}$, and $(e_i)_{i=1}^n \subseteq X$ be orthonormal. The sequence is complete if, and only if, $((e_i)_{i=1}^n)_{i=1}^\perp = \{\varnothing\}$.

**Proof** By Proposition 13.17, $((e_i)_{i=1}^n)_{i=1}^\perp = \{\varnothing\} \iff ((e_i)_{i=1}^n)_{i=1}^\perp = \{\varnothing\} 
\iff \text{span}((e_i)_{i=1}^n)_{i=1}^\perp = \{\varnothing\} 
\iff \text{span}((e_i)_{i=1}^n) = X$. This completes the proof of the proposition. \hfill \Box

**Example 13.34** Let $I := [-1, 1] \subseteq \mathbb{R}$, $I := ((I, |\cdot|), \mathcal{B}, \mu)$ be the complete metric measure subspace of $\mathbb{R}$, $Z := L_2(I, \mathbb{K})$, $z_i \in L_2(I, \mathbb{K}) =: \mathcal{Z}$ be defined by $z_i(t) = t^i$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$. It is easy to show that $([z_i])_{i=0}^\infty \subseteq \mathcal{Z}$ is a linearly independent set. Then, the Gram-Schmidt Procedure 13.25 can be applied to this sequence to produce an orthonormal sequence $([e_i])_{i=0}^\infty \subseteq \mathcal{Z}$, where $e_i \in \mathcal{Z}$, $\forall i \in \mathbb{Z}_+$, are $i$th-order polynomials. Let $p_i \in \mathcal{Z}$ be defined by $p_i(t) = \frac{z_i(t)}{2^{i+1}}$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$, and $P_i \in \mathcal{Z}$ be defined by $P_i(t) = p_i^{(i)}(t)$, $\forall t \in I$, $\forall i \in \mathbb{Z}_+$. Then, we will show that $e_i = \sqrt{\frac{2i+1}{2}} P_i$, $\forall i \in \mathbb{Z}_+$, and $([e_i])_{i=0}^\infty \subseteq \mathcal{Z}$ is a complete orthonormal sequence. As defined, $(P_i)_{i=0}^\infty \subseteq \mathcal{Z}$ is the well-known Legendre polynomials.

We will first show that $([P_i])_{i=0}^\infty$ is an orthogonal set. $\forall n, m \in \mathbb{Z}_+$, we have

$$\langle [P_n], [P_m] \rangle = \int_{-1}^1 P_n(t)P_m(t) \, dt = \int_{-1}^1 p_n^{(n)}(t)p_m^{(m)}(t) \, dt$$

$$= p_n^{(n-1)}(t)p_m^{(m)}(t) \bigg|_{-1}^1 - \int_{-1}^1 p_n^{(n-1)}(t)p_m^{(m+1)}(t) \, dt$$

$$= -\int_{-1}^1 p_n^{(n-1)}(t)p_m^{(m+1)}(t) \, dt$$

$$\vdots$$

$$= (-1)^n \int_{-1}^1 p_n^{(0)}(t)p_m^{(m+n)}(t) \, dt$$

where the first equality follows from Example 13.8; the third equality follows from Integration by Parts Theorem 12.89; the fourth equality follows the fact $p_n^{(i)}(x) = 0$, $\forall i \in \mathbb{Z}_+$ with $0 \leq i < n$, $x \in \{-1, 1\}$; and the last
equality follows from the recursive application of the Integration by Parts Theorem 12.89. Thus, if \( m \neq n \), we may without loss of generality, assume that \( m < n \). Then, we have \( p_{m+1}(m+n) = \vartheta_z \) since \( p_m \) is a polynomial of order \( 2m < m + n \). Then, \( \langle [P_n], [P_m] \rangle = 0 \). Hence, \( ([P_i])_{i=0}^\infty \) is an orthogonal set.

Next, we show that \( \| [P_i] \|_2^2 = \sqrt{\frac{2}{2n+1}} \), \( \forall i \in \mathbb{Z}_+ \). Clearly, \( P_0(t) = 1 \), \( \forall t \in I \) and \( \| [P_0] \|_2^2 = \langle [P_0], [P_0] \rangle = 2 = \sqrt{\frac{2}{2}} \). Recursively, assume that \( \| [P_{n-1}] \|_2^2 = \sqrt{\frac{2}{2n-1}} \), for some \( n \in \mathbb{Z}_+ \). \( \text{Then,} \)

\[
\| [P_n] \|_2^2 = \langle [P_n], [P_n] \rangle = (-1)^n \int_{-1}^1 p_n^{(0)}(t)p_n^{(2n)}(t) \, dt
\]

\[
= \frac{(2n)!}{2^{2n}n!} \int_{-1}^1 (1 - t^2)^n \, dt
\]

\[
= \frac{(2n)!}{2^{2n}n!} \left( (1 - t^2)^n \right|_{-1}^1 + \frac{2n(2n)!}{2^{2n}n!} \int_{-1}^1 t^2(1 - t^2)^{n-1} \, dt
\]

\[
= \frac{2n(2n)!}{2^{2n}n!} \left( \int_{-1}^1 (1 - t^2)^{n-1} \, dt - \int_{-1}^1 (1 - t^2)^n \, dt \right)
\]

\[
= (2n - 1) \| [P_{n-1}] \|_2^2 - 2n \| [P_n] \|_2^2
\]

where the second equality follows from the previous paragraph; the third equality follows from the definition of \( p_n \in \mathbb{Z} \); the fourth equality follows from Integration by Parts Theorem 12.89; the fifth equality follows from Proposition 11.92; and the last equality follows from the third equality. Therefore, we have \( \| [P_n] \|_2^2 = \sqrt{\frac{2}{2n-1}} \). This completes the induction process.

Clearly, \( P_i \) is an \( i \)th order polynomial with positive leading coefficient, then \( e_i \) is also an \( i \)th order polynomial with positive leading coefficient. By the above, we have \( \{ [e_i]\}_{i=0}^\infty \) is orthonormal and span \( \{ e_0, \ldots, e_n \} \) = span \( \{ z_0, \ldots, z_n \} \), \( \forall n \in \mathbb{Z}_+ \). Then, \( \{ [e_i]\}_{i=0}^\infty \) is the sequence generated by the Gram-Schmidt Procedure on \( \{ [z_i]\}_{i=0}^\infty \).

Finally, we will show that \( \{ [e_i]\}_{i=0}^\infty \subseteq \mathbb{Z} \) is complete by Proposition 13.33. All we need to show is that \( \forall z \in \mathbb{Z} \), with \( \langle [z], [z_0] \rangle = 0 \), \( \forall i \in \mathbb{Z}_+ \), we have \( [z] = \vartheta_z \). Since \( \mu(I) = 2 < \infty \), by Cauchy-Schwarz Inequality, we have \( \int_{-1}^1 \mathcal{P} \circ z \, dt \leq (\int_{-1}^1 \mathcal{P} \circ z \, dt)^{1/2} (\int_{-1}^1 1 \, dt)^{1/2} = \sqrt{2} \| z \|_2 < \infty \). \( \text{Then,} \)

\[
z \in \mathbb{L}_1(\mathbb{I}, \mathbb{K}). \text{ Define } F : I \to \mathbb{K} \text{ by } F(t) = \int_{-1}^t z(\tau) \, d\tau, \forall t \in I. \text{ By Proposition 12.75, } F \text{ is absolutely continuous and therefore continuous. Since } [z] \perp [z_0], \text{ then } F(-1) = 0 = F(1) = \langle [z], [z_0] \rangle. \text{ By Proposition 11.37, } F \text{ is } \mathcal{B} \text{-measurable. Then, } F \in \mathbb{Z}. \forall i \in \mathbb{Z}_+ \text{, we have}
\]

\[
\langle [F], [z_i] \rangle = \int_{-1}^1 F(t)z_i \, dt = \int_{-1}^1 F(t)z_i \, dt = 0
\]

where the second equality follows from Integration by Parts Theorem 12.89. Hence, \( [F] \perp ([z_i])_{i=0}^\infty \). Since \( F \) is continuous, by Bernstein Approximation
13.5. APPROXIMATION IN HILBERT SPACES

Theorem 11.203, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \) polynomial \( Q : I \to \mathbb{K} \), such that \( \| F - Q \|_{C(I, \mathbb{K})} < \epsilon \). Then,

\[
\| [F] \|_2^2 = \int_{-1}^{1} |F(t)|^2 \, dt = \int_{-1}^{1} F(t)(F(t) - Q(t)) \, dt \leq \| [F] \|_2 \| [F(t) - Q(t)] \|_2 \leq \| [F] \|_2 \sqrt{2\epsilon}
\]

where the second equality follows from the fact that \( |F| \perp ([z_i])_{i=0}^\infty \); and the first inequality follows from Cauchy-Schwarz Inequality. Then, \( \| [F] \|_2 \leq \sqrt{2\epsilon} \). By the arbitrariness of \( \epsilon \), we have \( \| [F] \|_2 = 0 \). Then \( [F] = 0 \). Since \( F \) is continuous, then \( F(t) = 0, \forall t \in I \). Then, by Fundamental Theorem of Calculus I 12.86, \( z = F(t) \) a.e. in \( I \). Then, \( [z] = 0 \).

Then, by Proposition 13.33, \( ([e_i])_{i=0}^\infty \) is complete.

Proposition 13.35 Let \( S_1 = [0, 2\pi] \subset \mathbb{R} \) and \( z \in C(S_1, \mathbb{C}) =: \mathbb{Z} \). Assume that \( z(0) = z(2\pi) \). Let \( M = \{ \tilde{z} \in \mathbb{Z} \mid \tilde{z}(x) = \cos(nx), \forall x \in S_1, \forall n \in \mathbb{Z}_+ \}; \) or \( \tilde{z}(x) = \sin(nx), \forall x \in S_1, \forall n \in \mathbb{N} \} \subset \mathbb{Z} \). Let \( \mathcal{M} = \text{span}(M) \subset \mathbb{Z} \). Then, \( z \in \mathcal{M} \).

Proof Let \( S_1 \subset \mathbb{R}^2 \) be the unit circle as defined in Example 12.104. Define a mapping \( \Psi : S_1 \to S_1 \) by \( \Psi(x) = (\cos(x), \sin(x)), \forall x \in S_1 \). Clearly, \( \Psi \) is surjective and continuous. Note that \( S_1 \) is compact and \( S_1 \) is Hausdorff. It is obvious that we may define a function \( \Phi : S_1 \to \mathbb{C} \) such that \( \Phi \circ \Psi = z \). By Proposition 5.18, we have \( \Phi \) is continuous. Note that \( S_1 \) is closed and bounded in \( \mathbb{R}^2 \), then \( S_1 \) is compact by Proposition 5.40. Hence, \( \Phi \in C(S_1, \mathbb{C}) \). Let \( \Phi_r := \text{Re} \circ \Phi \in C(S_1, \mathbb{R}) \) and \( \Phi_i := \text{Im} \circ \Phi \in C(S_1, \mathbb{R}) \).

By Tietze Extension Theorem 3.57, there exists \( \Phi_r \in C(I^2, \mathbb{R}) \) and \( \Phi_i \in C(I^2, \mathbb{R}) \), where \( I = [-1, 1] \subset \mathbb{R} \), such that \( \Phi_r|_{S_1} = \Phi_r \) and \( \Phi_i|_{S_1} = \Phi_i \). Let \( \Phi := \Phi_r + i\Phi_i \in C(I^2, \mathbb{C}) \). Then, \( \Phi|_{S_1} = \Phi \). By Bernstein Approximation Theorem 11.203, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), there exists a polynomial \( P : I^2 \to \mathbb{C} \) such that \( |\tilde{z}(y_1, y_2) - P(y_1, y_2)| < \epsilon, \forall y_1, y_2 \in I^2 \). Then, \( |z(x) - P \circ \Psi(x)| = |\Phi \circ \Psi(x) - P \circ \Psi(x)| < \epsilon, \forall x \in S_1 \). Note that \( P \circ \Psi \in \mathcal{M} \), since, \( \forall r, \gamma, \theta \in \mathbb{R} \),

\[
\begin{align*}
(\sin(\theta))^2 &= \frac{1}{2} (1 - \cos(2\theta)); & (\cos(\theta))^2 &= \frac{1}{2} (1 + \cos(2\theta)); \\
\sin(\gamma) \cos(\theta) &= \frac{1}{2} (\sin(\gamma + \theta) + \sin(\gamma - \theta)); \\
\sin(\gamma) \sin(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) - \cos(\gamma + \theta)); \\
\cos(\gamma) \cos(\theta) &= \frac{1}{2} (\cos(\gamma - \theta) + \cos(\gamma + \theta))
\end{align*}
\]

Hence, \( z \in \mathcal{M} \), by the arbitrariness of \( \epsilon \). This completes the proof of the proposition. \( \Box \)
Example 13.36  Let $\mathbb{I} := (\{\pi-\pi, | \cdot |\}, \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$; $\mathcal{Z} := L_2(\mathbb{I}, \sqrt{\pi} \mathbb{K})$ be the Hilbert space as defined in Example 13.11. Then, the inner product on $\mathcal{Z}$, $\langle \cdot, \cdot \rangle_\mathcal{Z}$, is defined by $\langle [g], [h] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t) \overline{h(t)} \, dt, \forall g, h \in \mathcal{Z}$, where $\mathcal{Z} := L_2(\mathbb{I}, \sqrt{\pi} \mathbb{K})$. Let $(e_n)_{n=0}^\infty \in \mathcal{Z}$ be defined by $e_0(t) = \frac{1}{\sqrt{\pi}}$, $\forall t \in \mathbb{I}$, $e_{2n-1}(t) = \sin(nt), \forall t \in \mathbb{I}$, $e_{2n}(t) = \cos(nt)$, $\forall t \in \mathbb{I}$, $\forall n \in \mathbb{N}$; and $\forall f \in \mathcal{Z}$. We will show that $a_{2n} := \langle [f], [e_{2n}] \rangle_\mathcal{Z} \in \mathbb{K}$, $\forall n \in \mathbb{Z}_+$; $a_{2n-1} := \langle [f], [e_{2n-1}] \rangle_\mathcal{Z} \in \mathbb{K}$, $\forall n \in \mathbb{N}$, $(\{e_n\})_{n=0}^\infty \subseteq \mathcal{Z}$ is a complete orthonormal sequence, and $\{f\} = \sum_{n=0}^\infty a_n [e_n]$ in $\mathcal{Z}$. Then, $\sum_{n=0}^\infty a_n e_n$ is called the Fourier series for $f$.

By the finiteness of measure $\mu$, we have $L_2(\mathbb{I}, \sqrt{\pi} \mathbb{K}) \subseteq L_1(\mathbb{I}, \sqrt{\pi} \mathbb{K})$. Then, $f$ is absolutely integrable over $\mathbb{I}$. By Propositions 7.23, 11.38, and 11.39, $f_{e_n}$ is $\mathcal{B}$-measurable and therefore absolutely integrable over $\mathbb{I}$, $\forall n \in \mathbb{Z}_+$. Hence, $a_n \in \mathbb{K}$, $\forall n \in \mathbb{Z}_+$. It is straightforward to check that $(\{e_n\})_{n=0}^\infty$ is an orthonormal sequence in $\mathcal{Z}$. We will show that it is complete by Proposition 13.33. Fix any $[z] \in (\{[e_n]\}_{n=0}^\infty)_{\perp}$, where $z \in \mathcal{Z}$. Then, $z$ is $\mathcal{B}$-measurable and absolutely integrable over $\mathbb{I}$. Define $F : \mathbb{R}_{\pi-\pi} \rightarrow \mathbb{K}$ by $F(t) = \int_{-\pi}^{t} z(s) \, ds$, $\forall t \in \mathbb{R}_{\pi-\pi}$. By Proposition 12.75, $F$ is absolutely continuous. Clearly $F(-\pi) = 0$. Since $[z] \perp [e_0]$, then $F(\pi) = 0$, $\forall n \in \mathbb{N}$,

$$\langle [F], [e_{2n-1}] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \sin(nt) \, dt = -\frac{1}{n\pi} \cos(nt) F(t) \bigg|_{t=-\pi}^{t=\pi} + \frac{1}{n\pi} \int_{-\pi}^{\pi} \cos(nt) z(t) \, dt = 0$$

where the second equality follows from Integration by Parts Theorem 12.89; and the last equality follows from the preceding discussion and the fact $[z] \perp [e_{2n}], \forall n \in \mathbb{N}$,

$$\langle [F], [e_{2n}] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \cos(nt) \, dt = \frac{1}{n\pi} \sin(nt) F(t) \bigg|_{t=-\pi}^{t=\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} \sin(nt) z(t) \, dt = 0$$

where the second equality follows from Integration by Parts Theorem 12.89; and the last equality follows from the preceding discussion and the fact $[z] \perp [e_{2n-1}]$. Hence, $[F] \perp (\{[e_i]\}_{i=1}^\infty)$.

Fix any $\epsilon \in (0, \infty) \subseteq \mathbb{R}$. Since $F \in C(\mathbb{R}_{\pi-\pi}, \mathbb{K}) := \mathcal{Y}$ and $F(-\pi) = F(\pi) = 0$, by Proposition 13.35, $\exists Q \in \text{span} (\{\tilde{e}_i\}_{i=0}^\infty) \subseteq \mathcal{Y}$ such that $\|F - Q\|_{\mathcal{Y}} < \frac{\epsilon}{3\sqrt{2(1+\|F\|_{\mathcal{Z}})^2}}$, where $\tilde{e}_i$ is the continuous extension of $e_i$ from $\mathbb{I}$ to $\mathbb{R}_{\pi-\pi}, \forall i \in \mathbb{Z}_+$. Let $Q := \sum_{i=0}^\infty \alpha_i \tilde{e}_i$, where $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, m$. Then, we have

$$\| [F]_1 \|_\mathcal{Z}^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \overline{F(t)} \, dt$$
\[ \begin{align*} &= \frac{1}{\pi} \int_{-\pi}^{\pi} F(t)F(t - Q(t)) \, dt + \frac{1}{\pi} \int_{-\pi}^{\pi} F(t)\overline{c(t)} \, dt \\ &\leq \left| \frac{1}{\pi} \int_{-\pi}^{\pi} F(t)F(t - Q(t)) \, dt + \frac{1}{\pi} \int_{-\pi}^{\pi} F(t)\overline{c(t)} \, dt \right| \\ &\leq \|[F]_2\| \|[(F - Q)]_1\|_z + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| tF(t) \frac{\|t\|_{I_{S_{\pi}}} - \|\overline{t}\|_{I_{S_{\pi}}}}{2\pi} \int_{-\pi}^{\pi} t\alpha(t) \, dt \right| \\ &\leq \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \int_{-\pi}^{\pi} t\alpha(t) \, dt \right| \\ \end{align*} \]

where the second equality follows from the previous paragraph; the third equality follows from the fact that the left-hand-side is a nonnegative real number; the second inequality follows from the Cauchy-Schwarz Inequality and Integration by Parts Theorem 12.89; and the last inequality follows from the preceding discussion.

Define \( f_1 : r_{-\pi,\pi} \to \mathbb{K} \) by \( f_1(t) = t, \forall t \in r_{-\pi,\pi} \). Then, \( f_1 \in \tilde{L}_2(I, \mathbb{K}) \). Let \( g_1 : r_{-\pi,\pi} \to S_1 \) be defined by \( g_1(x) = (\cos(x), \sin(x)) \in \mathbb{S}_1, \forall x \in r_{-\pi,\pi} \), where \( S_1 \) is defined in Example 12.104. By Example 12.105, \( g_1 \) is an isomereasure between \( I \) and \( S_1 \). Then, \( f_1 \circ g_1^{-1} \in \tilde{L}_2(S_1, \mathbb{K}) =: \tilde{Y}_1 \). By Proposition 11.182, \( \exists f_2 \in \tilde{L}_2(S_1, \mathbb{K}) \) such that \( f_2 \in C(S_1, \mathbb{K}) =: W \) and \( \|f_1 \circ g_1^{-1} - f_2\|_{\tilde{Y}_1} < \frac{\epsilon^2}{3(1 + |\alpha_0|\|z\|_z)} \). Then, by the proof of Proposition 13.35, \( \exists h \in W \) with \( h \) being a polynomial in two coordinates, such that \( \|f_2 - h\|_W < \frac{\epsilon^2}{3(1 + |\alpha_0|\|z\|_z)} \). Then \( h \circ g_1 \in \text{span}((e_i)_{i=0}^{\infty}) \). Since \([z] \perp \text{span}((e_i)_{i=0}^{\infty})\), we have

\[ \begin{align*} \|[F]_1\|_2^2 &\leq \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \int_{-\pi}^{\pi} t\alpha(t) \, dt \right| \\ &= \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \left| \int_{-\pi}^{\pi} t\alpha(t) \, dt - \int_{-\pi}^{\pi} h(g_1(t))z(t) \, dt \right| \\ &= \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \left| \int_{-\pi}^{\pi} (f_1 - h \circ g_1)z(t) \, dt \right| \\ &\leq \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \|[z]\|_z \|[f_1 - h \circ g_1]\|_z \\ &\leq \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \|[z]\|_z (\|[f_1 - f_2 \circ g_1]\|_z + \|[f_2 \circ g_1 - h \circ g_1]\|_z) \\ &= \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \|[z]\|_z \frac{1}{\sqrt{\pi}} \left( \|[f_1 \circ g_1^{-1} - f_2 \circ g_1^{-1}]\|_z + \|[f_2 \circ g_1 - h \circ g_1]\|_z \right) \\ &< \frac{\epsilon^2}{3} + \frac{\epsilon^2}{3} + \left| \frac{\alpha_0}{\sqrt{2\pi}} \right| \|[z]\|_z \frac{1}{\sqrt{\pi}} \sqrt{2\pi} \|[f_2 - h]\|_W < \epsilon^2 \end{align*} \]
where the first equality follows from the fact that \([z] \perp [e_n], \forall n \in \mathbb{Z}_+\); the second equality follows from Proposition 11.89; the second inequality follows from Cauchy-Schwarz Inequality; the third inequality follows from Definition 7.1; the third equality follows from Proposition 12.12; and the last two inequalities follow from the preceding discussion. By the arbitrariness of \(\epsilon\), we have \([F|_I] = \vartheta_Z\). Then, \([F|_I] = 0\) a.e. in \(I\) Since \(F\) is absolutely continuous, then it is continuous by Proposition 12.59, and then \(F(t) = 0, \forall t \in \mathbb{T}_{-\pi, \pi}\) (otherwise, we can always find a nonempty interval that is a subset of \(I\) such that \(F\) is nonzero on the interval.) Then, by Fundamental Theorem of Calculus I, Theorem 12.86, \(z(t) = 0\) a.e. \(t \in I\). Hence, \([z] = \vartheta_Z\).

We have shown that \((\{e_n\}_{n=0}^\infty)^\perp = \{\vartheta_Z\}\). Then, by Proposition 13.33, \(\text{span}(\{e_n\}_{n=0}^\infty) = \mathbb{Z},\) and \((\{e_n\}_{n=0}^\infty)\) is a complete orthonormal sequence. Then \([f] = \sum_{n=0}^\infty a_n [e_n]\) in \(\mathbb{Z}\).

Given independent vectors \(y_1, \ldots, y_n\) generating a subspace \(M\) of a Hilbert space \(X\), we wish to find \(m_0 \in M\) that minimizes \(\|x - m_0\|\), for some given \(x \in X\). Rather than solving the normal equation (13.1), we may employ the Gram-Schmidt orthogonalization procedure together with Proposition 13.31. First use the Gram-Schmidt Procedure to obtain orthonormal sequence \((e_i)_{i=1}^n\). Then, \(m_0 = \sum_{i=1}^n (x, e_i) e_i\). Thus, the optimization problem is easily solved once the vectors \(y_i\)'s are orthogonalized.

Since solution to the approximation problem is equivalent to solution to the normal equation. It is clear that the Gram-Schmidt procedure can be interpreted as a procedure for inverting the Gram matrix.

It is interesting to note that the Gram-Schmidt procedure may itself be viewed as an approximation problem.

\[
e_1 = \frac{y_1}{\|y_1\|}; \quad e_k = \frac{y_k - \sum_{i=1}^{k-1} (y_k, e_i) e_i}{\|y_k - \sum_{i=1}^{k-1} (y_k, e_i) e_i\|}, \quad \forall k > 1
\]

The vector \(y_k - \sum_{i=1}^{k-1} (y_k, e_i) e_i\) is the optimal error for the problem \(\min_{m \in \text{span}(\{y_1, \ldots, y_{k-1}\})} \|y_k - m\|\). Thus, the Gram-Schmidt procedure consists of solving a series of minimum norm approximation problems by use of the Projection Theorem. Alternatively, the minimum norm approximation of \(x\) can be found by applying the Gram-Schmidt procedure to the sequence \((y_1, \ldots, y_n, x)\). The optimal error \(x - m_0\) is found at the last step.

13.6 Other Minimum Norm Problems

In approximation problems, finite dimensionality of the subspace allows the reduction of the problem into a finite-dimensional normal equation, which leads to a feasible computation procedure. In many important and interesting practical problems, the subspace is not finite-dimensional. In
such problems, it is generally not possible to reduce the problem to a finite dimensional normal equation. However, there is an important class of such problems that can be reduced by the projection theorem to finite-dimensional equations. The motivation for this class of problems is illustrated in Figure 13.2. In this class of problems, the subspace $M^\perp$ is finite-dimensional, which leads to feasible computation procedures.

Next, we state a modified version of the Projection Theorem 13.13.

**Theorem 13.37 (Restatement of Projection Theorem)** Let $\mathcal{X}$ be a Hilbert space, $M \subseteq \mathcal{X}$ be a closed subspace, $x, y \in \mathcal{X}$, $V = y + M$ be a closed linear variety. Then, there is a unique vector $v_0 \in V$ such that $\|x - v_0\| = \min_{v \in V} \|x - v\|$. A necessary and sufficient condition that $v_0 \in V$ is the unique minimizing vector is that $(x - v_0) \perp M$.

**Proof** This is straightforward, and is therefore omitted. □
A point of caution is necessary here: \((x - v_0) \perp M\) but not \((x - v_0) \perp V\).

The concept of the above theorem is illustrated in Figure 13.3.

A special type of linear variety is

\[ V = \{ v \in \mathbb{X} \mid \langle v, y_i \rangle = c_i, \ i = 1, \ldots, n \} \]

where \(y_1, \ldots, y_n \in \mathbb{X}\) are fixed vectors and \(c_1, \ldots, c_n \in \mathbb{K}\) are fixed scalars. Let \(M = \text{span}\{y_1, \ldots, y_n\}\). When \(c_1 = \cdots = c_n = 0\), then \(V = M^\perp\). For arbitrary \(c_i\)'s, \(V\) is equal to \(y + M^\perp\) assuming \(V \neq \emptyset\). A linear variety of this form is said to be of finite co-dimension, since \(M^\perp = M\) is finite-dimensional by Proposition 13.17.

**Theorem 13.38** Let \(\mathbb{X}\) be a Hilbert space over \(\mathbb{K}\), \(n \in \mathbb{N}\), \(y_1, \ldots, y_n \in \mathbb{X}\), \(c_1, \ldots, c_n \in \mathbb{K}\), \(V = \{ v \in \mathbb{X} \mid \langle v, y_i \rangle = c_i, \ i = 1, \ldots, n \} \neq \emptyset, x \in \mathbb{X}\). Then, the following equation

\[
\begin{bmatrix}
\langle x, y_1 \rangle - c_1 \\
\vdots \\
\langle x, y_n \rangle - c_n
\end{bmatrix}
= (\text{Gram}(y_1, \ldots, y_n))^\prime
\begin{bmatrix}
\beta_1 \\
\vdots \\
\beta_n
\end{bmatrix}
\tag{13.2}
\]

admits at least one solution \((\beta_1, \ldots, \beta_n) \in \mathbb{K}^n\). Any such solution implies that \(v_0 = x - \sum_{i=1}^n \beta_i y_i \in V\) is the unique minimizing vector for \(\min_{v \in V} \|x - v\|\).
13.6. OTHER MINIMUM NORM PROBLEMS

Proof  Let $M := \text{span} \left( \{ y_1, \ldots, y_n \} \right)$. By Theorem 7.36 and Proposition 4.39, $M$ is a closed subspace of $X$. Then, $M^\perp = \{ \bar{x} \in X \mid \langle \bar{x}, y_i \rangle = 0, \; i = 1, \ldots, n \}$. Fix any $y_0 \in V \neq \emptyset$, we have $V = y_0 + M^\perp$. By Theorem 13.37, $\exists v_0 \in V$ such that $\| x - v_0 \| = \min_{v \in V} \| x - v \|$ and a necessary and sufficient condition for any $v \in V$ to be $v_0$ is that $(x - v_0) \perp M^\perp$. Then, $(x - v_0) \in M^{\perp \perp} = M$, by Proposition 13.17. Then, $\exists (\beta_1, \ldots, \beta_n) \in \mathbb{K}^n$ such that $x - v_0 = \sum_{i=1}^n \beta_i y_i$. Then, $v_0 = x - \sum_{i=1}^n \beta_i y_i$. Since $v_0 \in V$, we must have $\langle v_0, y_i \rangle = c_i, \; i = 1, \ldots, n$. This leads to (13.2).

Any solution $(\beta_1, \ldots, \beta_n) \in \mathbb{K}^n$ to (13.2), we have $v_0 \in V$ and $(x - v_0) \in M = M^{\perp \perp}$. By Theorem 13.37, $v_0$ is the unique minimizing vector for $\min v \in V \| x - v \|$. This completes the proof of the theorem. \hfill \square

Much of the discussion in the above can be generalized from linear varieties to convex sets.

Theorem 13.39  Let $X$ be a real Hilbert space, $x \in X$, $K \subseteq X$ be a nonempty closed convex subset. Then, there is a unique vector $k_0 \in K$ such that $\| x - k_0 \| = \min_{k \in K} \| x - k \|$. Furthermore, a necessary and sufficient condition for $k_0 \in K$ being the unique minimizing vector is that $\langle x - k_0, k - k_0 \rangle \leq 0$, $\forall k \in K$.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{projection_to_convex_set.png}
\caption{Projection to a convex set.}
\end{figure}

The main idea of this theorem is illustrated in Figure 13.4, which shows that the angle between $x - k_0$ and $k - k_0$ is greater than or equal to $90^\circ$.

Proof  First, we show the existence of $k_0$. Let $\delta := \inf_{k \in K} \| x - k \| \in [0, \infty) \subset \mathbb{R}$ since $K \neq \emptyset$. Then, $\exists (k_i)_{i=1}^\infty \subseteq K$ such that $\lim_{i \in \mathbb{N}} \| x - k_i \| = \infty$. Then, $\exists (k_i)_{i=1}^\infty \subseteq K$ such that $\lim_{i \in \mathbb{N}} \| x - k_i \| = \infty$. This completes the proof of the theorem.
\[ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall i \in \mathbb{N} \text{ with } n_0 \leq i, \text{ we have } \delta \leq \|x - k_i\| < \sqrt{\delta^2 + \epsilon^2}/4. \]

\[ \forall i, j \in \mathbb{N} \text{ with } n_0 \leq i \leq j, \text{ by Parallelogram Law 13.3, } \| (x - k_i) - (x - k_j) \|^2 + \| (x - k_i) + (x - k_j) \|^2 = 2 \| x - k_i \|^2 + 2 \| x - k_j \|^2. \]

This implies that \[ \| k_i - k_j \|^2 = 2 \| x - k_i \|^2 + 2 \| x - k_j \|^2 - 4 \| x - (k_i + k_j)/2 \|^2 < 4(\delta^2 + \epsilon^2/4) - 4\delta^2 = \epsilon^2 , \]

where the inequality follows from the fact that \( K \) is convex and \( (k_i + k_j)/2 \in K \). Hence, \( (k_i)_{i=1}^\infty \subseteq K \) is Cauchy sequence. By Proposition 4.39, \( K \) is complete and \( \lim_{i \in \mathbb{N}} k_i = k_0 \in K \). By Propositions 3.66 and 7.21, \( \delta = \lim_{i \in \mathbb{N}} \| x - k_i \| = \| x - k_0 \| . \)

Next, we show the uniqueness of \( k_0 \). Let \( \tilde{k} \in K \) be such that \( \delta = \| x - \tilde{k} \| \). The sequence \( (k_i)_{i=1}^\infty := (k_0, \tilde{k}, k_0, \tilde{k}, \ldots) \subseteq K \) satisfies \( \lim_{i \in \mathbb{N}} \| x - k_i \| = \delta \). By the proof of the existence, \( (k_i)_{i=1}^\infty \) is convergent. Then, we must have \( \tilde{k} = k_0 \). Hence, \( k_0 \) is unique.

Next, we show that \( \langle x - k_0, k - k_0 \rangle \leq 0, \forall k \in K \) by an argument of contradiction. Suppose this is not true. Then, \( \exists k_1 \in K \) such that

\[ \langle x - k_0, k_1 - k_0 \rangle =: \lambda > 0. \]

Consider the vector \( k_\alpha := (1 - \alpha)k_0 + \alpha k_1 \), where \( \alpha \in (0, 1) \subset \mathbb{R} \) is to be determined. Since \( K \) is convex, then \( k_\alpha \in K \).

Note that \( \| x - k_\alpha \|^2 = \| (x - k_0) - \alpha (k_1 - k_0) \|^2 = \| x - k_0 \|^2 - 2\alpha \langle x - k_0, k_1 - k_0 \rangle + \alpha^2 \| k_1 - k_0 \|^2 = \delta^2 - 2\alpha \lambda + \alpha^2 \| k_1 - k_0 \|^2 \)

Then, for sufficiently small \( \alpha \in (0, 1) \subset \mathbb{R} \), we have \( \| x - k_0 \|^2 < \delta^2 \), which contradicts with the definition of \( \delta \). Hence, \( \langle x - k_0, k - k_0 \rangle \leq 0, \forall k \in K \).

Finally, we show that if \( \tilde{k} \in K \) satisfies \( \langle x - \tilde{k}, k - \tilde{k} \rangle \leq 0, \forall k \in K \), then \( \| x - \tilde{k} \| = \min_{k \in K} \| x - k \| \), \( \forall k \in K \), we have \( \| x - k \|^2 = \| (x - \tilde{k}) - (k - \tilde{k}) \|^2 = \| x - \tilde{k} \|^2 - 2 \langle x - \tilde{k}, k - \tilde{k} \rangle + \| k - \tilde{k} \|^2 \geq \| x - \tilde{k} \|^2 \).

This completes the proof of the theorem. \( \square \)

### 13.7 Positive Definite Operators on Hilbert Spaces

Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{R} \). In Chapter 10, we defined the symmetric operators \( \mathcal{S}_X \) to be \( \mathcal{B}_{22}(\mathcal{X}, \mathbb{R}) \subseteq \mathcal{B}(\mathcal{X}, \mathcal{X}^*) \). By Riesz-Frèchet Theorem 13.15, \( \mathcal{X}^* = \mathcal{X} \). Then, \( \mathcal{S}_X \subseteq \mathcal{B}(\mathcal{X}, \mathcal{X}) \). Fix any \( A \in \mathcal{S}_X \), \( \forall x, y \in \mathcal{X} \), we have

\[ \langle y, Ax \rangle = A(x)(y) = A(y)(x) = \langle x, Ay \rangle = \langle Ax, y \rangle = \langle y, A^* x \rangle , \]

where the first equality follows since \( A(x) \) is a linear functional on \( \mathcal{X} \) and by Riesz-Frèchet Theorem 13.15; the second equality follows since \( A \in \mathcal{B}_{22}(\mathcal{X}, \mathbb{R}) \); the third equality follows from the same argument as the first equality; the fourth equality follows from Proposition 13.22; and the last equality follows from Definition 13.1 and \( \mathcal{X} \) is a real Hilbert space. By the arbitrariness of \( x \) and \( y \), we have \( A^* = A \), and \( A \) is Hermitian. By Definition 13.21, \( A = A^* = A^t \) since \( \mathcal{X} \) is a real Hilbert space. Here, we will generalize the definition of \( \mathcal{S}_X, \mathcal{S}_{\pm} X, \mathcal{S}_{\text{psd}} X, \mathcal{S}_{\text{psd}} X, \) and \( \mathcal{S}_{\text{psd}} X \) to the case where \( \mathcal{X} \) is a complex Hilbert space.
13.7. POSITIVE DEFINITE OPERATORS ON HILBERT SPACES

Definition 13.40 Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$ and $A \in B(\mathcal{X}, \mathcal{X})$. We will write $A \in \mathcal{S}_X$, if $A$ is Hermitian. We will write $A \in \mathcal{S}_{+X}$ if $A \in \mathcal{S}_X$ and $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $\langle x, Ax \rangle \geq m \| x \|^2$, $\forall x \in \mathcal{X}$. We will write $A \in \mathcal{S}_{\text{pos}} X$ if $A \in \mathcal{S}_X$ and $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $\langle x, Ax \rangle \geq m \| x \|^2$, $\forall x \in \mathcal{X}$. We will write $A \in \mathcal{S}_{-X}$ if $A \in \mathcal{S}_X$ and $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $\langle x, Ax \rangle \leq -m \| x \|^2$, $\forall x \in \mathcal{X}$. We will write $A \in \mathcal{S}_{\text{neg}} \mathcal{X}$ if $A \in \mathcal{S}_X$ and $\exists m \in (0, \infty) \subset \mathbb{R}$ such that $\langle x, Ax \rangle \leq -m \| x \|^2$, $\forall x \in \mathcal{X}$.

Proposition 13.41 Let $\mathcal{X}$ be a Hilbert space, $A \in B(\mathcal{X}, \mathcal{X})$ be Hermitian (that is, $A \in \mathcal{S}_X$). Assume that $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\langle x, Ax \rangle \geq \delta \| x \|^2$, $\forall x \in \mathcal{X}$. Then, $A$ is bijective and $A^{-1} \in B(\mathcal{X}, \mathcal{X})$. Furthermore, $A^{-1}$ is Hermitian and $\forall x \in \mathcal{X}$, we have $\langle x, A^{-1} x \rangle \geq \frac{\delta}{\| x \|_{\| x \|}} \| x \|^2$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$.

Proof By the assumption, $\forall x \in \mathcal{X}$ with $x \neq \vartheta_X$, $A(x) \neq \vartheta_X$ since $\langle x, Ax \rangle \geq \delta \| x \|^2 > 0$. Hence, $\mathcal{N}(A) = \{ \vartheta_X \}$ and $A$ is injective. By Proposition 13.22, $(\mathcal{R}(A))^\perp = \mathcal{N}(A^*) = \mathcal{N}(A) = \{ \vartheta_X \}$, where the second equality follows from the fact that $A$ is Hermitian. This shows that $((\mathcal{R}(A))^\perp)^\perp = \mathcal{X}$. By Proposition 13.17, $\overline{\mathcal{R}(A)} = \mathcal{X}$. Hence, $\mathcal{R}(A)$ is dense in $\mathcal{X}$. Clearly, $\vartheta_X \in \mathcal{R}(A)$. $\forall x \in \mathcal{X}$ with $x \neq \vartheta_X$, there exists $(x_i)_{i=1}^\infty \subseteq \mathcal{R}(A)$ such that $x = \lim_{i \to \infty} x_i$. Without loss of generality, we may assume that $x_i \neq \vartheta_X$, $\forall i \in \mathbb{N}$. Let $x_i = A(x_i)$, $\forall i \in \mathbb{N}$, where $\bar{x}_i \in \mathcal{X}$. Clearly, $\bar{x}_i \neq \vartheta_X$ since $x_i \neq \vartheta_X$, $\forall i \in \mathbb{N}$. Then, we have

$$\| \bar{x}_i \| \geq \| \bar{x}_i \| = \| \bar{x}_i, A\bar{x}_i \| \geq \delta \| \bar{x}_i \|^2; \quad \forall i \in \mathbb{N}$$

where the first inequality follows from the Cauchy-Schwarz Inequality. This implies that $\| x_i \| \geq \delta \| \bar{x}_i \|$ and $\| \bar{x}_i \| \leq \| x_i \| / \delta$, $\forall i \in \mathbb{N}$. Since $(x_i)_{i=1}^\infty$ is convergent, then there exists $c \in (0, \infty) \subset \mathbb{R}$ such that $\| x_i \| \leq c$, $\forall i \in \mathbb{N}$. Then, $\| \bar{x}_i \| \leq c / \delta := c_1 \in (0, \infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$. This shows that $(\bar{x}_i)_{i=1}^\infty \subseteq \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_X, c_1) =: S_1 \subseteq \mathcal{X} = \mathcal{X}^\ast$, where the last equality follows from Riesz-Frèchet Theorem 13.15. By Alaoglu Theorem 7.122, $S_1 \subseteq \mathcal{X}^\ast$ is weak* compact. By Propositions 5.22 and 5.26, $(\bar{x}_i)_{i=1}^\infty$ has a cluster point $\bar{x} \in \mathcal{X}^\ast = \mathcal{X}$ in weak* topology. Since $\mathcal{X}$ is reflexive, the weak* topology on $\mathcal{X}^\ast = \mathcal{X}$ is identical to the weak topology on $\mathcal{X}$. By Proposition 7.123, $A \in B(\mathcal{X}, \mathcal{X})$ is continuous as a mapping $A : \mathcal{X}_{\text{weak}} \to \mathcal{X}_{\text{weak}}$. By Proposition 3.66, $(x_i)_{i=1}^\infty = (A(x_i))_{i=1}^\infty$ has a cluster point $A(\bar{x}) \in \mathcal{X}_{\text{weak}}$ in weak topology. By Proposition 7.116, $\mathcal{X}_{\text{weak}}$ is completely regular, then it is Hausdorff by Proposition 3.61. By Proposition 3.65, $A(\bar{x}) = x$, since $x = \lim_{i \to \infty} x_i$ and $x = \lim_{i \to \infty} x_i$ weakly. This shows that $x \in \mathcal{R}(A)$. By the arbitrariness of $x$, we have $\mathcal{R}(A) = \mathcal{X}$ and $A$ is surjective.

Hence, $A$ is bijective. By Open Mapping Theorem 7.103, $A^{-1} \in B(\mathcal{X}, \mathcal{X})$. Then, $A^{-1}$ is Hermitian by Proposition 13.22 and the fact that $A$ is Hermitian.

$\forall x \in \mathcal{X}$, let $\bar{x} = A^{-1}(x)$. We have

$$\langle x, A^{-1} x \rangle = \langle A\bar{x}, \bar{x} \rangle = \| \bar{x} \| \geq \delta \| x \|^2$$
Then, $x = A\bar{x}$ and $\|x\| \leq \|A\|_{B(X,X)} \|\bar{x}\|$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|A\|_{B(X,X)} = 0$; Case 2: $\|A\|_{B(X,X)} > 0$.

Case 1: $\|A\|_{B(X,X)} = 0$, we have $\|x\| = 0$, $\forall x \in X$. Hence, $X = \{0\}$.

Then, $\langle x, A^{-1}x \rangle = 0 \geq \frac{\delta}{\varepsilon + \|A\|_{B(X,X)}^2} \|x\|^2$, $\forall \varepsilon \in (0, \infty) \subset R$.

Case 2: $\|A\|_{B(X,X)} > 0$. Then, we have

$$\langle x, A^{-1}x \rangle \geq \frac{\delta}{\varepsilon + \|A\|_{B(X,X)}^2} \|x\|^2,$$

$$\geq \frac{\delta}{\varepsilon + \|A\|_{B(X,X)}^2} \|x\|^2, \quad \forall \varepsilon \in (0, \infty) \subset R$$

Hence, we have shown, in both cases, that $\langle x, A^{-1}x \rangle \geq \frac{\delta}{\varepsilon + \|A\|_{B(X,X)}^2} \|x\|^2 \geq 0$. This completes the proof of the proposition.

**Proposition 13.42** Let $X$ be a Hilbert space over $K$ and $A \in S_X$. Then, the following statements hold.

(i) $\forall x \in X$, $\langle x, Ax \rangle \in R$.

(ii) $S_X$ is a closed subset of $B(X,X)$. $\forall A_1, A_2 \in S_X$, $\forall \lambda \in R$, we have $A_1 + A_2 \in S_X$ and $\lambda A_1 \in S_X$. If $K = R$, $S_X$ is a closed subspace of $B(X,X)$, and hence a Banach subspace. If $K = C$, no further conclusion can be made about $S_X$ except that it is a closed convex cone in $B(X,X)$.

(iii) $S_{+X}$ and $S_{-X}$ are open subsets of $S_X$ (in the subset topology of $S_X \subseteq B(X,X)$); and $S_{+X} = -S_{-X}$.

(iv) $S_{psd\,X}$ and $S_{nsd\,X}$ are closed convex cones in $S_X$ (in the subset topology of $S_X \subseteq B(X,X)$); and $S_{psd\,X} = -S_{nsd\,X}$.

(v) $A \in S_{+X}$ if, and only if, $A^{-1} \in S_{+X}$.

(vi) $A \in S_{-X}$ if, and only if, $A^{-1} \in S_{-X}$.

(vii) The interior of $S_{psd\,X}$ relative to $S_X$ is $S_{+X}$, and the interior of $S_{nsd\,X}$ relative to $S_X$ is $S_{-X}$.

(viii) $A \in S_{+X}$, $B, C \in S_{psd\,X}$, $\alpha \in (0, \infty) \subset R$, and $\beta \in [0, \infty) \subset R$, implies that $A + B \in S_{+X}$, $B + C \in S_{psd\,X}$, $\alpha A \in S_{+X}$, and $\beta B \in S_{psd\,X}$.

(ix) Let $Y$ be a Hilbert space over $K$, and $B \in B(X,Y)$. Then, $BB^* \in S_{psd\,Y}$ and $B^*B \in S_{psd\,X}$. 
13.7. POSITIVE DEFINITE OPERATORS ON HILBERT SPACES 695

Proof 

(i) \( \forall x \in \mathcal{X}, \langle x, Ax \rangle = \langle A^* x, x \rangle = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle} \), where the first equality follows from Proposition 13.22; the second equality follows from \( A \in \mathcal{S}_\mathcal{X} \); and the last equality follows from Definition 13.1. Hence, \( \langle x, Ax \rangle \in \mathbb{R} \).

(ii) \( \forall A \in \overline{\mathcal{S}_\mathcal{X}}, \) by Proposition 4.13, there exists \( (A_i)_{i=1}^\infty \subseteq \mathcal{S}_\mathcal{X} \) such that \( \lim_{i \to \infty} A_i = A \) in \( \mathcal{B}(\mathcal{X}, \mathcal{X}) \). Then, \( \forall x, y \in \mathcal{X}, \) we have \( \langle Ax, y \rangle = \lim_{i \to \infty} \langle A_i x, y \rangle = \lim_{i \to \infty} \langle x, A_i^* y \rangle = \lim_{i \to \infty} \langle x, A_i y \rangle = \langle x, Ay \rangle = \langle A^* x, y \rangle \), where the first equality follows from Propositions 7.65, 13.4, 3.12, and 3.66; the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.22; the fourth inequality follows from the previous discussion and Definition 13.1; the third inequality follows from Propositions 7.65, 13.4, 3.12, and 3.66; and the last equality follows from Proposition 13.22. By the arbitrariness of \( x \) and \( y \), we have \( A = A^* \) and therefore \( A \in \mathcal{S}_\mathcal{X} \). Thus, we have shown that \( \overline{\mathcal{S}_\mathcal{X}} \subseteq \mathcal{S}_\mathcal{X} \subseteq \overline{\mathcal{S}_\mathcal{X}} \).

By Proposition 3.3, \( \mathcal{S}_\mathcal{X} \) is a closed subset of \( \mathcal{B}(\mathcal{X}, \mathcal{X}) \).

\( \forall A_1, A_2 \in \mathcal{S}_\mathcal{X} \) and \( \forall \lambda \in \mathbb{R}, \) we have \( (A_1 + A_2)^* = A_1^* + A_2^* = A_1 + A_2 \), where the first equality follows from Proposition 13.22. Then, \( A_1 + A_2 \in \mathcal{S}_\mathcal{X} \).

In addition, \( (\lambda A_1)^* = \overline{\lambda} A_1 \), where the first equality follows from Proposition 13.22. Hence, \( A_1 \in \mathcal{S}_\mathcal{X} \).

Note that, the preceding paragraph is enough to conclude that \( \mathcal{S}_\mathcal{X} \) is a subspace if \( \mathcal{X} \) is a real Hilbert space. In this case, \( \mathcal{S}_\mathcal{X} \) is a closed subspace and therefore a Banach subspace of \( \mathcal{B}(\mathcal{X}, \mathcal{X}) \) by Proposition 4.39. When \( \mathbb{K} = \mathbb{C}, \) we only have \( \mathcal{S}_\mathcal{X} \) is a closed convex cone and hence a complete metric subspace of \( \mathcal{B}(\mathcal{X}, \mathcal{X}) \) by Proposition 4.39.

(iii) It is clear that \( \mathcal{S}_{+\mathcal{X}} = -\mathcal{S}_{-\mathcal{X}} \). We will show that \( \mathcal{S}_{+\mathcal{X}} \) is open in \( \mathcal{S}_\mathcal{X} \). Fix any \( B \in \mathcal{S}_{+\mathcal{X}} \). Then, \( \exists m \in (0, \infty) \subset \mathbb{R} \) such that \( \langle x, Bx \rangle \geq m \| x \|^2, \forall x \in \mathcal{X}. \) \( \forall C \in \mathcal{S}_\mathcal{X} \cap \mathcal{B}(\mathcal{X}, \mathcal{X}) \) \( (B, \frac{1}{m} m), \) we have \( \langle x, Cx \rangle \in \mathbb{R}, \forall x \in \mathcal{X}, \) by (i). Then, \( \langle x, Cx \rangle \geq \langle x, Bx \rangle - |\langle x, Bx \rangle - \langle x, Cx \rangle| \geq m \| x \|^2 - |\langle x, (B - C)x \rangle| \geq m \| x \|^2 - \| x \| \| (B - C)x \| \geq m \| x \|^2 - \| x \| \| B - C \| \| x \| \geq m \| x \|^2 - \frac{m}{4} \| x \|^2 = \frac{m}{4} \| x \|^2, \forall x \in \mathcal{X}, \) where the second inequality follows from the previous discussion and Definition 13.1; the third inequality follows from Cauchy-Schwarz Inequality; the fourth inequality follows from Proposition 7.64; and the fifth inequality follows from the fact that \( \| B - C \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} < \frac{1}{4} m \). By the arbitrariness of \( x, C \in \mathcal{S}_{+\mathcal{X}} \). By the arbitrariness of \( C, \) we have \( \mathcal{B}(\mathcal{X}, \mathcal{X}) (B, \frac{1}{m} m) \cap \mathcal{S}_\mathcal{X} \subseteq \mathcal{S}_{+\mathcal{X}} \). Hence, \( \mathcal{S}_{+\mathcal{X}} \) is a closed subset of \( \mathcal{S}_\mathcal{X} \). By the relation \( \mathcal{S}_{+\mathcal{X}} = -\mathcal{S}_{-\mathcal{X}} \), we have \( \mathcal{S}_{-\mathcal{X}} \) is also an open subset of \( \mathcal{S}_\mathcal{X} \).

(iv) It is clear that \( \mathcal{S}_{\text{psd} \mathcal{X}} = -\mathcal{S}_{\text{psd} \mathcal{X}} \). We will show that \( \mathcal{S}_{\text{psd} \mathcal{X}} \) is a closed subset of \( \mathcal{S}_\mathcal{X} \). \( \forall B \in \overline{\mathcal{S}_{\text{psd} \mathcal{X}}}, \) by Proposition 4.13, there exists \( (B_i)_{i=1}^\infty \subseteq \mathcal{S}_{\text{psd} \mathcal{X}} \) such that \( \lim_{i \to \infty} B_i = B \) in \( \mathcal{B}(\mathcal{X}, \mathcal{X}) \). By (ii), \( B \in \mathcal{S}_\mathcal{X} \). \( \forall x \in \mathcal{X}, \) \( \langle x, Bx \rangle = \lim_{i \to \infty} \langle x, B_i x \rangle \geq 0, \) where the equality follows from Propositions 7.65, 13.4, 3.12, and 3.66; and the inequality follows from \( B_i \in \mathcal{S}_{\text{psd} \mathcal{X}} \). By the arbitrariness of \( x, B \in \mathcal{S}_{\text{psd} \mathcal{X}} \). Thus, we have shown that \( \overline{\mathcal{S}_{\text{psd} \mathcal{X}}} \subseteq \mathcal{S}_{\text{psd} \mathcal{X}} \subseteq \overline{\mathcal{S}_{\text{psd} \mathcal{X}}}. \) By Proposition 3.3, \( \mathcal{S}_{\text{psd} \mathcal{X}} \) is a closed subset of \( \mathcal{S}_\mathcal{X} \).
\( \forall A_1, A_2 \in S_{\text{psd}, X}, \forall \alpha \in [0, 1] \subset \mathbb{R}. \) Let \( B := \alpha A_1 + (1 - \alpha) A_2. \) By (ii), \( B \in S_X. \) \( \forall x \in X, (x, Bx) = (x, \alpha A_1 x + (1 - \alpha) A_2 x) = (x, \alpha A_1 x) + (1 - \alpha) (x, A_2 x) = \alpha (x, A_1 x) + (1 - \alpha) (x, A_2 x) \geq 0 = 0 \|x\|^2, \) where the second and third equalities follow from Definition 13.1; and the inequality follows from \( A_1, A_2 \in S_{\text{psd}, X}. \) By the arbitrariness of \( x, \) we have \( B \in S_{\text{psd}, X}. \) Hence, \( S_{\text{psd}, X} \) is convex.

Obviously, \( \partial B(\lambda x, x) \in S_{\text{psd}, X} \). \( \forall A \in S_{\text{psd}, X}, \forall \lambda \in [0, \infty) \subset \mathbb{R}, \) we have \( \lambda A \in S_X \) by (ii). \( \forall x \in X, (x, \lambda A_1 x) = \lambda (x, A_1 x) \geq 0, \) where the first equality follows from Definition 13.1 and the fact \( \lambda \in \mathbb{R}; \) and the inequality follows from \( \lambda \geq 0 \) and \( A_1 \in S_{\text{psd}, X}. \) By the arbitrariness of \( x, \) we have \( \lambda A_1 \in S_{\text{psd}, X}. \) Thus, \( S_{\text{psd}, X} \) is a cone with vertex at origin in \( S_X \subseteq B(X, X). \) Hence, \( S_{\text{psd}, X} \) is a closed convex cone in \( S_X. \)

By the relation \( S_{\text{psd}, X} = -S_{\text{nsd}, X} \), we have \( S_{\text{nsd}, X} \) is a closed convex cone in \( S_X. \)

(v) This follows directly from Proposition 13.41 and Definition 13.40.

(vi) This follows immediately from (v) and (iii).

(vii) Clearly, \( S_{+}, X \) is an open subset of \( S_X \) and is contained in \( S_{\text{psd}, X} \), which is a closed subset of \( S_X. \) Let \( P \) be the interior of \( S_{\text{psd}, X} \) relative \( S_X. \) Then, \( P \supseteq S_+^X. \) We will show that \( P \subseteq S_+^X. \) This will imply that \( P = S_+^X. \) Suppose \( P \nsubseteq S_+^X. \) Then, there exists \( B \in P \setminus S_+^X. \) \( B \in P \) implies that \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( B(B(\lambda x, x), (B, \delta) \cap S_+^X \subseteq P \subseteq S_{\text{psd}, X}. \) Since \( B \notin S_+^X, \) then \( \exists x_0 \in X \) such that \( (x_0, Bx_0) < \frac{\delta}{2} \|x_0\|^2. \) Clearly, \( x_0 \neq \partial_\lambda. \) Consider the operator \( \hat{B} := B - \frac{\delta}{2} \text{id}_X. \) Clearly, \( \hat{B} \in S_{B(\lambda x, x)} (B, \delta). \) By Proposition 13.22, \( \hat{B}^* = B^* - \frac{\delta}{2} \text{id}_X = B - \frac{\delta}{2} \text{id}_X = \hat{B}. \) Hence, \( \hat{B} \in S_X. \)

By our earlier discussion, we must have \( \hat{B} \in S_{\text{psd}, X}. \) Yet, \( (x_0, Bx_0) = (x_0, Bx_0) - \left( \frac{\delta}{2} \right) (x_0, \text{id}_X x_0) < \frac{\delta}{2} \|x_0\|^2 + \frac{\delta}{2} \|x_0\|^2 = -\frac{\delta}{2} \|x_0\|^2 < 0, \) where the first equality follows from Definition 13.1; the first inequality follows from our earlier discussion; and the last equality follows from the fact \( x_0 \neq \partial_\lambda. \) This contradicts with the fact that \( \hat{B} \in S_{\text{psd}, X}. \) Hence, the hypothesis does not hold. \( P \subseteq S_+^X. \) This proves that \( P = S_+^X. \)

The statement that the interior of \( S_{\text{nsd}, X} \) relative to \( S_X \) is \( S_{-}, X \) can be proved by an argument that is similar to the preceding paragraph.

(viii) This follows immediately from Definition 13.40.

(ix) Clearly, \( B^* \in B(y, X). \) Then, \( BB^* \in B(y, y). \) By Proposition 13.22, we have \( (BB^*)^* = (B^*)^* B^* = BB^*. \) Hence, \( BB^* \in S_Y. \) \( \forall y \in Y, \) \( \langle y, BB^*y \rangle = \langle B^*y, B^*y \rangle X = \|B^*y\|_X^2 \geq 0, \) where the first equality follows from Proposition 13.22; and the second equality and the inequality follow from Proposition 13.2. Hence, \( BB^* \in S_{\text{psd}, Y}. \)

By similar arguments, we have \( B^* B \in S_{\text{psd}, X}. \)

This completes the proof of the proposition. \( \square \)

Based on the preceding proposition, when \( X \) is a real Hilbert space, we will let \( S_{\text{psd}, X} \) be the usual positive cone in \( S_X, \) which is a real Banach space. In this case, \( \forall A, B \in S_X, \) we will write \( A \geq B \) or \( B \leq A \) if \( A - B \in S_{\text{psd}, X}; \)
we will write \( A > B \) or \( B < A \) if \( A - B \in \mathcal{S}_+ \).

## 13.8 Pseudoinverse Operator

**Proposition 13.43** Let \( \mathcal{X} \) be a Hilbert space, \( M \subseteq \mathcal{X} \) be a closed subspace. Then, we may define a mapping \( P: \mathcal{X} \to M \) by \( P(x) = x_0, \forall x \in \mathcal{X} \), where \( x_0 \) is the unique solution to the \( \| x - x_0 \| = \min_{m \in M} \| x - m \| \), by the Classical Projection Theorem 13.13. \( P \) is said to be the projection operator of \( \mathcal{X} \) to \( M \). Then, we have the following properties for the projection operator.

(i) \( P \in \mathfrak{B}(\mathcal{X}, \mathcal{X}) \).

(ii) \( \| P \| \mathfrak{B}(\mathcal{X}, \mathcal{X}) \leq 1 \).

(iii) \( P^2 = P \) (idempotent)

(iv) \( P^* = P \) (Hermitian)

(v) \( \forall A \in \mathfrak{B}(\mathcal{X}, \mathcal{X}) \) that satisfies (iii) and (iv) is a projection operator.

**Proof**

(i) By Theorem 13.19, \( \mathcal{X} = M \oplus M^\perp \). Then, \( \forall x \in \mathcal{X} \), there exists a unique pair \( x_1 \in M \) and \( x_2 \in M^\perp \) such that \( x_1 + x_2 = x \). Then, \( x_1 = Px \). We will show that \( P \) is a linear operator. \( \forall x, y \in \mathcal{X}, \forall \lambda \in \mathbb{K} \), we have \( x = x_1 + x_2 \) and \( y = y_1 + y_2 \) with \( x_1, y_1 \in M \) and \( x_2, y_2 \in M^\perp \). Then, \( x + y = x_1 + y_1 + x_2 + y_2 \) with \( x_1 + y_1 \in M \) and \( x_2 + y_2 \in M^\perp \). Hence, \( P(x + y) = x_1 + y_1 = P(x) + P(y) \). \( \lambda x = \lambda x_1 + \lambda x_2 \) with \( \lambda x_1 \in M \) and \( \lambda x_2 \in M^\perp \). Then, \( P(\lambda x) = \lambda x_1 = \lambda P(x) \). Hence, \( P \) is linear.

Next, we show that \( P \) is bounded. \( \| P \| \mathfrak{B}(\mathcal{X}, \mathcal{X}) = \sup_{x \in \mathcal{X}, \| x \| \leq 1} \| Px \| \).

\[
\| Px \|^2 = \| x_1 \|^2 = \langle x_1, x_1 \rangle = \langle x - x_2, x - x_2 \rangle = \langle x, x \rangle - \langle x, x_2 \rangle - \langle x_2, x \rangle + \langle x_2, x_2 \rangle = \| x \|^2 - \langle x_2, x \rangle + \langle x_2, x_2 \rangle = \| x \|^2 - \| x_2 \|^2 \leq \| x \|^2,
\]

where the first equality follows from the definition of \( P \); the second equality follows from Proposition 13.2; the third through fifth equalities follows from basic algebra; the sixth equality follows from the fact \( x_1 \in M \) and \( x_2 \in M^\perp \). Hence, \( \| P \| \mathfrak{B}(\mathcal{X}, \mathcal{X}) \leq 1 \). Hence, \( P \) is bounded. Then, \( P \in \mathfrak{B}(\mathcal{X}, \mathcal{X}) \).

(ii) This follows from the preceding paragraph.

(iii) This is obvious.

(iv) \( \forall x, y \in \mathcal{X}, \langle x, Py \rangle = \langle P^* x, y \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x_1, y_1 \rangle = \langle x_1, y_1 + y_2 \rangle = \langle Px, y \rangle \), where the first equality follows from Proposition 13.22; \( x = x_1 + x_2, y = y_1 + y_2, x_1, y_1 \in M \), and \( x_2, y_2 \in M^\perp \). By the arbitrariness of \( x \) and \( y \), we have \( P^* = P \).

(v) Let \( A \in \mathfrak{B}(\mathcal{X}, \mathcal{X}) \) satisfies (iii) and (iv). Let \( M := \mathcal{R}(A) \). Then, \( M^\perp = \mathcal{N}(A^* \mathfrak{N}(A)) \), by Proposition 13.22 and (iv). By Theorem 13.19, \( \mathcal{X} = M \oplus (M^\perp)^\perp \). By Proposition 13.17, \( (M^\perp)^\perp = \mathcal{N}(A) \). Let \( P : \mathcal{X} \to
$\mathcal{R}(A)$ be the projection operator. We will show that $P = A$. \forall x, y \in X,$ we have

$$\langle Ax, y - Ay \rangle = \langle Ax, y \rangle - \langle Ax, Ay \rangle = \langle Ax, y \rangle - \langle A^*Ax, y \rangle = \langle Ax, y \rangle - \langle Ax, y \rangle = 0$$

where the second equality follows from Proposition 13.22; the third equality follows from (iv); and the fourth equality follows from (iii). Thus, $Ax \perp y - Ay.$ By the arbitrariness of $x$, we have $y - Ay \in M^\perp$. Note that $Ay \in \mathcal{R}(A).$ Then, we have $Ay = Py.$ By the arbitrariness of $y$, we have $A = P.$

This completes the proof of the proposition. \hfill \Box

**Proposition 13.44** Let $X$ and $Y$ be Hilbert spaces over $\mathbb{K}$, $A \in B(X, Y)$, and $y \in Y$. Assume that $\mathcal{R}(A) \subseteq Y$ is closed. Then, there exists a unique vector $x_0 \in X$ such that $x_0 = \arg \min_{x \in X} \|Ax - y\|_Y$, where $x_0$ satisfies $\|Ax_1 - y\|_Y = \min_{x \in X} \|Ax - y\|_Y$. We will denote $x_0 = A^\dagger y$, where $A^\dagger : Y \to X$ is called the pseudoinverse of $A$. Furthermore, we have the following properties for $A^\dagger$.

(i) $A^\dagger \in B(Y, X)$.

(ii) $(A^\dagger)^\dagger = A$.

(iii) $(A^*)^\dagger = (A^\dagger)^*$.

(iv) $A^\dagger A A^\dagger = A^\dagger$.

(v) $AA^\dagger A = A$.

(vi) $(A^\dagger A)^* = A^\dagger A$.

(vii) $(AA^\dagger)^* = AA^\dagger$.

(viii) $A^\dagger = (A^*A)^\dagger A^*$.

(ix) $A^\dagger = A^* (AA^*)^{-1}$.

(x) $A^\dagger = \lim_{\epsilon \to 0^+} A^* (\epsilon \text{id}_Y + AA^*)^{-1} = \lim_{\epsilon \to 0^+} (\epsilon \text{id}_X + A^*A)^{-1} A^*$.

(xi) $A$ can be uniquely expressed as $A = P_1 A_1 P_2$, where $P_1 : Y \to \mathcal{R}(A)$ and $P_2 : X \to (\mathcal{N}(A))^\perp$ are projection operators, and $A_1 \in B((\mathcal{N}(A))^\perp, \mathcal{R}(A))$ is bijective with $A_1^{-1} \in B(\mathcal{R}(A), (\mathcal{N}(A))^\perp)$.

(xii) $AA^\dagger = P_1$ and $A^\dagger A = P_2$.

**Proof** Since $\mathcal{R}(A)$ is closed. Then, by the Classical Projection Theorem 13.13, $\exists y_0 \in \mathcal{R}(A)$ such that $\|y - y_0\|_Y = \min_{y \in \mathcal{R}(A)} \|y - y_0\|_Y$, furthermore, $y_0 \in \mathcal{R}(A)$ is defined to be the unique vector such that
13.8. PSEUDOINVERSE OPERATOR

\( y - y_0 \in (R(A))^\perp \). Then, the set of \( x_1 \)'s defined in the proposition statement is given by \( N(A) + x_1 \), where \( x_1 \) is any vector such that \( Ax_1 = y_0 \). Then, by the Classical Projection Theorem 13.13 and Proposition 13.22, \( \exists! x_0 \in (N(A))^\perp = R(A^*) \) such that \( \| x_0 \|_X = \min_{x \in N(A)} \| x_1 - x \|_X \). This proves the existence and uniqueness of \( x_0 \in X \) for any given \( y \in Y \).

Thus, \( A^† \) is well-defined.

(i) By Theorem 13.19, \( X = N(A) \oplus (N(A))^\perp =: M \oplus M^⊥ \). Define the projection mapping \( P_2 : X \to M^⊥ \). \( \forall x \in X \), there is unique pair \( x_1 \in M \) and \( x_2 \in M^⊥ \) such that \( x = x_1 + x_2 = x_1 + P_2x \). Then, \( Ax = AP_2x \). This allows us to define a mapping \( A_r : M^⊥ \to N := R(A) \) by \( A_r(x_2) = A(x_2) \), \( \forall x_2 \in M^⊥ \). Clearly, \( A_r \in B(M^⊥, N) \) and \( A_r \) is surjective. It is easy to show that \( N(A_r) = \{ \theta \}_{M^⊥} \) since the domain of \( A_r \) is \( M^⊥ \). Then, \( A_r \) is bijective. Note that \( N \) is a closed subspace of \( Y \) by assumption. Then, by Proposition 4.39, \( N \) is a Banach subspace of \( Y \). By Proposition 13.17, \( M^⊥ \) is a closed subspace of \( X \). Then, by Proposition 4.39, \( M^⊥ \) is a Banach subspace of \( X \). Then, \( A_{r}^{-1} \in B(N, M^⊥) \) by Open Mapping Theorem 7.103.

Then, \( A = A_rP_2 \). By Theorem 13.19, \( Y = N \oplus N^⊥ \). Define \( P_1 : Y \to N \) to be the projection mapping. Clearly, \( P_1 \in B(Y, N) \). \( \forall y \in Y, y_0 = P_1y \). Let \( x_1 := A_{r}^{-1}y_0 \in M^⊥ \subseteq X \). Then, \( Ax_1 = A_rP_2A_{r}^{-1}y_0 = A_rA_{r}^{-1}y_0 = y_0 \).

Then, \( x_0 = P_2x_1 = x_1 = A_{r}^{-1}P_1y \). Hence, \( A^† = A_{r}^{-1}P_1 = P_2A_{r}^{-1}P_1 \in B(Y, X) \).

(xi) This follows immediately from the preceding discussion in (i).

(ii) For \( A \in B(X, Y) \), \( A_r \in B(M^⊥, N) \), \( A = A_rP_2 = P_1A_rP_2 \), and \( A^† = A_{r}^{-1}P_1 = P_2A_{r}^{-1}P_1 \). Then, \( (A^†)^† = P_1A_rP_2 = A \).

(iii) \( (A^†)^* = (A_{r}^{-1}P_1)^* = (P_2A_{r}^{-1}P_1)^* = P_1^*(A_{r}^{-1})^*P_2^* = P_1A_{r}^{-1}P_2 \), where the first two equalities follow from (xi); the third equality follows from Proposition 13.22; and the last equality follows from Proposition 13.43. Now, \( N(A^*)^⊥ = R(A) = N \) and \( R(A^*) = N(A)^⊥ = M^⊥ \), by Proposition 13.22. Therefore, \( (A^*)^† = (P_2A_{r}^*P_2)^† = (P_2A_{r}^*P_2)^† = P_1(A_{r}^{-1})^{-1}P_2 = P_1A_{r}^{-1}P_2 \), where the first equality follows from (xi); the second equality follows from Proposition 13.43; the third equality follow from (xi); and the last equality follows from Proposition 13.22. Hence, \( (A^†)^* = (A^*)^† \).

(xii) By (xi), we have \( A^†A = P_2A_{r}^{-1}P_1P_2 = P_2A_{r}^{-1}A_rP_2 = P_2 \), where the second and third equalities follow from the fact that \( P_1 \) and \( P_2 \) are projection operators. Similarly, \( AA^† = P_1A_rP_2A_{r}^{-1}P_1 = P_1A_rA_{r}^{-1}P_1 = P_1 \), where the second and third equalities follow from the fact that \( P_1 \) and \( P_2 \) are projection operators.

(iv) By (xi) and (xii), we have \( A^†AA^† = P_2A_{r}^{-1}P_1P_2 = P_2A_{r}^{-1}P_1A_rP_2 = P_2 \), where the second equality follows from Proposition 13.43.

(v) By (xi) and (xiii), we have \( AA^†A = P_1P_2 = P_1A_rP_2 = A \), where the second equality follows from Proposition 13.43.

(vi) By (xi) and (xii), we have \( (A^†)^* = P_2 = P_2 = A^† \), where the second equality follows from Proposition 13.43.

(vii) By (xi) and (xii), we have \( (AA^†)^* = P_1^* = P_1 = AA^† \), where the second equality follows from Proposition 13.43.
(viii) By (xi), we have \( A^*A = (P_1A_rP_2)^*P_1A_rP_2 = P_2A_2^*A_1P_1A_rP_2 \), where the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.43 and the fact \( P_1 \) is a projection operator. Then, we have \( (A^*)^1A^* = P_2(A_2^*A_r)^{-1}P_2(P_1A_rP_2)^* = P_2A_2^{-1}A_r^{-*}P_2 A_2^*P_1 = P_2A_2^{-1}A_r^{-*}A_1P_1 = P_2A_2^{-1}A_1P_1 = A_r^{-1}P_1 = A_1^* \), where the first equality follows from (xi) applied to the operator \( A_r^*A_r \); the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.43; the fourth equality follows from Proposition 13.43; the fifth equality follows from (xi). The proof of (viii); the second equality follows from Proposition 13.43. Then, \( P_2 \) is a projection operator; the first equality follows from Theorem 13.19 and the third equality follows from Proposition 13.43 and the fact \( A_r \) is invertible, and then \( P_2 \) is bijective, and then \( P_2P_2^* = P_2 \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), where the equality follows from the fact \( P_2 \) is a projection operator; and the last equality follows from (xi).

(ix) By (xi), we have \( AA^* = P_1A_rP_2(P_1A_rP_2)^* = P_1A_rP_2^*A_2^*P_1 = P_1A_rA_1^*P_1 \), where the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.43 and the fact \( P_2 \) is a projection operator. Then, we have \( A^*(AA^*)^1 = (P_1A_rP_2)^*P_1(A_rA_2^*)^{-1}P_1 = P_2A_2^*P_1^*P_1A_r^{-*}A_r^{-1}P_1 = P_2A_2^*P_1^*P_1A_r^{-*}A_1^{-1}P_1 = P_2A_2^*A_r^*A_1^{-1}P_1 = P_2A_2^{-1}A_r^{-1}P_1 = A_1^* \), where the first equality follows from (xi) applied to the operator \( A_r^*A_r \); the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.43; the fourth equality follows from Proposition 13.43; the fifth equality follows from (xi). The proof of (viii); the second equality follows from Proposition 13.43. Then, \( P_2 \) is a projection operator; the first equality follows from Theorem 13.19 and the proof of (viii); the second equality follows from Proposition 13.43. Then, \( \epsilon \text{id}_X + A^*A)^{-1} = P_2 \epsilon \text{id}_M P_2 + P_2(\epsilon \text{id}_{M^*} + A_r^*A_r)^{-1}P_2, \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), where the equality follows from the fact \( P_2 \) is a projection operator; and the last equality follows from (xi).

(x) Note that \( \epsilon \text{id}_X + A^*A = \epsilon(\text{id}_X + P_2) + P_2A_2^*A_rP_2 = \epsilon P_2 \text{id}_M P_2 + P_2(\epsilon \text{id}_{M^*} + A_r^*A_r)P_2, \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), where \( P_2 : X \to M \) is the projection operator; the first equality follows from Proposition 13.43 and the proof of (viii); the second equality follows from Proposition 13.43. Then, \( \epsilon \text{id}_X + A^*A)^{-1} = P_2 \epsilon \text{id}_M P_2 + P_2(\epsilon \text{id}_{M^*} + A_r^*A_r)^{-1}P_2, \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), where the equality follows from the fact \( P_2 \) is a projection operator; and the last equality follows from (xi).

This completes the proof of the proposition. \( \square \)

**Proposition 13.45** Let \( X \) and \( Y \) be Hilbert spaces over \( \mathbb{K} \), and \( B \in B(X, Y) \). If \( B \) is surjective, then \( BB^* \in S_+ \). If \( B \) is injective and \( \mathcal{R}(B^*) \subseteq X \) is closed, then \( BB^* \in S_+ \).

**Proof** If \( B \) is surjective, then \( \mathcal{R}(B) = Y \) and it is a closed set in \( Y \). By Proposition 13.44, we have \( B = P_1B_rP_2 \), where \( P_1 : Y \to \mathcal{R}(B) \) and \( P_2 : X \to (\mathcal{N}(B))^\perp \) are projection operators, and \( B_r \in B((\mathcal{N}(B))^\perp, \mathcal{R}(B)) \) is bijective. Clearly, \( P_1 = \text{id}_Y \). Then, \( BB^* = B_rB_r^* \) is bijective, and then \( (BB^*)^{-1} \in B(Y, Y) \) by Open Mapping Theorem 7.103. Fix any \( \forall y \in Y \). Let \( \tilde{y} := (BB^*)^{-1}y \). This yields \( \langle y, (BB^*)^{-1}y \rangle_y = \langle BB^*\tilde{y}, \tilde{y} \rangle_Y = \langle B^*\tilde{y}, B^*\tilde{y} \rangle_X = \|B^*\tilde{y}\|^2_X \). Define \( \bar{y} := B^*\tilde{y} = B^*(BB^*)^{-1}y = B^1y \in X \), where the second equality
follows from Proposition 13.44. This leads to \( B\bar{x} = BB^\dagger y = P_1y = y \), where the second equality follows from the proof for (vi) of Proposition 13.44. This yields \( \|y\|_Y \leq \|B\|_{B(X,Y)} \|\bar{x}\|_X \). This implies that

\[
\langle y, (BB^*)^{-1}y \rangle_Y = \|\bar{x}\|_X^2 \geq \frac{\|B\|_{B(X,Y)}^2}{\epsilon + \|B\|_{B(X,Y)}^2} \|y\|_Y^2 \geq \frac{1}{\epsilon + \|B\|_{B(X,Y)}^2} \|y\|_Y^2,
\]

where the first equality follows from the preceding discussion; the first inequality holds for any \( \epsilon \in (0, \infty) \subset \mathbb{R} \); and the the last inequality follows the preceding discussion. This yields that \( (BB^*)^{-1} \in S_+y \) (clearly, \( (BB^*)^{-1} \in S_9 \)). By (v) of Proposition 13.42, \( BB^* \in S_+x \).

If \( B \) is injective and \( \mathcal{R}(B^*) \subseteq X \) is closed, then \( \mathcal{N}(B) = \{ \varnothing_X \} \). By Proposition 13.22, we have \( \mathcal{R}(B^*) = (\mathcal{N}(B))^\perp = X \). Hence, \( B^* \) is surjective. By the preceding paragraph, we have \( B^*(B^*)^* = B^*B \in S_+X \), where the equality follows from Proposition 13.22.

This completes the proof of the proposition. \( \Box \)

### 13.9 Spectral Theory of Linear Operators

**Definition 13.46** Let \( X \) be a normed linear space over \( K \), \( A \in B(X,X) \), \( \lambda \in K \), and \( x_0 \in X \) with \( x_0 \neq \varnothing_X \). If \( \lambda x_0 = Ax_0 \), or equivalently \( (\lambda \text{id}_X - A)x_0 = \varnothing_X \), we will say that \( \lambda \) is an eigenvalue of \( A \), \( x_0 \) is an eigenvector of \( A \) associated with the eigenvalue \( \lambda \). If \( (\lambda \text{id}_X - A)^k x_0 \neq \varnothing_X \) and \( (\lambda \text{id}_X - A)^k x_0 = \varnothing_X \), for some \( k \in \mathbb{N} \) with \( k \geq 2 \), we will say that \( x_0 \) is a generalized eigenvector of grade \( k \) of \( A \) associated with the eigenvalue \( \lambda \).

**Proposition 13.47** Let \( X \) be a Hilbert space over \( K \), \( Q \in B(X,X) \) be Hermitian, and \( A \in B(X,X) \). Then, the following statements hold.

(i) If \( \lambda \in K \) is an eigenvalue of \( Q \), then \( \lambda \in \mathbb{R} \);

(ii) If \( S_\gamma \subseteq X \) be an linearly independent set of eigenvectors of \( Q \) associated with eigenvalue \( \lambda_\gamma \in \mathbb{R} \), \( \forall \gamma \in \Gamma \), and \( (\lambda_\gamma)_\gamma \in \Gamma \) is pairwise distinct, then the set \( S := \bigcup_{\gamma \in \Gamma} S_\gamma \) is a linearly independent set;

(iii) If \( x_i \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda_i \in K \), \( i = 1,2 \), and \( \lambda_1 \neq \lambda_2 \), then \( \langle x_1, x_2 \rangle = 0 \) and \( x_1 \perp x_2 \).

(iv) \( Q \) has no generalized eigenvectors of grade \( k \geq 2 \) associated with any eigenvalue of \( Q \), (i.e., \( Q \) has only eigenvectors but not generalized eigenvectors);

(v) If \( \mathcal{R}(Q) \subseteq X \) is closed, then \( 0 \in K \) is not an eigenvalue of \( Q \) if, and only if, \( Q \) is bijective and \( Q^{-1} \in B(X,X) \);

(vi) If \( Q \in S_+X \) then \( \exists m \in (0, \infty) \subset \mathbb{R} \) such that \( \forall \lambda \in \mathbb{R} \) with \( \lambda \) being an eigenvalue of \( Q \), we have \( \lambda \leq m \);

(vii) If \( \lambda \in K \) is an eigenvalue of \( A \), then \( |\lambda| \leq \|A\|_{B(X,X)} \|\text{id}_X\|_{B(X,X)} \leq \|A\|_{B(X,X)} \).
(viii) If $R(Q) \subseteq X$ is closed, then $Q \in S_{psd,X}$ implies that all of the eigenvalues of $Q$ are nonnegative.

**Proof**

(i) Let $x_0 \in X$ be the eigenvector of $Q$ associated with the eigenvalue $\lambda$. Then, $\lambda x_0 = Qx_0$. This implies that $\langle \lambda x_0, x_0 \rangle = \langle x_0, Qx_0 \rangle = \langle Q^*x_0, x_0 \rangle = \langle Qx_0, x_0 \rangle = \langle \lambda x_0, x_0 \rangle = \lambda \|x_0\|^2$, where the first equality follows from Proposition 13.2; the second equality follows from Definition 13.1; the third equality follows from the fact that $x_0$ is an eigenvector of $Q$ associated with the eigenvalue $\lambda$; the fourth equality follows from Proposition 13.22; the fifth equality follows from the assumption that $Q$ is Hermitian; the sixth equality follows from the fact that $x_0$ is an eigenvector of $Q$ associated with the eigenvalue $\lambda$; and the seventh equality follows from Definition 13.1 and Proposition 13.2. Note that $x_0 \neq \emptyset_X$ since it is an eigenvector, then $\|x_0\|^2 > 0$. Thus, the above implies that $\lambda = \lambda$. Hence, $\lambda \in \mathbb{R}$.

(ii) We will use mathematical induction on $n$ to show that there doesn't exist an $n \in \mathbb{Z}_+$ such that $\exists x_0 \in S$ and $\exists x_1, \ldots, x_n \in S \setminus \{x_0\}$, which are distinct, and $\exists \alpha_1, \ldots, \alpha_n \in K \setminus \{0\}$ such that $x_0 = \sum_{i=1}^{n} \alpha_i x_i$.

1° $n = 0$. Suppose that the result doesn't hold, then $\exists x_0 \in S$ such that $x_0 = \emptyset_X$. Note that $x_0 \in S_{\gamma_0}$, for some $\gamma_0 \in \Gamma$. Then, $x_0$ is an eigenvector of $Q$ associated with eigenvalue $\lambda_{\gamma_0}$. By Definition 13.46, $x_0 \neq \emptyset_X$. This is a contradiction. This case is proved.

2° Assume the claim holds for $n \leq k \in \mathbb{Z}_+$.

3° Consider the case $n = k + 1 \in \mathbb{N}$. Suppose that the result doesn’t hold. Then, $\exists x_0 \in S$ and $\exists x_1, \ldots, x_n \in S \setminus \{x_0\}$, which are distinct, and $\exists \alpha_1, \ldots, \alpha_n \in K \setminus \{0\}$ such that $x_0 = \sum_{i=1}^{n} \alpha_i x_i$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $x_0, x_i \in S_{\gamma_i}$ for some $\gamma_i \in \Lambda$; Case 2: $x_0 \in S_{\gamma_0}$ with $\gamma_0 \in \Gamma$, and $\exists \alpha_i \in \{1, \ldots, n\}$ such that $x_i \in S_{\gamma_i}$ with $\gamma_i \in \Gamma$ and $\gamma_i \neq \gamma_0$.

Case 1: $x_0, x_1, \ldots, x_n \in S_{\gamma_0}$ for some $\gamma_0 \in \Gamma$. Then, $x_0, x_1, \ldots, x_n$ are distinct vectors in $S_{\gamma_0}$. By assumption, $S_{\gamma_0}$ is a linearly independent set. Then, $x_0 = \sum_{i=1}^{n} \alpha_i x_i$ contradicts with the fact that $S_{\gamma_0}$ being a linearly independent set.

Case 2: $x_0 \in S_{\gamma_0}$ with $\gamma_0 \in \Gamma$, and $\exists \alpha_i \in \{1, \ldots, n\}$ such that $x_{i_1} \in S_{\gamma_{i_1}}$ with $\gamma_{i_1} \in \Gamma$ and $\gamma_{i_1} \neq \gamma_0$. Without loss of generality, assume that $x_i \in S_{\gamma_i}$, $i = 1, \ldots, n$, where $\gamma_i \in \Gamma$. Then, $\emptyset_X = (\lambda_{\gamma_0} \text{id}_X - Q)x_0 = \sum_{i=1}^{n} (\lambda_{\gamma_0} \text{id}_X - Q)(\alpha_i x_i) = \sum_{i=1}^{n} \alpha_i (\lambda_{\gamma_0} - \lambda_{\gamma_i}) x_i$. By the preceding discussion and (i), $0 \neq \lambda_{\gamma_0} - \lambda_{\gamma_i} \in \mathbb{R}$. Then, we have $\emptyset_X = \sum_{i=1}^{n} \alpha_i (\lambda_{\gamma_0} - \lambda_{\gamma_i}) x_i$. Rearranging the equation, we have $x_{i_1} = \sum_{i=1, i \neq i_1}^{n} \alpha_i x_i - \frac{(\lambda_{\gamma_0} - \lambda_{\gamma_i}) \alpha_i}{(\lambda_{\gamma_0} - \lambda_{\gamma_i}) \alpha_i} x_i$. This contradicts with the inductive assumption. Hence, the result holds in this case as well.

This completes the inductive process. Hence, $S$ is a linearly independent set.
(iii) By (i), \( \lambda_1, \lambda_2 \in \mathbb{R} \). We have \( \lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = (Qx_1, x_2) = \langle x_1, Q^* x_2 \rangle = \langle x_1, Qx_2 \rangle = \langle x_1, \lambda_2 x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle \), where the first equality follows from Definition 13.1; the second equality follows from the fact that \( x_1 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda_1 \); the third equality follows from Proposition 13.22; the fourth equality follows from the fact that \( Q \) is Hermitian; the fifth equality follows from the fact that \( x_2 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda_2 \); the six equality follows from Definition 13.1; and the last equality follows from the fact \( \lambda_2 \in \mathbb{R} \). Since \( \lambda_1 \neq \lambda_2 \), then \( \langle x_1, x_2 \rangle = 0 \) and \( x_1 \perp x_2 \).

(iv) Suppose the result doesn’t hold. Let \( x_0 \in \mathcal{X} \) be a generalized eigenvector of grade \( k \geq 2 \) of \( Q \) associated with eigenvalue \( \lambda \in \mathbb{R} \). Then, \( B^{k-1}x_0 \neq \vartheta_X \) and \( B^kx_0 = \vartheta_X \), where \( B := \lambda \text{id}_X - Q \). Let \( x_{k-2} := B^{k-2}x_0 \) and \( x_{k-1} := B^{k-1}x_0 \). Then, \( B^2x_{k-2} = \vartheta_X \). This implies that \( 0 = \langle x_{k-2}, B^2x_{k-2} \rangle = \langle B^2x_{k-2}, Bx_{k-2} \rangle = \langle Bx_{k-2}, Bx_{k-2} \rangle = \langle x_{k-1}, x_{k-1} \rangle = \| x_{k-1} \|^2 \), where the second equality follows from Proposition 13.22; the third equality follows from the assumption that \( Q \) is Hermitian; and the fifth equality follows from Proposition 13.2. Then, \( x_{k-1} = \vartheta_X \). This contradicts with \( B^{k-1}x_0 \neq \vartheta_X \). Hence, the result holds.

(v) “Necessity” Since \( 0 \) is not an eigenvalue of \( Q \), then \( \mathcal{N}(Q) = \{ \vartheta_X \} \). By Proposition 13.22, \( \mathcal{R}(Q) = (\mathcal{N}(Q^*))^\perp = (\mathcal{N}(Q))^\perp = \mathcal{X} \). Then, \( Q \) is bijective and invertible. By Open Mapping Theorem 7.103, \( Q^{-1} \in \mathcal{B}(\mathcal{X}, \mathcal{X}) \).

“Sufficiency” Since \( Q \) is bijective, then \( \mathcal{N}(Q) = \{ \vartheta_X \} \). Then, \( 0 \in \mathcal{K} \) is not an eigenvalue of \( Q \), since there is no eigenvector associated with \( 0 \).

(vi) By Definition 13.40, \( \exists m \in (0, \infty) \subset \mathbb{R} \) such that \( \langle x, Qx \rangle \geq m \| x \|^2 \), \( \forall x \in \mathcal{X} \). Let \( x_0 \in \mathcal{X} \) be an eigenvector of \( Q \) associated with the eigenvalue \( \lambda \in \mathbb{R} \). Then, we have \( \lambda \| x_0 \|^2 = \langle x_0, \lambda x_0 \rangle = \langle x_0, Qx_0 \rangle \geq m \| x_0 \|^2 \). Since \( x_0 \) is an eigenvector, then \( x_0 \neq \vartheta_X \) and \( \| x_0 \| > 0 \). Then, we have \( \lambda \geq m \).

(vii) Fix any \( \lambda \in \mathcal{K} \) with \( |\lambda| > \| A \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} \| \text{id}_X \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} \geq 0 \). Then, \( \| \text{id}_X - (\lambda \text{id}_X - A) \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} \| \text{id}_X^{-1} \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} = \frac{1}{|\lambda|} \| A \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} \| \text{id}_X \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} < 1 \). By Proposition 9.55, \( \text{id}_X - \frac{1}{\lambda}A \) is bijective and admits continuous inverse. Then, \( \lambda \text{id}_X - A \) is bijective and admits continuous inverse since \( \lambda \neq 0 \). This implies that \( \mathcal{N}(\lambda \text{id}_X - A) = \{ \vartheta_X \} \). Hence, \( \lambda \) is not an eigenvalue of \( A \). Hence, the result holds noting \( \| \text{id}_X \|_{\mathcal{B}(\mathcal{X}, \mathcal{X})} \leq 1 \).

(viii) Since \( Q \in \mathcal{S}_{\text{res}, \mathcal{X}} \), then \( \langle x, Qx \rangle \geq 0 \), \( \forall x \in \mathcal{X} \). By (i), all eigenvalues of \( Q \) are real. Suppose that the result doesn’t hold. Then, \( \exists \lambda \in (-\infty, 0) \subset \mathbb{R} \) and \( \exists x_0 \in \mathcal{X} \) such that \( x_0 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda \). Then, \( 0 \leq \langle x_0, Qx_0 \rangle = \langle x_0, \lambda x_0 \rangle = \lambda \| x_0 \|^2 < 0 \), where the first equality follows from the hypothesis; and the second equality follows from Definition 13.1 and Proposition 13.2; and the last inequality follows from \( x_0 \neq \vartheta_X \) since it is an eigenvector. This is a contradiction. Hence, all eigenvalues of \( Q \) are nonnegative.

This completes the proof of the proposition. \( \square \)
Proposition 13.48 Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{K} \), \( A \in \text{B}(\mathcal{X}, \mathcal{X}) \) be a Hermitian operator. Then, \( \lambda := \sup_{x \in \mathcal{X}, \|x\| \leq 1} \langle Ax, x \rangle = \|A\|_{\text{B}(\mathcal{X}, \mathcal{X})} \).

Proof Note that \( |\langle Ax, x \rangle| \leq \|Ax\| \|x\| \leq \|A\|_{\text{B}(\mathcal{X}, \mathcal{X})} \|x\|^2 \), \( \forall x \in \mathcal{X} \), where the first inequality follows from Cauchy-Schwarz Inequality; and the second inequality follows from Proposition 7.64. Hence, \( \lambda \leq \|A\|_{\text{B}(\mathcal{X}, \mathcal{X})} \).

Fix any \( u \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1) \), let \( v_1 := \alpha u + \alpha^{-1} Au \) and \( v_2 := \alpha u - \alpha^{-1} Au \), where \( \alpha \in (0, \infty) \subset \mathbb{R} \). Note also that

\[
\|Au\|^2 = \langle Au, Au \rangle = \langle A^* Au, u \rangle = \langle A^2 u, u \rangle = \langle u, A^2 u \rangle = \frac{1}{4}(\langle Av_1, v_1 \rangle - \langle Av_2, v_2 \rangle) \leq \frac{1}{4}(\lambda \|v_1\|^2 + \lambda \|v_2\|^2) = \frac{1}{2}\lambda(\alpha^2 \|u\|^2 - \alpha^{-2} \|Au\|^2) \leq \frac{1}{2}\lambda(\alpha^2 - \alpha^{-2} \|Au\|^2)
\]

where the first equality follows from Proposition 13.2; the second equality follows from Proposition 13.22; the third equality follows from the fact \( A \) is Hermitian; the fourth equality follows from Proposition 13.22 and \( A \) being Hermitian; the fifth equality follows from straightforward algebra; the first inequality follows from the definition of \( \lambda \); the sixth equality follows from straightforward algebra; and the last inequality follows from \( u \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1) \).

If \( \|Au\| = 0 \), then \( \|Au\| \leq \lambda \). On the other hand, if \( \|Au\| > 0 \), let \( \alpha^2 := \|Au\| \), we have \( \|Au\|^2 \leq \frac{1}{2}\lambda^2 \|Au\| \), which leads to \( \|Au\| \leq \lambda \). In both cases, we have \( \|Au\| \leq \lambda \), \( \forall u \in \overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1) \). Hence, \( \|A\| \leq \lambda \).

This completes the proof of the proposition. \( \square \)

Definition 13.49 Let \( \mathcal{X} \) be a Banach space over \( \mathbb{K}_1 \) and \( \mathcal{Y} \) be a Banach space over \( \mathbb{K}_2 \), \( D \subseteq \mathcal{X} \), \( f : D \rightarrow \mathcal{Y} \) be continuous. \( f \) is said to be a compact operator if \( \forall \varrho \in (0, \infty) \subset \mathbb{R} \), \( f(\overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, \varrho)) \subseteq \mathcal{Y} \) is a compact set.

It is then clear that if \( f : \mathcal{X} \rightarrow \mathcal{Y} \) is a compact operator, then, \( \forall (x_i)_{i=1}^{\infty} \subseteq \mathcal{X} \), which is bounded, then \( (f(x_i))_{i=1}^{\infty} \subseteq \mathcal{Y} \) has a subsequence converges to some \( y_0 \in \mathcal{Y} \), by Borel-Lebesgue Theorem 5.37.

Proposition 13.50 Let \( \mathcal{X} \) be a Banach space over \( \mathbb{K} \) and \( \mathcal{Y} \) be a Banach space over \( \mathbb{K} \), \( A \in \text{B}(\mathcal{X}, \mathcal{Y}) \). \( A \) is a compact operator if, and only if, \( A(\overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)) \subseteq \mathcal{Y} \) is a compact set. We will denote the set of such \( A \)'s by \( \mathcal{K}(\mathcal{X}, \mathcal{Y}) := \{ A \in \text{B}(\mathcal{X}, \mathcal{Y}) \mid A \text{ is a compact operator} \} \).

Proof

“Necessity” is immediate.

“Sufficiency” Let \( \varTheta := A(\overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, 1)) \subseteq \mathcal{Y} \). Then, by \( A \) being a compact operator, we have \( \overline{\varTheta} \) is a compact set. \( \forall r \in (0, \infty) \subset \mathbb{R} \), \( A(\overline{\mathcal{B}}_{\mathcal{X}}(\vartheta_{\mathcal{X}}, r)) = r\overline{\varTheta} \subseteq \mathcal{Y} \) since \( A \) is a linear operator. Then, by Proposition 7.102, \( A(\overline{\mathcal{B}}_{\mathcal{X}}(x_0, r)) = r\overline{\varTheta} \) is sequentially compact since \( \varTheta \) is. By Borel-Lebesgue Theorem 5.37, \( A(\overline{\mathcal{B}}_{\mathcal{X}}(x_0, r)) \) is compact. Hence, \( A \) is a compact operator.

This completes the proof of the proposition. \( \square \)
Proposition 13.51 Let \( X, Y, \) and \( Z \) be Banach space over \( K. \) Then, the following statements holds.

(i) Let \( A \in K(X, Y). \) \( \forall C \in B(Z, X), \forall D \in B(Y, Z), \) we have \( AC \in K(Z, Y) \) and \( DA \in K(X, Z). \)

(ii) (Fredholm Alternative) Let \( A \in K(X, X). \) \( \forall \lambda \in K \) with \( \lambda \neq 0, \) we have \( E := (\lambda \text{id}_X - A)(F) \subseteq X \) is a closed set, \( \forall F \subseteq X \) that is a closed and bounded set. (The operator \( \lambda \text{id}_X - A \) is called the Fredholm alternative of \( A. \))

Proof

(i) Note that

\[
E_1 := (AC)(B_z(\vartheta_z, 1)) \subseteq A(\overline{B_X}(\vartheta_X, \|C\|_{B(Z, X)})) =: F_1 \subseteq Y
\]

where the set containment follows from the fact that \( C \in B(Z, X). \) By Definition 13.49, \( F_1 \) is a compact. Then, \( E_1 \) is a compact set by Proposition 5.5. Hence, \( AC \in K(Z, Y) \) by Proposition 13.50.

Note that \( G_1 := (DA)(\overline{B_X}(\vartheta_X, 1)) \subseteq D(A(\overline{B_X}(\vartheta_X, 1))) =: D(F_2) \subseteq Z. \) By Proposition 13.50, \( F_2 \) is compact. By Proposition 5.7, \( D(F_2) \) is compact. By Proposition 5.5, \( D(F_2) = D(F_2) \) and hence \( D(F_2) \) is compact. Then, \( G_1 \)

is compact by Proposition 5.5. Hence, \( DA \in K(X, Z) \) by Proposition 13.50.

(ii) Fix any \( y_0 \in \overline{E}. \) By Proposition 4.13, \( \exists (y_n)_{n=1}^{\infty} \subseteq E \) such that

\[
y_0 = \lim_{n \to \infty} y_n.
\]

Then, by the definition of \( E, \) there exists \( (x_n)_{n=1}^{\infty} \subseteq F \subseteq \overline{B_X}(\vartheta_X, m), \) for some \( m \in \mathbb{N} \) since \( F \) is a bounded set, such that

\[
y_n = (\lambda \text{id}_X - A)x_n = \lambda x_n - Ax_n, \ \forall n \in \mathbb{N}.
\]

Since \( A \) is a compact operator, by Definition 13.49, we have \( A(\overline{B_X}(\vartheta_X, m)) \subseteq X \) is a compact set. By Borel-Lebesgue Theorem 5.37, there exists a subsequence \( (x_{n_k})_{k=1}^{\infty} \) such that

\[
\lim_{k \to \infty} x_{n_k} = \bar{x} \in A(\overline{B_X}(\vartheta_X, m)).
\]

Thus, we have \( \lambda \lim_{k \to \infty} x_{n_k} = \lambda \lim_{k \to \infty} y_{n_k} + \lim_{k \to \infty} Ax_{n_k} = y_0 + \bar{x} \in X. \) Since \( \lambda \neq 0, \) we have \( \lim_{k \to \infty} Ax_{n_k} = \lambda^{-1}(y_0 + \bar{x}) =: x_0 \in F, \) where the last set element operation follows since \( F \) is a closed subset of \( X. \) Then, we have \( (\lambda \text{id}_X - A)(x_0) = \lim_{k \to \infty}(\lambda \text{id}_X - A)(x_{n_k}) = \lim_{k \to \infty}(\lambda x_{n_k} - Ax_{n_k}) = \lambda x_0 - \bar{x} = y_0. \) Hence, \( y_0 \in (\lambda \text{id}_X - A)(F) = E. \) By the arbitrariness of \( y_0, \) we have \( \overline{E} \subseteq E \subseteq \overline{E}. \)

Hence, \( E = \overline{E} \) and \( E \) is closed by Proposition 3.3. \( \square \)

Theorem 13.52 (Spectral Theory) Let \( X \) and \( Y \) be Hilbert space over \( K, A \in B(X, Y), \) and \( Q \in S_X. \) Then, the following results hold.

(i) Let \( A \in K(X, Y). \) Then, \( A \neq \vartheta_{B(X, Y)} \) if, and only if, \( \|A\|_{B(X, Y)} =: \sigma > 0. \) In this case, \( \exists x \in X \) with \( \|x\|_X = 1 \) such that \( \|Ax\|_Y = \sigma. \)

Let \( y := \frac{1}{\|x\|_Y} Ax \in Y, \) then \( \|y\|_Y = 1 \) and \( A = \sigma y x^* + A, \) where \( A \in K(X, Y) \) and \( \tilde{A} = P_1 AP_2, \) \( P_1 : Y \to \text{span}\{\{y\}\} =: N, \) \( P_2 : X \to \text{span}\{\{x\}\} =: M, \) \( P_3 = \text{id}_y - P_1 : Y \to N^+, \) and \( P_2 = \text{id}_X - P_2 : X \to M^+ \) are projection operators.
(ii) \( A \in K(X, Y) \) if, and only if, \( \exists \tilde{n} \in \mathbb{Z}^+ \cup \{ \infty \} \) and \( \exists \{x_i\}_{i=1}^{\tilde{n}} \subseteq X \), which is an orthonormal sequence, and \( \exists \{y_i\}_{i=1}^{\tilde{n}} \subseteq Y \), which is an orthonormal sequence, and \( \sigma =: \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_i \in (0, \infty) \subseteq \mathbb{R} \), \( \forall i \in \mathbb{N} \) with \( 1 \leq i < \tilde{n} + 1 \), such that \( A = \sum_{i=1}^{\tilde{n}} \sigma_i y_i x_i^* \), and in case \( \tilde{n} = \infty \), we have \( \lim_{i \in \mathbb{N}} \sigma_i = 0 \).

(iii) Let \( Q \in S_X \cap K(X, X) \). Then, \( Q \neq 0 \) if, and only if, \( Q \) admits an eigenvector \( x_0 \in X \) with \( \| x_0 \| = 1 \) associated with the eigenvalue \( \lambda \in \mathbb{R} \setminus \{ 0 \} \) with \( |\lambda| = \| Q \|_{(X, X)} \in (0, \infty) \subseteq \mathbb{R} \). Then, \( Q = \lambda x_0 x_0^* + \tilde{Q} \), where \( \tilde{Q} = \tilde{P}_1 \tilde{P}_1^* : X \to \text{span}(\{x_0\}) =: M \), and \( \tilde{P}_1 : X \to M^\perp \) are projection operators with \( \tilde{P}_1^* \tilde{P}_1 = \text{id}_X \).

(iv) \( Q \in S_X \cap K(X, X) \) if, and only if, \( \exists \in \mathbb{Z}^+ \cup \{ \infty \} \), \( Q \) has \( n \) eigenvectors \( x_i \in X \) with \( \| x_i \| = 1 \), each associated with an eigenvalue \( \lambda_i \in \mathbb{R} \setminus \{ 0 \} \), and \( \| Q \|_{(X, X)} = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_i| \in (0, \infty) \subseteq \mathbb{R} \), \( i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \), where \( \{x_i\}_{i=1}^n \subseteq X \) is an orthonormal sequence, and \( Q = \sum_{i=1}^n \lambda_i x_i x_i^* \); and in case \( n = \infty \), we have \( \lim_{i \in \mathbb{N}} |\lambda_i| = 0 \).

(v) Let \( Q \in S_X \cap K(X, X) \) and the results of (iv) hold. \( \forall \lambda \in \mathbb{K} \) with \( \lambda \neq 0 \) is an eigenvalue of \( Q \) if, and only if, \( \exists \{x_i\}_{i=1}^n \subseteq X \) with \( 1 \leq i < n + 1 \) such that \( \lambda x_i = \lambda x_i = \cdots = \lambda x_i = 0 \). In this case, \( \exists m \in \mathbb{N} \) with \( m < n + 1 \) such that \( \sin (N(\lambda \text{id}_X - A)) = m = \sin (\bigcup_{i=1}^m \sin ((\lambda \text{id}_X - A)^*) \), which is called the multiplicity of the eigenvalue \( \lambda \) for \( Q \), and there is exactly \( \exists \lambda \) such that \( \lambda = \lambda \).

(vi) Let \( A \in K(X, Y) \) and the results of (ii) hold. Then, the values \( \{\sigma_i\}_{i=1}^\tilde{n} \) are uniquely defined that is independent of the choice of \( \{y_i\}_{i=1}^\tilde{n} \) and \( \{x_i\}_{i=1}^\tilde{n} \). These \( \{\sigma_i\}_{i=1}^\tilde{n} \) is said to be the singular values of the operator \( A \). The formula \( A = \sum_{i=1}^\tilde{n} \sigma_i y_i x_i^* \) is said to be the singular value decomposition of the operator \( A \in K(X, Y) \).

In the above theorem, we have established that \( A \in K(X, Y) \) if, and only if, \( A \) admits the infinite-dimensional version of the singular value decomposition for matrices. \( \sigma_i \)s are the singular values, \( 1 \leq i < \tilde{n} + 1 \), \( \{y_i\}_{i=1}^\tilde{n} \) replaces the unitary matrix \( U \) and \( \{x_i\}_{i=1}^\tilde{n} \) replace the unitary matrix \( V \). For a Hermitian compact operator \( Q \), it can always be diagonalized by its eigenvectors of all of its nonzero eigenvalues. Its singular value decomposition is given in terms of its eigenvectors and eigenvalues, where \( y_i = \text{sgn}(\lambda_i)x_i \) and \( x_i \) is the eigenvector associated with eigenvalue \( \lambda_i \neq 0 \) with \( |\lambda_i| = 1 \) and \( \forall i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \).

**Proof**

(i) Let \( A \in K(X, Y) \). Note that \( \| A \|_{B(X, Y)} = \sup_{x \in X, \| x \| = 1} \| Ax \|_Y = \sigma > 0 \) if, and only if, \( A \neq 0 \). In case of \( \sigma > 0 \), then, \( \exists \{x_n\}_{n=1}^\infty \subseteq B_X (\partial_X, 1) \) such that \( \lim_{n \in \mathbb{N}} \| A x_n \|_Y = \sigma \). Then, \( \{A x_n\}_{n=1}^\infty \subseteq A(B_X (\partial_X, 1)) =: E \subseteq Y \). By \( A \in K(X, Y) \) and Definition 13.49, we have \( E \) is compact. Then, \( E \) is sequentially compact by Borel-Lebesgue Theorem 5.37. Thus, there exists a subsequence \( \{A x_{n_k}\}_{k=1}^\infty \) of \( \{A x_n\}_{n=1}^\infty \) such
Claim 13.52.1

Suppose $\langle \sigma \rangle \subseteq \mathbb{R}$ and $\lim_{k \to \infty} \alpha_k = \sigma > 0$. By Alaoglu Theorem 7.122 and Riesz-Fréchet Theorem 13.15, $\mathcal{B}_X (\mathcal{O}_X, 1) =: F \subseteq X$ is weakly compact (since $X$ is a Hilbert space). Then, $(\alpha_k)_{k=1}^{\infty} \subseteq F$ has the Bolzano-Weierstrass property by Propositions 5.22 and 5.26. Then, $\exists \tilde{z} \in F$ such that $(\alpha_k)_{k=1}^{\infty}$ has $\tilde{z}$ as a weak cluster point. By Proposition 7.123, $A \in \mathcal{B}(X, Y)$ implies that $A : X_{\text{weak}} \to Y_{\text{weak}}$ is continuous. Then, the sequence $(A\alpha_k)_{k=1}^{\infty}$ has a weak cluster point at $A\tilde{z} \in Y_{\text{weak}}$, by Proposition 3.66. By $\lim_{k \to \infty} A\alpha_k = \tilde{y}$, we have $\lim_{k \to \infty} A\alpha_k = \tilde{y}$ weakly. Then, by Proposition 7.116, $\tilde{y}$ weak is Hausdorff, which further implies, by Proposition 3.65, $\tilde{y} = A\tilde{z}$. Thus, we have $A\tilde{z} = \tilde{y}$.

Let $x := \tilde{z}$ and $y := \frac{1}{\sigma} A\tilde{z} = \frac{1}{\sigma} \tilde{y}$. Then, we have $\|x\|_X = 1 = \|y\|_Y$. Since $X$ is a Hilbert space, we have $x^* \in X^*$, $\|x^*\|_X = 1$, and $\langle x^*, x \rangle = \langle x, x \rangle_X = 1$. Let $A := A - \sigma x^*$. By Theorem 13.19, we have $P_1 + P_1 = \text{id}_Y$ and $P_2 + P_2 = \text{id}_X$. For $\forall \tilde{x} \in X$, we have $A\tilde{x} = A\tilde{x} - \sigma y \langle \tilde{x}, x \rangle_X = A(P_2\tilde{x} + P_2\tilde{x} - \sigma y \langle \tilde{x}, x \rangle_X) = A(P_2\tilde{x} - \sigma y \langle \tilde{x}, x \rangle_X) = A(P_2\tilde{x}, x)_X$, where the first equality follows from the definition of $A$; and the last two equalities follow from the preceding discussion. Note that $P_2\tilde{x} \in M = \text{span} \{x \}$, and then $P_2\tilde{x} = \tilde{\alpha} x$, for some $\tilde{\alpha} \in \mathbb{K}$. Then, $A\tilde{x} = \tilde{\alpha} Ax + A(P_2\tilde{x} - \sigma y \langle \tilde{x}, x \rangle_X) - \sigma y \langle P_2\tilde{x}, x \rangle_X = \tilde{\alpha} \sigma y + A(P_2\tilde{x} - \sigma y \langle \tilde{x}, x \rangle_X = A\tilde{P}_x\tilde{x}, x \rangle_X = A\tilde{x}_x$, $\forall \tilde{x} \in X$, where the second equality follows from the fact $\sigma y = A\tilde{x} = A\tilde{z}$; and the last equality follows from Propositions 13.22 and 13.43 and the fact that $P_2\tilde{x} = \tilde{\alpha} x$. By the arbitrariness of $\tilde{x}$, we have $A = \tilde{A} P_2$.

Note that $A = \sigma y x^* + \tilde{A} = \sigma y x^* + \tilde{A} P_2 + \tilde{P}_1 \tilde{A} P_2$, where the first equality follows from the definition of $A$; the second equality follows from the preceding paragraph; and the last equality follows from $P_1 + P_1 = \text{id}_Y$. We need the following result.

Claim 13.52.1 $P_1 \tilde{A} P_2 = \tilde{A}_{\text{B}(M^+, N)}$.

Proof of claim: We will prove this by an argument of contradiction. Suppose $P_1 \tilde{A} P_2 \neq \tilde{A}_{\text{B}(M^+, N)}$. Then, $\|P_1 \tilde{A} P_2 \|_{\text{B}(X, Y)} = \|P_1 \tilde{A} P_2 \|_{\text{B}(M^+, N)} =: \beta > 0$. Then, $\exists s \in M^+$ with $\|s\|_X = 1$ such that $\|P_1 \tilde{A} P_2 s\|_Y > \frac{\beta}{2}$.

Note that $P_1 \tilde{A} P_2 s \in N$ and hence $P_1 \tilde{A} P_2 t = ty$, for some $t \in \mathbb{K}$ with $t \neq 0$.

Choose $\alpha \in \mathbb{K}$, $\|A(x + \alpha s)\|_Y = \|\sigma y + \alpha \tilde{A} P_2 s\|_Y = \|(\sigma + \alpha \tilde{t}) s\|_Y$ and $\alpha = \frac{1}{\tilde{t}} \in (0, \infty) \subset \mathbb{R}$.

Then, $\|A(x + \alpha s)\|_Y^2 = (\sigma + |t|^2 \delta)^2 y^2 + \tilde{t}^2 \sigma |P_1 \tilde{A} P_2 s\|_Y^2 = (\sigma + |t|^2 \delta)^2 + |t|^2 \delta^2 \|P_1 \tilde{A} P_2 s\|_Y^2 \geq (\sigma + |t|^2 \delta)^2 = \sigma^2 + \frac{|t|^2 \delta^2}{\sigma^2} > \sigma^2 (1 + \frac{|t|^2 \delta^2}{\sigma^2}) > \sigma^2 (1 + \frac{|t|^2 \delta^2}{\sigma^2})^2 > \sigma^2 (1 + \frac{|t|^2 \delta^2}{\sigma^2})^2 > \sigma^2 (1 + \frac{|t|^2 \delta^2}{\sigma^2})^2$, where the second equality follows from $P_1 y = \tilde{y}$ and Proposition 13.2. Note that $\|x + \alpha s\|_X^2 = \|x + \tilde{t} \delta s\|_X^2 = \|x\|_X^2 + \|\tilde{t} \delta s\|_X^2 = \|x\|_X^2 + \|\tilde{t} \delta s\|_X^2 = 1 + |t|^2 \delta^2$. These implies that $\|A(x + \alpha s)\|_Y^2 > \sigma^2 \|x + \alpha s\|_X^2$. 

\( \alpha s \leq \| x \|^2 \) and \( \| A \|_{B(X,Y)} > \sigma \). This is a contradiction. Hence, the hypothesis is invalid. So, \( P_1 A P_2 = \emptyset_{B(M^+,N)} \). This completes the proof of the claim. \( \square \)

Then, we have \( A = \sigma y x^* + P_1 A P_2 \) and hence \( \bar{A} = \bar{P}_1 A \bar{P}_2 \). By Proposition 13.51, \( \bar{A} \in K(X,Y) \). Furthermore, \( \| \bar{A} \|_{B(X,Y)} \leq \| \bar{P}_1 \|_{B(Y,Y)} \| A \|_{B(X,Y)} \).

\( \cdot \| \bar{P}_2 \|_{B(X,X)} \leq \| A \|_{B(X,Y)} = \sigma \). Clearly, \( \bar{A} \in K(M^+,N^+) \). This proves (i).

(ii) “Necessity” Let \( A \in K(X,Y) \). Then, (i) holds. Repeat (i) for \( A \). For notational consistency, we let \( \sigma_1 = \sigma, N_0 := \{ \emptyset \}, N_1 := N, M_0 := \{ \emptyset \}, M_1 := M, x_1 := x, y_1 := y, P_{1,1} := P_1, P_{1,2} := P_2, P_{1,2} := P_2, A_1 := A, \) and \( A_2 := \bar{A} \). Recursively, assume that we have completed \( i \) steps and arrived at \( A_{i+1} = P_{i,1} A P_{i,2}, (x_j^i)_{j=1}^i \subseteq X \) is an orthonormal sequence, \( (y_j^i)_{j=1}^i \subseteq Y \) is an orthonormal sequence, \( P_{j,1} : N_{j-1}^i \to \text{span} \{ \{ y_j \} \} \) and \( P_{j,2} : M_{j-1}^i \to \text{span} \{ \{ x_j \} \} \) are projection operators, \( P_{j,1} : N_{j-1}^i \to N_j^i := (\text{span} \{ (y_1, \ldots, y_j) \})^\perp \) and \( P_{j,2} : M_{j-1}^i \to M_j^i := (\text{span} \{ (x_1, \ldots, x_j) \})^\perp \) are projection operators, \( P_{j,1} + P_{j,2} = \mathbb{1}_{N_j^i}, P_{j,2} + P_{j,2} = \mathbb{1}_{M_j^i}, \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_i \geq 0, \sigma_i = \| A_i \|_{B(X,Y)}, j = 1, \ldots, i, \) and \( A = \sum_{j=1}^i \sigma_j y_j x_j^i + A_{i+1} \). Consider \( \sigma_{i+1} := \| A_{i+1} \|_{B(X,Y)} \). If \( \sigma_{i+1} = 0 \), then, we have \( n = i, A_{i+1} = \emptyset_{B(X,Y)} \) and \( A = \sum_{j=1}^i \sigma_j y_j x_j^i \). The desired result holds. In case \( \sigma_{i+1} > 0 \), we apply (i) to \( A_{i+1} \). By (i), \( \exists \gamma_{i+1} \subseteq M_{\geq i}^i \) and \( y_{i+1} := \frac{1}{\sigma_{i+1}} A x_{i+1} \in N_{\geq i}^i \) with \( \| x_{i+1} \|_{\mathcal{X}} = 1 = \| y_{i+1} \|_{\mathcal{Y}} \) such that \( A_{i+1} = \sigma_{i+1} y_{i+1} x_{i+1}^i + A_{i+2} \), where \( A_{i+2} = \bar{P}_{i,1} x_{i+1} A_{i+2}, P_{i,1} : N_{\geq i}^i \to \text{span} \{ \{ y_{i+1} \} \} \) and \( P_{i,2} : M_{\geq i}^i \to \text{span} \{ \{ x_{i+1} \} \} \) are projection operators, 

\( P_{i,1} + P_{i,2} = \mathbb{1}_{N_{\geq i}^i}, P_{i,2} + P_{i,2} = \mathbb{1}_{M_{\geq i}^i} \) and \( P_{i,2} = \mathbb{1}_{M_{\geq i}}^i - P_{i,2} : M_{\geq i}^i \to M_{\geq i}^i := (\text{span} \{ (x_1, \ldots, x_{i+1}) \}) \) are projection operators. Clearly, \( (x_j^i)_{j=1}^i \) is orthonormal and so is \( (y_j^i)_{j=1}^i \). Clearly, \( \sigma = \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{i+1} > 0 \).

Recursively, either the above process stops in finite number of steps, say \( n \in \mathbb{N} \), with \( \sigma_{n+1} = 0 \), then, (ii) holds with a finite \( n \); or the above process continues indefinitely, and we have \( n = \infty \) and \( \sigma_i > 0, \forall i \in \mathbb{N} \). In the latter case, clearly \( (x_j^i)_{j=1}^i \) and \( (y_j^i)_{j=1}^i \) are orthonormal sequences. To complete the proof of (ii) for this case, we need the following result.

**Claim 13.52.2** \( \lim_{i \in \mathbb{N}} \sigma_i = 0 \).

**Proof of claim:** Clearly, \( \lim_{i \in \mathbb{N}} \sigma_i := \bar{\sigma} \in [0, \infty) \subset \mathbb{R} \). We will show the claim by an argument of contradiction. Suppose \( \lim_{i \in \mathbb{N}} \sigma_i = \bar{\sigma} > 0 \). Then, we have \( Ax_i = \sigma_i y_i, \forall i \in \mathbb{N} \), \( (x_i^i)_{i=1}^\infty \subseteq \mathcal{X} (\emptyset, 1) \), but \( (\sigma_i y_i)_{i=1}^\infty \subseteq \mathcal{E} \) doesn’t have any convergent subsequence since \( (y_i)_{i=1}^\infty \) is orthonormal. This contradicts with the assumption that \( A \) is a compact operator and \( E \) is sequentially compact. Hence, the hypothesis is invalid. Hence, we must have \( \bar{\sigma} = 0 = \lim_{i \in \mathbb{N}} \sigma_i \). This completes the proof of the claim. \( \square \)

Then, \( A_{n+1} = A - \sum_{i=1}^n \sigma_i y_i x_i^i, \forall n \in \mathbb{N} \). We have \( \lim_{n \in \mathbb{N}} \| A_{n+1} \|_{B(X,Y)} = \lim_{n \in \mathbb{N}} \sigma_{n+1} = 0 \). This implies that \( A = \sum_{i=1}^\infty \sigma_i y_i x_i^i \). This completes
the necessity part of the proof.

“Sufficiency” Let there \( \exists \bar{n} \in \mathbb{Z}_+ \cup \{ \infty \} \) and \( \exists (x_i)_{i=1}^{\bar{n}} \subseteq X \), which is an orthonormal sequence, and \( \exists (y_i)_{i=1}^{\bar{n}} \subseteq Y \), which is an orthonormal sequence, and \( \sigma =: \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\bar{n}} \in (0, \infty) \subseteq \mathbb{R}, \forall i \in \mathbb{N} \) with \( 1 \leq i < \bar{n} + 1 \), such that \( A = \sum_{i=1}^{\bar{n}} \sigma_i y_i x_i^* \), and in case \( \bar{n} = \infty \), we have \( \lim_{i \in \mathbb{N}} \sigma_i = 0 \). Fix any \( (\bar{x}_i)_{i=1}^{\bar{n}} \subseteq F \), we will show that there exists a subsequence \( (A\bar{x}_{i_k})_{k=1}^{\infty} \subseteq E \) that converges to some \( y_0 \in Y \), or equivalently, that the subsequence \( (A\bar{x}_{i_k})_{k=1}^{\infty} \subseteq E \) is a Cauchy sequence. This shows that \( E \) is sequentially compact. By Proposition 13.50, \( A \in K(X, Y) \). This then completes the sufficiency part of the proof. We need the following intermediate result.

Claim 13.52.3 There exists a subsequence \( (\bar{x}_{i_k})_{k=1}^{\infty} \) of \( (\bar{x}_i)_{i=1}^{\bar{n}} \) such that \( \lim_{k \in \mathbb{N}} \langle \bar{x}_{i_k}, x_s \rangle_X = \alpha_s \in F, \forall s \in \mathbb{N} \) with \( 1 \leq s < \bar{n} + 1 \).

Proof of claim: Note that \( ((\bar{x}_i, x_1)_X)_{i=1}^{\infty} \subseteq F \) and \( F \subseteq K \) is compact. Then, there exists a subsequence \( (\bar{x}_{i_k})_{k=1}^{\infty} \) such that \( \lim_{k \in \mathbb{N}} \langle \bar{x}_{i_k}, x_1 \rangle_X = \alpha_1 \in F \). Note that \( ((\bar{x}_{i_k}, x_2)_X)_{k=1}^{\infty} \subseteq F \). Then, there exists a sub-subsequence \( (\bar{x}_{i_{k_j}})_{j=1}^{\infty} \) such that \( \lim_{j \in \mathbb{N}} \langle \bar{x}_{i_{k_j}}, x_2 \rangle_X = \alpha_2 \in F \). Clearly, \( \lim_{j \in \mathbb{N}} \langle \bar{x}_{i_{k_j}}, x_1 \rangle_X = \alpha_1 \in F \). Note that \( ((\bar{x}_{i_{k_j}}, x_3)_X)_{j=1}^{\infty} \subseteq F \). Then, there exists a sub-sub-subsequence \( (\bar{x}_{i_{k_{j_l}}} x_3)_X = \alpha_3 \in F \). Clearly, \( \lim_{k \in \mathbb{N}} \langle \bar{x}_{i_{k_{j_l}}}, x_s \rangle_X = \alpha_s \in F, s = 1, 2, 3 \). Recursively, continue this process till \( \bar{n} \).

If \( \bar{n} \) is finite, we may obtain (after finite recursions) a subsequence \( (\bar{x}_{i_{k_{j_l}}})_{k=1}^{\infty} \) of \( (\bar{x}_i)_{i=1}^{\infty} \) such that \( \lim_{k \in \mathbb{N}} \langle \bar{x}_{i_{k_{j_l}}}, x_s \rangle_X = \alpha_s \in F, \forall s \in \mathbb{N} \) with \( 1 \leq s \leq \bar{n} \).

If \( \bar{n} = \infty \), we may obtain (after countably infinite recursions) infinitely many subsequences by the construction procedure of the second to last paragraph. Then, we take the Cantor’s diagonal sequence \( (\bar{x}_{i_1}, \bar{x}_{i_2}, \bar{x}_{i_3}, \ldots) \) and denote it as \( (\bar{x}_{i_k})_{k=1}^{\infty} \), which is a subsequence of \( (\bar{x}_i)_{i=1}^{\infty} \). Then, we have \( \lim_{k \in \mathbb{N}} \langle \bar{x}_{i_k}, x_s \rangle_X = \alpha_s \in F, \forall s \in \mathbb{N} \).

This completes the proof of the claim. \( \square \)

Then, we have the subsequence \( (A\bar{x}_{i_k})_{k=1}^{\infty} \subseteq E \subseteq Y \). Note that \( A\bar{x}_{i_k} = \sum_{s=1}^{\bar{n}} \sigma_s y_s \langle \bar{x}_{i_k}, x_s \rangle_X \). We will show that this is a Cauchy sequence. Fix any \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists n_1 \in \mathbb{N} \) with \( 1 \leq n_1 < \bar{n} + 1 \), such that \( \sum_{s=1}^{\bar{n}} \sigma_s^2 \geq \sigma_{n_1+1} \geq \sigma_{n_1+2} \geq \cdots \geq 0 \). Then, \( \exists n_2 \in \mathbb{N} \) such that \( \| \langle \bar{x}_{i_k}, x_s \rangle_X - \alpha_s \| < \frac{\epsilon^2\sqrt{2\sigma_{n_1}(\sigma_{n_1}+1)}}{2\sigma_{n_1}(\sigma_{n_1}+1)} \), \( \forall s \in \mathbb{N} \) with \( 1 \leq s \leq n_1, \forall k \in \mathbb{N} \) with \( k \geq n_2 \). \( \forall k, j \in \mathbb{N} \) with \( k, j \geq n_2 \), we have

\[
\| A\bar{x}_{i_k} - A\bar{x}_{i_j} \|_y^2 = \| \sum_{i=1}^{\bar{n}} \sigma_i y_i \langle \bar{x}_{i_k} - \bar{x}_{i_j}, x_s \rangle_X \|_y^2
= \sum_{s=1}^{\bar{n}} \sigma_s^2 \| \langle \bar{x}_{i_k} - \bar{x}_{i_j}, x_s \rangle_X \|_y^2
\]
\[
= \sum_{s=1}^{n_1} \sigma_s^2 \left| \langle \tilde{x}_{i_k} - \tilde{x}_{i_j}, x_s \rangle \right|^2 + \sum_{s=n_1+1}^{\tilde{n}} \sigma_s^2 \left| \langle \tilde{x}_{i_k} - \tilde{x}_{i_j}, x_s \rangle \right|^2 \\
\leq \sum_{s=1}^{n_1} \sigma_s^2 \left( \frac{\epsilon}{\sqrt{2n_1(1 + \sigma)}} \right)^2 + \frac{\epsilon^2}{8} \sum_{s=n_1+1}^{\tilde{n}} \left| \langle \tilde{x}_{i_k} - \tilde{x}_{i_j}, x_s \rangle \right|^2 \\
\leq \frac{\epsilon^2}{2} + \frac{\epsilon^2}{8} \| \tilde{x}_{i_k} - \tilde{x}_{i_j} \|_X^2 \leq \epsilon^2
\]

where the second equality follows from the fact that \( (y_i)_{i=1}^{\tilde{n}} \) is orthonormal; the first inequality follows from the assumption and the preceding discussion; and the second inequality follows from Bessel’s Inequality, Proposition 13.30. This proves that \( (A\tilde{x}_{i_k})_{k=1}^{\infty} \subseteq E \subseteq Y \) is a Cauchy sequence. This completes the sufficiency proof.

(iii) Let \( Q \in S_X \cap K(X, X) \). “Necessity” \( \|Q\|_{B(X,X)} > 0 \) since \( Q \neq \vartheta_{B(X,X)} \). By Proposition 13.48, \( \exists \left( u_i \right)_{i=1}^{\infty} \subseteq X \) with \( \|u_i\| = 1 \) such that \( \lim_{i \in \mathbb{N}} \|Q(u_i), u_i\| = \|Q\|_{B(X,X)} > 0 \). Since \( Q \) is a compact operator, then \( \left( Q(u_i) \right)_{i=1}^{\infty} \subseteq \overline{Q(B_X(\vartheta_X,1))} \), which is a compact set. By Borel-Lebesgue Theorem 5.37, the set \( \overline{Q(B_X(\vartheta_X,1))} \subseteq X \) is sequentially compact. Then, there exists a subsequence \( \left( Q(u_{i_k}) \right)_{k=1}^{\infty} \subseteq \overline{Q(B_X(\vartheta_X,1))} \) such that \( \lim_{k \in \mathbb{N}} Q(u_{i_k}) = u_0 \). Then, \( \lim_{i \in \mathbb{N}} \|Q(u_{i_k}), u_{i_k}\| = \|Q\|_{B(X,X)} \). Let \( \lambda = \lim_{i \in \mathbb{N}} \langle Q(u_{i_k}), u_{i_k} \rangle \in \mathbb{K} \) by possibly selecting a further convergent subsequence, we have \( \lambda = \|Q\|_{B(X,X)} \). Note that \( \lambda = \lim_{k \in \mathbb{N}} \langle Q(u_{i_k}), u_{i_k} \rangle = \lim_{k \in \mathbb{N}} \langle u_{i_k}, Q^*(u_{i_k}) \rangle = \lim_{k \in \mathbb{N}} \langle u_{i_k}, Q(u_{i_k}) \rangle = \lim_{k \in \mathbb{N}} \langle Q(u_{i_k}), u_{i_k} \rangle = \lambda \), where the second equality follows from Proposition 13.22: the third equality follows from \( Q \) being Hermitian; and the fourth equality follows from Definition 13.1. Hence, we have \( \lambda \in \mathbb{R} \setminus \{0\} \). (\( \lambda = \|Q\|_{B(X,X)} \) or \( \lambda = -\|Q\|_{B(X,X)} \).) Note also that \( 0 \leq \|Q(u_{i_k}) - \lambda u_{i_k}\|^2 = \langle Q(u_{i_k}) - \lambda u_{i_k}, Q(u_{i_k}) - \lambda u_{i_k} \rangle = \lambda^2 \langle u_{i_k}, u_{i_k} \rangle + \|Q(u_{i_k})\|^2 - 2\lambda \langle Q(u_{i_k}), u_{i_k} \rangle \)). The first inequality follows from Proposition 13.2; the second inequality follows from simple algebra and Proposition 13.22; and the last equality follows from \( Q \) is Hermitian. Then, we have \( 0 = \lim_{k \in \mathbb{N}} \|Q(u_{i_k}) - \lambda u_{i_k}\| = \lim_{k \in \mathbb{N}} \|u_0 - \lambda u_{i_k}\| \). Hence, \( \lim_{k \in \mathbb{N}} u_{i_k} = \frac{1}{\lambda} u_0 \). By continuity of \( Q \), we have \( Q\left(\frac{1}{\lambda} u_0\right) = \lim_{k \in \mathbb{N}} Q(u_{i_k}) = u_0 \). Furthermore, \( \|\frac{1}{\lambda} u_0\| = \lim_{k \in \mathbb{N}} \|u_{i_k}\| = 1 \). Hence, \( \frac{1}{\lambda} u_0 =: x_0 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda \).

Let \( \text{span} \{x_0\} =: M \subseteq X \). It is a closed subspace in \( X \) by Theorem 7.36 and Proposition 4.39. By Theorem 13.19, \( X = M \oplus M^\perp \). Define \( P_1 : X \to M \) and \( P_\perp : X \to M^\perp \) be projection operators. We have \( P_1 + P_\perp = \text{id}_X \). This implies that \( Q = (P_1 + P_\perp)Q(P_1 + P_\perp) = P_1QP_1 + P_\perp QP_\perp + P_1QP_1 + P_\perp QP_ \perp \). Note that \( P_1 = x_0 x_0^* \). \( \forall x \in \)}
13.9. SPECTRAL THEORY OF LINEAR OPERATORS

\[ X, P_1(x) = x_0 \langle \langle x_0^*, x \rangle \rangle = x_0 \langle x, x_0 \rangle. \]  
This implies that \( P_1Q_1(x) = \langle x, x_0 \rangle P_1Q(x) = \langle x, x_0 \rangle P_1(\lambda x_0) = \lambda \langle x, x_0 \rangle x_0 = \lambda x_0 \langle \langle x_0^*, x \rangle \rangle \) and \( P_1Q_1(x) = \langle x, x_0 \rangle P_1(\lambda x_0) = \vartheta_x. \) Thus, \( Q(x) = \lambda x_0 \langle \langle x_0^*, x \rangle \rangle + P_1Q_1(x) + P Q \bar{Q}_1(x). \) Hence, \( Q = \lambda x_0 x_0^* + \bar{P} Q P_1 + \bar{Q} P \bar{Q}_1 \). By Propositions 13.22 and 13.43, we have \( Q = Q^* = \lambda x_0 x_0^* + \bar{P} Q^* \bar{P}_1 + \bar{Q} P \bar{Q}_1 \). Hence, we have \( Q = \lambda x_0 x_0^* + \bar{P} Q \bar{Q}_1 \).

"Sufficiency" Let \( Q \) admit an eigenvector \( x_0 \) associated with the eigenvalue \( \lambda \in \mathbb{R} \setminus \{0\} \) with \( \| \lambda \| = \| Q \| \in \mathbb{R} \setminus \{0\} \). Then, \( \langle x_0, Q x_0 \rangle = \langle x_0, \lambda x_0 \rangle = \lambda \| x_0 \| ^2 \neq 0. \) Hence, \( Q \neq \vartheta \).

This proves the (iii).

(iv) "Necessity" If \( Q = \vartheta \), the result is trivial with \( n = 0. \) Consider the case \( Q \neq \vartheta \). By (iii), \( Q = \lambda x_0 x_0^* + Q_2 \), where \( Q_2 = \bar{P} Q P_1 \). Denote \( \lambda_1 := \lambda \in \mathbb{R} \setminus \{0\} \) with \( \| \lambda_1 \| = \| Q \| \in \mathbb{R} \setminus \{0\} \), \( x_1 := x_0 \), \( Q_1 := Q \), and \( P_{1,1} := \bar{P}_1 \). Now, we will repeat (iii) for the operator \( Q_2 \). Clearly, \( Q_2 \in \mathfrak{X} \). We claim that

**Claim 13.52.4** Any eigenvector \( x_2 \in \mathfrak{X} \) with \( x_2 \| = 1 \) of \( Q_2 \) associated with an eigenvalue \( \lambda_2 \in \mathbb{R} \setminus \{0\} \) must be an eigenvector of \( Q_1 \) associated with the eigenvalue \( \lambda_2 \), and \( x_2 \perp x_1 \).

**Proof of claim:** \( Q_2 x_2 = \lambda_2 x_2 \). Since \( Q_2 = \bar{P}_1 Q \bar{P}_1 \), then \( \lambda_2 \langle x_1, x_2 \rangle = \langle x_1, \lambda_2 x_2 \rangle = \langle x_1, \bar{P}_1 Q \bar{P}_1 x_2 \rangle = \langle \bar{P}_1 x_1, Q \bar{P}_1 x_2 \rangle = 0 \), where the first equality follows from Definition 13.1 and the fact that \( \lambda_2 \in \mathbb{R} \); the second equality follows from our assumption that \( x_2 \) is an eigenvector of \( Q_2 \) associated with the eigenvalue \( \lambda_2 \); the third equality follows from the expression for \( Q_2 \) and Proposition 13.22; the fourth equality follows from Proposition 13.43; and the last equality follows from \( \bar{P}_1(\lambda x_0) = \vartheta_\lambda \). Since \( \lambda_2 \neq 0 \), then \( x_1 \perp x_2 \). Then, \( x_2 \in M^\perp \). It is then easy to check that \( Q(2) = Q_1(x_2) = \lambda x_0 \langle \langle x_0^*, x_2 \rangle \rangle + Q_2(x_2) = \lambda x_0 \langle x_2, x_0 \rangle + \lambda_2 x_2 = \lambda x_0 \langle x_2, x_1 \rangle + \lambda_2 x_2 = \lambda x_2 x_1 \), where the first equality follows from our notation; the second equality follows from the expression for \( Q \); the third equality follows from Riesz-Fréchet Theorem 13.15; the fourth equality follows from our notation; and the last equality follows from the preceding discussion. Hence, \( x_2 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda_2 \). This completes the proof of the claim. \( \square \)

Repeat (iii) until \( Q_{n+1} = \vartheta \) or indefinitely. After any step \( k \in \mathbb{N} \) (with \( k < n \)), we assume that we have found orthonormal sequence \( (x_i)_{i=1}^k \subseteq \mathfrak{X} \), \( x_i \) is an eigenvector of \( Q_i \) (and \( Q \)) associated with the eigenvalue \( \lambda_i \in \mathbb{R} \setminus \{0\} \) with \( \| \lambda_i \| = \| Q_i \| \in (0, \infty) \subseteq \mathbb{R} \), \( i = 1, \ldots, k \), and \( Q = \sum_{j=1}^k \lambda_j x_j x_j^* + Q_{k+1} \), \( i = 1, \ldots, k \). At step \( k + 1 \), since \( Q_{k+1} = \vartheta \), then \( \| Q_{k+1} \| \in (0, \infty) \subseteq \mathbb{R} \). Furthermore,
Proof of claim: $x_{k+1} = \lambda_{k+1} x_{k+1} x_{k+1}^* + \tilde{P}_{1,k+1} Q_{k+1} \tilde{P}_{1,k+1}$. Now, by Claim 13.52.4, $x_{k+1}$ is an eigenvector of $Q_k$ associated with the eigenvalue $\lambda_{k+1}$, and $x_{k+1} \perp x_k$. Recursively, by Claim 13.52.4, $x_{k+1}$ is an eigenvector of $Q_i$, $i = k - 1, \ldots, 1$, associated with the eigenvalue $\lambda_{k+1}$, and $x_{k+1} \perp x_i$. Hence, the sequence $(x_i)_{i=1}^{\infty}$ is an orthonormal sequence. Thus, either the process stops at $n$, for some $n \in \mathbb{Z}^+$, when $Q_{n+1} = \vartheta_{B(X,X)}$, or we have $n = \infty$ and an orthonormal sequence $(x_i)_{i=1}^{\infty}$. In the latter case, $Q = \sum_{j=1}^i \lambda_j x_j x_j^* + Q_{i+1}$, $\forall i \in \mathbb{N}$, where $Q_{i+1} := \tilde{P}_{i,i} Q_{i} \tilde{P}_{i,i}$. Clearly, $Q_{i+1} \in \mathcal{S}_X$, $\forall i \in \mathbb{N}$.

$|\lambda_i| = \|Q_i\|_{B(X,X)} \geq \|Q_i\|_{B(X,X)} \|\tilde{P}_{i,i}\|^2_{B(X,X)} \geq \|\tilde{P}_{1,i} Q_{1,i} Q_{i+1}\|_{B(X,X)} = \|Q_{i+1}\|_{B(X,X)} = |\lambda_{i+1}| > 0$, $\forall i \in \mathbb{N}$ with $1 \leq i < n$. In case of $n = \infty$. Suppose that $\lim_{i \in \mathbb{N}} |\lambda_i| > 0$. Then, $Q x_i = \lambda_i x_i$, $\forall i \in \mathbb{N}$. The sequence $(Q x_i)_{i=1}^{\infty} = (\lambda_i x_i)_{i=1}^{\infty}$ does not have a converging subsequence. This contradicts with the assumption that $Q$ is a compact operator. Therefore, $\lim_{i \in \mathbb{N}} |\lambda_i| = 0$. Then, $\lim_{i \in \mathbb{N}} \|Q - \sum_{j=1}^i \lambda_j x_j x_j^*\|_{B(X,X)} = \lim_{i \in \mathbb{N}} \|Q - \sum_{j=1}^\infty \lambda_j x_j x_j^*\|_{B(X,X)} = \lim_{i \in \mathbb{N}} |\lambda_{i+1}| = 0$, where the first two equalities follow from the recursive construction; and the last equality follows from the preceding discussion. Hence, $Q = \sum_{i=1}^\infty \lambda_i x_i x_i^*$. This completes the necessity part of the proof.

"Sufficiency" By the expression for $Q = \sum_{i=1}^\infty |\lambda_i| (\text{sgn}(\lambda_i) x_i) x_i^*$ and (ii), $Q \in K(X,X) \cap \mathcal{S}_X$. This completes the sufficiency part of the proof.

(v) Let $Q \in \mathcal{S}_X \cap K(X,X)$. Fix any $\lambda \in \mathbb{K}$ with $\lambda \neq 0$. If there exists an $i_0 \in \mathbb{N}$ with $1 \leq i_0 < n + 1$ such that $\lambda = \lambda_{i_0}$. By (iv), $Q = \sum_{i=1}^n \lambda_i x_i x_i^*$. It is easy to show that $Q(x_{i_0}) = \lambda_{i_0} x_{i_0}$. Then, $x_{i_0}$ is an eigenvector of $Q$ associated with the eigenvalue $\lambda = \lambda_{i_0}$, and $\lambda$ is an eigenvalue of $Q$.

On the other hand, if $\lambda$ is an eigenvalue of $Q$. By Proposition 13.47, $\lambda \in \mathbb{R} \setminus \{0\}$. Then, $M := \mathcal{N}(\lambda \text{id}_X - Q) \supset \{ \vartheta_X \}$. We need the following intermediate results.

Claim 13.52.5 $M = \mathcal{N}(\lambda \text{id}_X - Q)$ is finite dimensional.

Proof of claim: Suppose $M$ is infinite dimensional. Then, $\exists (\bar{x}_i)_{i=1}^{\infty} \subseteq M$ that is orthonormal. $Q \bar{x}_i = \lambda_i \bar{x}_i$, $\forall i \in \mathbb{N}$. The sequence $(Q \bar{x}_i)_{i=1}^{\infty} = (\lambda_i \bar{x}_i)_{i=1}^{\infty}$ doesn’t have a convergent subsequence. This contradicts with the assumption $Q \in K(X,X)$. Hence, the hypothesis is invalid. $M$ must be finite dimensional.

Claim 13.52.6 $M \supset \{ \vartheta_X \}$ implies that $\lambda = \lambda_{i_0}$ for some $i_0 \in \mathbb{N}$ with $1 \leq i_0 < n + 1$.

Proof of claim: Let $N := \text{span}((x_i)_{i=1}^n)$, and $P : X \to N$ and $\bar{P} : X \to N^\perp$ be projection operators. By Theorem 13.19, we have $P + \bar{P} = \text{id}_X$, $\forall x \in N(\lambda \text{id}_X - Q) = M$ with $x \neq \vartheta_X$. $x = Px + \bar{P}x = \bar{x} + \bar{\bar{x}}$. $\lambda \bar{x} = Qx$ implies that $\lambda \bar{x} + \lambda \bar{\bar{x}} = \sum_{i=1}^n \lambda_i x_i (x,x_i) \in N$. This implies that $\lambda \bar{x} \in N \cap N^\perp$ and hence $\bar{x} = \vartheta_X$. Hence, we have $\lambda \bar{x} = \sum_{i=1}^n \lambda_i x_i (\bar{x},x_i) = \lambda \sum_{i=1}^n x_i (\bar{x},x_i)$, where the last equality follows from $\bar{x} \in N$ and $(x_i)_{i=1}^n$ is
orthonormal. Therefore, we must have \( \lambda \langle \bar{x}, x_i \rangle = \lambda_i \langle \bar{x}, x_i \rangle \), \( \forall i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \). Then, \( \lambda = \lambda_i \), \( \forall i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \) and \( \langle \bar{x}, x_i \rangle \neq 0 \).

Since \( \vartheta_X \neq x = \bar{x} \in N \), \( \exists i_0 \in \mathbb{N} \) with \( 1 \leq i_0 < n + 1 \) such that \( \langle \bar{x}, x_{i_0} \rangle \neq 0 \). Then, \( \lambda = \lambda_{i_0} \).

This completes the proof of the claim. \( \square \)

By the preceding two claims, we have proved that if \( \lambda \in \mathbb{K} \) with \( \lambda \neq 0 \) is an eigenvalue of \( Q \), then \( \lambda = \lambda_{i_0} \) for some \( i_0 \in \mathbb{N} \) with \( 1 \leq i_0 < n + 1 \).

Let \( \lambda \in \mathbb{R} \setminus \{0\} \) be an eigenvalue of \( Q \). By Claim 13.52.5, \( M \) is finite dimensional. Let \( \dim M =: m \in \mathbb{N} \). We will show that there are exactly \( m \) \( \lambda_i \)'s such that \( \lambda = \lambda_i \).

Let \( M \) admit an orthonormal basis \( \{\bar{x}_i\}_{i=1}^m \). By Claim 13.52.6 and its proof, \( \lambda = \lambda_i \), \( \forall i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \) and \( \langle \bar{x}_j, x_i \rangle \neq 0 \), for some \( j \in \{1, \ldots, m\} \), and \( \langle \bar{x}_i \rangle_{i=1}^m \subseteq N = \text{span}(\{x_i\}_{i=1}^n) \). Then, there are at least \( m \) \( \lambda_i \)'s such that \( \lambda = \lambda_i \). Then, there must be exactly \( m \) such \( \lambda_i \)'s. This is because \( x_i \) is an eigenvector of \( Q \) associated with \( \lambda_i \), \( \forall i \in \mathbb{N} \) with \( 1 \leq i < n + 1 \). Then, \( x_i \in M \) whenever \( \lambda_i = \lambda \). Those \( x_i \)'s for which \( \lambda_i \)'s form a linearly independent subset of \( M \). \( M \) has dimension \( m \), and therefore such \( \lambda_i \)'s can’t be more than \( m \). This shows that there are exactly \( m \) \( \lambda_i \)'s such that \( \lambda = \lambda_i \).

We need the following intermediate result.

**Claim 13.52.7** Let \( \lambda \in \mathbb{R} \setminus \{0\} \) be an eigenvalue of \( Q \). Then,

\[
\dim \left( \mathcal{N}(\lambda \text{id}_X - Q) \right) = \dim \left( \bigcup_{s=1}^{\infty} \mathcal{N}(\lambda \text{id}_X - Q^s) \right) \in \mathbb{N}
\]

**Proof of claim:** We will show that \( \mathcal{N}(\lambda \text{id}_X - Q^{s+1}) \setminus \mathcal{N}(\lambda \text{id}_X - Q^s) = \emptyset, \forall s \in \mathbb{N} \). Suppose we have \( x_0 \in \mathcal{N}(\lambda \text{id}_X - Q^{s+1}) \setminus \mathcal{N}(\lambda \text{id}_X - Q^s) \), for some \( s \in \mathbb{N} \). Then, \( (\lambda \text{id}_X - Q^{s+1})x_0 = \vartheta_X \) and \( (\lambda \text{id}_X - Q^s)x_0 =: \xi \neq \vartheta_X \).

Clearly, \( (\lambda \text{id}_X - Q)\xi = \vartheta_X \) and \( \xi \in M \). But, \( \xi \in \mathcal{R}(\lambda \text{id}_X - Q) \subseteq \mathcal{R}(\lambda \text{id}_X - Q) = ( (\mathcal{R}(\lambda \text{id}_X - Q)^\perp )^\perp = (\mathcal{N}(\lambda \text{id}_X - Q^s))^\perp = M^\perp, \)

where the first equality follows from Proposition 13.17, and the second equality follows from Proposition 13.22. Thus, \( \xi \in M \cap M^\perp \) and hence \( \xi = \vartheta_X \). This contradicts with \( \xi \neq \vartheta_X \). Hence, the hypothesis is invalid. Therefore, we must have \( \mathcal{N}(\lambda \text{id}_X - Q^{s+1}) \setminus \mathcal{N}(\lambda \text{id}_X - Q^s) = \emptyset, \forall s \in \mathbb{N} \). Clearly, \( \mathcal{N}(\lambda \text{id}_X - Q^{s+1}) \supseteq \mathcal{N}(\lambda \text{id}_X - Q^s), \forall s \in \mathbb{N} \). Then, we must have \( \mathcal{N}(\lambda \text{id}_X - Q) = \mathcal{N}(\lambda \text{id}_X - Q^s), \forall s \in \mathbb{N} \). Then, the result holds, by Claim 13.52.5. This completes the proof of the claim. \( \square \)

This completes the proof for (v).

(vi) Let \( A \in K(\mathcal{X}, \mathcal{Y}) \) and the result of (ii) hold. Then, \( A^*A \in S_{\text{ped} \mathcal{X}} \cap K(\mathcal{X}, \mathcal{X}) \) by Propositions 13.42 and 13.51. Then, we have \( A^*A = \sum_{i=1}^{\bar{n}} \sigma_i^2 x_i x_i^* \).

By (v), \( \sigma_i^2 \)'s are the nonzero eigenvalues of \( A^*A \), and is independent of the choice of \( \{y_i\}_{i=1}^{\bar{n}} \) and \( \{x_i\}_{i=1}^{\bar{n}} \). Hence, \( \sigma_i = \sqrt{\lambda_i}, \forall i \in \mathbb{N} \) with \( 1 \leq i < \bar{n} + 1 \), where \( (\lambda_i)_{i=1}^{\bar{n}} \) are the nonzero eigenvalues of \( A^*A \)}
(including multiplicity), which is independent of the choice of \((y_i)_{i=1}^n\) and 
\((x_i)_{i=1}^n\).

This completes the proof of the theorem. \(\square\)

\textbf{Example 13.53} \ Let \(\mathcal{H}\) be a separable Hilbert space over \(\mathbb{K}\), \(\mathcal{Y}\) be a finite-dimensional Hilbert space over \(\mathbb{K}\), \(I_i \subseteq \mathbb{R}\) be compact intervals, \(\mathcal{I}_1 := \langle (I_i, | \cdot |), \mathcal{B}_i, \mu_i \rangle\) be the compact finite metric measure subspace of \(\mathbb{R}\), \(i = 1, 2\), \(\mathcal{H}_1 := L_2(I_1, \mathcal{X})\) and \(\mathcal{H}_2 := L_2(I_2, \mathcal{Y})\) be Hilbert spaces as defined in Example 13.11, \(K : I_2 \times I_1 \rightarrow B(\mathcal{X}, \mathcal{Y})\) be a continuous function. Define the mapping \(T : \mathcal{H}_1 \rightarrow \mathcal{H}_2\) by \(T(h) = |T(h)|\) and \(T : L_2(I_1, \mathcal{X}) \rightarrow L_2(I_2, \mathcal{Y})\) defined by

\[
T(h)(t) = \int_{I_1} K(t, s)h(s) \, d\mu_1(s); \quad \forall t \in I_2, \forall h \in L_2(I_1, \mathcal{X})
\]

We will show that \(\hat{T} \in B(\mathcal{H}_1, \mathcal{H}_2)\) and is a compact operator.

\[\forall [h] \in \mathcal{H}_1, \forall t \in I_2, \text{ we have } K(t, \cdot)h(\cdot) : I_1 \rightarrow \mathcal{Y} \text{ is } B_1\text{-measurable, by Propositions 11.37, 11.39, 7.126, and 7.65. By Proposition 5.29, }\exists M \in [0, \infty) \subseteq \mathbb{R}\text{ such that } \|K(t, s)\|_{B(\mathcal{X}, \mathcal{Y})} \leq M, \forall (t, s) \in I_2 \times I_1. \]

Then, \[\forall [h] \in \mathcal{H}_1, \forall t \in I_2, \forall s \in I_1, \text{ we have } \int_{I_1} \|K(t, s)h(s)\| \, d\mu_1(s) \leq \int_{I_1} M \|h(s)\|_{\mathcal{X}} \, d\mu_1(s) \leq M \sqrt{\mu_1(I_1)} \|h\|_{\mathcal{X}_{t,s}} < \infty, \text{ where the third inequality follows from Cauchy-Schwarz Inequality. By Proposition 11.89, } T(h)(t) \in \mathcal{Y} \text{ and } \|T(h)(t)\|_{\mathcal{Y}} \leq M \sqrt{\mu_1(I_1)} \|h\|_{\mathcal{X}_{t,s}}, \forall t \in I_2, \forall [h] \in \mathcal{H}_1, \text{ and the value } T(h)(t) \text{ is dependent only on } [h]. \]

Next, we will show that \(T(h) \in C(I_2, \mathcal{Y})\). By Proposition 5.39, \(K\) is uniformly continuous. \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta(\epsilon) \in (0, \infty) \subseteq \mathbb{R}\) such that \(\forall (t_1, s_1), (t_2, s_2) \in I_2 \times I_1\) with \(|(t_1, s_1) - (t_2, s_2)| < \delta(\epsilon)\), we have \(\|K(t_1, s_1) - K(t_2, s_2)\|_{B(\mathcal{X}, \mathcal{Y})} < \epsilon. \forall [h] \in \mathcal{H}_1, \forall t_1, t_2 \in I_2 \text{ with } |t_1 - t_2| < \delta(\epsilon) \), we have \(\|T(h)(t_1) - T(h)(t_2)\|_{\mathcal{Y}} = \|\int_{I_1} (K(t_1, s) - K(t_2, s))h(s) \, d\mu_1(s)\|_{\mathcal{Y}} \leq \int_{I_1} \|K(t_1, s) - K(t_2, s)\|_{B(\mathcal{X}, \mathcal{Y})} \|h(s)\|_{\mathcal{X}} \, d\mu_1(s) \leq \int_{I_1} \|K(t_1, s) - K(t_2, s)\|_{B(\mathcal{X}, \mathcal{Y})} \|h(s)\|_{\mathcal{X}} \, d\mu_1(s) \leq \int_{I_1} \sqrt{\mu_1(I_1)} \|h(s)\|_{\mathcal{X}} \, d\mu_1(s) \leq \int_{I_1} \sqrt{\mu_1(I_1)} \|h\|_{\mathcal{X}_{t,s}} \text{, where the first equality and inequality follow from Proposition 11.89; the third inequality follows from the preceding discussion; and the fourth inequality follows from Cauchy-Schwarz Inequality. Then, by Proposition 11.37, } T(h) \text{ is } B_2\text{-measurable. Furthermore, } T(h) \in L_2(I_2, \mathcal{Y}) \text{ since } \int_{I_2} \|T(h)(t)\|_{\mathcal{Y}}^2 \, d\mu_2(t) = \int_{I_2} M^2 \mu_1(I_1) \|h\|_{\mathcal{X}_{t,s}}^2 \, d\mu_2(t) = M^2 \mu_1(I_1) \|h\|_{\mathcal{X}_{t,s}}^2 < \infty. \text{ Then, } \|T(h)\|_{L_2(\mathcal{H}_1, \mathcal{H}_2)} = \|T(h)\|_{\mathcal{Y}_{t,s}} \leq M \sqrt{\mu_1(I_1)} \|h\|_{\mathcal{X}_{t,s}} < \infty \text{ and } T(h) \text{ is } B_1\text{-measurable. Hence, } \hat{T} \in B(\mathcal{H}_1, \mathcal{H}_2). \}

Finally, we will show that \(\hat{T}\) is a compact operator. Let \((y_i)_{i=1}^n \subseteq \mathcal{Y}\) be a set of orthonormal basis vectors of \(\mathcal{Y}\), where \(n \in \mathbb{Z}_+\) is the dimension of \(\mathcal{Y}\).
13.9. SPECTRAL THEORY OF LINEAR OPERATORS

Fix any \( ([h_i])_{i=1}^{\infty} \subseteq \mathcal{B}_{\mathcal{H}}(\theta_{\mathcal{H}1}, 1) := F_1 \). By Alaoglu Theorem 7.122 and Riesz-Fréchet Theorem 13.15, \( F_1 \) is weakly compact. By Propositions 5.22 and 5.26, \( \exists [h_0] \in F_1 \) such that it is a weak cluster point of \( ([h_i])_{i=1}^{\infty} \).

Let \( v_0 := T(h_0) \in L_2(\mathcal{I}_2, y) \). By the preceding paragraph, we have \( v_i := T(h_i) \in \mathcal{C}(\mathcal{I}_2, y), \forall i \in \mathbb{Z}_+ \), furthermore, \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \forall t_1, t_2 \in \mathcal{I}_2 \) with \( |t_1 - t_2| < \delta \left( \frac{\sqrt{\mu_1(t_1)}}{1 + \sqrt{\mu_1(t_1)}} \right) \), we have \( \|T(h_i)(t_1) - T(h_i)(t_2)\|_y < \varepsilon \).

Let \( t_0 := 0 \in \mathbb{Z}_+ \). Fix any \( n \in \mathbb{N} \), let \( \delta := \delta \left( \frac{2^{-n/3}}{1 + \sqrt{\mu_1(t_1)}} \right) > 0 \). Then, exists \( n \in \mathcal{I}_1 \) and \( (t_i)_{i=1}^{\hat{n}} \subseteq \mathcal{I}_2 \) with \( \min \mathcal{I}_2 = t_1 < t_2 < \cdots < t_{\hat{n}} = \max \mathcal{I}_2 \) and \( t_j - t_{j-1} < \delta, \forall j = 2, \ldots, \hat{n} \). Define \( x_{n,j,k} \in \mathcal{H}_1^\mathcal{I}_1 \) by \( (x_{n,j,k}, [h]) := \left( \int_{\mathcal{I}_1} K(t_j, s) \delta_0(s) d\mu_1(s), y_k \right) \in \mathcal{K}, \forall [h] \in \mathcal{H}_1, \forall j = 1, \ldots, \hat{n}, \forall k = 1, \ldots, n \). It is easy to see that \( x_{n,j,k} \)'s are well-defined. Let \( O_n := \{ [h] \in \mathcal{H}_1 : \|(x_{n,j,k}, h - h_0)\|_y < 2^{-n/3}, j = 1, \ldots, \hat{n}, k = 1, \ldots, n \} \subseteq \mathcal{H}_1 \).

Clearly, \( [h_0] \in O_n \) and \( O_n \) is a weakly open subset of \( \mathcal{H}_1 \) by Proposition 7.116. Since \( [h_0] \) is a weak cluster point of \( ([h_i])_{i=1}^{\infty} \), then there exists \( v_{n} \in \mathcal{I}_1 \) with \( i_n > t_{n-1} \), such that \( [h_{i_n}] \in O_n \). Thus, we have obtained a subsequence \( (T([h_{i_n}]))_{n=1}^{\infty} \subseteq \bar{T}(F_1) \subseteq \mathcal{H}_2 \). By \( [h_{i_n}] \in O_n \), we have \( \|T(h_{i_n})(t_j) - v_0(t_j), y_k\| = \|T(h_{i_n})(t_j) - T(h_0)(t_j), y_k\| = \|x_{n,j,k}, h_{i_n} - h_0\| < 2^{-n/3}, j = 1, \ldots, \hat{n}, k = 1, \ldots, n \). This implies that \( \|T(h_{i_n})(t_j) - v_0(t_j)\|_y = (\sum_{k=1}^{\hat{n}} |\langle T(h_{i_n})(t_j) - v_0(t_j), y_k\rangle|)^{1/2} < 2^{-n/3} \).

Thus, \( \forall t_0 \in 1, \ldots, \hat{n} - 1 \) such that \( t_{j_0} \leq t \leq t_{j_0+1} \). This yields \( \|T(h_{i_n})(t) - v_0(t)\|_y \leq \|T(h_{i_n})(t) - T(h_{i_n})(t_{j_0})\|_y + \|T(h_{i_n})(t_{j_0}) - v_0(t_{j_0})\|_y + \|v_0(t_{j_0}) - v_0(t)\|_y \leq 2^{-n/3} + 2^{-n/3} + 2^{-n/3} = 2^{-n} \). Thus, we have \( \|T(h_{i_n}) - v_0\|_{C(\mathcal{I}_2, y)} < 2^{-n} \), and hence \( \|T([h_{i_n}]) - [v_0]\|_{\mathcal{H}_2} < 2^{-n} \nu_2(I_2) \).

This shows that \( \lim_{n \in \mathbb{N}} T([h_{i_n}]) = [v_0] \) in \( \mathcal{H}_2 \). Thus, we have shown that \( \forall (v_0)_{i=1}^{\infty} \subseteq \bar{T}(F_1) \), there exists a subsequence \( (v_{i_n})_{n=1}^{\infty} \) that converges to an element of \( \bar{T}(F_1) \). This shows that \( \bar{T}(F_1) \) is sequentially compact. By Borel-Lebesgue Theorem 5.37, \( \bar{T}(F_1) \) is compact and closed. Hence, \( \bar{T}(F_1) = \bar{T}(F_1) \) is compact. By Proposition 13.50, \( \bar{T} \) is a compact operator.

This completes this example.

\[ \text{(ii)} \text{ Assume that } z(0) = 0. \text{ Let } M = \{ \tilde{z} \in \mathbb{Z} \mid \tilde{z}(x) = \sin\left(\frac{(n - \frac{\pi}{2})x}{2}\right), \forall x \in \mathcal{I}, \forall n \in \mathbb{N} \} \subseteq \mathbb{Z}. \text{ Let } M = \text{span } (M) \subseteq \mathbb{Z}. \text{ Then, } z \in \overline{M}. \]
Proof

(i) This follows immediately from Corollary 7.58 and Proposition 13.35.

(ii) Clearly, we have only to prove this for $T = \frac{\pi}{2}$. For general $T \in (0, \infty) \subset \mathbb{R}$, we may work with $\tilde{z} \in C(\mathbb{R}_0, 2; \mathbb{K})$ with $\tilde{z}(t) := z(\frac{2\pi t}{T}), \forall t \in \mathbb{R}_0, 2$. Let $T = \frac{\pi}{2}$. Let $\bar{I} := [\pi, \pi, \tau] \subset \mathbb{R}$ and $\bar{I} := (\bar{I}, | \cdot |)$ be the complete metric subspace of $\mathbb{R}$. Define $h \in C(\bar{I}, \mathbb{K}) := \mathcal{Y}$ by

$$h(t) = \begin{cases} 
  z(2T - t) & \text{if } T < t \leq 2T \\
  z(t) & \text{if } 0 \leq t \leq T \\
  -z(-t) & \text{if } -T \leq t < 0 \\
  -z(2T + t) & \text{if } -2T \leq t < -T
\end{cases}, \quad \forall t \in \bar{I}$$

By Theorem 3.11, $h \in \mathcal{Y}$. Clearly, $h(-t) = -h(t), \forall t \in \bar{I}$. Hence, $h$ is an odd function. Clearly, $h(\pi) = z(0) = 0 = h(-\pi)$. Hence, by (i), $h \in \mathcal{M}$, where $\mathcal{M} := \text{span} (\tilde{M})$ and $\tilde{M} := \{ \tilde{z} \in \mathcal{Y} : \tilde{z}(x) = \cos(nx), \forall x \in \bar{I}, \forall n \in \mathbb{Z}_+; \text{ or } \tilde{z}(x) = \sin(nx), \forall x \in \bar{I}, \forall n \in \mathbb{N} \} \subseteq \mathcal{Y}$. Let $S_n \in C(\bar{I}, \mathbb{K})$ be defined by $S_n(t) := \frac{1}{\pi} \sum_{i=1}^{2n} a_i \cos(it) + b_i \sin(it)), \forall t \in \bar{I}, \forall n \in \mathbb{Z}_+,$ be the partial sum of the Fourier series for $h$, where $a_i := \frac{1}{\pi} \int_{-\pi}^{\pi} h(t) \cos(it) dt \in \mathbb{K}, \forall i \in \mathbb{Z}_+$; and $b_i := \frac{1}{\pi} \int_{-\pi}^{\pi} h(t) \sin(it) dt \in \mathbb{K}, \forall i \in \mathbb{N}$. It is straightforward to show that $a_i = 0, \forall i \in \mathbb{Z}_+$, since $h$ is an odd function. Now, we will show that $b_{2k} = 0, \forall k \in \mathbb{N}$. Fix any $k \in \mathbb{N}$. We have

$$b_{2k} = \frac{1}{\pi} \int_{-\pi}^{\pi} h(t) \sin(2kt) dt$$

$$= \frac{1}{\pi} \left( \int_{-\pi}^{\pi} -z(\pi + t) \sin(2kt) dt + \int_{-\pi}^{0} -z(-t) \sin(2kt) dt \right)$$

$$+ \int_{0}^{\pi} z(t) \sin(2kt) dt + \int_{\pi}^{2\pi} z(\pi - t) \sin(2kt) dt$$

$$= \frac{1}{\pi} \left( -\int_{0}^{\pi} z(\tilde{t}) \sin(2k\tilde{t} - 2k\pi) d\tilde{t} + \int_{0}^{\pi} z(\tilde{t}) \sin(-2k\tilde{t}) d\tilde{t} \right)$$

$$+ \int_{0}^{\pi} z(t) \sin(2kt) dt - \int_{\pi}^{2\pi} z(t) \sin(2k\pi - 2k\tilde{t}) d\tilde{t}$$

$$= \frac{1}{\pi} \left( -\int_{0}^{\pi} z(\tilde{t}) \sin(2k\tilde{t} d\tilde{t} + \int_{0}^{\pi} z(\tilde{t}) \sin(2k\tilde{t}) d\tilde{t} \right)$$

$$+ \int_{0}^{\pi} z(t) \sin(2kt) dt - \int_{\pi}^{2\pi} z(t) \sin(2k\tilde{t}) d\tilde{t} = 0$$

where the second equality follows from Definition 12.71; the third equality follows from Change of Variable Theorem 12.91 and $\tilde{t} = t + \pi, \tilde{t} = -t$, and $\tilde{t} = \pi - t$; and fourth equality follows from Definition 12.71. Thus, $S_n(t) = \sum_{k=1}^{\lfloor \frac{n+1}{2} \rfloor} b_{2k-1} \sin((2k - 1)t), \forall t \in \bar{I}, \forall n \in \mathbb{Z}_+$. By Fejér’s Theorem 38.12.
of Bartle (1976), the Cesàro means of the Fourier series for \( h \) converges to \( h \) in \( Y \) (working with the real part and the imaginary part separately when \( K = \mathbb{C} \)), that is \( \lim_{n \to \infty} K_n = h \) in \( Y \), where \( K_n \in Y \) is defined by \( K_n = \frac{1}{n} \sum_{i=0}^{n-1} S_i, \forall n \in \mathbb{N} \). Since \( S_i \)'s only involve terms with \( \sin((2k-1)t) \), then so is \( K_n \). Hence, we have \( h \in \mathcal{M} \), where \( \mathcal{M} := \text{span} \left( \bar{M} \right) \) and \( \bar{M} := \{ \bar{z} \in Y \mid \bar{z}(x) = \sin((2k-1)x), \forall x \in I, \forall k \in \mathbb{N} \} \subseteq Y \).

Note that \( z = h|_I \). Then, \( z \in \mathcal{M} \), where \( \mathcal{M} \) is as defined in (ii). This completes the proof of the proposition. \( \square \)

**Proposition 13.55** Let \( X \) and \( Y \) be Hilbert space over \( K \) and \( A \in K(X, Y) \). Then, the following statements hold.

1. \( \forall \lambda \in (0, \infty) \subset K \), \( \lambda \) is an eigenvalue of \( AA^* \) with a corresponding eigenvector \( y \) if, and only if, \( \lambda \) is an eigenvalue of \( A^*A \) with a corresponding eigenvector \( A^*y \).
2. \( AA^* \in \mathcal{S}_{\text{psd}}Y \) and \( A^*A \in \mathcal{S}_{\text{psd}}X \) are both compact operators.
3. \( A^* \in K(Y, X) \).
4. There exists \( \bar{n} \in \mathbb{Z}_+ \cup \{ \infty \} \) such that \( A^*A \) admits eigenvectors \( x_i \in X \) with \( \| x_i \|_X = 1 \), each associated with an eigenvalue \( \lambda_i \in (0, \infty) \subset K \), \( i = 1, \ldots, \bar{n} \), and \( (x_i)_{i=1}^{\bar{n}} \) is an orthonormal sequence, and \( A^*A = \sum_{i=1}^{\bar{n}} \lambda_i x_i x_i^* \) and \( \lim_{n \to \infty} \lambda_i = 0 \) if \( \bar{n} = \infty \). Then, \( AA^* \) admits eigenvectors \( y_i := \frac{1}{\| Ax_i \|_Y} Ax_i \in Y \) with \( \| y_i \|_Y = 1 \), each associated with an eigenvalue \( \lambda_i \in (0, \infty) \subset K \), \( i = 1, \ldots, \bar{n} \), and \( (y_i)_{i=1}^{\bar{n}} \) is an orthonormal sequence, and \( AA^* = \sum_{i=1}^{\bar{n}} \lambda_i y_i y_i^* \).
5. \( \| AA^* \|_{B(Y, Y)} = \| A^*A \|_{B(X, X)} \) = \( \| A \|_{B(X, Y)}^2 \).

**Proof**

(i) “Necessity” We have \( \lambda y = AA^*y \) and \( \| y \| > 0 \). Then, \( \lambda^*A^* = A^*(\lambda y) = A^*AA^*y = (A^*A)(A^*y) \). Now, \( y \neq \vartheta_y \) and \( \lambda \neq 0 \), implies that \( \lambda y \neq \vartheta_y \) and then \( AA^*y \neq \vartheta_X \). This shows that \( A^*y \) is an eigenvector of \( A^*A \) associated with eigenvalue \( \lambda \).  

“Sufficiency” This follows from symmetry and the “Necessity” proof.

(ii) By (xi) of Proposition 13.42, \( AA^* \in \mathcal{S}_{\text{psd}}Y \) and \( A^*A \in \mathcal{S}_{\text{psd}}X \). By Proposition 13.51, \( A^*A \) and \( AA^* \) are compact operators.

(iii) By Spectral Theory Theorem 13.52, \( \exists \bar{n} \in \mathbb{Z}_+ \cup \{ \infty \} \) and \( \exists (x_i)_{i=1}^{\bar{n}} \subseteq X \), which is an orthonormal sequence, and \( \exists (y_i)_{i=1}^{\bar{n}} \subseteq Y \), which is an orthonormal sequence, and \( \sigma =: \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_i \in (0, \infty) \subset K \), \( \forall i \in \mathbb{N} \) with \( 0 \leq i - 1 < \bar{n} \), such that \( A = \sum_{i=1}^{\bar{n}} \sigma_i x_i x_i^* \) and in case \( \bar{n} = \infty \), we have \( \lim_{n \to \infty} \sigma_i = 0 \). Then, \( A^* = \sum_{i=1}^{\bar{n}} \sigma_i x_i y_i^* \). By Spectral Theory Theorem 13.52, \( A^* \in K(Y, X) \).

(iv) \( A^*A = (\sum_{i=1}^{\bar{n}} \sigma_i x_i x_i^*) (\sum_{i=1}^{\bar{n}} \sigma_i y_i y_i^*) = \sum_{i=1}^{\bar{n}} \sigma_i^2 x_i x_i^* \), where the second equality follows from the fact that \( (y_i)_{i=1}^{\bar{n}} \) is orthonormal. It is easy.
to check that $\sigma_i^2 x_i = A^* A x_i, \forall i \in \mathbb{N}$ with $1 \leq i \leq n + 1$. Similarly, $AA^* = \sum_{i=1}^{n} \sigma_i^2 y_i y_i^*$. By (i) of Spectral Theory Theorem 13.52, we have $y_i = \frac{1}{\sigma_i} A x_i = \frac{1}{|A^* x_i|} A x_i$. Then, (iv) follows.

(v) By Spectral Theory Theorem 13.52, we have $\| A A^* \|_{B(Y,Y)} = \sigma_1^* = \| A A^* \|_{B(X,X)} = \| A \|_{B(X,Y)}^2$.

This completes the proof of the proposition. \qed

**Proposition 13.56 (Fredholm Alternative)** Let $X$ be a Hilbert space over $\mathbb{K}$, $A \in \mathcal{S} \cap K(X,X)$, and $\lambda \in \mathbb{R} \setminus \{0\}$. Then, the following statements hold:

(i) $\mathcal{N} (\lambda \text{id}_X - A) = \{ 0 \}$ if, and only if, $\mathcal{R} (\lambda \text{id}_X - A) = X$.

(ii) $\mathcal{R} (\lambda \text{id}_X - A) = (\mathcal{N} (\lambda \text{id}_X - A))^\perp$.

**Proof**

(i) “Sufficiency” Let $\mathcal{R} (\lambda \text{id}_X - A) = X$. By Proposition 13.22, $\mathcal{N} (\lambda \text{id}_X - A) = (\mathcal{R} (\lambda \text{id}_X - A))^\perp = \mathcal{N} (\lambda \text{id}_X - A) = \{ 0 \}$.

“Necessity” Let $\mathcal{N} (\lambda \text{id}_X - A) = \{ 0 \}$. By Proposition 13.22, $\mathcal{N} (\lambda \text{id}_X - A) = (\mathcal{R} (\lambda \text{id}_X - A))^\perp$. This further implies that $\mathcal{R} (\lambda \text{id}_X - A) = (\mathcal{R} (\lambda \text{id}_X - A))^\perp = \mathcal{N} (\lambda \text{id}_X - A)$. Fix any $x_0 \in X = \mathcal{R} (\lambda \text{id}_X - A)$. Then, there exists $(x_i)_{i=1}^\infty \subseteq X$ such that $\lim_{i \to \infty} (\lambda \text{id}_X - A)x_i = x_0$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\exists r \in [0, \infty] \subseteq \mathbb{R}$ such that $(x_i)_{i=1}^\infty \subseteq \overline{B}_X (\vartheta_X, r)$; Case 2: $\forall r \in [0, \infty] \subseteq \mathbb{R}$, $\exists i \in \mathbb{N}$ such that $\| x_i \| > r$.

Case 1: $\exists r \in [0, \infty] \subseteq \mathbb{R}$ such that $(x_i)_{i=1}^\infty \subseteq \overline{B}_X (\vartheta_X, r)$. By $A \in \mathcal{K}(X, X)$ and Definition 13.49, there exists a subsequence $(x_{i_k})_{k=1}^\infty$ such that $(Ax_{i_k})_{k=1}^\infty$ converges to $\bar{x} \in \mathbb{X}$. Then, we have $\lim_{k \to \infty} \lambda \text{id}_X x_{i_k} = \lim_{k \to \infty} (\lambda \text{id}_X - A)(x_{i_k}) = \lim_{k \to \infty} \lambda \text{id}_X (x_{i_k}) - \lim_{k \to \infty} A x_{i_k} = \lambda \bar{x} - \bar{x} = x_0$. Hence, $x_0 \in \mathcal{R} (\lambda \text{id}_X - A)$.

Case 2: $\forall r \in [0, \infty] \subseteq \mathbb{R}$, $\exists i \in \mathbb{N}$ such that $\| x_i \| > r$. By possibly further restrict to a subsequence if necessary, we may assume, without loss of generality, $\exists i \in \mathbb{N}$ such that $\| x_i \| =: \lim_{i \to \infty} r_i = \infty$. Then, we have $\lim_{i \to \infty} \lambda \text{id}_X x_{i_k} = \lim_{i \to \infty} \lambda \text{id}_X (1/r_i x_{i_k}) = \lambda \bar{x}$. Note that $\| \frac{1}{r_i} x_{i_k} \| = 1, \forall i \in \mathbb{N}$. By $A \in \mathcal{K}(X, X)$ and Definition 13.49, there exists a subsequence $(\frac{1}{r_{i_k}} x_{i_k})_{k=1}^\infty$ such that $\lim_{k \to \infty} A (\frac{1}{r_{i_k}} x_{i_k}) = \bar{x} \in \mathbb{X}$. Then, $\lim_{k \to \infty} \lambda \frac{1}{r_{i_k}} x_{i_k} = \lim_{k \to \infty} \lambda \text{id}_X (\frac{1}{r_{i_k}} x_{i_k}) = \lambda \bar{x}$ and $A (\frac{1}{r_{i_k}} x_{i_k}) = \lim_{k \to \infty} A (\frac{1}{r_{i_k}} x_{i_k}) = \bar{x}$. We can conclude that $\bar{x} \in \mathcal{N} (\lambda \text{id}_X - A) = \{ 0 \}$. Hence, $\bar{x} = \vartheta_X$. But $\| \frac{1}{r_{i_k}} x_{i_k} \| = \lim_{k \to \infty} \| \frac{1}{r_{i_k}} x_{i_k} \| = 1 > 0$. This is a contradiction! Hence, this case is impossible.

In the above, we have shown that $x_0 \in \mathcal{R} (\lambda \text{id}_X - A)$. By the arbitrariness of $x_0$, we have $\mathcal{R} (\lambda \text{id}_X - A) = X$.
13.9. SPECTRAL THEORY OF LINEAR OPERATORS

This completes the proof of (i).

(ii) Let $\mathcal{N}(\lambda \text{id}_X - A) =: M$. If $M = \{ \varnothing_X \}$, then by (i) we have $\mathcal{R}(\lambda \text{id}_X - A) = X = M^\perp$. This case is proved.

If $M \supset \{ \varnothing_X \}$, then $\lambda$ is an eigenvalue of $A$. By Spectral Theory Theorem 13.52, $M$ is finite dimensional. By Proposition 13.22, we have $(\mathcal{R}(\lambda \text{id}_X - A))^\perp = \mathcal{N}(\lambda \text{id}_X - A) = M$. Thus, by Proposition 13.17, we have $\mathcal{R}(\lambda \text{id}_X - A) = M^\perp$. Fix any $x_0 \in M^\perp$. Then, there exists $(\tilde{x}_i)_{i=1}^\infty \subseteq X$ such that $\lim_{i \to \infty} (\lambda \text{id}_X - A)\tilde{x}_i = x_0$. Note that $M$ is finite dimensional implies that $M$ is closed by Theorem 7.36 and Proposition 4.39.

Let $P : X \to M$ and $\tilde{P} : X \to M^\perp$ be projection operators. By Theorem 13.19, we have $P + \tilde{P} = \text{id}_X$. Then, $x_0 = \lim_{i \to \infty} (\lambda \text{id}_X - A)\tilde{x}_i = \lim_{i \to \infty} (\lambda \text{id}_X - A)(P\tilde{x}_i + \tilde{P}\tilde{x}_i) = \lim_{i \to \infty} (\lambda \text{id}_X - A)(\tilde{x}_i)$. Hence, the sequence $(\tilde{x}_i)_{i=1}^\infty := (\tilde{P}\tilde{x}_i)_{i=1}^\infty \subseteq M^\perp$ satisfies $x_0 = \lim_{i \to \infty} (\lambda \text{id}_X - A)\tilde{x}_i$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\exists r \in [0, \infty) \subseteq \mathbb{R}$ such that $(\tilde{x}_i)_{i=1}^\infty \subseteq B_X(\varnothing_X, r)$; Case 2: $\forall r \in [0, \infty) \subseteq \mathbb{R}$, there exists $i_0 \in \mathbb{N}$ such that $\| \tilde{x}_{i_0} \| > r$.

Case 1: $\exists r \in [0, \infty) \subseteq \mathbb{R}$ such that $(\tilde{x}_i)_{i=1}^\infty \subseteq B_X(\varnothing_X, r)$. By $A \in K(X, X)$, there exists a subsequence $(\tilde{x}_{i_k})_{k=1}^\infty$ such that $\lim_{k \to \infty} A\tilde{x}_{i_k} = \tilde{x} \in X$. Then, we have $\lim_{k \to \infty} \lambda \tilde{x}_{i_k} = \lim_{k \to \infty} (\lambda \text{id}_X - A)\tilde{x}_{i_k} = x_0 + \tilde{x} =: \lambda \tilde{x}_0$. This implies that $\mathcal{R}(\lambda \text{id}_X - A)\tilde{x}_0 = \lim_{k \to \infty} (\lambda \text{id}_X - A)\tilde{x}_{i_k} = x_0 \in \mathcal{R}(\lambda \text{id}_X - A)$.

Case 2: $\forall r \in [0, \infty) \subseteq \mathbb{R}$, there exists $i_0 \in \mathbb{N}$ such that $\| \tilde{x}_{i_0} \| > r$. By possibly further restricting to a subsequence, we may, without loss of generality, assume that $\lim_{i \to \infty} \| \tilde{x}_i \| =: \lim_{i \to \infty} r_i = \infty$. Then, we have $\lim_{i \to \infty} (\lambda \text{id}_X - A)(\frac{1}{r_i}\tilde{x}_i) = \varnothing_X$. By $A \in K(X, X)$ and Definition 13.49, there exists a subsequence $(\frac{1}{r_{i_k}}\tilde{x}_{i_k})_{k=1}^\infty \subseteq M^\perp$ such that $\lim_{k \to \infty} A(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) =: \tilde{x} \in X$. This implies that $\lim_{k \to \infty} \lambda(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) = \lim_{k \to \infty} (\lambda \text{id}_X - A)(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) + \lim_{k \to \infty} A(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) = \varnothing_X + \tilde{x} = \tilde{x} \in M^\perp$.

Furthermore, $\| \frac{1}{r_{i_k}}\tilde{x}_{i_k} \| = \lim_{k \to \infty} \| \frac{1}{r_{i_k}}\tilde{x}_{i_k} \| = 1$ and $\tilde{x} \neq \varnothing_X$. Note that $A\tilde{x} = \lim_{k \to \infty} A(\lambda(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) = \lambda \lim_{k \to \infty} A(\frac{1}{r_{i_k}}\tilde{x}_{i_k}) = \lambda \tilde{x}$. This implies that $\tilde{x}$ is an eigenvector of $A$ associated with the eigenvalue $\lambda$ and $\tilde{x} \in M$. Thus, we have $\tilde{x} \in M \cap M^\perp$ and hence $\tilde{x} = \varnothing_X$. This is a contradiction with $\tilde{x} \neq \varnothing_X$. Thus, this case is impossible.

Based on the preceding analysis, we have $x_0 \in \mathcal{R}(\lambda \text{id}_X - A)$. By the arbitrariness of $x_0$, we have $M^\perp = \mathcal{R}(\lambda \text{id}_X - A)$.

This completes the proof of the proposition. □

Example 13.57 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \subseteq \mathcal{B}(\mathbb{R}^m)$ be a nonempty open rectangle, $X := ((\Omega, \cdot, \cdot), \mathcal{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}^m$, $\mathcal{Y}$ be a separable Hilbert space over $\mathbb{K}$ with $\mathcal{Y}^*$ being separable. Fix $x_0 \in \Omega$. Then, $W_{2,1,x_0}(X, \mathcal{Y})$ (as defined in Example 12.106) is a Hilbert space over $\mathbb{K}$ with inner product
defined by
\[
\langle u, v \rangle_{W^{2,1, x_0}(X, Y)} := \int_{\Omega} (u(x), v(x))_y \, d\mu(x) + \sum_{j \leq J, J \neq \emptyset} \int_{\pi_j(\Omega)} \langle u_{J,x_0}(s), v_{J,x_0}(s) \rangle_y \, d\mu_j(s) \in K
\]

\forall u, v \in W^{2,1, x_0}(X, Y), where \( u_{J,x_0} \)'s are the stream functions of \( u \) with respect to \( x_0 \), \( v_{J,x_0} \)'s are the stream functions of \( v \) with respect to \( x_0 \). (Notations are as defined in Example 12.106.)

It is easy to show that \( \langle \cdot, \cdot \rangle_{W^{2,1, x_0}(X, Y)} \) defines an inner product on \( W^{2,1, x_0}(X, Y) \). Then, \( W^{2,1, x_0}(X, Y) \) is a pre-Hilbert space. By Example 12.106, it is a Hilbert space.

Then, by Riesz-Fréchet Theorem 13.15, \( W^{2,1, x_0}(X, Y) \) is reflexive and its dual also a Hilbert space. \( \forall f \in (W^{2,1, x_0}(X, Y))^*, \) there exists a unique \( u_0 \in W^{2,1, x_0}(X, Y) \) such that \( f(u) = \langle u, u_0 \rangle_{W^{2,1, x_0}(X, Y)}, \) \( \forall u \in W^{2,1, x_0}(X, Y) \), and \( \| f \|_{(W^{2,1, x_0}(X, Y))^*} = \| u_0 \|_{W^{2,1, x_0}(X, Y)} \). \( f \) is denoted to be \( u_0^* \).

The tools of this chapter are applicable to \( W^{2,1, x_0}(X, Y) \). ∎
Chapter 14

Probability Theory

14.1 Fundamental Notions

Definition 14.1 Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a finite measure space with \( P(\Omega) = 1 \). Then, it is said to be a probability measure space. Let \( Y \) be a separable Banach space over \( K \). A \( Y \)-valued random variable is a \( \mathcal{B} \)-measurable function \( x : \Omega \rightarrow Y \). When \( Y = \mathbb{R} \), we will simply say \( x \) is a random variable. The integral \( \int_\Omega x \, dP \) is said to be the expectation of \( x \), when it makes sense, which will be denoted by \( E(x) \).

Definition 14.2 Let \( (\Omega, \mathcal{B}) \) be a measurable space, \( Y \) be a topological space, \( \mathcal{F} \subseteq \mathcal{B} \), \( x : \Omega \rightarrow Y \) be \( \mathcal{B} \)-measurable, and \( x_n : \Omega \rightarrow Y \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). We will denote the \( \sigma \)-algebra generated by \( \mathcal{F} \) by \( \sigma(\mathcal{F}) \subseteq \mathcal{B} \); and denote the \( \sigma \)-algebra generated by \( \{ E \in \mathcal{B} \mid E = x_{inv}(O), O \in \mathcal{O}_Y \} \) by \( \sigma(x) \subseteq \mathcal{B} \); and denote the \( \sigma \)-algebra generated by \( \{ E \in \mathcal{B} \mid E = x_{inv}(O), O \in \mathcal{O}_Y, n \in \mathbb{N} \} \) by \( \sigma((x_n)_{n=1}^{\infty}) \subseteq \mathcal{B} \). Clearly, \( \sigma(x) \) is the smallest \( \sigma \)-algebra on which \( x \) is measurable; and \( \sigma((x_n)_{n=1}^{\infty}) \) is the smallest \( \sigma \)-algebra on which \( x_n, \forall n \in \mathbb{N} \), are measurable.

Proposition 14.3 Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( n \in \mathbb{N}, F_1, \ldots, F_n \in \mathcal{B} \). Then,

\[
P(\bigcup_{i=1}^n F_i) = \sum_{l=1}^n (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq n} P(\bigcap_{j=1}^l F_{i_j})
\]

(14.1)

Proof We will prove this using mathematical induction on \( n \).

1° \( n = 1 \). This is straightforward. \( n = 2 \). Then, by countable additivity of measure, we have

\[
P(F_1 \cup F_2) = P((F_1 \setminus F_2) \cup F_2) = P(F_1 \setminus F_2) + P(F_2).
\]

Note that \( P(F_1) = P(F_1 \setminus F_2) + P(F_1 \cap F_2) \). This implies that \( P(F_1 \cup F_2) = P(F_1) + P(F_2) - P(F_1 \cap F_2) \). This case is proved.

2° Assume the result holds for \( n \leq k \in \{2, 3, \ldots\} \).
3° Consider the case $n = k + 1 \in \{3, 4, \ldots \}$. Then, we have

$$P\left(\bigcup_{i=1}^{k+1} F_i\right) = P\left(\bigcup_{i=1}^{k} F_i \cup F_{k+1}\right) = P\left(\bigcup_{i=1}^{k} F_i\right) + P(F_{k+1}) - P\left(\bigcup_{i=1}^{k} (F_i \cap F_{k+1})\right)$$

$$= \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} F_{i_j}) + P(F_{k+1})$$

$$- \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} (F_{i_j} \cap F_{k+1}))$$

$$= \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} F_{i_j}) + P(F_{k+1})$$

$$+ \sum_{l=1}^{k+1} (-1)^l \sum_{1 \leq i_1 < \cdots < i_l \leq k+1} P(\bigcap_{j=1}^{l} F_{i_j})$$

This completes the induction process and therefore the proof of the proposition. □

**Lemma 14.4 (Fatou’s Lemma)** Let $\Omega := (\Omega, B, P)$ be a probability measure space, and $(E_n)_{n=1}^{\infty} \subseteq B$. Define $F_i := \bigcap_{m=1}^{\infty} \bigcup_{m=n}^{\infty} E_m = \{\omega \in \Omega \mid \omega \in E_i \text{ for infinitely many numbers of } i \in \mathbb{N}\}$, and $F_2 := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} E_m = \{\omega \in \Omega \mid \omega \in E_i, \text{ eventually for all } i \in \mathbb{N} \text{ with } i \geq n, n \in \mathbb{N}\}$. Then, $F_2 \subseteq F_1$, and $P(F_2) \leq \liminf_{n \in \mathbb{N}} P(E_n) \leq \limsup_{n \in \mathbb{N}} P(E_n) \leq P(F_1)$.

**Proof** \(\forall \omega \in F_2, \exists n_\omega \in \mathbb{N}, \omega \in \bigcup_{m=n}^{\infty} E_m, \forall n \in \mathbb{N} \text{ with } n_\omega \leq n\). Then, $\omega \in \bigcup_{m=n}^{\infty} E_m, \forall n \in \mathbb{N}$. Hence, $\omega \in F_1$. By the arbitrariness of $\omega$, we have $F_2 \subseteq F_1$.

Note that $\chi_{F_2, \omega} = \liminf_{n \in \mathbb{N}} \chi_{E_n, \omega} \leq \limsup_{n \in \mathbb{N}} \chi_{E_n, \omega} = \chi_{F_1, \omega}$. Then, by Fatou’s Lemma 11.80, we have $P(F_2) = \int_{\Omega} \chi_{F_2, \omega} dP \leq \liminf_{n \in \mathbb{N}} \int_{\Omega} \chi_{E_n, \omega} dP = \liminf_{n \in \mathbb{N}} P(E_n) \leq \limsup_{n \in \mathbb{N}} P(E_n)$. Let $G_n := \bigcup_{m=n}^{\infty} E_m, \forall n \in \mathbb{N}$. Then, $G_n \subseteq G_{n+1}, \forall n \in \mathbb{N}$. By Proposition 11.5, $P(F_1) = P(\bigcap_{n=1}^{\infty} G_n) = \liminf_{n \in \mathbb{N}} P(G_n) \geq \limsup_{n \in \mathbb{N}} P(E_n)$. This completes the proof of the lemma. □

**Definition 14.5** Let $\Omega := (\Omega, B, P)$ be a probability measure space, $\Lambda$ be a set, and $\mathcal{B}_\Lambda \subseteq B$ be a $\sigma$-algebra on $\Omega$, $\forall \lambda \in \Lambda$. We will say that the collection...
of $\sigma$-algebras $(\mathcal{B}_\lambda)_{\lambda \in \Lambda}$ is independent if, $\forall n \in \mathbb{N}$, $\forall \lambda_1, \ldots, \lambda_n \in \Lambda$, which are distinct, $\forall E_l \in \mathcal{B}_{\lambda_l}$, $l = 1, \ldots, n$, we have $P(\bigcap_{l=1}^n E_l) = \prod_{l=1}^n P(E_l)$.

A collection of random variables (possibly Banach space valued) $(x_\lambda)_{\lambda \in \Lambda}$ are said to be independent if the collection of $\sigma$-algebras $(\sigma(x_\lambda))_{\lambda \in \Lambda}$ is independent.

**Lemma 14.6 (Borel-Cantelli)** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$. Then, the following statements hold.

(i) Assume that $\sum_{n=1}^\infty P(E_n) < \infty$. Then $P(\bigcap_{n=1}^\infty \bigcup_{m=n}^\infty E_m) := P(F_1) = 0$.

(ii) Assume that the collection of $\sigma$-algebras $(\sigma(\emptyset, E_i, \Omega))_{i=1}^\infty$ is independent, and $\sum_{n=1}^\infty P(E_n) = \infty$. Then, $P(F_1) = 1$.

**Proof**

(i) $P(F_1) = \lim_{n \in \mathbb{N}} P(\bigcup_{m=n}^\infty E_m) \leq \lim_{n \in \mathbb{N}} \sum_{m=n}^\infty P(E_m) = 0$, where the first equality follows from the proof of Lemma 14.4; and the inequality follows from Proposition 11.6; and the last equality follows from the assumption.

(ii) Note that $\Omega \setminus F_1 = \bigcup_{n=1}^\infty \bigcap_{m=n}^\infty (\Omega \setminus E_m)$. Let $p_k := P(E_k)$, $\forall k \in \mathbb{N}$. Then, $P(\bigcap_{m=n}^\infty (\Omega \setminus E_m)) = \lim_{k \in \mathbb{N}} P(\bigcap_{m=n}^k (\Omega \setminus E_m)) = \lim_{k \in \mathbb{N}} \prod_{m=n}^k P(\Omega \setminus E_m) = \lim_{k \in \mathbb{N}} \prod_{m=n}^k (1 - p_m) = \prod_{m=n}^\infty (1 - p_m)$, $\forall n \in \mathbb{N}$, where the first equality follows from Proposition 11.5; the second equality follows from the independence assumption; and the third equality follows from Definition 11.2. $\forall \alpha \in \mathbb{R}$, $1 + \alpha \leq \exp(\alpha)$. Then, $(1 - p_m) \leq \exp(-p_m)$, $\forall m \in \mathbb{N}$. This implies that $0 \leq \prod_{m=n}^\infty (1 - p_m) \leq \exp(-\sum_{m=n}^\infty p_m) = 0$. Then, we have $P(\bigcap_{m=n}^\infty (\Omega \setminus E_m)) = 0$, $\forall n \in \mathbb{N}$. This implies $0 \leq P(\Omega \setminus F_1) \leq \sum_{n=1}^\infty P(\bigcap_{m=n}^\infty (\Omega \setminus E_m)) = 0$, where the second inequality follows from Proposition 11.6. Hence, $P(F_1) = 1$.

This completes the proof of the lemma. \(\square\)

**Definition 14.7** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $x \in L_1(\Omega, \mathcal{Y})$, $\mathcal{B} \subseteq \mathcal{B}$ be a $\sigma$-algebra on $\Omega$, and $\mathcal{P} := P|_{\mathcal{B}}$. By Proposition 11.13, $\Omega := (\Omega, \mathcal{B}, \mathcal{P})$ is a probability measure space. Define $\nu : \mathcal{B} \to \mathcal{Y}$ by $\nu(E) = \int_E x \, d\mathcal{P}$, $\forall E \in \mathcal{B}$. By Proposition 11.16, $\nu$ is a finite $\mathcal{Y}$-valued measure on $(\Omega, \mathcal{B})$ with $\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ x \, d\mathcal{P}$, $\forall E \in \mathcal{B}$. By Definition 11.166 and Proposition 11.167, $x = \frac{d\nu}{d\mathcal{P}}$ a.e. in $\Omega$. By Proposition 11.107, $\nu := \nu|_{\mathcal{B}}$ is a $\mathcal{Y}$-valued pre-measure on $(\Omega, \mathcal{B})$. By Definition 11.99, we have $\mathcal{P} \circ \nu \leq (\mathcal{P} \circ \nu)|_{\mathcal{B}}$. Then, $(\Omega, \mathcal{B}, \nu)$ is a finite $\mathcal{Y}$-valued measure space. $\forall E \in \mathcal{B}$, $\mathcal{P}(E) = 0$, we have $\nu(E) = 0$. Hence, $\mathcal{P} \circ \nu \ll \mathcal{P}$. Let there exist $f : \Omega \to \mathcal{Y}$ that is the Radon-Nikodym derivative of $\nu$ with respect to $\mathcal{P}$, where $f$ is $\mathcal{B}$-measurable. The function $f$ is unique as described in Proposition 11.167. $f = \frac{d\nu}{d\mathcal{P}}$ a.e. in $\Omega$, and $f \in L_1(\Omega, \mathcal{Y})$.

Then, $[f] \in L_1(\Omega, \mathcal{Y})$ is said to be the conditional expectation of $x$ given $\mathcal{B}$.
and is denoted by $E(x|\tilde{B})$. $f \in E(x|\tilde{B})$ is a $\mathcal{Y}$-valued random variable and is said to be a version of the conditional expectation. Then,

$$
\int_E x \, dP = \nu(E) = \tilde{\nu}(E) = \int_E f \, d\tilde{P} = \int_E f \, dP; \quad \forall E \in \tilde{B}(14.2a)
$$

$$
\int_E P \circ f \, d\tilde{P} = P \circ \tilde{\nu}(E) = \int_E P \circ f \, dP \leq P \circ \nu(E) = \int_E P \circ x \, dP
$$

where the last equality in (14.2a) follows from Proposition 11.72; the first equality in (14.2b) follows from Definition 11.166; the second equality in (14.2b) follows from Proposition 11.72; the first inequality in (14.2b) follows from Definition 11.99; and the last equality in (14.2b) follows from Proposition 11.116.

In the above, if $\mathcal{Y}$ is a separable reflexive Banach space with $\mathcal{Y}^*$ being separable, then $E(x|\tilde{B}) \in L_1(\Omega, \mathcal{Y})$ exists by Radon-Nikodym Theorem 11.171.

**Proposition 14.8** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $x \in L_1(\Omega, \mathcal{Y})$, $\tilde{B} \subseteq \mathcal{B}$ be a $\sigma$-algebra on $\Omega$, $\tilde{\Omega} := (\tilde{\Omega}, \tilde{B}, \tilde{P} := P|_{\tilde{B}})$, and $f \in L_1(\tilde{\Omega}, \mathcal{Y})$. Assume that $f$ satisfies (14.2a), then $E(x|\tilde{B})$ exists and $f \in E(x|\tilde{B})$.

**Proof** Let $\nu$ and $\tilde{\nu}$ be as defined in Definition 14.7. By Proposition 11.116, we may define the $\mathcal{Y}$-valued measure $\tilde{\nu}$ with kernel $f$ on $(\tilde{\Omega}, \tilde{B})$. $\tilde{\nu}$ is finite since $f \in L_1(\tilde{\Omega}, \mathcal{Y})$. By (14.2a), $\tilde{\nu}(E) = \nu(E), \forall E \in \tilde{B}$. Therefore, $\tilde{\nu} = \nu$. By Definition 11.166 and Proposition 11.116, $f = \frac{df}{dP}$ a.e. in $\tilde{\Omega}$. Hence, $E(x|\tilde{B})$ exists and $f \in E(x|\tilde{B})$. This completes the proof of the proposition. □

**Proposition 14.9** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $\mathcal{W} \subseteq \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be a separable subspace, $\tilde{\mathcal{B}} \subseteq \mathcal{B}$ and $\mathcal{B} \subseteq \mathcal{B}$ be independent $\sigma$-algebras on $\Omega$, $\tilde{\Omega} := (\tilde{\Omega}, \tilde{\mathcal{B}}, \tilde{P} := P|_{\tilde{B}})$ and $\Omega := (\Omega, \mathcal{B}, P := P|_{\mathcal{B}})$ be probability measure spaces, $x \in L_1(\tilde{\Omega}, \mathcal{Y})$ be $\tilde{\mathcal{B}}$-measurable, $A \in L_1(\Omega, \mathcal{W})$ be $\mathcal{B}$-measurable. Assume that $g := \|A(\cdot)\|_{\mathcal{B}(\mathcal{Y}, \mathcal{Z})} \|x(\cdot)\|_{\mathcal{Y}}$ on $\Omega \to [0, \infty) \subset \mathbb{R}$ is integrable over $\Omega$. Then, $Ax \in L_1(\tilde{\Omega}, \mathcal{Y})$ and $E(Ax) = E(A)E(x)$.

**Proof** First, we will consider the special case where $A$ and $x$ are simple functions. Let $A$ admit the canonical representation $A = \sum_{i=1}^n A_i \chi_{E_i}, \Omega$, where $n \in \mathbb{Z}^+$, $A_1, \ldots, A_n \in \mathcal{W}$ are distinct and none equal to $\nu|_{\mathcal{B}(\mathcal{Y}, \mathcal{Z})}$, and $E_1, \ldots, E_n \in \tilde{\mathcal{B}}$ are nonempty and pairwise disjoint. Let $x$ admit the canonical representation $x = \sum_{j=1}^m x_j \chi_{F_j}, \Omega$, where $m \in \mathbb{Z}^+$, $x_1, \ldots, x_m \in \mathcal{Y}$ are distinct and none equal to $\tilde{\nu}|_{\mathcal{Y}}$ and $F_1, \ldots, F_m \in \tilde{\mathcal{B}}$ are nonempty and
pairwise disjoint. Then, clearly, $Ax \in \bar{L}_1(\Omega, \mathcal{F})$ is $\mathcal{B}$-measurable. This leads to

$$E(Ax) = \int_{\Omega} A_x d\mathcal{P} = \lim_{n \to \infty} \int_{\Omega} \phi_n \psi_n d\mathcal{P} = \lim_{n \to \infty} E(\phi_n, \psi_n)$$

where the second equality follows from Proposition 11.92; the third equality follows from Proposition 11.75; the fourth equality follows from the independence assumption; and the last equality follows from Proposition 11.75. Hence, the result holds in the special case.

Now, consider the general case. Since $A \in \bar{L}_1(\Omega, \mathcal{F})$, by Proposition 11.66, there exists a sequence of simple functions $(\phi_n)_{n=1}^{\infty}$ with $\phi_n : \Omega \to \mathcal{F}$ being $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, such that $\lim_{n \to \infty} \phi_n = A$ a.e. in $\Omega$ and $\|\phi_n(\omega)\|_{B(y, \mathcal{F})} \leq \|A(\omega)\|_{B(y, \mathcal{F})}$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. Since $x \in \bar{L}_1(\Omega, \mathcal{F})$, again by Proposition 11.66, there exists a sequence of simple functions $(\psi_n)_{n=1}^{\infty}$ with $\psi_n : \Omega \to \mathcal{F}$ being $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, such that $\lim_{n \to \infty} \psi_n = x$ a.e. in $\Omega$ and $\|\psi_n(\omega)\|_{\mathcal{F}} \leq \|x(\omega)\|_{\mathcal{F}}$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. By Propositions 7.65, 11.39, and 11.38, $Ax$ is $\mathcal{B}$-measurable and $\phi_n, \psi_n$ is $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. By the assumption, $Ax \in \bar{L}_1(\Omega, \mathcal{F})$. Clearly, $\lim_{n \to \infty} \phi_n = A$ a.e. in $\Omega$ and $\lim_{n \to \infty} \psi_n = x$ a.e. in $\Omega$. This implies that $\lim_{n \to \infty} \phi_n \psi_n = Ax$ a.e. in $\Omega$. $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, $\|\phi_n(\omega)\psi_n(\omega)\|_{\mathcal{F}} \leq \|\phi_n(\omega)\|_{B(y, \mathcal{F})} \|\psi_n(\omega)\|_{\mathcal{F}} \leq \|A(\omega)\|_{B(y, \mathcal{F})} \|x(\omega)\|_{\mathcal{F}} = g(\omega)$. Then, we have

$$E(Ax) = \int_{\Omega} Ax d\mathcal{P} = \lim_{n \to \infty} \int_{\Omega} \phi_n \psi_n d\mathcal{P} = \lim_{n \to \infty} E(\phi_n, \psi_n)$$

where the second equality follows from Lebesgue Dominated Convergence Theorem 11.91; the fourth equality follows from the special case; the sixth equality follows from Proposition 11.72; the seventh equality follows from Lebesgue Dominated Convergence Theorem 11.91; and the eighth equality follows from Proposition 11.72.

This completes the proof of the proposition. □

**Proposition 14.10** Let $\Omega := (\Omega, \mathcal{B}, \mathcal{P})$ be a probability measure space, $\mathcal{B} \subseteq \mathcal{G}$ and $\mathcal{B} \subseteq \mathcal{G}$ be $\sigma$-algebras on $\Omega$, $\mathcal{F} \subseteq \mathcal{B}$ be a $\pi$-system on $\Omega$, $\mathcal{I} \subseteq \mathcal{G}$
be a $\pi$-system on $\Omega$. Assume that $\sigma(\hat{I}) = \hat{\mathcal{B}}$, $\sigma(\tilde{I}) = \tilde{\mathcal{B}}$, and $\forall \hat{E} \in \hat{I}$, $\forall \tilde{E} \in \tilde{I}$, we have $P(\hat{E} \cap \tilde{E}) = P(\hat{E})P(\tilde{E})$. Then, $\mathcal{B}$ and $\hat{\mathcal{B}}$ are independent $\sigma$-algebras.

**Proof** Define $\mathcal{F} := \{ \hat{E} \in \hat{\mathcal{B}} \mid \forall \tilde{E} \in \tilde{I}, P(\hat{E} \cap \tilde{E}) = P(\hat{E})P(\tilde{E}) \}$. Clearly, $\tilde{I} \subseteq \mathcal{F}$. We will show that $\mathcal{F}$ is a monotone class on $\Omega$. Then, by Monotone Class Lemma 12.19, $\mathcal{F} = \sigma(\tilde{I}) = \hat{\mathcal{B}}$.

(i) $\emptyset, \Omega \in \tilde{I} \subseteq \mathcal{F}$. (ii) Fix any $E_1, E_2 \in \mathcal{F}$ with $E_1 \subseteq E_2$. $\forall \tilde{E} \in \tilde{I}$, we have $P((E_2, E_1) \cap \tilde{E}) = P((E_2, E_1) \setminus (E_1 \cap \tilde{E})) = P(E_2 \cap \tilde{E}) - P(E_1 \cap \tilde{E}) = P(E_2)P(\tilde{E}) - P(E_1)P(\tilde{E}) = P(E_2 \setminus E_1)P(\tilde{E})$, where the first equality follows from the fact that $E_1 \cap \tilde{E} \subseteq E_2 \cap \tilde{E}$ and both sets are in $\mathcal{B}$: the third equality follows from the fact that $E_1, E_2 \in \mathcal{F}$; and the fourth equality follows from Proposition 2.5; the third equality follows from countable additivity of measure; the fourth equality follows from (ii); and the last equality follows from countable additivity of measure. Hence, by the arbitrariness of $\tilde{E} \in \tilde{I}$, we have $P(\bigcup_{i=1}^{\infty} E_i) = P((\bigcup_{i=2}^{\infty} E_i \setminus E_{i-1}) \cap \tilde{E}) = \sum_{i=2}^{\infty} (P(E_i \setminus E_{i-1}) \cap \tilde{E}) = \sum_{i=2}^{\infty} P(E_i \setminus E_{i-1})P(\tilde{E}) + P(E_1)P(\tilde{E}) = P(\tilde{E})(\sum_{i=2}^{\infty} P(E_i \setminus E_{i-1}) + P(E_1)) = P(\tilde{E}) \mathcal{P}(\bigcup_{i=1}^{\infty} E_i)$, where the first two equalities follow from Proposition 2.5; the third equality follows from countable additivity of measure; the fourth equality follows from (ii); and the last equality follows from countable additivity of measure. Define $\mathcal{G} := \{ \hat{E} \in \hat{\mathcal{B}} \mid \forall \tilde{E} \in \tilde{I}, P(\hat{E} \cap \tilde{E}) = P(\hat{E})P(\tilde{E}) \}$. Clearly, $\tilde{I} \subseteq \mathcal{G}$. We will show that $\mathcal{G}$ is a monotone class on $\Omega$. Then, by Monotone Class Lemma 12.19, $\mathcal{G} = \sigma(\tilde{I}) = \hat{\mathcal{B}}$.

(i) $\emptyset, \Omega \in \tilde{I} \subseteq \mathcal{G}$. (ii) Fix any $E_1, E_2 \in \mathcal{G}$ with $E_1 \subseteq E_2$. $\forall \tilde{E} \in \tilde{I}$, we have $P((E_2, E_1) \cap \tilde{E}) = P((E_2, E_1) \setminus (E_1 \cap \tilde{E})) = P(E_2 \cap \tilde{E}) - P(E_1 \cap \tilde{E}) = P(E_2)P(\tilde{E}) - P(E_1)P(\tilde{E}) = P(E_2 \setminus E_1)P(\tilde{E})$, where the first equality follows from the fact that $E_1 \cap \tilde{E} \subseteq E_2 \cap \tilde{E}$; the second equality follows from the fact that $E_1, E_2 \in \mathcal{G}$; and the fourth equality follows from the fact that $E_1 \subseteq E_2$. Hence, by the arbitrariness of $\tilde{E} \in \tilde{I}$, we have $P(\bigcup_{i=1}^{\infty} E_i) = P((\bigcup_{i=2}^{\infty} E_i \setminus E_{i-1}) \cap \tilde{E}) = \sum_{i=2}^{\infty} P(E_i \setminus E_{i-1})P(\tilde{E}) + P(E_1)P(\tilde{E}) = P(\tilde{E})(\sum_{i=2}^{\infty} P(E_i \setminus E_{i-1}) + P(E_1)) = P(\tilde{E}) \mathcal{P}(\bigcup_{i=1}^{\infty} E_i)$, where the first two equalities follow from Proposition 2.5; the third equality follows from countable additivity of measure; the fourth equality follows from (ii); and the last equality follows from countable additivity of measure. By the arbitrariness of $\tilde{E} \in \tilde{I}$, we have $\bigcup_{i=1}^{\infty} E_i \in \mathcal{G}$. This proves that $\mathcal{G}$ is a monotone class on $\Omega$. 

726

CHAPTER 14. PROBABILITY THEORY
This proves that $\forall \hat{E} \in \mathcal{B}$ and $\forall \check{E} \in \mathcal{B}$, we have $P(\hat{E} \cap \check{E}) = P(\hat{E})P(\check{E})$. Hence, $\mathcal{B}$ and $\mathcal{B}$ are independent $\sigma$-algebras.

This completes the proof of the proposition. \hfill $\Box$

**Proposition 14.11** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{A}$ be a separable Banach space over $\mathbb{K}$, $x \in L_1(\Omega, \mathcal{A})$, $(x_n)_{n=1}^{\infty} \subseteq L_1(\Omega, \mathcal{A})$, $\mathcal{B} \subseteq \mathcal{B}$ be a $\sigma$-algebra on $\Omega$, $\Omega := (\Omega, \mathcal{B}, \mathcal{P} := P|_\mathcal{B})$, $f \in E(x|\mathcal{B})$, and $f_n \in E(x_n|\mathcal{B})$, $\forall n \in \mathbb{N}$. Then, the following statements hold.

(a) $E(f) = E(x)$.

(b) Let $x$ be $\mathcal{B}$-measurable. Then $[x] = E(x|\mathcal{B}) \in L_1(\Omega, \mathcal{A})$.

(c) (Linearity) Let $\alpha, \beta \in \mathbb{K}$. Then $E(\alpha x_1 + \beta x_2|\mathcal{B}) = \alpha E(x_1|\mathcal{B}) + \beta E(x_2|\mathcal{B})$.

(d) Let $F \subseteq \mathcal{A}$ be nonempty closed convex set, and $x : \Omega \to F$. Then, $f(\omega) \in F$ a.e. $\omega \in \Omega$.

(e) (Monotone Convergence) Let $\mathcal{A} = \mathbb{R}$, $0 \leq x_n(\omega) \leq x_{n+1}(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, and $\lim_{n \in \mathbb{N}} x_n(\omega) = x(\omega)$ a.e. $\omega \in \Omega$. Then $f_n \leq f_{n+1}$ a.e. in $\Omega$ and $\lim_{n \in \mathbb{N}} f_n = f$ a.e. in $\Omega$.

(f) (Fatou) Let $\mathcal{A} = \mathbb{R}$, $0 \leq x_n(\omega) < \infty$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$, and $\bar{f} \in E(\liminf_{n \in \mathbb{N}} x_n|\mathcal{B})$. Then $\bar{f} \leq \liminf_{n \in \mathbb{N}} f_n$ a.e. in $\Omega$.

(g) (Dominated Convergence) Let $g_n : \Omega \to [0, \infty) \subseteq \mathbb{R}$ be integrable over $\Omega$, $\forall n \in \mathbb{N}$, $g : \Omega \to [0, \infty) \subseteq \mathbb{R}$ be integrable over $\Omega$, that satisfies

(i) $\lim_{n \in \mathbb{N}} x_n = \bar{x}$ a.e. in $\Omega$, where $\bar{x} : \Omega \to \mathcal{A}$ is $\mathcal{B}$-measurable;

(ii) $\mathcal{P} \circ x_n(\omega) \leq g_n(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$;

(iii) $\lim_{n \in \mathbb{N}} g_n = g$ a.e. in $\Omega$, and $E(g) = \lim_{n \in \mathbb{N}} E(g_n) < \infty$.

Then, $\bar{x} \in L_1(\Omega, \mathcal{A})$, $E(\bar{x}|\mathcal{B})$ exists, $\lim_{n \in \mathbb{N}} f_n = \bar{f}$ a.e. in $\Omega$, $\lim_{n \in \mathbb{N}} \int_{\Omega} \mathcal{P} \circ f_n \, d\mathcal{P} = \int_{\Omega} \mathcal{P} \circ \bar{f} \, d\mathcal{P} < \infty$, and $\lim_{n \in \mathbb{N}} \int_{\Omega} \mathcal{P} \circ (f_n - \bar{f}) \, d\mathcal{P} = 0$, where $\bar{f} \in E(\bar{x}|\mathcal{B})$.

(h) Let $\mathcal{Z}$ be a separable Banach space over $\mathbb{K}$, $\mathcal{W} \subseteq \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be a separable subspace, $A : \Omega \to \mathcal{W}$ be $\mathcal{B}$-measurable, and $\|A(\cdot)\|_{B(\mathcal{Y}, \mathcal{Z})} \|x(\cdot)\|_\mathcal{Y} : \Omega \to [0, \infty) \subseteq \mathbb{R}$ be integrable over $\Omega$. Then, $E(Ax|\mathcal{B})$ exists and $Af \in E(Ax|\mathcal{B})$.

(i) (Jensen’s Inequality) Let $\mathcal{K} = \mathbb{R}$, $F \subseteq \mathcal{A}$ be a nonempty closed convex set, $x : \Omega \to F$, $c : F \to \mathbb{R}$ be a convex functional with its epigraph $\{c, F\}$ being closed, $c \circ x : \Omega \to \mathbb{R}$ being absolutely integrable over $\Omega$, and $\bar{f} \in E(c \circ x|\mathcal{B})$. Then $c \circ f \leq \bar{f}$ a.e. in $\Omega$. 

\[ \int_{\Omega} c(\bar{f}) \, dP = \int_{\Omega} c \circ f \, dP \leq \int_{\Omega} \bar{f} \, dP. \]
(j) (Tower Property) Let \( \mathcal{B} \subseteq \mathcal{B} \) be a \( \sigma \)-algebra on \( \Omega \). Then \( E(f|\mathcal{B}) = E(x|\mathcal{B}) \), whenever one of them exists.

(k) (Role of Independence) Let \( \mathcal{B} \subseteq \mathcal{B} \) be a \( \sigma \)-algebra on \( \Omega \), \( \mathcal{B} \) and \( \sigma(\mathcal{B} \cup \sigma(x)) \) are independent. Then \( f \in E(x|\sigma(\mathcal{B} \cup \mathcal{B})) \).

**Proof**

(a) \( E(f) = \int f \ d\nu = \int f \ d\hat{\nu} = \int x \ d\nu = E(x) \), where the first equality follows from the definition of \( E \); the second equality follows from \( f \) is \( \mathcal{B} \)-measurable and Proposition 11.72; the third equality follows from the definition of \( E(x|\mathcal{B}) \) and the fact \( \Omega \in \mathcal{B} \); and the last equality follows from the definition of \( E \).

(b) Let \( \nu \) and \( \hat{\nu} \) be the \( \mathcal{Y} \)-valued measures defined in Definition 14.7. For every \( \forall E \in \mathcal{B} \), we have \( \mathcal{P} \circ \hat{\nu}(E) = \int_E \mathcal{P} \circ x \ d\hat{\nu} = \int_E \mathcal{P} \circ x \ d\nu < \infty \), where the first equality follows from Proposition 11.116; and the second equality follows from the fact that \( \mathcal{P} \circ x \) is \( \mathcal{B} \)-measurable and Proposition 11.72. Furthermore, we have \( \nu(E) = \int_E x \ d\nu = \int_E x \ d\hat{\nu} = \hat{\nu}(E) \), where the second equality follows from the fact that \( x \) is \( \mathcal{B} \)-measurable and Proposition 11.72. By Proposition 14.8, we have \( x \in E(x|\mathcal{B}) \).

(c) This is a direct consequence of (iii) of Proposition 11.168.

(d) By the assumption, \( x(\omega) \in F, \forall \omega \in \Omega \). By the first three paragraphs of the proof for Jensen’s Inequality Theorem 11.98, which clearly applies to the case when \( \mathcal{Y} \) is a complex Banach space as well, we have \( \forall E \in \mathcal{B} \) with \( \hat{\nu}(E) > 0 \), \( \frac{1}{\hat{\nu}(E)} \int_E f \ d\hat{\nu} = \frac{1}{\hat{\nu}(E)} \int_E x \ d\hat{\nu} \in F \), where the equality follows from \( f \in E(x|\mathcal{B}) \). Hence, \( \frac{1}{\hat{\nu}(E)} \int_E f \ d\hat{\nu} \in F \), \( \forall E \in \mathcal{B} \) with \( \hat{\nu}(E) > 0 \). Suppose that the result does not hold. Then, we have \( E := \{ \omega \in \Omega \ | \ f(\omega) \notin F \} \in \mathcal{B} \) and \( \hat{\nu}(E) > 0 \). Note that \( \mathcal{Y} \setminus F \) is an open set. By Proposition 4.38, \( \mathcal{Y} \setminus F \) is separable. By Proposition 4.4, \( \mathcal{Y} \setminus F \) is second countable. Then, \( \exists (B_n)_{n=1}^{\infty} \) such that \( B_n \) is an open ball in \( \mathcal{Y}, \mathcal{B}_n \subseteq \mathcal{Y} \setminus F, \forall n \in \mathcal{N}, \) and \( \mathcal{Y} \setminus F = \bigcup_{n=1}^{\infty} B_n \). Let \( E_n := f_{inv}(B_n) \in \mathcal{B} \). Then, \( \hat{\nu}(E) = \hat{\nu}(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \hat{\nu}(E_n) \). Since \( \hat{\nu}(E) > 0 \), we have \( \exists n_0 \in \mathcal{N} \) such that \( \hat{\nu}(E_{n_0}) > 0 \). Then, \( f(\omega) \in B_{n_0}, \forall \omega \in E_{n_0} \). By the first three paragraphs of the proof for Jensen’s Inequality Theorem 11.98 (applying to possibly complex Banach spaces), \( \frac{1}{\hat{\nu}(E_{n_0})} \int_{E_{n_0}} f \ d\hat{\nu} \in B_{n_0} \) since \( B_{n_0} \) is a nonempty convex set. This contradicts our earlier conclusion that \( \frac{1}{\hat{\nu}(E_{n_0})} \int_{E_{n_0}} f \ d\hat{\nu} \in F \subseteq \mathcal{Y} \setminus B_{n_0} \). Hence, the hypothesis does not hold. Hence, (d) holds.

(e) By (c), we have \( f_{n+1} - f_n \in E(x_{n+1} - x_n|\mathcal{B}), \forall n \in \mathcal{N} \). since \( x_{n+1} - x_n(\omega) \geq 0, \forall \omega \in \Omega, \forall n \in \mathcal{N} \), then, by (d), \( f_{n+1} - f_n \geq 0 \) a.e. in \( \Omega \), \( \forall n \in \mathcal{N} \). Then, \( f_n \leq f_{n+1} \) a.e. in \( \Omega \). \( \int \mathcal{P} \circ (f - f_n) \ d\nu = \int \mathcal{P} \circ (f_n - f) \ d\nu \leq \int \mathcal{P} \circ (x_n - x) \ d\nu = \int (x_n - x) \ d\nu \), where the first equality follows from Proposition 11.72; and the inequality follows from (14.2b). By Monotone Convergence Theorem 11.81, \( \int x \ d\nu = \lim_{n \in \mathcal{N}} \int x_n \ d\nu \). By Definition 14.7, \( f_n, f \in L_1(\Omega, \mathbb{R}), \forall n \in \mathcal{N} \). Then, we have \( \lim_{n \in \mathcal{N}} f_n = f \) in \( L_1(\Omega, \mathbb{R}) \). Hence, \( f = \lim_{n \in \mathcal{N}} f_n \) a.e. in \( \Omega \).
14.1. FUNDAMENTAL NOTIONS

\( \forall n \in \mathbb{N}, \) define \( \tilde{x}_n : \Omega \to \mathbb{R} \) by \( \tilde{x}_n(\omega) := \inf_{m \in \mathbb{N}, n \leq m} x_m(\omega) \in [0, \infty) \subset \mathbb{R}, \forall \omega \in \Omega. \) Then, \( \tilde{x}_n(\omega) \leq \tilde{x}_{n+1}(\omega), \forall \omega \in \Omega, \) and \( \tilde{x}_n \) is integrable over \( \Omega, \forall n \in \mathbb{N}. \) By monotonicity, we have \( \lim_{n \in \mathbb{N}} \tilde{x}_n = \liminf_{n \in \mathbb{N}} x_n =: \bar{x} : \Omega \to [0, \infty] \subset \mathbb{R}. \) Since \( \bar{f} \in E(\bar{x}|\mathcal{B}) \) exists, then \( \tilde{\bar{x}} \in L_1(\Omega, \mathcal{B}). \) Thus, we may have \( \tilde{x} \in [\tilde{\bar{x}}] \) where \( \tilde{x} \in L_1(\Omega, \mathbb{R}). \) This leads to \( \liminf_{n \in \mathbb{N}} x_n = \tilde{x} \) a.e. in \( \Omega. \) By \( \mathcal{Y} = \mathbb{R} \) and Radon-Nikodym Theorem 11.171, \( E(\tilde{x}_n|\mathcal{B}) \) exists, \( \forall n \in \mathbb{N}. \) By (e), we have \( \tilde{f} = \liminf_{n \in \mathbb{N}} f_n \) a.e. in \( \Omega, \) where \( f_n \in E(\tilde{x}_n|\mathcal{B}), \forall n \in \mathbb{N}. \) By (e), we have \( \tilde{f} \leq f_m \) a.e. in \( \Omega, \forall n, m \in \mathbb{N} \) with \( n \leq m. \) Hence, \( \tilde{f} \leq \liminf_{n \in \mathbb{N}} f_n \) a.e. in \( \Omega. \)

By Lebesgue Dominated Convergence Theorem 11.91, \( \tilde{x} \) is absolutely integrable over \( \Omega, 0 \leq \int_\Omega \mathcal{P} \circ \tilde{x} \, d\mathcal{P} = \lim_{n \in \mathbb{N}} \int_\Omega \mathcal{P} \circ x_n \, d\mathcal{P} \leq \lim_{n \in \mathbb{N}} E(g_n) = E(g) < \infty, \) and \( E(\tilde{x}) = \lim_{n \in \mathbb{N}} E(x_n) \in \mathcal{Y}. \) This implies that \( \tilde{x} \in L_1(\Omega, \mathcal{Y}) \) and \( \tilde{x} - x_n \in L_1(\Omega, \mathcal{Y}). \) Note that \( \mathcal{P} \circ (x_n - \tilde{x}) \leq g_n + g \) a.e. in \( \Omega, \lim_{n \in \mathbb{N}} (x_n - \tilde{x}) = 0 \) a.e. in \( \Omega. \) Then, by Lebesgue Dominated Convergence Theorem 11.91, we have \( \lim_{n \in \mathbb{N}} \int_\Omega \mathcal{P} \circ (x_n - \tilde{x}) \, d\mathcal{P} = 0. \) This yields that \( \liminf_{n \in \mathbb{N}} x_n = \tilde{x} \in L_1(\Omega, \mathcal{Y}), \) and \( (|x_n|)_{n=1}^{\infty} \subseteq L_1(\Omega, \mathcal{Y}) \) is a Cauchy sequence. By (c), \( f_n - f_m \in E((x_n - x_m)|\mathcal{B}) \). By (14.2b), \( \int_\Omega \mathcal{P} \circ (f_n - f_m) \, d\mathcal{P} \leq \int_\Omega \mathcal{P} \circ (x_n - x_m) \, d\mathcal{P}, \forall n, m \in \mathbb{N}. \) Then, \( (|f_n|)_{n=1}^{\infty} \subseteq L_1(\Omega, \mathcal{Y}) \) is a Cauchy sequence. By Example 11.179, \( L_1(\Omega, \mathcal{Y}) \) is a Banach space. Then, \( \exists f \in L_1(\Omega, \mathcal{Y}) \) such that \( \lim_{n \in \mathbb{N}} f_n = f \) in \( L_1(\Omega, \mathcal{Y}). \) Hence, \( \lim_{n \in \mathbb{N}} f_n = f \) a.e. in \( \Omega \) and \( \lim_{n \in \mathbb{N}} \int_\Omega \mathcal{P} \circ (f_n - f) \, d\mathcal{P} = 0. \) By the continuity of norm, we have \( \lim_{n \in \mathbb{N}} \int_\Omega \mathcal{P} \circ f_n \, d\mathcal{P} = \int_\Omega \mathcal{P} \circ f \, d\mathcal{P}. \)

Now, we will show that \( \tilde{f} \in E(\tilde{x}|\mathcal{B}). \) \( \forall E \in \tilde{\mathcal{B}}, \) we have \( \int_E \tilde{f} \, d\mathcal{B} = \lim_{n \in \mathbb{N}} \int_E f_n \, d\mathcal{B} = \lim_{n \in \mathbb{N}} \int_E x_n \, d\mathcal{B} = \int_E \tilde{x} \, d\mathcal{B}, \) where the first equality follows from the fact that \( \tilde{f} \) is \( \mathcal{B} \)-measurable and Proposition 11.72; the second equality follows from Lebesgue Dominated Convergence Theorem 11.91; the third equality follows from \( f_n \in E(x_n|\mathcal{B}); \) and the fourth equality follows from Lebesgue Dominated Convergence Theorem 11.91. By Proposition 14.8, we have \( f \in E(\tilde{x}|\mathcal{B}). \)

(h) First, we will prove the result when \( A \) is a simple function. Let \( A \) admit the canonical representation \( \sum_{i=1}^{m} A_i \chi_{E_i, \Omega}, \) where \( m \in \mathbb{Z}_+, A_i \)'s are distinct elements of \( \mathcal{W} \) that none equal to \( \emptyset_{\mathcal{B}(\mathcal{Y}, \mathcal{Z})}, E_i \)'s are pairwise disjoint and nonempty. Since \( A \) is \( \mathcal{B} \)-measurable, we have \( E_i \in \tilde{\mathcal{B}}, i = 1, \ldots, m. \) By Propositions 7.65, 11.39, and 11.38, \( A \) is \( \mathcal{B} \)-measurable, and \( A \) is \( \mathcal{B} \)-measurable. By \( A \) being a simple function, we have \( Af \in L_1(\Omega, \mathcal{Y}). \) By the assumption, \( Af \in L_1(\Omega, \mathcal{Y}). \) \( \forall E \in \tilde{\mathcal{B}}, \) \( \int_E Af \, d\mathcal{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i f \, d\mathcal{P} = \sum_{i=1}^{m} A_i \int_{E \cap E_i} f \, d\mathcal{P} = \sum_{i=1}^{m} A_i \int_{E \cap E_i} x \, d\mathcal{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i x \, d\mathcal{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i \chi_{E_i, \Omega} \, d\mathcal{P} = \int_\Omega Af \, d\mathcal{P}, \) where the first equality follows from Proposition 11.72; the second through fourth equalities follow from Proposition 11.92; and the fifth equality follows from \( f \in E(\tilde{x}|\mathcal{B}); \) and the sixth through eighth equalities follow from Proposition 11.92. By Proposition 14.8, we have \( E(Af|\mathcal{B}) \) exists and \( Af \in E(Af|\mathcal{B}). \)
Next, we prove the general case. By Proposition 11.66, there exists a sequence of simple functions \((\phi_n)_{n=1}^{\infty}\) with \(\phi_n : \Omega \to \mathcal{W}\) being \(\mathcal{B}\)-measurable, such that \(\lim_{n \to \infty} \phi_n = A\) a.e. in \(\Omega\) and \(\|\phi_n(\omega)\|_{B(y,z)} \leq \|A(\omega)\|_{B(y,z)}\), \(\forall \omega \in \Omega, \forall n \in \mathbb{N}\). By Propositions 7.65, 11.39, and 11.38, \(Af\) is \(\mathcal{B}\) measurable, \(\phi_n f\) is \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), and \(Ax\) is \(\mathcal{B}\)-measurable.

Let \(g : \Omega \to [0, \infty) \subset \mathbb{R}\) be given by \(g(\omega) = \|A(\omega)\|_{B(y,z)} \|x(\omega)\|_y\), \(\forall \omega \in \Omega\), and \(g_n = g, \forall n \in \mathbb{N}\). Then, by (g), we have \(E(Ax|\mathcal{B})\) exists and \(\bar{f} = \lim_{n \to \mathbb{N}} \phi_n f = Af\) a.e. in \(\Omega\), where \(\bar{f} \in E(Ax|\mathcal{B})\).

By Proposition 8.25, \(c\) is lower semicontinuous. By Definition 3.14, \(\forall \alpha \in \mathbb{R}, c_{\text{inv}}(\alpha, +\infty) = \{y \in \mathcal{Y} \mid c(y) > \alpha\} \subseteq \mathcal{B}(\mathcal{Y})\), where \(\mathcal{F}\) is the subset topology of \((\mathcal{Y}, \mathcal{O}_y)\) on the set \(F\). By Proposition 11.35, \(c\) is \(\mathcal{B}(\mathcal{Y})\)-measurable. By Proposition 11.34, \(c \circ f\) is \(\mathcal{B}\)-measurable, and \(c \circ x\) is \(\mathcal{B}\)-measurable.

By (d), \(f(\omega) \in F\) a.e. \(\omega \in \tilde{\Omega}\). Thus, \(c \circ f(\omega) \in \mathbb{R}\) a.e. \(\omega \in \tilde{\Omega}\). Fix any \(\omega \in f_{\text{inv}}(F)\). By Proposition 8.32, we have \(c(f(\omega)) = \sup_{y_* \in F_{\text{conj}}} (\langle y_*, f(\omega) \rangle - e_{\text{conj}}(y_*))\), where \(e_{\text{conj}} : F_{\text{conj}} \to \mathbb{R}\) is the conjugate functional to \(e\). Then, \(\exists y_{**} \in F_{\text{conj}}\) such that \(c(f(\omega)) - \epsilon \leq (\langle y_{**}, f(\omega) \rangle - e_{\text{conj}}(y_{**})) \leq (\langle y_{**}, f(\omega) \rangle - (\langle y_{**}, y \rangle + c(y)) = c(y)) \leq (\langle y_{**}, f(\omega) - y \rangle + c(y)), \forall y \in F\), where the second inequality follows from Definition 8.27. Take \(y = x(\omega)\) then \(c(f(\omega)) - \epsilon \leq (\langle y_{**}, f(\omega) - x(\omega) \rangle + c(x(\omega)))\) a.e. in \(\Omega\). Thus, by (d), (h), and (c) and taking conditional expectation of the preceding equation given \(\mathcal{B}\), \(c(f(\omega)) - \epsilon \leq (\langle y_{**}, f(\omega) - f(\omega) \rangle + f(\omega))\) a.e. \(\omega \in \tilde{\Omega}\), which implies that \(c(f(\omega)) - \epsilon \leq f(\omega)\) a.e. \(\omega \in \tilde{\Omega}\). By the arbitrariness of \(\epsilon\), we have \(c \circ f \leq f\) a.e. in \(\tilde{\Omega}\).

By (i) Proposition 14.48, \(E(f|\mathcal{B})\) exists and \(\tilde{f} \in E(f|\mathcal{B})\). Assume that \(E(f|\mathcal{B})\) exists and \(\tilde{f} \in E(f|\mathcal{B})\). Clearly, \(\tilde{f} \in L_1(\tilde{\Omega}, \mathcal{Y})\). \(\forall E \in \mathcal{B}\), we have \(\int_E \tilde{f} d\mathcal{P} = \int_E f d\mathcal{P} = \int_E x d\mathcal{P}\), where the first equality follows from \(\tilde{f} \in E(f|\mathcal{B})\); and the second equality follows from \(f \in E(x|\mathcal{B})\). By the arbitrariness of \(E\) and Proposition 14.8, we have \(\tilde{f} \in E(f|\mathcal{B})\). Hence, we have \(E(x|\mathcal{B}) = E(f|\mathcal{B}) = E(E(x|\mathcal{B})|\mathcal{B})\).

On the other hand, assume \(E(x|\mathcal{B})\) exists and \(\tilde{f} \in E(f|\mathcal{B})\). Clearly, \(\tilde{f} \in L_1(\tilde{\Omega}, \mathcal{Y})\). \(\forall E \in \mathcal{B}\), we have \(\int_E f d\mathcal{P} = \int_E x d\mathcal{P} = \int_E f d\mathcal{P}\), where the first equality follows from \(\tilde{f} \in E(x|\mathcal{B})\); and the second equality follows from \(f \in E(x|\mathcal{B})\). By the arbitrariness of \(E\) and Proposition 14.8, we have \(\tilde{f} \in E(f|\mathcal{B})\). Hence, we have \(E(x|\mathcal{B}) = E(f|\mathcal{B}) = E(E(x|\mathcal{B})|\mathcal{B})\). This establishes (i).

(k) \(\forall \bar{E} \in \mathcal{B}\) and \(\forall \bar{E} \in \hat{\mathcal{B}}\), we have \(\int_{\bar{E} \cap \hat{E}} x d\mathcal{P} = \int_{\bar{E}} x \chi_{\hat{E}, \Omega} \bar{\chi}_{\hat{E}, \Omega} d\mathcal{P} = P(\bar{E}) \int_{\hat{E}} x \chi_{\hat{E}, \Omega} d\mathcal{P} = P(\bar{E}) \int_{\hat{E}} f d\mathcal{P} = P(\bar{E}) \int_{\hat{E}} f \chi_{\hat{E}, \Omega} d\mathcal{P} = \int_{\bar{E} \cap \hat{E}} f d\mathcal{P}\), where the first equality follows from Proposition 11.92; the second equality follows from the independence of \(\hat{\mathcal{B}}\) and \(\sigma(\bar{\mathcal{B}} \cup \sigma(x))\) and Proposition 14.9; the fourth equality follows from \(f \in E(x|\mathcal{B})\); and the fifth equality follows from Proposition 14.9. By assumption, \(x \in L_1(\tilde{\Omega}, \mathcal{Y})\).
and \( f \in \bar{L}_1(\Omega, \gamma) \). Then, by Proposition 11.116, we may define a finite \( \gamma \)-valued measure \( \nu \) on \((\Omega, \mathcal{B})\) by \( \nu(E) := \int_E x \, dP \), \( \forall E \in \mathcal{B} \); and we may define a finite \( \gamma \)-valued measure \( \tilde{\nu} \) on \((\Omega, \sigma(\mathcal{B} \cup \tilde{\mathcal{B}}))\) by \( \tilde{\nu}(E) := \int_E f \, dP \), \( \forall E \in \sigma(\mathcal{B} \cup \tilde{\mathcal{B}}) \). 

By Proposition 11.107 and Definition 14.1, \( \hat{\nu} := \nu|_{\sigma(\mathcal{B} \cup \tilde{\mathcal{B}})} \) is a finite \( \gamma \)-valued measure on \( \sigma(\mathcal{B} \cup \tilde{\mathcal{B}}) \). We have show in the preceding

\[ \hat{\nu}(E) = \tilde{\nu}(E), \forall E \in \mathcal{F}, \text{ where } \mathcal{F} := \left\{ E \in \mathcal{B} \mid \exists \hat{E} \in \tilde{\mathcal{B}}, \exists \tilde{E} \in \mathcal{B} \exists \mathcal{F} \ni E = \hat{E} \cap \tilde{E} \right\} \].

It is straightforward to show that \( \mathcal{F} \) is a \( \pi \)-system on \( \Omega \) and \( \mathcal{F} \supseteq \mathcal{B} \cup \tilde{\mathcal{B}} \). Therefore, \( \hat{\nu} = \tilde{\nu} \) by Lemma 12.103. Note that \( f \) is \( \sigma(\mathcal{B} \cup \tilde{\mathcal{B}}) \)-measurable. Then, by Proposition 14.8, \( f \in E(x|\sigma(\mathcal{B} \cup \tilde{\mathcal{B}})) \).

This completes the proof of the proposition.

\[ \square \]

**Proposition 14.12** Let \( \lambda_i := (\Omega_i, \mathcal{B}_i, P_i) \) be a probability measure space, \( i = 1, 2, \lambda' := \lambda_1 \times \lambda_2 =: (\Omega, \mathcal{B}, P) \) be the product probability measure space, \( \gamma \) be a finite-dimensional Banach space over \( K \), \( x : \lambda_1 \times \lambda_2 \to \gamma \) be \( \mathcal{B} \)-measurable, \( x \in L_1(\lambda, \gamma) \), and \( \exists M \in [0, \infty) \subset \mathbb{R} \), such that \( \| x(\omega_1, \omega_2) \| \gamma \leq M \), \( \forall (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 = \Omega \). Define \( g : \Omega \to \gamma \) by

\[ g(\hat{\omega}_1, \hat{\omega}_2) := \int_{\Omega_1} x(\hat{\omega}_1, \omega_2) \, dP(\omega_1, \omega_2), \forall (\hat{\omega}_1, \hat{\omega}_2) \in \Omega \].

Then, the following statements hold.

(i) The collection of \( \sigma \)-algebras \( (\tilde{\mathcal{B}}_1, \mathcal{B}_2) \) is independent, where \( \tilde{\mathcal{B}}_i := \{ B_1 \times B_2 \in \mathcal{B} \mid B_j = \Omega_j, j = 1, 2 \text{ with } i \neq j \} \).

(ii) There exists \( g_1 : \Omega_1 \to \gamma \), such that \( g(\omega_1, \omega_2) = g_1(\omega_1), \forall (\omega_1, \omega_2) \in \Omega \), \( g_1 \in \bar{L}_1(\lambda', \gamma) \), and \( g \in E(x|\tilde{\mathcal{B}}_1) \)

**Proof**

(i) This holds by Proposition 12.21.

(ii) Clearly, \( g(\hat{\omega}_1, \hat{\omega}_2) = g_1(\hat{\omega}_1), \forall (\hat{\omega}_1, \hat{\omega}_2) \in \Omega \), where \( g_1(\hat{\omega}_1) := \int_{\Omega_1} x(\hat{\omega}_1, \omega_2) \, dP(\omega_1, \omega_2) = \int_{\Omega_1} \int_{\Omega_2} x(\hat{\omega}_1, \omega_2) \, dP_1(\omega_1) \, dP_2(\omega_2) = \int_{\Omega_2} x(\hat{\omega}_1, \omega_2) \, dP_1(\omega_1) \, dP_2(\omega_2) = \int_{\Omega_2} x(\hat{\omega}_1, \omega_2) \, dP_2(\omega_2) \, dP_1(\omega_1) \in \gamma, \forall \hat{\omega}_1 \in \Omega_1 \), where the first equality follows from Proposition 12.28 and Fubini’s Theorem 12.31 and the assumptions; the second equality follows from Propositions 11.92; the third equality follows from 11.75; and the belong to follows from Proposition 11.92. By Fubini’s Theorem 12.31, \( g_1 \) is \( B_1 \)-measurable and \( g_1 \in \bar{L}_1(\lambda_1, \gamma) \). Hence, \( g \) is \( \tilde{\mathcal{B}}_1 \)-measurable and \( g \in \bar{L}_1(\lambda, \gamma) \). \( \forall \in \tilde{\mathcal{B}}_1 \), we have \( E = E_1 \times \Omega_2 \), where \( E_1 \in \mathcal{B}_1 \). Then,

\[
\int_E x \, dP = \int_{\Omega_1 \times \Omega_2} x(\omega_1, \omega_2) \chi_{E_1 \times \Omega_2}(\omega_1, \omega_2) \, dP(\omega_1, \omega_2) \, dP_1(\omega_1)
\]

\[
= \int_{\Omega_1} \chi_{E_1 \times \Omega_2}(\omega_1) \int_{\Omega_2} x(\omega_1, \omega_2) \, dP_2(\omega_2) \, dP_1(\omega_1)
\]

\[
= \int_{\Omega_1} \chi_{E_1 \times \Omega_2}(\omega_1) g_1(\omega_1) \, dP_1(\omega_1) = \int g_1 \, dP_1
\]

\[
= \int \int_{\Omega_2} g(\omega_1, \omega_2) \, dP_2(\omega_2) \, dP_1(\omega_1) = \int g \, dP
\]
where the first equality follows from Proposition 11.92; the second equality follows from Fubini's Theorem 12.31; the third equality follows from Proposition 11.92; the fourth equality follows from the expression for $g_1$; and the last equality follows from Fubini's Theorem 12.31. By the arbitrariness of $E$ and Proposition 14.8, we have $g \in E(x|B_1)$.

This completes the proof of the proposition. □

**Theorem 14.13 (Fundamental Theorem on Modeling)** Let $\Lambda$ be a set, $\mathcal{X}_\alpha := (\mathcal{X}_\alpha, \mathcal{B}_\alpha, \mathbb{P}_\alpha)$ a probability measure space, $\forall \alpha \in \Lambda$, $\mathcal{F} := \{\Pi_{\alpha \in \Lambda} B_\alpha \subseteq X := \Pi_{\alpha \in \Lambda} X_\alpha | B_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda, \text{ and } B_\alpha = X_\alpha, \forall \alpha \in \Lambda \text{ except a finite number of } \alpha \text{'s} \}$, and $\mu : \mathcal{F} \to [0,1] \subset \mathbb{R}_+$ be defined by $\mu(\Pi_{\alpha \in \Lambda} B_\alpha) = \Pi_{\alpha \in \Lambda} \mathbb{P}_\alpha(B_\alpha) = \Pi_{\alpha \in \Lambda, B_\alpha \subseteq X_\alpha} \mathbb{P}_\alpha(B_\alpha), \forall \Pi_{\alpha \in \Lambda} B_\alpha \in \mathcal{F}$. Then, $\mathcal{F}$ is a semi-algebra on $X$ and there exists a unique probability measure space $(X, \mathcal{B} := \sigma(\mathcal{F}), \mathbb{P})$ such that $\mathbb{P}|_\mathcal{F} = \mu$. Furthermore, the collection of semi-algebras $(\mathcal{B}_\alpha)_{\alpha \in \Lambda}$ is independent, where $\mathcal{B}_\alpha := \{\Pi_{\alpha \in \Lambda} B_\lambda \subseteq \mathcal{F} | B_\lambda = X_\lambda, \forall \lambda \in \Lambda \text{ with } \lambda \neq \alpha \}, \forall \alpha \in \Lambda$.

**Proof** Note that the result holds if $\Lambda$ is a finite set, by Proposition 12.21. We need only to prove the result when $\Lambda$ is not a finite set.

We will first show that $\mathcal{F}$ is a semi-algebra on a finite set. Clearly, $\emptyset, X \in \mathcal{F}$. $\forall B := \Pi_{\alpha \in \Lambda} B_\alpha, C := \Pi_{\alpha \in \Lambda} C_\alpha \in \mathcal{F}, B \cap C = \Pi_{\alpha \in \Lambda} (B_\alpha \cap C_\alpha)$. Clearly, $B_\alpha \cap C_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda$, since $B_\alpha, C_\alpha \in \mathcal{B}_\alpha$. Furthermore, $B_\alpha \cap C_\alpha = X_\alpha, \forall \alpha \in \Lambda \text{ except a finite number of } \alpha \text{'s}$, since $B_\alpha = X_\alpha, \forall \alpha \in \Lambda \text{ except for a finite number of } \alpha \text{'s}$ and $C_\alpha = X_\alpha, \forall \alpha \in \Lambda \text{ except for a finite number of } \alpha \text{'s}$. Hence, $B \cap C \in \mathcal{F}$. Fix any $B := \Pi_{\alpha \in \Lambda} B_\alpha \in \mathcal{F}$. Then, $B_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda$, and $B_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_B, \text{ where } \Lambda_B = \{\alpha_1, \ldots, \alpha_n\} \subseteq \Lambda, n \in \mathbb{N}$. Then, $X \setminus B = \bigcup_{j=1}^n \bigcup_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} C_{j,i_1,\ldots,i_j,\alpha}$ and $C_{j,i_1,\ldots,i_j,\alpha} = \{X_\alpha \setminus B_\alpha | \alpha \in \{\alpha_1, \ldots, \alpha_j\}\}$. Clearly, $C_{j,i_1,\ldots,i_j,\alpha}$'s are pairwise disjoint. Hence, $\mathcal{F}$ is a semi-algebra on $X$.

We will next show that $\mu$ satisfies the assumptions of Proposition 11.32.

(i) Fix any $B \in \mathcal{F}$, any $n \in \mathbb{N}_+$, and any pairwise disjoint $(C_1)_{i=1}^n \subseteq \mathcal{F}$ with $B = \bigcup_{i=1}^n C_i$. Let $B := \Pi_{\alpha \in \Lambda} B_\alpha, \text{ where } B_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda, \text{ and } B_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_B, \text{ with } \Lambda_B \subseteq \Lambda$ being a finite set. Fix $i \in \{1, \ldots, n\}$, let $C_i := \Pi_{\alpha \in \Lambda} C_{i,\alpha}$, where $C_{i,\alpha} \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda, \text{ and } C_{i,\alpha} = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_B$ being a finite set. Then, let $\Lambda := \Lambda_B \cup (\bigcup_{i=1}^n \Lambda_i), \text{ which is a finite subset of } \Lambda$. Then, $B_\alpha = X_\alpha = C_{i,\alpha}, \forall \alpha \in \Lambda \setminus \Lambda_B, \text{ with } \Lambda_B \subseteq \Lambda$ being a finite set. Then, $\mu(B) := \Pi_{\alpha \in \Lambda} \mathbb{P}_\alpha(B_\alpha) = \sum_{i=1}^n \Pi_{\alpha \in \Lambda} \mathbb{P}_\alpha(C_{i,\alpha}) = \sum_{i=1}^n \mu(C_i)$, where the second equality follows from Proposition 12.21.

(ii) Fix any $B \in \mathcal{F}$ and any pairwise disjoint $(C_i)_{i=1}^\infty \subseteq \mathcal{F}$ with $B = \bigcup_{i=1}^\infty C_i$. Let $B := \Pi_{\alpha \in \Lambda} B_\alpha, \text{ where } B_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda, \text{ and } B_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_B$ being a finite set. Fix $i \in \mathbb{N}$, let $C_i := \Pi_{\alpha \in \Lambda} C_{i,\alpha}$, where $C_{i,\alpha} \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda, \text{ and } C_{i,\alpha} = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_i$ being a finite set. Then, $\Lambda := \Lambda_B \cup (\bigcup_{i=1}^\infty \Lambda_i), \text{ which is a countable subset of } \Lambda$. Then, $B_\alpha = X_\alpha = C_{i,\alpha}, \forall \alpha \in \mathbb{N}, \forall \alpha \in \Lambda \setminus \Lambda.$
14.1. FUNDAMENTAL NOTIONS

We will show that $\mu(B) \leq \sum_{i=1}^{\infty} \mu(C_i)$ by an argument of contradiction. Suppose $\mu(B) > \sum_{i=1}^{\infty} \mu(C_i)$. Clearly, $0 \leq \mu(B) \leq 1$. Then, \( \exists \varepsilon_0 \in (0, \infty) \subset \mathbb{R} \) such that $\mu(B) > \varepsilon_0 + \sum_{i=1}^{\infty} \mu(C_i), \forall n \in \mathbb{N}$. Let $\hat{\Lambda}_n := \Lambda_B \cup (\bigcup_{i=1}^{n} \Lambda_i), \forall n \in \mathbb{N}$, which is a finite subset of $\Lambda$. Let the elements of $\hat{\Lambda} := \{\alpha_1, \ldots, \alpha_{r(1)}, \alpha_{r(1)+1}, \ldots, \alpha_{r(2)}, \ldots\}$, where the $r(i)'s$ are defined such that $\hat{\Lambda}_n = \{\alpha_1, \ldots, \alpha_{r(1)}, \ldots, \alpha_{r(n-1)+1}, \ldots, \alpha_{r(n)}\}, \forall n \in \mathbb{N}$.

Fix any $n \in \mathbb{N}$. By Proposition 12.21, we have $Y_n := \prod_{\alpha \in \hat{\Lambda}_n} X_\alpha := (Y_{n,1}, \mathcal{B}_{n,1}, P_{n,1})$ is a probability measure space, on which the collection of $\sigma$-algebras $\{\mathcal{B}_{\alpha, n, 1}\}_{\alpha \in \hat{\Lambda}_n}$ is independent, where $\mathcal{B}_{\alpha, n, 1} := \{\prod_{\beta \in \hat{\Lambda}_n} \hat{\mathcal{B}}_\beta \mid \hat{\mathcal{B}}_\alpha = X_\alpha, \forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha\}\}$, $\forall \alpha \in \hat{\Lambda}_n$. Let $\hat{B}_{n,1} := \prod_{\alpha \in \hat{\Lambda}_n} \mathcal{B}_\alpha$ and $\hat{C}_{i,n,1} := \prod_{\alpha \in \hat{\Lambda}_n} \mathcal{C}_{i,\alpha}, \forall i \in \{1, \ldots, n\}$. Clearly, $\hat{B}_{n,1} \supseteq \bigcup_{i=1}^{n} \hat{C}_{i,n,1}$ and the sets in the union are pairwise disjoint, and all sets involved are in $\mathcal{B}_{n,1}$. Then, we have $P_{n,1}(\hat{B}_{n,1}) = \mu(B) > \varepsilon_0 + \sum_{i=1}^{n} \mu(C_i) = \varepsilon_0 + \sum_{i=1}^{n} P_{n,1}(\hat{C}_{i,n,1})$, where the equalities follow from the fact that $\Lambda_B \subseteq \Lambda_n$ and $\Lambda_i \subseteq \Lambda_n, \forall i \in \{1, \ldots, n\}$. Define, $H_{n,1} := \hat{B}_{n,1} \setminus \bigcup_{i=1}^{n} \hat{C}_{i,n,1}$. Then, $P_{n,1}(H_{n,1}) > \varepsilon_0$. Define $H_{n,1} := \hat{B}_{n,1} \setminus \bigcup_{i=1}^{n} \hat{C}_{i,n,1}$. Clearly, $H_{n,1} = H_{n,1} \times \prod_{\alpha \in \hat{\Lambda}_n \setminus \hat{\Lambda}_n} X_\alpha$. Obviously, $H_{n,1} \supseteq H_{n+1,1}$ and $\bigcap_{n=1}^{\infty} H_{n,1} = \emptyset$.

By Radon-Nikodym Theorem 11.169 and Definition 14.7, $E_{n,1}(\chi_{H_{n,1}, Y_{n,1}} | \mathcal{B}_{\alpha, n, 1})$ exists, where $\mathcal{B}_{n,1}$ is with respect to measure $P_{n,1}$. Define $g_{n,1} : Y_{n,1} = \prod_{\alpha \in \hat{\Lambda}_n} X_\alpha \rightarrow [0, 1] \subset \mathbb{R}$ by, $\forall (\hat{\omega}_1, \ldots, \hat{\omega}_{r(\alpha)}) \in Y_{n,1}$,

$$g_{n,1}(\hat{\omega}_1, \ldots, \hat{\omega}_{r(\alpha)}) = \int_{Y_{n,1}} \chi_{H_{n,1}, Y_{n,1}}(\hat{\omega}_1, \omega_2, \ldots, \omega_{r(\alpha)}) \, dP_{n,1}(\omega_1, \ldots, \omega_{r(\alpha)})$$

By Proposition 14.12, $g_{n,1} \in E_{n,1}(\chi_{H_{n,1}, Y_{n,1}} | \mathcal{B}_{\alpha, n, 1}), g_{n,1}(\hat{\omega}_1, \ldots, \hat{\omega}_{r(\alpha)}) = g_{n,1}(\hat{\omega}_1), \forall (\hat{\omega}_1, \ldots, \hat{\omega}_{r(\alpha)}) \in Y_{n,1}$ and $g_{n,1} : X_\alpha \rightarrow [0, 1] \subset \mathbb{R}$ is $\mathcal{B}_{\alpha, n, 1}$ measurable. Since $P_{n,1}(H_{n,1}) > \varepsilon_0$, then $E_{\alpha,1}(g_{n,1}) = E_{n,1}(E_{n,1}(\chi_{H_{n,1}, Y_{n,1}} | \mathcal{B}_{\alpha, n, 1})) = E_{n,1}(\chi_{H_{n,1}, Y_{n,1}}) = P_{n,1}(H_{n,1}) > \varepsilon_0$, where $E_{\alpha,1}$ is with respect to $P_{\alpha,1}$; and the second equality follows from (a) of Proposition 14.11. This leads to the following inequalities

$$\varepsilon_0 < E_{\alpha,1}(g_{n,1}) \leq P_{\alpha,1}(\{\omega_1 \in X_\alpha \mid g_{n,1}(\omega_1) > 2^{-1} \varepsilon_0\}) + 2^{-1} \varepsilon_0 P_{\alpha,1}(\{\omega_1 \in X_\alpha \mid g_{n,1}(\omega_1) \leq 2^{-1} \varepsilon_0\}) \leq P_{\alpha,1}(\{\omega_1 \in X_\alpha \mid g_{n,1}(\omega_1) > 2^{-1} \varepsilon_0\}) + 2^{-1} \varepsilon_0$$

Thus, we have $P_{\alpha,1}(\{\omega_1 \in X_\alpha \mid g_{n,1}(\omega_1) > 2^{-1} \varepsilon_0\}) > 2^{-1} \varepsilon_0$.

Now, since $H_{n,1} \supseteq H_{n+1,1}$, by the definition of $g_{n,1}$, Proposition 11.92, and Fubini's Theorem 12.31, we have $0 \leq g_{n+1,1}(\omega_1) \leq g_{n,1}(\omega_1), \forall \omega_1 \in X_\alpha$, (which requires us to work in the product measure space $\mathcal{Y}_{n+1,1}$ that includes both $H_{n,1} \times \prod_{i=r(n)+1}^{n} X_\alpha$ and $H_{n+1,1}$). Therefore, we have
\( P_{\alpha_1}(\Gamma_{n,1}) > 2^{-1}\epsilon_0 \), where \( \Gamma_{n,1} := \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \tilde{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \} \). Furthermore, \( \Gamma_{n,1} \supseteq \Gamma_{n+1,1} \). By Proposition 11.5, we have
\[
P_{\alpha_1}(\bigcap_{n=1}^{\infty} \Gamma_{n,1}) = \lim_{n \to \infty} P_{\alpha_1}(\Gamma_{n,1}) \geq 2^{-1}\epsilon_0 > 0.
\]
Then, there exists \( \omega_{\alpha_1} \in X_{\alpha_1} \) such that \( \omega_{\alpha_1} \in \bigcap_{n=1}^{\infty} \Gamma_{n,1} \). This implies that \( \tilde{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \), \( \forall n \in \mathbb{N} \). By the definition of \( \tilde{g}_{n,1} \), we have
\[
\int_{Y_{n,1}} \chi_{\hat{H}_{n,1,Y_{n,1}}}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) d\hat{P}_{n,1}(\omega_{\alpha_1}, \ldots, \omega_{\alpha_{r(n)}}) > 2^{-1}\epsilon_0, \forall n \in \mathbb{N}.
\]
Fix any \( n \in \mathbb{N} \). By Proposition 12.21, we have \( Y_{n,2} := \prod_{\alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\}} X_{\alpha} =: (Y_{n,2}, \tilde{\mathcal{B}}_{n,2}, \tilde{\mathbf{P}}_{n,2}) \) is a probability measure space, on which the collection of \( \sigma \)-algebras \( \left( \tilde{\mathcal{B}}_{\alpha,n,2} \right)_{\alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\}} \) is independent, where \( \tilde{\mathcal{B}}_{\alpha,n,2} := \{ \prod_{\alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\}} B_\alpha \in \tilde{\mathcal{B}}_{n,2} \mid B_\alpha = X_\alpha, \forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha_1, \alpha_2\} \}, \forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\} \). Let \( \tilde{\mathcal{B}}_{n,2} := \{ \prod_{\alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\}} B_\alpha \mid B_\alpha \in \tilde{\mathcal{B}}_{\alpha,n,2}, \forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha_1, \alpha_2\} \}, \forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha_1\} \). Clearly, \( \tilde{\mathcal{B}}_{n,2} \supseteq \bigcup_{i=1}^{n} \tilde{C}_{i,n,2} \) and the sets in the union are pairwise disjoint, and all sets involved are in \( \tilde{\mathcal{B}}_{n,2} \), by Proposition 12.25. Then, we have
\[
\tilde{P}_{n,2}(\tilde{B}_{n,2}) = \int_{Y_{n,2}} \chi_{\tilde{B}_{n,2,Y_{n,2}}}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) d\tilde{P}_{n,2}(\omega_{\alpha_1}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,1}} \chi_{\tilde{B}_{n,1,Y_{n,1}}}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) d\tilde{P}_{n,1}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,1}} \left( \chi_{\hat{H}_{n,1,Y_{n,1}} + \sum_{i=1}^{n} \chi_{\tilde{C}_{i,n,2,Y_{n,1}}}}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) \right) d\tilde{P}_{n,1}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,1}} \sum_{i=1}^{n} \chi_{\tilde{C}_{i,n,1,Y_{n,1}}}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) d\tilde{P}_{n,1}(\omega_{\alpha_1}, \omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) + 2^{-1}\epsilon_0
\]
\[
= 2^{-1}\epsilon_0 + \int_{Y_{n,2}} \sum_{i=1}^{n} \chi_{\tilde{C}_{i,n,2,Y_{n,2}}}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) d\tilde{P}_{n,2}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= 2^{-1}\epsilon_0 + \sum_{i=1}^{n} \tilde{P}_{n,2}(\tilde{C}_{i,n,2})
\]
where the first equality follows from Proposition 11.75; the second equality follows from the definition of \( \tilde{B}_{n,2} \), \( X_{\alpha_1} \) is a probability measure space,
and Fubini’s Theorem 12.31; the third equality follows from the fact that $B_{n,1} = H_{n,1}$ and the sets in the union are pairwise disjoint; the fourth equality follows from Proposition 11.92; the inequality follows from the conclusion of the previous paragraph; the fifth equality follows from the definition of $C_{i,n,2}$'s and Fubini’s Theorem 12.31; and the last equality follows from Proposition 11.75. Thus, we have

$$P_{n,2}(H_{n,2}) := P_{n,2}(B_{n,2} \setminus (\bigcup_{i=1}^{n} C_{i,n,2})) > 2^{-1} \epsilon_0.$$  Define $H_{n,2} := \{(\hat{n},\omega) \in \Lambda \setminus \{\alpha_1\} \in \prod_{\alpha \in \Lambda \setminus \{\alpha_1\}} X_\alpha \mid (\hat{\omega})_{\alpha} \in H_{n,1}, \text{ where } \hat{\omega} = \omega_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_1\}, \hat{\omega}_{\alpha_1} = \omega_{\alpha_1} \}$. It is obvious that $H_{n,2} = \{(\hat{\omega})_{\alpha} \in \Lambda \setminus \{\alpha_1\} \in X_\alpha \mid (\hat{\omega})_{\alpha} = \omega_{\alpha}, \forall \alpha \in \Lambda \setminus \{\alpha_1\}, \hat{\omega}_{\alpha_1} = \omega_{\alpha_1} \}$. Clearly, $H_{n,2} = \hat{H}_{n,2} \times \prod_{\alpha \in \Lambda \setminus \{\alpha_1\}} X_\alpha$. Therefore, we have $H_{n,2} = \hat{H}_{n,2}$.

By Radon-Nikodym Theorem 11.169 and Definition 14.7, $\hat{E}_{n,2}(\chi_{\hat{H}_{n,2},Y_{n,2}}|\hat{B}_{a,2,n,2})$ exists, where $E_{n,2}$ is with respect to measure $P_{n,2}$. Define $g_{n,2} : [0,1] \subset \mathbb{R}$ by

$$g_{n,2}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) = \int_{Y_{n,2}} \chi_{\hat{H}_{n,2},Y_{n,2}}(\omega_{\alpha_2}, \omega_{\alpha_3}, \ldots, \omega_{\alpha_{r(n)}}) d\hat{P}_{n,2}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}})$$

$\forall (\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) \in Y_{n,2}$. By Proposition 14.12, $g_{n,2} \in E_{n,2}(\chi_{\hat{H}_{n,2},Y_{n,2}}|\hat{B}_{a,2,n,2})$, $g_{n,2}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) = \hat{g}_{n,2}(\omega_{\alpha_2})$, $\forall (\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) \in Y_{n,2}$, and $g_{n,2} : X_{\alpha_2} \to [0,1] \subset \mathbb{R}$ is $B_{\alpha_2}$-measurable. Since $\hat{P}_{n,2}(H_{n,2}) > 2^{-1} \epsilon_0$, then $E_{a_2}(g_{n,2}) = E_{n,2}(E_{n,2}(\chi_{\hat{H}_{n,2},Y_{n,2}}|\hat{B}_{a,2,n,2})) = E_{n,2}(\chi_{\hat{H}_{n,2},Y_{n,2}}) = P_{n,2}(H_{n,2}) > 2^{-1} \epsilon_0$, where $E_{a_2}$ is with respect to $P_{a_2}$, and the second equality follows from (a) of Proposition 14.11. This leads to the following inequalities

$$2^{-1} \epsilon_0 < E_{a_2}(g_{n,2}) \leq P_{a_2}(\{\omega_{\alpha_2} \in X_{\alpha_2} \mid \hat{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2} \epsilon_0 \})$$

$$+ 2^{-2} \epsilon_0 P_{a_2}(\{\omega_{\alpha_2} \in X_{\alpha_2} \mid \hat{g}_{n,2}(\omega_{\alpha_2}) \leq 2^{-2} \epsilon_0 \})$$

$$\leq P_{a_2}(\{\omega_{\alpha_2} \in X_{\alpha_2} \mid \hat{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2} \epsilon_0 \}) + 2^{-2} \epsilon_0$$

Thus, we have $P_{a_2}(\{\omega_{\alpha_2} \in X_{\alpha_2} \mid \hat{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2} \epsilon_0 \}) > 2^{-2} \epsilon_0$.

Now, since $H_{n,2} \supseteq H_{n+1,2}$, by the definition of $g_{n,2}$, Proposition 11.92, and Fubini’s Theorem 12.31, we have $0 \leq \hat{g}_{n+1,2}(\omega_{\alpha_2}) \leq \hat{g}_{n,2}(\omega_{\alpha_2})$, $\forall \omega_{\alpha_2} \in X_{\alpha_2}$, (which requires us to work in the product measure space $\gamma_{n+1,2}$ that includes both $\hat{H}_{n,2} \times \prod_{i=n+1}^{\infty} X_{\alpha_i} \text{ and } H_{n+1,2}$). Therefore, we have $P_{a_2}(\Gamma_{n,2}) > 2^{-2} \epsilon_0$, where $\Gamma_{n,2} := \{\omega_{\alpha_2} \in X_{\alpha_2} \mid \hat{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2} \epsilon_0 \}$. Furthermore, $\Gamma_{n,2} \supseteq \Gamma_{n+1,2}$. By Proposition 11.5, we have $P_{a_2}(\bigcap_{n=1}^{\infty} \Gamma_{n,2}) \geq 2^{-2} \epsilon_0 > 0$. Then, $\exists \omega_{\alpha_2} \in X_{\alpha_2}$ such that $\hat{\omega}_{\alpha_2} \in \bigcap_{n=1}^{\infty} \Gamma_{n,2}$. This implies that $\hat{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2} \epsilon_0$, $\forall n \in \mathbb{N}$. By the definition
of \(\bar g_{n,2}\), we have 
\[
\int_{Y_{n,2}} \chi_{H_{n,2},Y_{n,2}}(\omega_{\alpha_2}, \omega_{\alpha_3}, \ldots, \omega_{\alpha_{r(n)}}) \, d\hat P_{n,2}(\omega_{\alpha_2}, \ldots, \omega_{\alpha_{r(n)}}) > 2^{-2} \epsilon_0, \quad \forall n \in \mathbb{N}.
\]

Recursively, assume we have completed index \(l \in \mathbb{N}\). Fix any \(n \in \mathbb{N}\). By Proposition 12.21, we have 
\[
(Y_{n,l+1}, \tilde B_{n,l+1}, P_{n,l+1}) \text{ is a probability measure space, on which the collection of } \sigma\text{-algebras } \left(\tilde B_{\alpha,n,l+1}\right)_{\alpha \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\}} \text{ is independent, where}
\]
\[
\tilde B_{\alpha,n,l+1} := \left\{ \prod_{\alpha \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\}} \tilde B_\alpha \in \tilde B_{n,l+1} \mid \tilde B_\alpha = X_\alpha, \forall \alpha \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\} \right\}, \quad \forall \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\}.
\]

Let 
\[
\tilde B_{n,l+1} := \left\{ \prod_{\alpha \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\}} \tilde B_\alpha \right\}, \quad \forall \in \tilde \Lambda_n \setminus \{\alpha_1, \ldots, \alpha_l\}.
\]

Clearly, \(\tilde B_{n,l+1} \supseteq \bigcup_{i=1}^n \tilde C_{i,n,l+1}\) and the sets in the union are pairwise disjoint, and all sets involved are in \(\tilde B_{n,l+1}\), by Proposition 12.25. Then, we have

\[
P_{n,l+1}(\tilde B_{n,l+1})
\]
\[
= \int_{Y_{n,l+1}} \chi_{\tilde B_{n,l+1},Y_{n,l+1}}(\omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) \, d\tilde P_{n,l+1}(\omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,l+1}} \chi_{\tilde B_{n,l},Y_{n,l}}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) \, d\tilde P_{n,l}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,l+1}} \left( \chi_{H_{n,l},Y_{n,l}} + \sum_{i=1}^n \chi_{\tilde C_{i,n,l},Y_{n,l}} \right)(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
d\tilde P_{n,l}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= \int_{Y_{n,l+1}} \chi_{\tilde B_{n,l},Y_{n,l}}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) \, d\tilde P_{n,l}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
+ \int_{Y_{n,l+1}} \sum_{i=1}^n \chi_{\tilde C_{i,n,l},Y_{n,l}}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) \, d\tilde P_{n,l}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
> 2^{-l} \epsilon_0 + \int_{Y_{n,l+1}} \sum_{i=1}^n \chi_{\tilde C_{i,n,l},Y_{n,l}}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) 
\]
\[
d\tilde P_{n,l}(\omega_{\alpha_1}, \omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
\[
= 2^{-l} \epsilon_0 + \int_{Y_{n,l+1}} \sum_{i=1}^n \chi_{\tilde C_{i,n,l+1},Y_{n,l+1}}(\omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}}) 
\]
\[
d\tilde P_{n,l+1}(\omega_{\alpha_{l+1}}, \ldots, \omega_{\alpha_{r(n)}})
\]
where the first equality follows from Proposition 11.75; the second equality follows from the definition of \( \bar{B}_{n,l+1} \), \( X_{\alpha_j} \) is a probability measure space, and Fubini’s Theorem 12.31; the third equality follows from the fact that \( \bar{B}_{n,l} = \bar{H}_{n,l} \cup (\bigcup_{\alpha \in \Lambda^l} \bar{C}_{\alpha}) \) and the sets in the union are pairwise disjoint; the fourth equality follows from Proposition 11.92; the inequality follows from the conclusion of the previous paragraph; the fifth equality follows from the definition of \( \bar{C}_{i,n,l+1} \)’s and Fubini’s Theorem 12.31; and the last equality follows from Proposition 11.75. Thus, we have \( P_{n,l+1}(\bar{H}_{n,l+1}) := P_{n,l+1}(\bar{B}_{n,l+1} \setminus (\bigcup_{\alpha \in \Lambda^l} \bar{C}_{\alpha})) > 2^{-t} \varepsilon_0 \). Define \( H_{n,l+1} := \{(\omega_\alpha)_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} \in \prod_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} X_\alpha | (\omega_\alpha)_{\alpha \in \Lambda} \in H_{n,1}, \text{where} \bar{\omega}_\alpha = \omega_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}, \bar{\omega}_{\alpha_j} = \bar{\omega}_{\alpha_j}, j = 1, \ldots, l\}. \) It is obvious that \( H_{n,l+1} = \{(\omega_\alpha)_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} \in \prod_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} X_\alpha | (\omega_\alpha)_{\bar{\alpha} \in \Lambda_\setminus \{\bar{\alpha}_1, \ldots, \bar{\alpha}_l\}} \in H_{n,1}, \text{where} \bar{\omega}_\alpha = \omega_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}, \bar{\omega}_{\alpha_j} = \bar{\omega}_{\alpha_j}, j = 1, \ldots, l\}. \) Clearly, \( H_{n,l+1} = H_{n,l+1} \times \prod_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} X_\alpha, H_{n,l+1} \supseteq H_{n+1,l+1}, \) and \( \bigcap_{n=1}^{\infty} H_{n,l+1} = \emptyset. \) Obviously, \( \bar{H}_{n,l+1} \in \bar{B}_{n,l+1}. \)

By Radon-Nikodym Theorem 11.169 and Definition 14.7, \( \mathcal{F}_{n,l+1}(\chi_{\mathbf{\bar{H}}_{n,l+1},Y_{n,l+1}}|\bar{B}_{\alpha_{l+1},n,l+1}) \) exists, where \( \mathcal{F}_{n,l+1} \) is with respect to measure \( P_{n,l+1} \). Define \( g_{n,l+1} : Y_{n,l+1} \times \prod_{\alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_l\}} X_\alpha \to [0,1] \subseteq \mathbb{R} \) by, \( \forall (\bar{\omega}_{\alpha_1}, \ldots, \bar{\omega}_{\alpha_r}) \in Y_{n,l+1}, \)

\[
g_{n,l+1}(\bar{\omega}_{\alpha_1}, \ldots, \bar{\omega}_{\alpha_r}) = \int_{Y_{n,l+1}} \chi_{H_{n,l+1},Y_{n,l+1}}(\bar{\omega}_{\alpha_1}, \bar{\omega}_{\alpha_2}, \ldots, \bar{\omega}_{\alpha_r}) dP_{n,l+1}(\bar{\omega}_{\alpha_1}, \ldots, \bar{\omega}_{\alpha_r})
\]

By Proposition 14.12, \( g_{n,l+1} \in \mathcal{F}_{n,l+1}(\chi_{H_{n,l+1},Y_{n,l+1}}|\bar{B}_{\alpha_{l+1},n,l+1}), \)

\[
g_{n,l+1}(\bar{\omega}_{\alpha_1}, \ldots, \bar{\omega}_{\alpha_r}) = \bar{g}_{n,l+1}(\bar{\omega}_{\alpha_1}), \forall (\bar{\omega}_{\alpha_1}, \ldots, \bar{\omega}_{\alpha_r}) \in Y_{n,l+1}, \]

and \( g_{n,l+1} : X_{\alpha_{l+1}} \to [0,1] \subseteq \mathbb{R} \) is \( \mathcal{B}_{\alpha_{l+1}} \)-measurable. Since \( P_{n,l+1}(\bar{H}_{n,l+1}) > 2^{-t} \varepsilon_0 \), then \( \mathbf{E}_{\alpha_{l+1}}(\bar{g}_{n,l+1}) = \mathcal{F}_{n,l+1}(\chi_{\mathbf{\bar{H}}_{n,l+1},Y_{n,l+1}}|\bar{B}_{\alpha_{l+1},n,l+1}) \)

\[
= \mathcal{F}_{n,l+1}(\chi_{\mathbf{\bar{H}}_{n,l+1},Y_{n,l+1}}|\bar{B}_{\alpha_{l+1},n,l+1}) = P_{n,l+1}(\bar{H}_{n,l+1}) > 2^{-t} \varepsilon_0, \text{where} \mathbf{E}_{\alpha_{l+1}} \text{ is with respect to} P_{\alpha_{l+1}} \text{ and the second equality follows from (a) of Proposition 14.11. This leads to the following inequalities}
\]

\[
2^{-t} \varepsilon_0 < \mathbf{E}_{\alpha_{l+1}}(\bar{g}_{n,l+1})
\]

\[
\leq P_{\alpha_{l+1}}(\{\omega_{\alpha_{l+1}} \in X_{\alpha_{l+1}} | \bar{g}_{n,l+1}(\omega_{\alpha_{l+1}}) > 2^{-t-1} \varepsilon_0\}) + 2^{-t} \varepsilon_0 P_{\alpha_{l+1}}(\{\omega_{\alpha_{l+1}} \in X_{\alpha_{l+1}} | \bar{g}_{n,l+1}(\omega_{\alpha_{l+1}}) \leq 2^{-t-1} \varepsilon_0\})
\]

\[
\leq P_{\alpha_{l+1}}(\{\omega_{\alpha_{l+1}} \in X_{\alpha_{l+1}} | \bar{g}_{n,l+1}(\omega_{\alpha_{l+1}}) > 2^{-t-1} \varepsilon_0\}) + 2^{-t-1} \varepsilon_0
\]

Thus, we have \( P_{\alpha_{l+1}}(\{\omega_{\alpha_{l+1}} \in X_{\alpha_{l+1}} | \bar{g}_{n,l+1}(\omega_{\alpha_{l+1}}) > 2^{-t-1} \varepsilon_0\}) > 2^{-t-1} \varepsilon_0. \)
Now, since \( H_{n,l+1} \supseteq H_{n+1,l+1} \), by the definition of \( \tilde{g}_{n,l+1} \), Proposition 11.92, and Fubini’s Theorem 12.31, we have \( 0 \leq \tilde{g}_{n,l+1} (\omega_{\alpha_{i+1}}) \leq g_{n,l+1} (\omega_{\alpha_{i+1}}) \), \( \forall \omega_{\alpha_{i+1}} \in X_{\alpha_{i+1}} \), (which requires us to work in the product measure space \( \mathcal{Y}_{n+1,l+1} \) that includes both \( H_{n,l+1} \times \prod_{i=r(n)+1}^{\infty} X_{\alpha_i} \) and \( \hat{H}_{n,l+1} \)). Therefore, we have \( P_{\alpha_{i+1}} (\Gamma_{n,l+1}) > 2^{-l-1} \epsilon_0 \), where \( \Gamma_{n,l+1} := \{ \omega_{\alpha_{i+1}} \in X_{\alpha_{i+1}} \mid g_{n,l+1} (\omega_{\alpha_{i+1}}) > 2^{-l-1} \epsilon_0 \} \). Furthermore, \( \Gamma_{n,l+1} \supseteq \Gamma_{n+1,l+1} \). By Proposition 11.5, we have \( P_{\alpha_{i+1}} (\bigcap_{n=1}^{\infty} \Gamma_{n,l+1}) = \lim_{n \to \infty} P_{\alpha_{i+1}} (\Gamma_{n,l+1}) \geq 2^{-l-1} \epsilon_0 > 0 \). Then, \( \exists \omega_{\alpha_{i+1}} \in X_{\alpha_{i+1}} \) such that \( \omega_{\alpha_{i+1}} \in \bigcap_{n=1}^{\infty} \Gamma_{n,l+1} \). This implies that \( \tilde{g}_{n,l+1} (\omega_{\alpha_{i+1}}) > 2^{-l-1} \epsilon_0 \), \( \forall n \in \mathbb{N} \). By the definition of \( \tilde{g}_{n,l+1} \), we have \( \int_{\mathcal{Y}_{n,l+1}} \chi_{H_{n,l+1}} (\omega_{\alpha_{i+1}}, \omega_{\alpha_{i+2}}, \ldots, \omega_{\alpha_{r(n)}}) d \hat{P}_{n,l+1} (\omega_{\alpha_{i+1}}, \ldots, \omega_{\alpha_{r(n)}}) > 2^{-l-1} \epsilon_0 \), \( \forall n \in \mathbb{N} \).

Thus, we may obtain \( \tilde{\omega}_{\alpha_i} \in \bigcap_{n=1}^{\infty} \Gamma_{n,i} \subseteq X_{\alpha_i} \), \( \forall i \in \mathbb{N} \) with \( \alpha_i \in \mathcal{A} \). By Axiom of Choice, there exists a \( \tilde{\omega} \in X = \prod_{\alpha \in \mathcal{A}} X_\alpha \) such that \( \pi_\alpha (\tilde{\omega}) = \tilde{\omega}_\alpha \), \( \forall i \in \mathcal{A} \). Fix any \( n \in \mathbb{N} \). \( \omega_{\alpha_{r(n)}} \in \bigcap_{m=1}^{\infty} \Gamma_{m,r(n)} \). This implies that \( \omega_{\alpha_{r(n)}} \in \Gamma_{n,r(n)} = \{ \omega_{\alpha_{r(n)}} \in X_{\alpha_{r(n)}} \mid \tilde{g}_{n,r(n)} (\omega_{\alpha_{r(n)}}) > 2^{-r(n)} \epsilon_0 \} \), and \( \tilde{g}_{n,r(n)} (\omega_{\alpha_{r(n)}}) > 2^{-r(n)} \epsilon_0 > 0 \). By definition of \( \tilde{g}_{n,r(n)} \), we have \( 0 < \hat{g}_{n,r(n)} (\omega_{\alpha_{r(n)}}) = \int_{\mathcal{Y}_{n,r(n)}} \chi_{\mathcal{H}_{n,r(n)}} (\omega_{\alpha_{r(n)}}) d \hat{P}_{n,r(n)} (\omega_{\alpha_{r(n)}}) = \omega_{\alpha_{r(n)}} (\omega_{\alpha_{r(n)}}) = 1 \), where the last equality follows since it is the value of an indicator function. Hence, \( \omega_{\alpha_{r(n)}} \in \hat{H}_{n,r(n)} \). By the definition of \( \hat{H}_{n,r(n)} \), we have \( (\omega_{\alpha_i})_{i \in \mathcal{A}} \in \hat{H}_{n,1} \). This yields \( \tilde{\omega} \in H_{n,1} \). By the arbitrariness of \( n \), we have \( \tilde{\omega} \in \bigcap_{n=1}^{\infty} H_{n,1} = \emptyset \). This is a contradiction. Hence, the hypothesis does not hold. We must have \( \mu (B) \leq \sum_{i=1}^{\infty} \mu (C_i) \).

Hence, (ii) of Proposition 11.33 holds.

By Proposition 11.32, \( \mu \) admits a unique extension to a measure \( \tilde{\mu} \) on the algebra on \( X, \mathcal{A} \), generated by \( \mathcal{F} \). Clearly, \( \tilde{\mu} (X) = \mu (X) = 1 \), since \( X \in \mathcal{F} \). Then, \( \tilde{\mu} \) is finite. By Carathéodory Extension Theorem 11.19, there exists a unique measure \( \mathcal{P} \) on the measurable space \( (X, \sigma (\mathcal{F})) \) that is an extension of \( \tilde{\mu} \). Then, we have a probability measure space \( (X, \mathcal{B} := \sigma (\mathcal{F}), \mathcal{P}) \) as we seek. It is straightforward to prove the collection of \( \sigma \)-algebras \( (\mathcal{B}_\alpha)_{\alpha \in \mathcal{A}} \) is independent by Definition 14.5. This completes the proof of the theorem.

Definition 14.14 Let \( \Omega := (\Omega, \mathcal{B}, \mathcal{P}) \) be a probability measure space, \( m \in \mathbb{N} \), \( x : \Omega \to \mathbb{R}^m \) be a \( \mathbb{R}^m \)-valued random variable. Define a mapping \( L_x : \mathcal{B} (\mathbb{R}^m) \to [0, 1] \subset \mathbb{R} \) by \( L_x (E) := \mathcal{P} (\{ x \in E \}) \), \( \forall E \in \mathcal{B} (\mathbb{R}^m) \). Then, \( L_x \) is a probability measure on the measurable space \( (\mathbb{R}^m, \mathcal{B} (\mathbb{R}^m)) \) and is said to be the law of \( x \). Since \( L_x \) is a probability measure, then we may define \( F : \mathbb{R}^m \to [0, 1] \subset \mathbb{R} \) by \( F (z) = L_x (\{ \bar{x} \in \mathbb{R}^m \mid \bar{x} \leq z \}) \), \( \forall z \in \mathbb{R}^m \). Then, by Proposition 12.51, \( F \) is a cumulative distribution function of \( L_x \). If \( L_x \ll \mu_{\mathbb{R}^m} \), we will let \( f = \frac{dL_x}{d\mu_{\mathbb{R}^m}} \) a.e. in \( \mathbb{R}^m \), and say \( f \) is the probability density function of \( L_x \).
By Proposition 12.75, we have 
\[ \int \forall \]
the third equality follows from Fubini’s Theorem 12.31. By 
\[ \forall \]
given by 
\[ \forall \]
Let 
\[ \forall \]
Example 14.15 Let \( I := (((0, 1), | \cdot |), B, \mu) \) be the finite measure subspace of \( \mathbb{R} \). Let \( x : (0, 1) \rightarrow [0, 1] \subset \mathbb{R} \) be a random variable given by \( x(\omega) = \omega, \forall \omega \in I \). Then, any cumulative distribution function \( F : \mathbb{R} \rightarrow [0, 1] \subset \mathbb{R} \) with \( F \) being of bounded variation, \( \lim_{z \rightarrow -\infty} F(z) = 0 \), \( \lim_{z \rightarrow -\infty} F(z) = 1 \), and \( T_F = 1 \), we seek a function \( h : (0, 1) \rightarrow \mathbb{R} \) such that the random variable \( y := h \circ x : I \rightarrow \mathbb{R} \) admits \( F \) as a cumulative distribution function of \( L_y \), which is the probability measure on the measurable space \( (\mathbb{R}, B_\mathbb{R}(\mathbb{R})) \). Since \( T_F = 1 \) and \( \lim_{z \rightarrow -\infty} F(z) = 0 \) and \( \lim_{z \rightarrow -\infty} F(z) = 1 \), then \( \forall z_1, z_2 \in \mathbb{R} \) with \( z_1 \leq z_2 \), we have \( \Delta_F(z_1, z_2) \geq 0 \) and hence \( F(z_1) \leq F(z_2) \). Since \( F \) is of bounded variation, then \( F \) is continuous on the right. We select \( h \) to be \( h(\alpha) = \inf \{ z \in \mathbb{R} \mid F(z) \geq \alpha \} \), \( \forall \alpha \in I \). Then, it is straightforward to prove that \( \mu(\{ \omega \in I \mid y(\omega) \leq z \}) = \mu(\{ \omega \in I \mid h(x(\omega)) \leq z \}) = \mu(\{ \omega \in I \mid \omega \leq F(z) \}) = F(z) \). Hence, the random variable \( y = h(x) \) admits the \( F \) as a cumulative distribution function for its law \( L_y \). Thus, for any cumulative distribution function \( F : \mathbb{R} \rightarrow [0, 1] \subset \mathbb{R} \) satisfying the stated assumptions, there exists a probability measure space \( I \) and a random variable \( y : I \rightarrow \mathbb{R} \) such that \( y \) admits \( F \) as a cumulative distribution function for its law \( L_y \). 

\[ \diamond \]

Proposition 14.16 Let \( \Omega := (\Omega, B, P) \) be a probability measure space, \( x_i : \Omega \rightarrow \mathbb{R}^{n_i} \) be an \( \mathbb{R}^{n_i} \)-valued random variable, \( n_i \in \mathbb{N} \), \( i = 1, 2 \), and \( x := (x_1, x_2) \) be an \( \mathbb{R}^n \)-valued random variable, \( n := n_1 + n_2 \), with probability density function \( f_x(z) \). Then, \( x_1 \) admits probability density function \( f_{x_1} : \mathbb{R}^{n_1} \rightarrow [0, \infty) \subset \mathbb{R} \), where \( f_{x_1}(z_1) := \int_{\mathbb{R}^{n_2}} f_x(z_1, z_2) \, d\mu_{B_{\mathbb{R}^{n_2}}}(z_2), \forall z_1 \in \mathbb{R}^{n_1} \) and \( z := (z_1, z_2) \) is partitioned as is for \( x \).

\[ \text{Proof} \]
Let \( L_{x_1} \) be the law for \( x_1 \). Then,
\[ L_{x_1}(E_1) = P(\{ (x_1, x_2) \in \mathbb{R}^{n_1+n_2} \mid x_1 \in E_1 \}) = \int_{E_1 \times \mathbb{R}^{n_2}} f_x(z) \, d\mu_{B_{\mathbb{R}^{n_2}}}(z) \]
\[ = \int_{E_1} \int_{\mathbb{R}^{n_2}} f_x(z_1, z_2) \, d\mu_{B_{\mathbb{R}^{n_2}}}(z_2) \, d\mu_{B_{\mathbb{R}^{n_1}}}(z_1) = \int_{E_1} f_{x_1}(z_1) \, d\mu_{B_{\mathbb{R}^{n_1}}}(z_1) \]
\( \forall E_1 \in B_\mathbb{R}(\mathbb{R}^{n_1}) \), where the first and second equalities follow from Definition 14.14; the third equality follows from Fubini’s Theorem 12.31. By Proposition 11.116, we have \( L_{x_1} \) is the measure with kernel \( f_{x_1} \) over \( \mathbb{R}^{n_1} \). By Proposition 12.75, we have \( \frac{dL_{x_1}}{d\mu_{B_{\mathbb{R}^{n_1}}}} = f_{x_1} \) a.e. in \( \mathbb{R}^{n_1} \), and hence \( x_1 \) admits probability density function \( f_{x_1} \). 
\[ \square \]
Proposition 14.17 Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $m \in \mathbb{N}$, $X$ be $\mathbb{R}^m$-valued random variable with probability density function $f_X : \mathbb{R}^m \to [0, \infty) \subset \mathbb{R}$, and $V \in \mathcal{B}(\mathbb{R}^m, \mathbb{R}^m)$ be invertible. Then, $Y := VX$ is a $\mathbb{R}^m$-valued random variable with probability density function $f_Y : \mathbb{R}^m \to [0, \infty) \subset \mathbb{R}$, where $f_Y(y) = \frac{1}{|\det(V)|} f_X(V^{-1}y)$.

Proof Let the law for $X$ be $L_X$ and the law for $Y$ be $L_Y$. Then, $orall E \in \mathcal{B}(\mathbb{R}^m)$,

$$L_Y(E) = P(\{\omega \in \Omega \mid Y(\omega) \in E\}) = P(\{\omega \in \Omega \mid X(\omega) \in V^{-1}(E)\})$$

$$= L_X(V^{-1}(E)) = \int_{V^{-1}(E)} f_X(x) \, d\mu_{Bm}(x)$$

$$= \int_E |\det(V^{-1})| f_X(V^{-1}y) \, d\mu_{Bm}(y)$$

$$= \int_E \frac{1}{|\det(V)|} f_X(V^{-1}y) \, d\mu_{Bm}(y) = \int_E f_Y(y) \, d\mu_{Bm}(y)$$

where the first equality follows from Definition 14.14; the second equality follows from the fact that $Y = VX$; the third equality follows from Definition 14.14; the fourth equality follows from Propositions 11.116 and 12.75; the fifth equality follows from the Change of Variable Theorem 12.91 with $y = Vx$; the sixth equality follows from properties of matrices; and the last equality follows from the definition of $f_Y$. Hence, we observe that $L_Y$ is the measure with kernel $f_Y$ over $\mathbb{R}^m$. By Propositions 11.116 and 12.75, we have $\frac{dL_Y}{d\mu_{Bm}} = f_Y$ a.e. in $\mathbb{R}^m$, and hence $Y$ admits probability density function $f_Y$.

This completes the proof of the proposition.

\[ \square \]

Proposition 14.18 Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $m \in \mathbb{N}$, $n_1, \ldots, n_m \in \mathbb{N}$, $x_i : \Omega \to \mathbb{R}^{n_i}$ be an $\mathbb{R}^{n_i}$-valued random variable with probability density function $f_i : \mathbb{R}^{n_i} \to [0, \infty) \subset \mathbb{R}$, $i = 1, \ldots, m$. Then, the random variables $x_1, \ldots, x_m$ are independent if, and only if, the $\mathbb{R}^n$-valued random variable $x := (x_1, \ldots, x_m) : \Omega \to \mathbb{R}^n$ with $n := \sum_{i=1}^m n_i \in \mathbb{N}$ admits probability density function $f : \mathbb{R}^n \to [0, \infty) \subset \mathbb{R}$, defined by $f(z) = \prod_{i=1}^m f_i(z_i)$, $\forall z := (z_1, \ldots, z_m) \in \mathbb{R}^n$ with $z_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, m$.

Proof “Sufficiency” By Proposition 11.39, $x$ is a $\mathbb{R}^n$-valued random variable. $\forall E_i \in \mathcal{B}(\mathbb{R}^{n_i})$, $i = 1, \ldots, m$, let $\bar{E}_i := \bigcap_{j=1}^n B_{j,i}$, where $B_{j,i} = \mathbb{R}^{n_j} \cap \{z_j \mid z_j \in \mathbb{R}^{n_j} \cap \{z_i \mid z_i \in \mathbb{R}^{n_i} \}$, $\forall j \notin i \in \bar{J} := \{1, \ldots, m\}$ with $j \neq i$. Then,

$$P(\bigcap_{i=1}^m \bar{E}_i) = P(\bigcap_{i=1}^m E_i) = P(\bar{E}) = \int_{\bar{E}} f(z) \, d\mu_{Bn}(z)$$

$$= \int_{\bar{E}} \prod_{i=1}^m f_i(z_i) \, d\mu_{Bn}(z_1, \ldots, z_m)$$

This completes the proof of the proposition.
14.2 GAUSSIAN RANDOM VARIABLES AND VECTORS

\[ \int_{E_1} \cdots \int_{E_m} \prod_{i=1}^{m} f_i(z_i) \, d\mu_{B_{n_m}}(z_m) \, \cdots \, d\mu_{B_{n_1}}(z_1) \]

\[ = \prod_{i=1}^{m} \int_{E_i} f_i(z_i) \, d\mu_{B_{n_i}}(z_i) = \prod_{i=1}^{m} P(E_i) \]

where the second equality follows from Definition 14.14; the third equality follows from Theorem 12.31; the fourth equality follows from Proposition 11.92; and the last equality follows from Definition 14.14. Hence, the random variables \( x_1, \ldots, x_m \) are independent.

“Necessity” By Proposition 11.39, \( x \) is a \( \mathbb{R}^n \)-valued random variable. Let \( L_x \) be the law of \( x \) and \( L_{x_i} \) be the law of \( x_i, i = 1, \ldots, m \). \( \forall E_i \in B_{\mathbb{R}^n}, i = 1, \ldots, m, \), let \( \bar{E}_i := \prod_{j=1, j \neq i}^{m} B_{j,i} \), where \( B_{j,i} = \mathbb{R}^n \), \( \forall j \in J := \{1, \ldots, m\} \) with \( j \neq i \). By the independence of random variables \( x_1, \ldots, x_m \), we have

\[ L_x(\bigcap_{i=1}^{m} \bar{E}_i) = L_x(E) = \prod_{i=1}^{m} L_{x_i}(\bar{E}_i) = \prod_{i=1}^{m} L_{x_i}(E_i) \]

\[ = \prod_{i=1}^{m} \int_{E_i} f_i(z_i) \, d\mu_{B_{n_i}}(z_i) = \int_{E} f(z) \, d\mu_{B_{n}}(z) \]

where the first equality follows from the independence assumption; the second equality follows from Definition 14.14; the third equality follows from the assumption that \( x_i \) admits probability density function \( f_i, i = 1, \ldots, m \); and the last equality follows from Theorem 12.31. By Lemma 12.103, \( L_x \) equals to the finite measure on \( \mathbb{R}^n \) with kernel \( f \). Then, the \( \mathbb{R}^n \)-valued random variable \( x \) admits probability density function \( f \), by Definition 14.14 and Proposition 12.75.

This completes the proof of the proposition.

14.2 Gaussian Random Variables and Vectors

Definition 14.19 Let \( \Omega := (\Omega, \mathcal{E}, P) \) be a probability measure space, \( m \in \mathbb{N} \), \( x : \Omega \to \mathbb{R}^m \) be an \( \mathbb{R}^m \)-valued random variable. We say that \( x \) is an \( \mathbb{R}^m \)-valued Gaussian (normal) random variable with mean \( \bar{x} \in \mathbb{R}^m \) and covariance \( K \in B(\mathbb{R}^m, \mathbb{R}^m) \), where \( K \in \mathcal{S}_+ \mathbb{R}^m \), if \( E(x) = \bar{x} \) and \( \mathbb{E}((x - \bar{x}) \otimes (x - \bar{x})) = K \) and a cumulative distribution function for the law \( L_x \) of \( x \) is

\[ F(z) = \lim_{k \to -\infty} \int_{\mathcal{E} \cap B_{n_m}(z, \frac{1}{\sqrt{2\pi}^m \det(K)^{1/2}})} \exp(-\frac{1}{2} \langle x - \bar{x}, K^{-1}(x - \bar{x}) \rangle_{\mathbb{R}^m}) \, d\mu_{B_{n}}(x) \]
∀z ∈ \( \mathbb{R}^m \); or the probability density function for \( L_x \) is \( f : \mathbb{R}^m \to [0, \infty) \subset \mathbb{R} \), defined by

\[
f(z) = \frac{1}{\sqrt{2\pi}^m (\det(K))^{1/2}} \exp\left(-\frac{1}{2} \left< x - \bar{x}, K^{-1}(x - \bar{x}) \right>_{\mathbb{R}^m}\right), \quad \forall z \in \mathbb{R}^m
\]

We will write \( x \sim N(\bar{x}, K) \) to denote that \( x \) is an \( \mathbb{R}^m \)-valued Gaussian random variable with mean \( \bar{x} \) and covariance \( K \).

**Example 14.20** Let \( m \in \mathbb{N} \), \( \bar{x} \in \mathbb{R}^m \) and \( K \in B(\mathbb{R}^m, \mathbb{R}^m) \), where \( K \in S_{+ \mathbb{R}^m} \). By Spectral Theory Theorem 13.52, there exists a unitary matrix \( V \in B(\mathbb{R}^m, \mathbb{R}^m) \) and a diagonal matrix \( D = \text{block diagonal} (\sigma_1^2, \ldots, \sigma_m^2) \in S_{+ \mathbb{R}^m} \) such that \( K = V' DV \). If \( x \) is an \( \mathbb{R}^m \)-valued Gaussian random variable with mean \( \bar{x} \) and covariance \( K \), then \( y = Vx \) is \( \mathbb{R}^m \)-valued Gaussian random variable with mean \( \bar{y} := V\bar{x} =: (\bar{y}_1, \ldots, \bar{y}_m) \) and covariance \( KV' = D \). Then, by Example 14.15 and Fundamental Theorem on Modeling 14.13, there exists a probability measure space \( \Omega := (\Omega, \mathcal{B}, P) \) and \( m \) independent random variables \( y_i : \Omega \to \mathbb{R} \), \( i = 1, \ldots, m \), such that \( y_i \) is a Gaussian normal random variable with mean \( \bar{y}_i \) and variance \( \sigma_i^2 \), \( i = 1, \ldots, m \). \( y_i \sim N(\bar{y}_i, \sigma_i^2) \). Then, by Proposition 14.18, we have the \( \mathbb{R}^m \)-valued random variable \( y := (y_1, \ldots, y_m) \) admits the probability density function \( f_y : \mathbb{R}^m \to [0, \infty) \subset \mathbb{R} \), given by

\[
f_y(z) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi} \sigma_i} \exp\left(-\frac{1}{2} \sigma_i^{-2} (z_i - \bar{y}_i)^2 \right)
\]

Thus, by Proposition 14.17, \( x = V'y \sim N(\bar{x}, K) \) as desired. \( \diamond \)

**Proposition 14.21** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( m \in \mathbb{N} \), \( x_i : \Omega \to \mathbb{R}^{n_i} \) be an \( \mathbb{R}^{n_i} \)-valued random variable, \( n_i \in \mathbb{N} \), \( i = 1, \ldots, m \), and \( x := (x_1, \ldots, x_m) \sim N(\bar{x}, K) \), where \( n := \sum_{i=1}^m n_i \in \mathbb{N} \), \( \bar{x} \in \mathbb{R}^n \) and \( K \in S_{+ \mathbb{R}^n} \). Then, \( x_1, \ldots, x_m \) are independent if, and only if, \( K \) is block diagonal \( (K_1, \ldots, K_m) \in S_{+ \mathbb{R}^n} \) with \( K_i \in S_{+ \mathbb{R}^{n_i}} \), \( i = 1, \ldots, m \).

**Proof** “Sufficiency” Note that \( x \) admits the probability density function \( f_x : \mathbb{R}^n \to \mathbb{R}_+ \), defined by, \( \forall z \in \mathbb{R}^n \),

\[
f_x(z) = \frac{1}{\sqrt{2\pi}^{n} (\det(K))^{1/2}} \exp\left(-\frac{1}{2} \left< z - \bar{x}, K^{-1}(z - \bar{x}) \right>_{\mathbb{R}^n}\right)
\]

\[
= \prod_{i=1}^m \frac{1}{\sqrt{2\pi}^{n_i} (\det(K_i))^{1/2}} \exp\left(-\frac{1}{2} \left< z_i - \bar{x}_i, K_i^{-1}(z_i - \bar{x}_i) \right>_{\mathbb{R}^{n_i}}\right)
\]
14.3. LAW OF LARGE NUMBERS

where \( z := (z_1, \ldots, z_m) \) and \( \bar{x} := (\bar{x}_1, \ldots, \bar{x}_m) \) are partitioned as is for \( x; \) and the second equality follows from the block diagonal structure of \( K. \) By Proposition 14.16, we have \( x_i \) admits probability density function \( f_{x_i} : \mathbb{R}^n \rightarrow [0, \infty) \subset \mathbb{R}, \) where

\[
f_{x_i}(z_i) := \frac{1}{\sqrt{2\pi}^{m_i}} \exp\left(-\frac{1}{2} \left( z_i - \bar{x}_i, K_i^{-1}(z_i - \bar{x}_i) \right)_{\mathbb{R}^{m_i}} \right)
\]

Clearly, we have \( f_z(z) = \prod_{i=1}^m f_{x_i}(z_i), \forall z = (z_1, \ldots, z_m) \in \mathbb{R}^n. \) By Proposition 14.18, \( x_1, \ldots, x_m \) are independent.

"Necessity" Let \( x_1, \ldots, x_m \) be independent. Then, \( K = E((x - \bar{x}) \otimes (x - \bar{x})) = \begin{bmatrix} K_{1,1} & \cdots & K_{1,m} \\ \vdots & \ddots & \vdots \\ K_{m,1} & \cdots & K_{m,m} \end{bmatrix}, \) where \( K_{i,j} := E((x_i - \bar{x}_i) \otimes (x_j - \bar{x}_j)), \forall 1 \leq i, j \leq m. \forall 1 \leq i, j \leq m \) with \( i \neq j, \) we have \( K_{i,j} = E((x_i - \bar{x}_i) \otimes (x_j - \bar{x}_j)) = E(x_i - \bar{x}_i) \otimes E(x_j - \bar{x}_j) = 0_{n_i \times n_j}, \) where the second equality follows from Proposition 14.9. Hence, \( K = \) block diagonal \( (K_{1,1}, \ldots, K_{m,m}). \)

This completes the proof of the proposition. \( \square \)

**Proposition 14.22** Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure space, \( m, p \in \mathbb{N}, \) \( (X_i)_{i=1}^p \) be random variables such that \( X := (X_1, \ldots, X_p) \sim N(\bar{x}, K), \) \( \bar{x} \in \mathbb{R}^p, K \in \mathbb{B}(\mathbb{R}^p, \mathbb{R}^p) \) with \( K \in S_+^{p \times p}, \) and \( (Y_i)_{i=1}^m \) be linear combinations of \( X_i \)'s, \( Y_i = \sum_{j=1}^p a_{i,j} X_j, \) \( i = 1, \ldots, m, \) where \( a_{i,j} \in \mathbb{R}, \forall i = 1, \ldots, m, \forall j = 1, \ldots, p. \) Assume that \( V := (a_{i,j})_{m \times p} \) is of full row rank. Then, \( Y := (Y_1, \ldots, Y_m) \) is a \( \mathbb{R}^m \)-valued random variable with distribution \( N(V \bar{x}, V K V'). \)

**Proof** Clearly \( m \leq p \) and \( \text{rank}(V) = m. \) Then, \( N(V) \) is \( (p - m) \)-dimensional. Let \( f_1, \ldots, f_{p-m} \) be an orthonormal basis of \( N(V), \) and \( F := [ f_1 \cdots f_{p-m} ]. \) Then, \( T := \begin{bmatrix} V \\ F K^{-1} \end{bmatrix} \) is an invertible matrix. By Proposition 14.17, \( Z := TX = \begin{bmatrix} Y \\ Z_2 \end{bmatrix} \) is an \( \mathbb{R}^p \)-valued random variable with distribution \( N(T \bar{x}, TK T'). \) Note that \( TK T' = \begin{bmatrix} V K V' & V F' \\ F V' & F K^{-1} F' \end{bmatrix} = \) block diagonal \( (VK V', F K^{-1} F'). \) By Proposition 14.21, \( Y \) and \( Z_2 \) are independent Gaussian random vectors with \( Y \sim N(V \bar{x}, V K V'). \) This completes the proof of the proposition. \( \square \)

### 14.3 Law of Large Numbers

**Theorem 14.23** (Strong Law of Large Numbers) Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure, \( (X_n)_{n=1}^\infty \) be a sequence of independent random variables, \( X_n : \Omega \rightarrow \mathbb{R}, \) with \( E(\bar{X}_n) = a \in \mathbb{R} \) and \( E((X_n - a)^2) \leq b \in \mathbb{R}_+, \forall n \in \mathbb{N}. \) Then, \( \lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^n X_i = a \) a.e. in \( \Omega. \)
14.4 Martingales Indexed by \( \mathbb{Z}_+ \)

**Definition 14.24** Let \( X := (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu) \) be a nonempty \( \sigma \)-finite topological measure space with a partial ordering \( \preceq \) defined on it, \( \Omega := (\Omega, \mathcal{B}, \mathcal{P}) \) be a probability measure space, and \( X \times \Omega := (\mathcal{X} \times \Omega, \mathcal{B} \times \mathcal{P}) \) be the \( \sigma \)-finite product measure space (as defined in Proposition 12.22), \( Y \) be a Banach space, and \( W : X \times \Omega \to Y \) be \( \mathcal{B} \)-measurable. Then, \( W \) is said to be a \( Y \)-valued stochastic process on \( X \times \Omega \). Then, \( W_t \) \( \Omega \to Y \) defined by \( W_t(\omega) = W(t, \omega) \), \( \forall \omega \in \Omega \), is a \( Y \)-valued random variable (i.e., \( \mathcal{B} \)-measurable by Proposition 12.28). We will sometimes write \((W_t)_{t \in X}\) to denote the stochastic process \( W \). When \( Y = \mathbb{R} \), we will simply say that \( W \) is a stochastic process on \( X \times \Omega \). \((B_x)_{x \in X}\) is said to be a filtration on \( \Omega \), if \( B_x \subseteq B_x \subseteq \mathcal{B} \) and \( B_x \) and \( B_z \) are sub-\( \sigma \)-algebras of \( \mathcal{B} \) on \( \Omega \), \( \forall x, z \in X \) with \( x \preceq z \). Then, the quadruple \((\Omega, \mathcal{B}, (B_x)_{x \in X}, \mathcal{P})\) is said to be a filtered probability measure space. We say that \((\tilde{W}_x)_{x \in X}\) is adapted to the filtration \((B_x)_{x \in X}\) if \( W_x \) is \( B_x \)-measurable, \( \forall x \in X \). The stochastic process \((\tilde{W}_x)_{x \in X}\) defines the natural filtration of \((W_x)_{x \in X}\), \((\tilde{B}_x)_{x \in X}\), where \( \tilde{B}_x := \sigma(\bigcup_{y \in X, y \preceq x} \sigma(W_y)) \), \( \forall x \in X \).

It is easy to check that \( \tilde{B}_x \subseteq \tilde{B}_x \subseteq \mathcal{B}, \forall x, z \in Z \) with \( x \preceq z \). Hence, the natural filtration of a stochastic process is indeed a filtration.

In this section and the next five sections, we will be primarily concerned with stochastic processes indexed by \( \mathbb{Z}_+ \). In this case, we will define \( \mathbb{N} := (\mathbb{Z}_+, \preceq, \mu) \) to be the topological measure space on \( \mathbb{Z}_+ \) with the counting measure and the partial ordering of \( \preceq \). (It is easy to check that it is a topological measure space, where the topology is given by the subset topology of \( \mathbb{R} \).) In this case, a \( Y \)-valued stochastic process \( W \) on \( \mathbb{N} \times \Omega \) adapted to some filtration \((B_n)_{n \in \mathbb{N}}\) is equivalent to that \( W_n : \Omega \to Y \) being \( B_n \)-measurable, \( \forall n \in \mathbb{N} \). In this case, we will simply say a \( Y \)-valued stochastic process \((W_n)_{n=0}^{\infty}\) on \( \Omega \) that is adapted to \((B_n)_{n=0}^{\infty}\) to save us from cumbersome notations.

**Definition 14.25** Let \( \Omega := (\Omega, \mathcal{B}, (B_n)_{n=0}^{\infty}, \mathcal{P}) \) be a filtered probability measure space, \((X_n)_{n=0}^{\infty}\) be a real-valued stochastic process on \( \Omega \). Then, the stochastic process \((X_n)_{n=0}^{\infty}\) is said to be a Martingale if
14.4. MARTINGALES INDEXED BY $\mathbb{Z}_+$

(i) $(X_n)_{n=0}^{\infty}$ is an adapted process (that is, it is adapted to the filtration $(\mathcal{F}_n)_{n=0}^{\infty}$);

(ii) $X_n \in L_1(\Omega, \mathbb{R})$, $\forall n \in \mathbb{Z}_+$;

(iii) $X_{n-1} \in E(X_n|\mathcal{F}_{n-1})$, $\forall n \in \mathbb{N}$.

The stochastic process $(X_n)_{n=0}^{\infty}$ is said to be a super Martingale if it satisfies (i) and (ii) and $X_{n-1}(\omega) \leq X_n(\omega)$ a.e. $\omega \in \Omega$, $\forall f_{n-1} \in E(X_n|\mathcal{F}_{n-1})$, $\forall n \in \mathbb{N}$.

The stochastic process $(X_n)_{n=0}^{\infty}$ is said to be a sub Martingale if it satisfies (i) and (ii) and $f_{n-1}(\omega) \geq X_{n-1}(\omega)$ a.e. $\omega \in \Omega$, $\forall f_{n-1} \in E(X_n|\mathcal{F}_{n-1})$, $\forall n \in \mathbb{N}$.

Clearly, a stochastic process is a super Martingale if, and only if, its negative process is a sub Martingale. A stochastic process is a Martingale if, and only if, it is a super Martingale and is a sub Martingale.

**Theorem 14.26** Let $\Omega := (\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0}^{\infty}, P)$ be a filtered probability measure space, $(X_n)_{n=0}^{\infty}$ be a super Martingale (or a Martingale), and $(C_n)_{n=1}^{\infty}$ be such that

(i) $C_n : \Omega \to [0, \alpha] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, where $\alpha \in \mathbb{R}_+$;

(ii) $C_n$ is $\mathcal{F}_{n-1}$-measurable, $\forall n \in \mathbb{N}$;

Then, the stochastic process $(Y_n)_{n=0}^{\infty}$ defined by $Y_0 = X_0$ and $Y_n = X_0 + \sum_{i=1}^{n-1} C_i (X_i - X_{i-1})$, $\forall n \in \mathbb{N}$, is a super Martingale (or a Martingale).

**Proof** Since, $\forall n \in \mathbb{N}$, $X_n$ is $\mathcal{F}_n$-measurable and $C_n$ is $\mathcal{F}_{n-1}$-measurable. Then, $C_n(X_n - X_{n-1})$ is $\mathcal{F}_n$-measurable by Propositions 11.38, 7.35, and 11.39 and Definition 14.25. Hence, $Y_n$ is $\mathcal{F}_n$-measurable, $\forall n \in \mathbb{Z}_+$.

Clearly, $|Y_n| = \left| \sum_{i=0}^{n-1} (C_i - C_{i+1})X_i + C_n X_n \right| \leq \sum_{i=0}^{n-1} 2\alpha |X_i| + \alpha |X_n|$, $\forall n \in \mathbb{Z}_+$. Then, by Definition 14.25, $X_n \in L_1(\Omega, \mathbb{R})$, $\forall n \in \mathbb{Z}_+$, implies that $Y_n \in L_1(\Omega, \mathbb{R})$ by Proposition 11.83.

Note that $E(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) = E(C_n(X_n - X_{n-1})|\mathcal{F}_{n-1}) = C_n E(X_n - X_{n-1}|\mathcal{F}_{n-1})$, $\forall n \in \mathbb{N}$, where the first equality follows from the definition of $Y_n$; the second equality follows from (h) of Proposition 14.11. This is less than or equal to zero almost everywhere if $(X_n)_{n=0}^{\infty}$ is a super Martingale; or is equal to zero almost everywhere if $(X_n)_{n=0}^{\infty}$ is a Martingale. Hence, $(Y_n)_{n=0}^{\infty}$ is a super Martingale, if $(X_n)_{n=0}^{\infty}$ is so; and it is a Martingale, if $(X_n)_{n=0}^{\infty}$ is a Martingale.

This completes the proof of the theorem. □

This result implies that one cannot beat the system in a gambling senario by varying his bet each time after observing the outcome of the past dice rolls.
Definition 14.27 Let $X$ be a nonempty topological subspace of $\mathbb{R}$, $\Omega := (\Omega, \mathcal{B}, (B_n)_{n=0}^{\infty}, P)$ be a filtered probability measure space, and $T : \Omega \to X$ be a random variable (i.e., it is $\mathcal{B}$-measurable). $T$ is said to be stopping time if $\{ \omega \in \Omega \mid T(\omega) \leq x \} \in \mathcal{B}_x, \forall x \in X$.

Example 14.28 Let $\Omega := (\Omega, \mathcal{B}, (B_n)_{n=0}^{\infty}, \mu)$ be a filtered probability measure space, $\mathcal{F}$ be a normed linear space over $\mathbb{K}$, $(X_n)_{n=0}^{\infty}$ be a $\mathcal{F}$-valued stochastic process, and $B \in \mathcal{B}_B(\mathcal{F})$. Then, $T : \Omega \to \mathbb{Z}_+ \cup \{ \infty \}$ defined by $T(\omega) := \inf\{ n \in \mathbb{Z}_+ \mid X_n(\omega) \in B \}$ is time of first entry of $(X_n)_{n=0}^{\infty}$ into set $B$ is a stopping time. This is because $\{ \omega \in \Omega \mid T(\omega) \leq n \} = \bigcup_{i=0}^{n} \{ \omega \in \Omega \mid X_i(\omega) \in B \} \in \mathcal{B}$.

Definition 14.29 Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{F} \subseteq \mathcal{B}$ be a sub-$\sigma$-algebra on $\Omega$, and $\Omega := (\Omega, \mathcal{B}, P := P|_{\mathcal{F}})$. $\forall F \in \mathcal{B}$, the condition probability of the event $F$ happens given $\mathcal{F}$ is denoted by $P(F|\mathcal{F})$, which is defined to be $E(\chi_{F, \Omega}|\mathcal{F}) \in L_1(\Omega, [0, 1])$.

Proposition 14.30 Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{F} \subseteq \mathcal{B}$ be a sub-$\sigma$-algebra on $\Omega$, and $\Omega := (\Omega, \mathcal{B}, P := P|_{\mathcal{F}})$. $\forall F \in \mathcal{B}$, let $P(F|\mathcal{F}) := E(\chi_{F, \Omega}|\mathcal{F}) \in L_1(\Omega, [0, 1])$. Then, we may define $\mu : B \to L_1(\Omega, [0, 1])$ by $\mu(F) = P(F|\mathcal{F})$. Then, $\mu$ is a finite $L_1(\Omega, [0, 1])$-valued measure on the measurable space $(\Omega, \mathcal{B})$ with $\mathcal{P} \circ \mu = P$.

Proof Clearly, $P$ is a finite measure on $(\Omega, \mathcal{B})$. (i) $\mu(\emptyset) = P(\emptyset|\mathcal{F}) = E(\chi_{\emptyset, \Omega}|\mathcal{F}) \in L_1(\Omega, [0, 1])$. Fix any $f \in \mu(\emptyset)$. Then, $f \in L_1(\Omega, [0, 1])$. $\forall E \in \mathcal{F}$, $0 = P(\emptyset) = \int_E \chi_{\emptyset, \Omega} dP = \int_E f dP$, where the first equality follows from $P$ being a measure on $(\Omega, \mathcal{B})$: the second equality follows from Definition 14.1; and the third equality follows from Definition 14.7. By Proposition 11.96, we have $f = 0$ a.e. in $\Omega$. Thus, $\mu(\emptyset) = \emptyset_{L_1(\Omega, [0, 1])}$.

(ii) $\forall E \in \mathcal{B}$, $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, $\mu(E) = P(E|\mathcal{F}) = E(\chi_{E, \Omega}|\mathcal{F}) = E(\sum_{i=1}^{\infty} \chi_{E_i, \Omega}|\mathcal{F}) = \sum_{i=1}^{\infty} \mu(E_i)$, where the first two equalities follow from Definition 14.12; the fourth equality follows from (e) of Proposition 14.11; and the last equality follows from Definition 14.29. Furthermore, $\infty > \|\mu(E)\|_{L_1(\Omega[0, 1])} = \int_\Omega |\mu(E)(\omega)| d\hat{P}(\omega) = \int_\Omega \mu(E)(\omega) d\hat{P}(\omega) = \int_\Omega \sum_{i=1}^{\infty} \mu(E_i)(\omega) d\hat{P}(\omega) = \sum_{i=1}^{\infty} \int_\Omega \mu(E_i)(\omega) d\hat{P}(\omega) = \sum_{i=1}^{\infty} \int_\Omega |\mu(E_i)(\omega)| d\hat{P}(\omega) = \sum_{i=1}^{\infty} \|\mu(E_i)\|_{L_1(\Omega[0, 1])}$, where the inequality follows from Definition 14.29; the second equality follows from the fact that $\mu(E)(\omega) \geq 0$ a.e. $\omega \in \Omega$; the third equality follows from the preceding discussion; the fourth equality follows from Lebesgue Dominated Convergence Theorem 11.91; and the fifth equality follows from the fact that $\mu(E_i)(\omega) \geq 0$ a.e. $\omega \in \Omega, \forall i \in \mathbb{N}$. Hence, $(\Omega, \mathcal{B}, \mu)$ is a $L_1(\Omega, [0, 1])$-valued pre-measure space, by Definition 11.99.

Note that, $\forall E \in \mathcal{B}$,

$\mathcal{P} \circ \mu(E)$.
14.4. MARTINGALES INDEXED BY $\mathbb{Z}_+$

\[ \sup_{n \in \mathbb{Z}_+} \left\{ \sum_{i=1}^{n} \| \mu(E_i) \|_{L_1((\Omega, [0,1]))} \right\} = \sup_{n \in \mathbb{Z}_+} \left\{ \sum_{i=1}^{n} \int_{\Omega} |\mu(E_i)(\omega)| \, d\hat{\mathcal{P}}(\omega) \right\} = \sup_{n \in \mathbb{Z}_+} \left\{ \sum_{i=1}^{n} \int_{\Omega} \mu(E_i)(\omega) \, d\hat{\mathcal{P}}(\omega) \right\} = \sup_{n \in \mathbb{Z}_+} \left\{ \int_{\Omega} \mu(E)(\omega) \, d\hat{\mathcal{P}}(\omega) \right\} = \int_{\Omega} \mu(E)(\omega) \, d\hat{\mathcal{P}}(\omega) = \int_{\Omega} \chi_{E, \Omega} \, d\mathcal{P} = \mathcal{P}(E) \]

where the first equality follows from Definition 11.99; the second equality follows from Example 11.179; the third equality follows from $\mu(E_i)(\omega) \geq 0$ a.e. $\omega \in \Omega$, $\forall i = 1, \ldots, n$; the fourth equality follows from Proposition 11.92; the seventh equality follows from (14.2a) of Definition 14.7. Hence, $\mathcal{P} = \mathcal{P} \circ \mu$.

Since $\mathcal{P}$ is a finite measure, then $(\Omega, \mathcal{B}, \mu)$ is a finite $L_1((\Omega, [0,1])$-valued measure space. This completes the proof of the proposition. \qed

\textbf{Theorem 14.31} Let $\Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^{\infty}, \mathcal{P})$ be a filtered probability measure space, $(X_n)_{n=0}^{\infty}$ be a real-valued adapted stochastic process, and $T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$ be a stopping time. Assume that $(X_n)_{n=0}^{\infty}$ is a super Martingale (or a Martingale), then the stopped stochastic process $(Y_n)_{n=0}^{\infty}$ defined by $Y_n(\omega) = X_{T(\omega)\land n}(\omega)$, $\forall n \in \mathbb{Z}_+$, is a super Martingale (or a Martingale).

\textbf{Proof} Define $(C_n)_{n=1}^{\infty}$ by $C_n : \Omega \to [0,1] \subset \mathbb{R}$ with $C_n(\omega) = \chi(T(\omega) \geq n), \forall \omega \in \Omega, \forall n \in \mathbb{N}$. Clearly, $C_n$ is $\mathcal{B}_{n-1}$-measurable. It is easy to see that $Y_n$ as defined in the statement of the theorem is exactly the same as that of Theorem 14.26. Then, the result follows immediately from Theorem 14.26. \qed

\textbf{Theorem 14.32 (Doob’s Optional Stopping)} Let $\Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^{\infty}, \mathcal{P})$ be a filtered probability measure space, $(X_n)_{n=0}^{\infty}$ be a super Martingale adapted to the filtration $(\mathcal{B}_n)_{n=0}^{\infty}$, $T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$ be a stopping time. Then, $X_T : \Omega \to \mathbb{R}$ defined by $X_T(\omega) := \begin{cases} X_{T(\omega)}(\omega) & \forall \omega \in \Omega \text{ with } T(\omega) < \infty \\ \lim_{n \to \infty} X_n(\omega) & \forall \omega \in \Omega \text{ with } T(\omega) = \infty \text{ and } \lim_{n \to \infty} X_n(\omega) \in \mathbb{R}, \forall \omega \\ 0 & \forall \omega \in \Omega \text{ with } T(\omega) = \infty \text{ and } \lim_{n \to \infty} X_n(\omega) \text{ doesn’t exist in } \mathbb{R} \end{cases}$
\( \in \Omega \), is absolutely integrable over \( \Omega \) and \( E(X_T) \leq E(X_0) \) if any of the following conditions hold.

(i) \( X_n(\omega) \geq 0 \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{N} \), and \( T(\omega) < \infty \) a.e. \( \omega \in \Omega \).

(ii) \( \exists N \in \mathbb{N} \) such that \( T(\omega) \leq N \), \( \forall \omega \in \Omega \).

(iii) \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( |X_n(\omega)| \leq M \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{Z}_+ \), and \( T(\omega) < \infty \) a.e. \( \omega \in \Omega \).

(iv) \( E(T) < \infty \) and \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( |X_n(\omega) - X_{n-1}(\omega)| \leq M \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{Z}_+ \).

If \( (X_n)_{n=0}^\infty \) is a Martingale adapted to the filtration \((\mathcal{B}_n)_{n=0}^\infty\), and any of the conditions (ii) — (iv) holds, then \( X_T \) is absolutely integrable over \( \Omega \) and \( E(X_T) = E(X_0) \).

**Proof**

We first consider the first paragraph of the theorem statements. By Propositions 11.51, 11.48, and 11.41, \( X_T \) is \( \mathcal{B} \)-measurable. Let (i) hold. Then, the stopped stochastic process \((Y_n)_{n=0}^\infty\) defined in Theorem 14.31 satisfies \( Y_n(\omega) \geq 0 \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{Z}_+ \). By Theorem 14.31, \((Y_n)_{n=0}^\infty\) is a super Martingale. By (i), \( T(\omega) < \infty \) a.e. \( \omega \in \Omega \). Then, \( \lim_{n \in \mathbb{N}} Y_n(\omega) = X_T(\omega) \) a.e. \( \omega \in \Omega \). Clearly, \( X_T(\omega) \geq 0 \), \( \forall \omega \in \Omega \). By Fatou’s Lemma 11.80, we have \( E(X_T) = \int \Omega X_T(\omega) \, d\mathbb{P}(\omega) \leq \liminf_{n \in \mathbb{N}} \int \Omega Y_n(\omega) \, d\mathbb{P}(\omega) = \liminf_{n \in \mathbb{N}} E(Y_n) = \liminf_{n \in \mathbb{N}} E(E(Y_n|\mathcal{B}_n-1)) \leq \liminf_{n \in \mathbb{N}} E(Y_n-1) \leq \cdots \leq E(Y_0) = E(X_0) \), where the first equality follows from Definition 14.1; the second equality follows from Definition 14.1; the third equality follows from Proposition 14.11; the second inequality follows from Theorem 14.31; and the last equality follows from the definition of \((Y_n)_{n=0}^\infty\) in Theorem 14.31. Hence, the result holds.

Let (ii) hold. Then, the stopped stochastic process \((Y_n)_{n=0}^\infty\) satisfies that \( Y_n(\omega) = X_T(\omega) \), \( \forall \omega \in \Omega \). Hence, by Theorem 14.31, \( X_T \) is absolutely integrable over \( \Omega \) and \( E(X_T) = E(Y_n) \leq E(Y_{n-1}) \leq \cdots \leq E(Y_0) = E(X_0) \). Hence, the result holds.

Let (iii) hold. Then, \((Y_n)_{n=0}^\infty\) satisfies that \( |Y_n(\omega)| \leq M \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{Z}_+ \), and \( \lim_{n \in \mathbb{Z}_+} Y_n(\omega) = X_T(\omega) \) a.e. \( \omega \in \Omega \). By Lebesgue Dominated Convergence Theorem 11.91, \( X_T \) is absolutely integrable over \( \Omega \) and \( E(X_T) = \lim_{n \in \mathbb{Z}_+} E(Y_n) \leq E(Y_0) = E(X_0) \). Hence, the result holds.

Let (iv) hold. Then, the stopped stochastic process \((Y_n)_{n=0}^\infty\) defined in Theorem 14.31 satisfies \( Y_n = X_0 + \sum_{i=1}^n C_i (X_n - X_{n-1}) \), \( \forall n \in \mathbb{N} \), where \((C_n)_{n=1}^\infty\) is as defined in the proof of Theorem 14.31. Then, \( |Y_n(\omega)| \leq |X_0(\omega)| + \sum_{i=1}^n \chi_{\{T(\omega) \geq n\}} \omega \left( |X_n(\omega) - X_{n-1}(\omega)| \leq |X_0(\omega)| + \sum_{i=1}^n M \chi_{\{T(\omega) \geq n\}} \omega \right) \), \( \forall n \in \mathbb{Z}_+ \), where the first inequality follows from the definition of \((C_n)_{n=1}^\infty\); and the second inequality follows from (iv). Then, \( E(|Y_n|) \leq E(|X_0|) + M \int \Omega T(\omega) \, d\mathbb{P}(\omega) = E(|X_0|) + M E(T) < \infty \), \( \forall n \in \mathbb{Z}_+ \), where the last inequality follows from (iv). Then, \( Y_n \) is absolutely integrable over \( \Omega \), \( \forall n \in \mathbb{Z}_+ \). Since \( E(T) < \infty \), then \( T(\omega) < \infty \), a.e.
Let (Ω, B, P) be a filtered probability measure space, and T : Ω → Z_+ ∪ {∞} be a stopping time. Assume that ∃N ∈ N and ∃ε ∈ (0, 1) ⊂ IR, ∀n ∈ N, P(ω ∈ Ω | T(ω) ≤ n + N) > ε a.e. ω ∈ Ω. Then, E(T) < ∞.

Proof
Note that

\[ E(T) = \int_\Omega T(\omega) \, dP(\omega) = \sum_{n=0}^{\infty} nP(\{\omega \in \Omega \mid T(\omega) = n\}) \]

\[ = \sum_{k=0}^{\infty} \sum_{n=kN}^{(k+1)N-1} nP(\{\omega \in \Omega \mid T(\omega) = n\}) \]

\[ \leq \sum_{k=0}^{\infty} \sum_{n=kN}^{(k+1)N-1} (k+1)N-1 \, P(\{\omega \in \Omega \mid T(\omega) = n\}) \]

\[ = \sum_{k=0}^{\infty} ((k+1)N-1) \, P(\{\omega \in \Omega \mid kN \leq T(\omega) < (k+1)N\}) \]

\[ \leq \sum_{k=0}^{\infty} ((k+1)N-1) \, P(\{\omega \in \Omega \mid T(\omega) > kN-1\}) \]

Let \( \alpha_k := P(\{\omega \in \Omega \mid T(\omega) \leq kN - 1\}) \in [0,1] \subset IR, \forall k \in N \). Then, \( \alpha_k = P(\{\omega \in \Omega \mid T(\omega) \leq kN - 1 - N\}) + P(\{\omega \in \Omega \mid kN - 1 - N < T(\omega) \leq kN - 1\}) = \alpha_{k-1} + P(\{\omega \in \Omega \mid kN - 1 - N < T(\omega) \}) \, P(\{\omega \in \Omega \mid T(\omega) \leq kN - 1\}) \). The last term is a conditional probability of an event given another event. We will treat this term in detail in the following paragraph.

Let \( E_1 := \{\omega \in \Omega \mid T(\omega) \leq kN - 1\} \in B \) and \( E_2 := \{\omega \in \Omega \mid kN - 1 - N < T(\omega) \} \in B_{kN-1-N} \). Let \( \mathcal{F} := \emptyset, E_2, \Omega \setminus E_2, \Omega \subseteq B_{kN-1-N} \), where \( \mathcal{F} \) is a σ-algebra on \( \Omega \). Then, the notion \( P(E_1 | \mathcal{F}) \) is well defined, which is a function on \( \Omega \) that is \( \mathcal{F} \)-measurable. Then, \( P(E_1 | \mathcal{F})(\omega) = P(E_1 | E_2)(\omega) \in [0,1] \subset IR, \forall \omega \in E_2 \), and \( P(E_1 | \mathcal{F})(\omega) = P(E_1 | \Omega \setminus E_2)(\omega) \in [0,1] \subset IR \).
∀ω ∈ Ω \ E_2. Then, we have the following identities:

\[
P(E_1|E_2) = \int_{E_2} \frac{P(E_1|\mathcal{F})(\omega)}{P(\omega)} \, d\hat{P}(\omega) = \frac{1}{P(E_2)} \int_{E_2} P(E_1|\mathcal{F})(\omega) \, d\hat{P}(\omega)
\]

\[
= \frac{1}{P(E_2)} \int_{E_2} E(\chi_{E_1,\Omega}|\mathcal{F})(\omega) \, d\hat{P}(\omega)
\]

\[
= \frac{1}{P(E_2)} \int_{E_2} E(\chi_{E_1,\Omega}|\mathcal{B}_{kN-1-N}|\mathcal{F})(\omega) \, d\hat{P}(\omega)
\]

\[
= \frac{1}{P(E_2)} \int_{E_2} E(\chi_{E_1,\Omega}|\mathcal{B}_{kN-1-N})(\omega) \, d\hat{P}(\omega) \geq \frac{1}{P(E_2)} \int_{E_2} \epsilon \, d\hat{P}
\]

\[
= \epsilon \frac{1}{P(E_2)} \hat{P}(E_2) = \epsilon
\]

where \( \hat{P} = P|_\mathcal{F} \) and \( \bar{P} = P|_{\mathcal{B}_{kN-1-N}} \); the first equality follows from the preceding discussion; the second equality follows from \( E_2 \in \mathcal{F} \); the third equality follows from the definition of \( P \) and Definition 14.29; the fourth equality follows from (j) of Proposition 14.11; the fifth equality follows from (14.2a) of Definition 14.7; the inequality follows from the assumption of the lemma; the sixth equality follows from \( E_2 \in \mathcal{B}_{kN-1-N} \); and the last equality follows from the definition of \( P \).

Then, we have \( \alpha_k \geq \alpha_{k-1} + (1 - \alpha_{k-1})\epsilon, \forall k \in \{2, 3, \ldots\} \). Therefore, \( P(\{\omega \in \Omega \mid T(\omega) > kN - 1\}) = 1 - \alpha_k \leq (1 - \alpha_{k-1})(1 - \epsilon), \forall k \in \{2, 3, \ldots\} \). Recursively, we have \( P(\{\omega \in \Omega \mid T(\omega) > N - 1\}) \leq P(\{\omega \in \Omega \mid T(\omega) > N - 1\})(1 - \epsilon)^{k-1} \leq (1 - \epsilon)^{k-1}, \forall k \in \{2, 3, \ldots\} \). Hence, \( E(T) \leq N - 1 + 2N - 1 + \sum_{k=2}^{\infty}((k+1)N-1)(1-\epsilon)^{k-1} < \infty \), where the last inequality follows from Proposition A.2.

This completes the proof of the lemma. □

**Definition 14.34** Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, P) \) be a filtered probability measure space, \((X_n)_{n=0}^\infty \) be an adapted \( \mathbb{R} \)-valued random process on \( \Omega \), and \( a, b \in \mathbb{R} \) with \( a < b \). For any \( N \in \mathbb{Z}_+ \) and any \( \omega \in \Omega \), we will say that the sample path of the random process \((X_n(\omega))_{n=0}^\infty \) has \( U_N(a,b)(\omega) \in \mathbb{Z}_+ \) upcrossings of the closed interval \([a,b] \subset \mathbb{R} \) by the time \( N \), if we can find \( k(\omega) := U_N(a,b)(\omega) \in \mathbb{Z}_+ \) that is the largest nonnegative integer with the property that: there exists \( 0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \leq N \) such that \( X_s(\omega) < a \) and \( X_t(\omega) > b \), \( i = 1, \ldots, k \). Clearly, \((U_N(a,b))_{N=0}^\infty \) is an adapted \( \mathbb{Z}_+ \)-valued random process, i.e., \( U_N(a,b) : \Omega \rightarrow \mathbb{Z}_+ \) and is \( \mathcal{B}_N \)-measurable, \( \forall N \in \mathbb{Z}_+ \).

**Lemma 14.35** (Doob’s Upcrossing Lemma) Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, P) \) be a filtered probability measure space, \((X_n)_{n=0}^\infty \) be an adapted \( \mathbb{R} \)-valued random process on \( \Omega \), \( a, b \in \mathbb{R} \) with \( a < b \), and \((U_N(a,b))_{N=0}^\infty \) be the adapted \( \mathbb{Z}_+ \)-valued random process defined in Definition 14.34. Assume that \((X_n)_{n=0}^\infty \) is a super Martingale. Then,

\[
(b - a)E(U_N(a,b)) \leq E(-(X_N - a) \wedge 0)
\]

(14.3)
14.4. MARTINGALES INDEXED BY $\mathbb{Z}_+$

\textbf{Proof} \quad \text{We will prove the lemma using Theorem 14.26. Define a $\mathbb{R}$-valued random process $(C_n)_{n=1}^{\infty}$ on $\Omega$ by: $C_1 := \chi_{\{\omega \in \Omega \mid X_0(\omega) < a\}} \omega$, $C_n := \chi_{\{\omega \in \Omega \mid C_{n-1}(\omega) = 1, X_{n-1}(\omega) \leq b\}} \omega + \chi_{\{\omega \in \Omega \mid C_{n-1}(\omega) = 0, X_{n-1}(\omega) < b\}} \omega$, \forall n \in \mathbb{N}$ with $n \geq 2$. Clearly, $C_n : \Omega \to [0,1] \subset \mathbb{R}$ is $\mathcal{B}_{n-1}$-measurable, \forall $n \in \mathbb{N}$. Essentially, $(C_n)_{n=1}^{\infty}$ is like a gambling strategy of wagering 1 coin after $X_{n-1} < a$ and then stop wagering after $X_{n-1} > b$, and so on so forth. Define the random process $(Y_n)_{n=0}^{\infty}$ as in Theorem 14.26: $Y_0 = X_0$, and $Y_n = \sum_{i=1}^{n} C_n(X_n - X_{n-1})$, \forall $n \in \mathbb{N}$. $Y_n$ is then the money the gambler has at time $n \in \mathbb{Z}_+$. Then, by Theorem 14.26, $(Y_n)_{n=0}^{\infty}$ is a super Martingale. It should be clear that}

$$Y_N - Y_0 \geq (b - a)U_N(a,b) + ((X_N - a) \wedge 0); \quad \forall \omega \in \Omega, \quad \forall N \in \mathbb{N} \tag{14.4}$$

This is because, for every upcrossing of $[a,b]$ by the time $N$, the winning is at least $(b-a)$, that explains the term $(b-a)U_N(a,b)$. After the last upcrossing before $N$, the winning can not be less than $((X_N - a) \wedge 0)$. Thus, (14.4) holds. Now, integrate both sides over $\Omega$, and we arrive at

$$0 \leq E(Y_N) - E(Y_0) \geq (b - a)E(U_N(a,b)) + E(((X_N - a) \wedge 0)).$$

This yields the desired inequality (14.3). This completes the proof of the lemma.

\[ \square \]

\textbf{Corollary 14.36} Let $\Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^{\infty}, \mathcal{P})$ be a filtered probability measure space, $(X_n)_{n=0}^{\infty}$ be an adapted $\mathbb{R}$-valued random process on $\Omega$, $a, b \in \mathbb{R}$ with $a < b$, and $(U_N(a,b))_{n=0}^{\infty}$ be the adapted $\mathbb{Z}_+$-valued random process defined in Definition 14.34. Assume that $(X_n)_{n=0}^{\infty}$ is a super Martingale and is bounded in $L_1(\Omega, \mathbb{R})$, \textit{i.e.}, $\sup_{n \in \mathbb{Z}_+} E(|X_n|) < \infty$. Then, $(b - a)E(U_\infty(a,b)) \leq |a| + \sup_{n \in \mathbb{Z}_+} E(|X_n|) < \infty$, where $U_\infty(a,b) := \lim_{N \in \mathbb{N}} U_N(a,b)$, and therefore $U_\infty(a,b)(\omega) < \infty$ a.e. $\omega \in \Omega$.

\textbf{Proof} \quad \text{Clearly, $U_N(a,b)(\omega)$ is monotonically nondecreasing in $N \in \mathbb{Z}_+$ and nonnegative, $\forall \omega \in \Omega$. Then, $U_\infty(a,b)(\omega) \in [0,\infty) \subset \mathbb{R}$. By Proposition 11.48, $U_\infty(a,b)$ is $\mathcal{B}$-measurable. By Fatou’s Lemma 11.80, we have $E(U_\infty(a,b)) \leq \liminf_{N \in \mathbb{N}} E(U_N(a,b))$. Then, by Doob’s Upcrossing Lemma 14.35, we have}

$$(b - a)E(U_\infty(a,b)) \leq \liminf_{N \in \mathbb{N}} (b - a)E(U_N(a,b))$$

\[ \leq \liminf_{N \in \mathbb{N}} E\left(-(\langle X_N - a \rangle \wedge 0)\right) = \liminf_{N \in \mathbb{N}} \int_{\Omega} -\langle X_N(\omega) - a \rangle \wedge 0 \, d\mathcal{P}(\omega) \]

\[ \leq \liminf_{N \in \mathbb{N}} \int_{\Omega} |X_N(\omega) - a| \, d\mathcal{P}(\omega) \leq \liminf_{N \in \mathbb{N}} \int_{\Omega} (|X_N(\omega)| + |a|) \, d\mathcal{P}(\omega) \]

\[ \leq |a| + \sup_{N \in \mathbb{Z}_+} E(|X_N|) < \infty \]

where the first inequality follows from Proposition 3.81; the second inequality follows from Lemma 14.35; the fifth inequality follows from Proposition 3.83. Thus, we conclude that $E(U_\infty(a,b)) < \infty$ and therefore $U_\infty(a,b)(\omega) < \infty$ a.e. $\omega \in \Omega$.\[ \square \]
This completes the proof of the corollary.

**Theorem 14.37 (Doob’s Forward Convergence Theorem)** Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, \mathbb{P}) \) be a filtered probability measure space, \((X_n)_{n=0}^\infty \) be an adapted \( \mathbb{R} \)-valued random process on \( \Omega \), and \( X_\infty : \Omega \to \mathbb{R} \) be defined by, \( \forall \omega \in \Omega \),

\[
X_\infty(\omega) := \left\{ \begin{array}{ll}
\lim_{n \to \infty} X_n(\omega) & \text{if } \lim_{n \to \infty} X_n(\omega) \in \mathbb{R} \\
0 & \text{if } \lim_{n \to \infty} X_n(\omega) \text{ does not exist in } \mathbb{R}
\end{array} \right.
\]

Assume that \((X_n)_{n=0}^\infty \) is a super Martingale and is bounded in \( L_1(\Omega, \mathbb{R}) \), i.e., \( \sup_{n \in \mathbb{Z}_+} E(|X_n|) < \infty \). Then, \( X_\infty = \lim_{n \to \infty} X_n(\omega) \) a.e. \( \omega \in \Omega \), \( X_\infty \) is \( \mathcal{B} \)-measurable, and \( E(|X_\infty|) \leq \sup_{n \in \mathbb{Z}_+} E(|X_n|) \).

**Proof** Define \( Y_\infty : \Omega \to \mathbb{R}_+ \) and \( Z_\infty : \Omega \to \mathbb{R}_+ \) by \( Y_\infty(\omega) := \limsup_{n \to \infty} X_n(\omega) \) and \( Z_\infty(\omega) := \liminf_{n \to \infty} X_n(\omega) \), \( \forall \omega \in \Omega \). By Proposition 11.40, \( Y_\infty \) and \( Z_\infty \) are \( \mathcal{B} \)-measurable. Define the set \( \Lambda := \{ \omega \in \Omega \mid Y_\infty(\omega) > Z_\infty(\omega) \} \in \mathcal{B} \). \( \forall \omega \in \Lambda \), \exists a, b \in \mathbb{Q} \) with \( a < b \) such that \( Y_\infty(\omega) > b > a > Z_\infty(\omega) \). Then, let \((U_n(a, b))_{n=0}^\infty \) be the adapted \( \mathbb{Z}_+ \)-valued random process for \((X_n)_{n=0}^\infty \) as defined in Definition 14.34. Then, \( U_\infty(a, b)(\omega) := \lim_{n \to \infty} U_n(a, b)(\omega) = \infty \). Thus, \( \Lambda \subseteq \bigcup_{n=0}^\infty \{ \omega \in \Omega \mid U_\infty(a, b)(\omega) = \infty \} =: \bigcup_{a < b} \Lambda_{a, b} \subseteq \mathcal{B} \). By Corollary 14.36, \( \mathbb{P}(\Lambda_{a, b}) = 0, \forall a, b \in \mathbb{R} \) with \( a < b \). Then, \( \mathbb{P}(\Lambda) = 0 \).

Define \( W_\infty : \Omega \to \mathbb{R}_+ \) by \( W_\infty(\omega) := \left\{ \begin{array}{ll}
\lim_{n \to \infty} X_n(\omega) & \omega \in \Omega \setminus \Lambda \\
0 & \omega \in \Lambda
\end{array} \right. \)
\( \forall \omega \in \Omega \). Then, \( W_\infty \) is \( \mathcal{B} \)-measurable by Propositions 11.40 and 11.41. By Fatou’s Lemma 11.80, we have

\[
E(|W_\infty|) = \int_\Omega |W_\infty| \ d\mathbb{P} \leq \liminf_{n \to \infty} \int_\Omega |X_n| \ d\mathbb{P} \leq \sup_{n \in \mathbb{Z}_+} E(|X_n|) < \infty
\]

Thus, we have \( W_\infty(\omega) \in \mathbb{R} \) a.e. \( \omega \in \Omega \). Clearly, \( W_\infty(\omega) = X_\infty(\omega) \) whenever \( W_\infty(\omega) \in \mathbb{R} \). Then, \( X_\infty = \lim_{n \to \infty} X_n(\omega) \) a.e. \( \omega \in \Omega \), \( X_\infty \) is \( \mathcal{B} \)-measurable, by Proposition 11.41, and \( E(|X_\infty|) = E(|W_\infty|) \leq \sup_{n \in \mathbb{Z}_+} E(|X_n|) \), where the equality follows from Proposition 11.89.

This completes the proof of the theorem.

**14.5 Banach Space Valued Martingales Indexed by \( \mathbb{Z}_+ \)**

**Definition 14.38** Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, \mathbb{P}) \) be a filtered probability measure space, \( \mathcal{Y} \) be a Banach space, \((X_n)_{n=0}^\infty \) be a \( \mathcal{Y} \)-valued stochastic process. Then, the stochastic process \((X_n)_{n=0}^\infty \) is said to be a \( \mathcal{Y} \)-valued Martingale if
14.5. BANACH SPACE VALUED MARTINGALES INDEXED BY $\mathbb{Z}_+$

(i) $(X_n)_{n=0}^\infty$ is an adapted process (that is, it is adapted to the filtration $(\mathcal{F}_n)_{n=0}^\infty$);

(ii) $X_n \in \bar{L}_1(\Omega, \mathbb{Y}_n)$, $\forall n \in \mathbb{Z}_+$, where $\mathbb{Y}_n$ is a separable Banach subspace of $\mathbb{Y}$;

(iii) $X_{n-1} \in E(X_n|\mathcal{F}_{n-1})$, $\forall n \in \mathbb{N}$.

**Theorem 14.39** Let $\Omega := (\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0}^\infty, \mathbb{P})$ be a filtered probability measure space, $\mathbb{Y}$ be a Banach space over $\mathbb{K}$, $(X_n)_{n=0}^\infty$ be a $\mathbb{Y}$-valued Martingale, and $(C_n)_{n=1}^\infty$ be such that

(i) $C_n : \Omega \to \mathcal{B}_{\mathbb{K}}(0, \alpha) \subset \mathbb{K}$, $\forall n \in \mathbb{N}$, where $\alpha \in \mathbb{R}_+$;

(ii) $C_n$ is $\mathcal{F}_{n-1}$-measurable, $\forall n \in \mathbb{N}$;

Then, the stochastic process $(Y_n)_{n=0}^\infty$ defined by $Y_0 = X_0$ and $Y_n = X_0 + \sum_{i=1}^n C_n(X_n - X_{n-1})$, $\forall n \in \mathbb{N}$, is a $\mathbb{Y}$-valued Martingale.

**Proof** Since, $\forall n \in \mathbb{N}$, $X_n$ is $\mathcal{F}_n$-measurable and $C_n$ is $\mathcal{F}_{n-1}$-measurable. Then, $C_n(X_n - X_{n-1})$ is $\mathcal{F}_n$-measurable by Propositions 11.38, 7.35, and 11.39 and Definition 14.38. Hence, $Y_n$ is $\mathcal{F}_n$-measurable, $\forall n \in \mathbb{Z}_+$.

Clearly, $\|Y_n(\omega)\|_\mathbb{Y} = \left\| \sum_{i=0}^{n-1} (C_i(\omega) - C_{i+1}(\omega))X_i(\omega) + C_n(\omega)X_n(\omega) \right\|_\mathbb{Y} \leq \sum_{i=0}^{n-1} 2\alpha \|X_i(\omega)\|_\mathbb{Y} + \alpha \|X_n(\omega)\|_\mathbb{Y}$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. Then, by Definition 14.38, $X_n \in \bar{L}_1(\Omega, \mathbb{Y})$, $\forall n \in \mathbb{Z}_+$, implies that $Y_n \in \bar{L}_1(\Omega, \mathbb{Y})$ by Proposition 11.83.

Note that $E(Y_n - Y_{n-1}|\mathcal{F}_{n-1}) = E(C_n(X_n - X_{n-1})|\mathcal{F}_{n-1}) = C_n E(X_n - X_{n-1}|\mathcal{F}_{n-1})$ = $\check{\mathcal{L}}_{\mathcal{F}_n}(\Omega_{n-1}, \mathbb{Y})$, $\forall n \in \mathbb{N}$, where the first equality follows from the definition of $Y_n$; the second equality follows from (h) of Proposition 14.11; $\Omega_{n-1} := (\Omega, \mathcal{F}_{n-1}, \mathbb{P}|_{\mathcal{F}_{n-1}} =: \mathbb{P}_{n-1}$); and the last equality follows from the fact that $(X_n)_{n=0}^\infty$ is a $\mathbb{Y}$-valued Martingale and (c) of Proposition 14.11. Hence, $(Y_n)_{n=0}^\infty$ is a $\mathbb{Y}$-valued Martingale.

This completes the proof of the theorem. \qed

**Theorem 14.40** Let $\Omega := (\Omega, \mathcal{F}, (\mathcal{F}_n)_{n=0}^\infty, \mathbb{P})$ be a filtered probability measure space, $\mathbb{Y}$ be a separable Banach space over $\mathbb{K}$, $(X_n)_{n=0}^\infty$ be a $\mathbb{Y}$-valued Martingale, and $T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$ be a stopping time. Then the stopped $\mathbb{Y}$-valued stochastic process $(Y_n)_{n=0}^\infty$ defined by $Y_n(\omega) = X_{T(\omega)\wedge n}(\omega)$, $\forall n \in \mathbb{Z}_+$, is a Martingale.

**Proof** Define $(C_n)_{n=1}^\infty$ by $C_n : \Omega \to [0, 1] \subset \mathbb{R}$ with $C_n(\omega) = \chi_{\{T(\omega)\geq n\}}(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. Clearly, $C_n$ is $\mathcal{F}_{n-1}$-measurable. It is easy to see that $Y_n$ as defined in the statement of the theorem is exactly the same as that of Theorem 14.39. Then, the result follows immediately from Theorem 14.39. \qed
Theorem 14.41 (Doob’s Optional Stopping) Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, P) \) be a filtered probability measure space, \( Y \) be a separable Banach space over \( \mathbb{K} \), \( (X_n)_{n=0}^\infty \) be a \( Y \)-valued Martingale adapted to the filtration \( (\mathcal{B}_n)_{n=0}^\infty \), \( T : \Omega \to \mathbb{Z}_+ \cup \{ \infty \} \) be a stopping time. Then, \( X_T : \Omega \to Y \) defined by, \( \forall \omega \in \Omega \)

\[
X_T(\omega) := \begin{cases} 
X_T(\omega)(\omega) & \forall \omega \in \Omega \text{ with } T(\omega) < \infty \\
\lim_{n \in \mathbb{N}} X_n(\omega) & \forall \omega \in \Omega \text{ with } T(\omega) = \infty \text{ and } \lim_{n \in \mathbb{N}} X_n(\omega) \in Y \\
\varnothing & \forall \omega \in \Omega \text{ with } T(\omega) = \infty \text{ and } \lim_{n \in \mathbb{N}} X_n(\omega) \text{ doesn’t exist in } Y
\end{cases}
\]

is absolutely integrable over \( \Omega \) and \( E(X_T) = E(X_0) \) if any of the following conditions hold.

(i) \( \exists N \in \mathbb{N} \) such that \( T(\omega) \leq N, \forall \omega \in \Omega \).

(ii) \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \|X_n(\omega)\|_Y \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{Z}_+ \), and \( T(\omega) < \infty \) a.e. \( \omega \in \Omega \).

(iii) \( E(T) < \infty \) and \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \|X_n(\omega) - X_{n-1}(\omega)\|_Y \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{N} \).

Proof By Propositions 11.51, 11.48, and 11.41, \( X_T \) is \( \mathcal{B} \)-measurable. Let (i) hold. Then, the stopped stochastic process \( (Y_n)_{n=0}^\infty \) defined in Theorem 14.40 satisfies that \( Y_N(\omega) = X_T(\omega), \forall \omega \in \Omega \). Hence, by Theorem 14.40, \( X_T \) is absolutely integrable over \( \Omega \) and \( E(X_T) = E(Y_N) = E(Y_{N-1}) = \cdots = E(Y_0) = E(X_0) \in Y \). Hence, the result holds.

Let (ii) hold. Then, \( (Y_n)_{n=0}^\infty \) satisfies that \( \|Y_n(\omega)\|_Y \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{Z}_+ \), and \( Y_n(\omega) = X_T(\omega) \) a.e. \( \omega \in \Omega \). By Lebesgue Dominated Convergence Theorem 11.91, \( X_T \) is absolutely integrable over \( \Omega \) and \( E(X_T) = \lim_{n \in \mathbb{Z}_+} E(Y_n) = E(Y_0) = E(X_0) \). Hence, the result holds.

Let (iii) hold. Then, the stopped stochastic process \( (Y_n)_{n=0}^\infty \) defined in Theorem 14.40 satisfies \( Y_n = X_0 + \sum_{i=1}^{n} C_i (X_i - X_{i-1}), \forall n \in \mathbb{N} \), where \( (C_n)_{n=1}^\infty \) is as defined in the proof of Theorem 14.40. Then, \( \|Y_n(\omega)\|_Y \leq \sum_{i=1}^{n} \chi(T(\omega) \geq n) \Omega \|X_i(\omega) - X_{i-1}(\omega)\|_Y + \sum_{i=1}^{n} M \chi(T(\omega) \geq n) \Omega \|X_0(\omega)\|_Y + MT(\omega), \forall n \in \mathbb{Z}_+ \), where the first inequality follows from the definition of \( (C_n)_{n=1}^\infty \); and the second inequality follows from (iii). Then, \( E(P \circ Y_n) \leq E(P \circ X_0) + M \int_{\lozenge} T(\omega) dP(\omega) = E(P \circ X_0) + M E(T) < \infty, \forall n \in \mathbb{Z}_+ \), where the last inequality follows from (iii). Then, \( Y_n \) is absolutely integrable over \( \Omega, \forall n \in \mathbb{Z}_+ \). Since \( E(T) < \infty \), then \( T(\omega) < \infty \) a.e. \( \omega \in \Omega \), and \( \lim_{n \in \mathbb{Z}_+} Y_n(\omega) = X_T(\omega) \) a.e. \( \omega \in \Omega \). By Lebesgue Dominated Convergence Theorem 11.91, \( E(X_T) = \lim_{n \in \mathbb{Z}_+} E(Y_n) = E(Y_0) = E(X_0) \in Y \), where the second equality follows from Theorem 14.40. Hence, the result holds.

This completes the proof of the theorem. \( \Box \)
Definition 14.42 Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $\mathcal{Y}$ be a separable Banach space, and $(X_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{Y}$-valued random variables. Let $T_n := \sigma(X_{n+1}, X_{n+2}, \ldots), \forall n \in \mathbb{N}$, and $T := \bigcap_{n=1}^{\infty} T_n$. The $\sigma$-algebra $T$ is called the tail $\sigma$-algebra of the sequence $(X_n)_{n=1}^{\infty}$.

Clearly, $T_n$ is the $\sigma$-algebra that is generated by $\mathcal{F}_n := \{X_{n+i}\text{inv}(O) \mid i \in \mathbb{N}, O \in \mathcal{O}_y\} \subseteq \mathcal{B}, \forall n \in \mathbb{N}$. Clearly, we have $T_n \supseteq T_{n+1}, \forall n \in \mathbb{N}$. It is straightforward to show that $T$ is a $\sigma$-algebra. Hence, the definition makes sense.

Theorem 14.43 (Kolmogorov 0-1 Law) Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $(\mathcal{Y}, \mathcal{Y})$ be a separable Banach space, $(X_n)_{n=1}^{\infty}$ be a sequence of independent $\mathcal{Y}$-valued random variables, and $T$ be the tail $\sigma$-algebra of $(X_n)_{n=1}^{\infty}$. Then, the following statements hold.

(i) $\forall E \in T$, we have $P(E) = 0$ or $P(E) = 1$.

(ii) Let $\mathcal{Z} := (Z, \rho)$ be a nonempty separable complete metric space and $\xi : \Omega \rightarrow \mathcal{Z}$ be $T$-measurable. Then, $\exists x_0 \in \mathcal{Z}$ such that $P(\{\omega \in \Omega \mid \xi(\omega) = x_0\}) = P(\xi(\{x_0\})) = 1$.

Proof

(i) Let $\mathcal{X}_n := \sigma(X_1, \ldots, X_n), \forall n \in \mathbb{N}$. We need the following result.

Claim 14.43.1 $\mathcal{X}_n$ and $\mathcal{T}_n$ are independent $\sigma$-algebras, $\forall n \in \mathbb{N}$, where $\mathcal{T}_n$ is as defined in Definition 14.42.

Proof of claim: Fix any $n \in \mathbb{N}$. Let $\mathcal{K} := \{F \in \mathcal{B} \mid \exists O_1, \ldots, O_n \in \mathcal{O}_y \exists r \in \mathbb{N}, \exists O_1, \ldots, O_r \in \mathcal{O}_y \exists F = \bigcap_{i=1}^{n} X_{n+i}\text{inv}(O_i)\} \subseteq \mathcal{X}_n$. Let $\mathcal{J} := \{F \in \mathcal{B} \mid \exists r \in \mathbb{N}, \exists O_1, \ldots, O_r \in \mathcal{O}_y \exists F = \bigcap_{i=1}^{r} X_{n+i}\text{inv}(O_i)\} \subseteq \mathcal{T}_n$. Clearly, $\mathcal{K}$ and $\mathcal{J}$ are $\pi$-systems on $\Omega$, $\mathcal{\sigma}(\mathcal{K}) = \mathcal{X}_n, \mathcal{\sigma}(\mathcal{J}) = \mathcal{T}_n$. $\forall E_1 \in \mathcal{K}, \forall E_2 \in \mathcal{J}$, by the assumption that $(X_n)_{n=1}^{\infty}$ are independent, we have $P(E_1 \cap E_2) = P(E_1)P(E_2)$. By Proposition 14.10, we have $\mathcal{X}_n$ and $\mathcal{T}_n$ are independent $\sigma$-algebras. This completes the proof of the claim.

Since $\mathcal{T} \subseteq \mathcal{T}_n, \forall n \in \mathbb{N}$, then $\mathcal{T}$ and $\mathcal{X}_n$ are independent $\sigma$-algebras.

Claim 14.43.2 $\mathcal{X}_\infty := \sigma(X_1, X_2, \ldots)$ and $\mathcal{T}$ are independent $\sigma$-algebras.

Proof of claim: Clearly, we have $\mathcal{X}_n \subseteq \mathcal{X}_\infty, \forall n \in \mathbb{N}$. Define $\mathcal{K}_\infty := \bigcup_{n=1}^{\infty} \mathcal{X}_n \subseteq \mathcal{X}_\infty$. Clearly, $\mathcal{K}_\infty$ is a $\pi$-system on $\Omega$ and $\mathcal{\sigma}(\mathcal{K}_\infty) = \mathcal{X}_\infty$. By the fact that $\mathcal{X}_n$ and $\mathcal{T}$ are independent $\sigma$-algebras, $\forall n \in \mathbb{N}$, we have $\forall E_1 \in \mathcal{K}_\infty, \forall E_2 \in \mathcal{T}, P(E_1 \cap E_2) = P(E_1)P(E_2)$, since $E_1 \in \mathcal{X}_n$, for some $n \in \mathbb{N}$. Then, by Proposition 14.10, $\mathcal{X}_\infty$ and $\mathcal{T}$ are independent $\sigma$-algebras. This completes the proof of the claim.

Note that $\mathcal{T} \subseteq \mathcal{X}_\infty$. By Claim 14.43.2, $\mathcal{T}$ and $\mathcal{X}$ are independent $\sigma$-algebras. $\forall E \in \mathcal{T}$, we have $P(E) = P(E \cap E) = P(E)P(E)$, where the last equality follows from Definition 14.5. Hence, $P(E) = 0$ or $P(E) = 1$. 

(ii) By (i), \( \forall \omega \in \Omega_B \), we have \( P(\xi_{n}(\Omega)) = 0 \) or \( P(\xi_{n}(\Omega)) = 1 \). Clearly, \( P(\xi_{n}(\Omega)) = P(\Omega) = 1 \). Fix any \( \bar{x} \in Z \neq \emptyset \). Then, by Proposition 11.7, we have \( \lim_{n \to \infty} P(\xi_{n}(\Omega)) = P(\bigcup_{n=1}^{\infty} \xi_{n}(\Omega)) = P(\xi_{n}(\Omega)) = 1 \). This implies that \( \exists \theta_0 \in \mathbb{N} \) such that \( P(\xi_{n}(\Omega)) > 0 \). By (i), we have \( P(\xi_{n}(\Omega)) = 1 \).

Let \( x_1 := \bar{x} \) and \( E_1 := B_\varepsilon(x_1, \theta_0) \subseteq Z \). By Proposition 4.38, \( E_1 \) is separable, and \( \exists D_1 \subseteq E_1 \) that is a countable dense subset of \( E_1 \). Then, \( 1 = P(\bigcup_{n \in D_1} \xi_{n}(\Omega)) \leq \sum_{n \in D_1} P(\xi_{n}(\Omega)) \). Hence, \( \exists \theta_2 \in D_1 \subseteq E_1 \) such that \( P(\xi_{n}(\Omega)) > 0 \). By (i), we have \( P(\xi_{n}(\Omega)) = 1 \).

Let \( E_2 := B_\varepsilon(x_2, \theta_2) \subseteq Z \). By Proposition 4.38, \( E_2 \) is separable, and \( \exists D_2 \subseteq E_2 \) that is a countable dense subset of \( E_2 \). Then, \( 1 = P(\bigcup_{n \in D_2} \xi_{n}(\Omega)) \leq \sum_{n \in D_2} P(\xi_{n}(\Omega)) \). Hence, \( \exists \theta_3 \in D_2 \subseteq E_2 \) such that \( P(\xi_{n}(\Omega)) > 0 \). By (i), we have \( P(\xi_{n}(\Omega)) = 1 \).

Recursively, we have \( E_n := B_\varepsilon(x_n, \theta_n) \), and \( \exists x_{n+1} \in E_n \) such that \( P(\xi_{n}(E_{n+1})) = 1 \), \( \forall n \in \mathbb{N} \). Thus, we have \( (x_n)_{n=1}^{\infty} \subseteq Z \) with \( \rho(x_n, x_{n+1}) < 3^{-n+1}\theta_n \), \( \forall n \in \mathbb{N} \). It is obvious that \( (x_n)_{n=1}^{\infty} \) is a Cauchy sequence. Since \( Z \) is complete, we have \( \lim_{n \to \infty} x_n =: x_0 \in Z \). It is a simple matter to show that \( x_0 \in B_\varepsilon(x_n, 3^{-n+2^{-2}}\theta_n) =: E_0 \subseteq E_n \), \( \forall n \in \mathbb{N} \). By the preceding discussion, the second equality follows from Proposition 2.5; the third equality follows from Proposition 11.5; and the last equality follows from \( E_n \subseteq E_{n+1} \) and then \( P(\xi_{n}(E_{n+1})) = 1 \), \( \forall n \in \mathbb{N} \).

This completes the proof of the theorem.

**Proposition 14.44** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( \mathcal{B} \) be a separable Banach space, \( (X_n)_{n=1}^{\infty} \) be a sequence of \( \mathcal{B} \)-valued random variables, and \( \mathcal{T} \) be the tail \( \sigma \)-algebra of \( (X_n)_{n=1}^{\infty} \). Then, the following statements hold.

(i) \( F_1 := \{ \omega \in \Omega \mid \lim_{n \to \infty} X_n(\omega) \in \mathcal{B} \} \in \mathcal{T} \).

(ii) \( F_2 := \{ \omega \in \Omega \mid \lim_{n \to \infty} \sum_{i=1}^{n} X_i(\omega) \in \mathcal{B} \} \in \mathcal{T} \).

(iii) \( F_3 := \{ \omega \in \Omega \mid \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \in \mathcal{B} \} \in \mathcal{T} \).

**Proof**

(i) By Proposition 11.51, \( F_1 = \bigcap_{j=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{l=1}^{\infty} \bigcap_{n=m}^{\infty} E_{j,l,n} \subseteq B \), where \( E_{j,l,n} := \{ \omega \in \Omega \mid \| X_j(\omega) - X_n(\omega) \| < 1/j \} \), \( \forall j, l, n \in \mathbb{N} \). This shows that \( F_1 \in \sigma(X_1, X_2, \ldots) \). Since \( \lim_{n \to \infty} X_n(\omega) \in \mathcal{B} \) is independent
14.6. CHARACTERISTIC FUNCTIONS

of $X_1, \ldots, X_n$, then $F_1 = \bigcap_{j=1}^{\infty} \bigcup_{m=0}^{\infty} \bigcap_{l=m}^{\infty} \bigcap_{n=m}^{\infty} E_{j,l,n} \in T_k$, $\forall k \in \mathbb{N}$, where $T_k$ is as defined in Definition 14.42. Hence, $F_1 \in \bigcap_{k \in \mathbb{N}} T_k = T$.

(ii) By (i), $F_2 = \bigcap_{j=1}^{\infty} \bigcup_{m=0}^{\infty} \bigcap_{l=m}^{\infty} \bigcap_{n=m}^{\infty} E_{j,l,n} \in B$, where $E_{j,l,n} := \{ \omega \in \Omega \mid \left\| \sum_{i=1}^{n} X_i(\omega) \right\|_y < 1/j \} \forall j, l, n \in \mathbb{N}$. This shows that $F_1 \in \sigma(X_1, X_2, \ldots)$. Since $\lim_{n \in \mathbb{N}} \sum_{i=1}^{n} X_i(\omega) \in \mathbb{Y}$ if and only $\lim_{n \in \mathbb{N}} \sum_{i=1}^{n} \omega_i(\omega) \in \mathbb{Y}$, $\forall k \in \mathbb{N}$, then $F_2 = \bigcap_{j=1}^{\infty} \bigcup_{m=0}^{\infty} \bigcap_{l=m}^{\infty} \bigcap_{n=m}^{\infty} E_{j,l,n} \in T_k$, $\forall k \in \mathbb{N}$. Hence, $F_2 \in \bigcap_{k \in \mathbb{N}} T_k = T$.

(iii) It is easy to see that $F_3 \in \sigma(X_1, X_2, \ldots)$. Let $F_3^{(k)} := \{ \omega \in \Omega \mid \lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \in \mathbb{Y} \} \forall k \in \mathbb{N}$. We will show that $F_3 = F_3^{(k)}$, $\forall k \in \mathbb{N}$. Fix any $k \in \mathbb{N}$. $\forall \omega \in F_3$, we have $\lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \in \mathbb{Y}$. Then, $\lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) = \lim_{n \in \mathbb{N}} \left( \frac{n-k}{n} \sum_{i=1}^{n-k} X_i(\omega) - \frac{1}{n} \sum_{i=1}^{k} X_i(\omega) \right) \in \mathbb{Y}$. Then, $\omega \in F_3^{(k)}$. By the arbitrariness of $\omega$, we have $F_3 \subseteq F_3^{(k)}$.

On the other hand, $\forall \omega \in F_3^{(k)}$, we have $\lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) \in \mathbb{Y}$. Then, $\lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) = \lim_{n \in \mathbb{N}} \left( \frac{n-k}{n} \sum_{i=1}^{n-k} X_i(\omega) - \frac{1}{n} \sum_{i=1}^{k} X_i(\omega) \right) \in \mathbb{Y}$. Then, $\omega \in F_3$. By the arbitrariness of $\omega$, we have $F_3^{(k)} \subseteq F_3$. Hence, $F_3 = F_3^{(k)}$.

It is clear that $F_3^{(k)} \in T_k$, $\forall k \in \mathbb{N}$. Then, we have $F_3 \in T_k$, $\forall k \in \mathbb{N}$, and hence $F_3 \in T$.

This completes the proof of the proposition.

14.6 Characteristic Functions

Definition 14.45 Let $\Omega := (\Omega, B, P)$ be a probability measure space, and $x : \Omega \rightarrow \mathbb{R}$ be a random variable with law $L_x$ and a cumulative distribution function $F : \mathbb{R} \rightarrow [0, 1] \subset \mathbb{R}$. The characteristic function of $x$ is $\Psi_x : \mathbb{R} \rightarrow \mathbb{C}$ defined by $\Psi_x(\theta) = E(\exp(i \theta x)) = E(\cos(\theta x)) + i E(\sin(\theta x)) = \int_{\Omega} \exp(i \theta x(\omega)) \, dP(\omega) = \int_{\mathbb{R}} \exp(i \theta x) \, dL_x(x) = \int_{\mathbb{R}} \exp(i \theta x) \, dF(x)$, $\forall \theta \in \mathbb{R}$.

By Lebesgue Dominated Convergence Theorem 11.91, the characteristic function of a random variable is well defined. The last two equalities in the above definition holds as deduced in the following. Fix any $\theta \in \mathbb{R}$. Let $f : \Omega \rightarrow \mathbb{C}$ be defined by $f(\omega) = \exp(i \theta x(\omega))$, $\forall \omega \in \Omega$, and $g : \mathbb{R} \rightarrow \mathbb{C}$ be defined by $g(x) = \exp(i \theta x)$, $\forall x \in \mathbb{R}$. By Proposition 11.69, let $I(\mathbb{C})$ be the integration system on $\mathbb{C}$. Then, $\Psi_x(\theta) = \int_{\Omega} f(\omega) \, dP(\omega) = \lim_{R \in I(\mathbb{C})} F_R$, where $(F_R)_{R \in I(\mathbb{C})}$ is the net as defined in Definition 11.70 for the integral $\int_{\Omega} f(\omega) \, dP(\omega)$. $\forall R \in \{ (\alpha, \mathbb{A}) \mid \alpha \in \Lambda \} \in I(\mathbb{C})$, $F_R = \sum_{\alpha \in \Lambda} y_{\alpha} P(\{ f_{\text{inv}}(U_{\alpha}) \}) = \sum_{\alpha \in \Lambda} y_{\alpha} P(\{ f_{\text{inv}}(f_{\text{inv}}(U_{\alpha})) \}) = \sum_{\alpha \in \Lambda} y_{\alpha} L_x(\{ f_{\text{inv}}(U_{\alpha}) \}) = \sum_{\alpha \in \Lambda} y_{\alpha} F_{\alpha}$, where the third equality follows from Definition 14.14, and $(F_R)_{R \in I(\mathbb{C})}$ is the net as defined in Definition 11.70.
for the integral \( \int_{\mathbb{R}} \tilde{f}(x) \, dL_x(x) \). Hence, we have \( \Psi_x(\theta) = \int_{\Omega} f(\omega) \, dP(\omega) = \int_{\mathbb{R}} \tilde{f}(x) \, dL_x(x) = \int_{\mathbb{R}} \exp(i\theta x) \, dL_x(x) \). By Lebesgue-Stieltjes integral Definition 12.57, we have \( \int_{\mathbb{R}} \exp(i\theta x) \, dL_x(x) = \int_{\mathbb{R}} \exp(i\theta x) \, dF(x) \).

**Proposition 14.46** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, and \( x : \Omega \to \mathbb{R} \) be a random variable with characteristic function \( \Psi_x : \mathbb{R} \to \mathbb{C} \). Then, the following statements hold.

(i) \( \Psi_x(0) = 1 \).
(ii) \( |\Psi_x(\theta)| \leq 1, \forall \theta \in \mathbb{R} \).
(iii) \( \Psi_x \) is continuous.
(iv) \( \Psi_{-\theta}(\theta) = \overline{\Psi_x(\theta)}, \forall \theta \in \mathbb{R} \).
(v) \( \Psi_{a\theta+b}(\theta) = \exp(ib\theta)\Psi_x(a\theta), \forall \theta \in \mathbb{R}, \forall a, b \in \mathbb{R} \).
(vi) If \( n \in \mathbb{N} \) with \( E(|x|^n) < \infty \), then, \( \Psi_x \) is \( C_n \), and \( \Psi_x^{(i)}(\theta) = E((ix)^i \exp(i\theta x)), \forall i = 1, \ldots, n \).

**Proof**

(i) \( \Psi_x(0) = E(1) = 1 \).
(ii) Note that \( |\exp(i\theta x(\omega))| = 1, \forall \omega \in \Omega, \forall \theta \in \mathbb{R} \). By Proposition 11.92, we have \( |\Psi_x(\theta)| \leq \int_{\Omega} |\exp(i\theta x(\omega))| \, dP(\omega) = 1, \forall \theta \in \mathbb{R} \).
(iii) \( \forall \theta_0 \in \mathbb{R} \), we have \( \lim_{\theta \to \theta_0} \exp(i\theta x(\omega)) = \exp(i\theta_0 x(\omega)), \forall \omega \in \Omega \). Then, by Lebesgue Dominated Convergence Theorem 11.91, \( \Psi_x(\theta_0) = \lim_{\theta \to \theta_0} \int_{\Omega} \exp(i\theta x(\omega)) \, dP(\omega) = \lim_{\theta \to \theta_0} \int_{\Omega} \exp(i\theta x(\omega)) \, dP(\omega) = \lim_{\theta \to \theta_0} \Psi_x(\theta) \).
By Proposition 3.74, \( \Psi_x \) is continuous at \( \theta_0 \). By the arbitrariness of \( \theta_0 \), we have \( \Psi_x \) is continuous.
(iv) \( \Psi_{-\theta}(\theta) = E(\exp(-i\theta x)) = E(\overline{\exp(i\theta x)}) = \overline{E(\exp(i\theta x))} = \overline{\Psi_x(\theta)}, \forall \theta \in \mathbb{R} \), where the second equality follows from the definition of exponential function; and the third equality follows from Proposition 11.92.
(v) \( \Psi_{a\theta+b}(\theta) = E(\exp(i\theta(ax+b))) = E(\exp(i\theta b)) \exp(i\theta ax) = \exp(i\theta b) \cdot E(\exp(i\theta ax)), \forall \theta \in \mathbb{R} \), where the second equality follows from the definition of exponential function; the third equality follows from Proposition 11.92; and the fourth equality follows from Definition 14.45.
(vi) Let \( n \in \mathbb{N} \) with \( E(|x|^n) < \infty \). Then, \( x \in \mathbb{L}_n(\Omega, \mathbb{R}) \). By Hölder’s Inequality Theorem 11.178, \( E(|x|^i) = \int_{\Omega} |x(\omega)|^i \, dP(\omega) \leq \left( \int_{\Omega} (|x(\omega)|^i)^{n/i} \, dP(\omega) \right)^{i/n} \left( \int_{\Omega} 1^{n/(n-i)} \, dP(\omega) \right)^{(n-i)/n} = (E(|x|^n))^{i/n} < \infty, \forall i = 1, \ldots, n-1 \). Hence, \( x \in \mathbb{L}_n(\Omega, \mathbb{R}), i = 1, \ldots, n \).

Fix any \( i = 1, \ldots, n \). By Lebesgue Dominated Convergence Theorem 11.91 and \( E(|x|^i) < \infty, E((ix)^i \exp(i\theta x)) \in \mathbb{C}, \forall \theta \in \mathbb{R}, \forall \theta_0 \in \mathbb{R} \), we have \( \lim_{\theta \to \theta_0} (ix)^i \exp(i\theta x(\omega)) = (ix)^i \exp(i\theta_0 x(\omega)), \forall \omega \in \Omega \). Then, by Lebesgue Dominated Convergence Theorem 11.91, \( \lim_{\theta \to \theta_0} E((ix)^i \cdot \exp(i\theta x)) = E((ix)^i \exp(i\theta_0 x)). \) By Proposition 3.74 and arbitrariness of \( \theta_0 \), \( E((ix)^i \exp(i\theta x)) \) as a function of \( \theta \) is continuous.
Clearly, \( \Psi_x(\theta) = E(\exp(i \theta x)) \), \( \forall \theta \in \mathbb{R} \). Inductively, assume that \( \Psi_x^{(i)}(\theta) = E((ix)^i \exp(i \theta x)) \), \( \forall \theta \in \mathbb{R} \), for some \( 0 \leq i < n \). Consider the case \( i + 1 \in \{1, \ldots, n\} \). Fix any \( \theta_0 \in \mathbb{R} \).

\[
\Psi^{(i+1)}_x(\theta_0) = \lim_{\theta \to \theta_0} \frac{\Psi^{(i)}_x(\theta) - \Psi^{(i)}_x(\theta_0)}{\theta - \theta_0} = \lim_{\theta \to \theta_0} \frac{E((ix)^i \exp(i \theta x)) - E((ix)^i \exp(i \theta_0 x))}{\theta - \theta_0} = \lim_{\theta \to \theta_0} E \left( (ix)^i \frac{\exp(i \theta x) - \exp(i \theta_0 x)}{\theta - \theta_0} \right) = \lim_{\theta \to \theta_0} E((ix)^{i+1} \exp(i \theta x)) = E((ix)^{i+1} \exp(i \theta_0 x))
\]

where the first equality follows from the definition of the derivative; the second equality follows from the inductive assumption; the third equality follows from Proposition 11.92; the fourth equality follows from Taylor’s Theorem 9.48 and \( \theta_1(\omega) \in (\theta \land \theta_0, \theta \lor \theta_0) \subset \mathbb{R} \), \( \forall \omega \in \Omega \); and last equality follows from Lebesgue Dominated Convergence Theorem 11.91 and Proposition 4.16.

Hence, we have \( \Psi^{(i)}_x(\theta) = E((ix)^i \exp(i \theta x)) \), \( \forall \theta \in \mathbb{R} \), \( \forall i = 0, \ldots, n \).

This completes the proof of the proposition.

\( \square \)

**Proposition 14.47** Let \( \Omega := (\Omega, \mathcal{B}, \mathcal{P}) \) be a probability measure, \( X \) and \( Y \) be two independent \( \mathbb{R} \)-valued random variables. Then, \( \Psi_{X+Y}(\theta) = \Psi_X(\theta)\Psi_Y(\theta) \), \( \forall \theta \in \mathbb{R} \).

**Proof** \( \forall \theta \in \mathbb{R} \), we have \( \Psi_{X+Y}(\theta) = E(\exp(i \theta (X + Y))) = E(\exp(i \theta X) \exp(i \theta Y)) = E(\exp(i \theta X))E(\exp(i \theta Y)) = \Psi_X(\theta)\Psi_Y(\theta) \), where the first equality follows from Definition 14.45; the second equality follows from Proposition 9.88; the third equality follows from Proposition 14.9; and the last equality follows from Definition 14.45. This completes the proof of the proposition.

\( \square \)

**Theorem 14.48 (Lévy’s Inversion Formula)** Let \( \Omega := (\Omega, \mathcal{B}, \mathcal{P}) \) be a probability measure, \( X : \Omega \to \mathbb{R} \) be random variables with law \( L : \mathcal{B}(\mathbb{R}) \to [0,1] \subset \mathbb{R} \), the cumulative distribution function \( F : \mathbb{R} \to [0,1] \subset \mathbb{R} \) as defined in Proposition 12.51, and the characteristic function \( \Psi_X : \mathbb{R} \to \mathbb{C} \). Then, the following statements hold.

(i) \( \forall a, b \in \mathbb{R} \) with \( a < b \), we have

\[
\lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{\exp(-i \theta a) - \exp(-i \theta b)}{i \theta} \Psi_X(\theta) d\theta = \frac{1}{2} L(\{a\}) + L(r_{a,b}^+) + \frac{1}{2} L(\{b\}) = \frac{1}{2} (F(b) + \lim_{t \to b^-} F(t)) - \frac{1}{2} F(a) + \lim_{t \to a^-} F(t) \quad (14.5)
\]
(ii) If \( \int \Psi_X(\theta) \, d\mu_B(\theta) < \infty \), then \( F \) is \( C_1 \) and \( DF(x) = \frac{1}{2\pi} \int \Psi_X(\theta) \cdot \exp(-i\theta x) \, d\mu_B(\theta) = f(x), \forall x \in \mathbb{R} \), where the function \( f : \mathbb{R} \to \mathbb{R} \) is the probability density function of \( X \).

(iii) Under the assumption of (ii), we have the duality result:

\[
\Psi_X(\theta) = \int_{\mathbb{R}} \exp(i\theta x) f(x) \, d\mu_B(x); \quad \forall x \in \mathbb{R} \quad (14.6a)
\]

\[
f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta x) \Psi_X(\theta) \, d\mu_B(\theta); \quad \forall \theta \in \mathbb{R} \quad (14.6b)
\]

Proof : We need the following intermediate result.

**Claim 14.48.1** For all \( u, v \in \mathbb{R} \) with \( u \leq v \), we have \( |\exp(iv) - \exp(iu)| \leq |v - u| \).

**Proof of claim:** Note that \( |\exp(iv) - \exp(iu)| = |\int_u^v i \exp(it) \, dt| \leq \int_u^v |i \exp(it)| \, dt = \int_u^v 1 \, dt = v - u \), where the first equality follows from Theorem 12.83; and the first inequality follows from Proposition 11.92. This completes the proof of the claim. \( \square \)

\( \forall a, b \in \mathbb{R} \) with \( a < b \). Fix \( T \in (0, \infty) \subset \mathbb{R} \). Then,

\[
C_T := \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{-T}^{T} \left| \frac{\exp(i\theta(x-a)) - \exp(i\theta(x-b))}{i\theta} \right| \, d\theta \right) \, dL(x)
\]

\[
\leq \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{-T}^{T} \left| \frac{\theta(x-a) - \theta(x-b)}{\theta} \right| \, d\theta \right) \, dL(x)
\]

\[
= \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{-T}^{T} (b-a) \, d\theta \right) \, dL(x) = \frac{(b-a)T}{\pi} < \infty
\]

where the first inequality follows from Claim 14.48.1. By Tonelli’s Theorem 12.29, \( \frac{\exp(i\theta(x-a)) - \exp(i\theta(x-b))}{i\theta} \) is absolutely integrable over \( \mathbb{R} \times (-T,T) \) with respect to measure \( \mu_B \times L \). Then, by Fubini’s Theorem 12.31, we have

\[
\frac{1}{2\pi} \int_{-T}^{T} \exp(-i\theta a) - \exp(-i\theta b) \Psi_X(\theta) \, d\theta
\]

\[
= \frac{1}{2\pi} \int_{-T}^{T} \exp(i\theta a) - \exp(i\theta b) \left( \int_{\mathbb{R}} \exp(i\theta x) \, dL(x) \right) \, d\theta
\]

\[
= \frac{1}{2\pi} \int_{\mathbb{R}} \left( \int_{-T}^{T} \exp(i\theta(x-a)) - \exp(i\theta(x-b)) \right) \, dL(x)
\]

Note that

\[
f(x, T) := \frac{1}{2\pi} \int_{-T}^{T} \exp(i\theta(x-a)) - \exp(i\theta(x-b)) \, d\theta
\]

\[
= \frac{1}{2\pi} \int_{-T}^{T} \frac{1}{i\theta} \left( \cos(\theta(x-a)) + i \sin(\theta(x-a)) - \cos(\theta(x-b)) \right)
\]
14.6. CHARACTERISTIC FUNCTIONS

\[ \sin(\theta(x-b)) \, d\theta \]

\[ = \frac{1}{2\pi} \int_{-T}^{T} \frac{\sin(\theta(x-a)) - \sin(\theta(x-b))}{\theta} \, d\theta \]

\[ = \frac{1}{\pi} \int_{0}^{\infty} \left( \frac{\sin(\theta(x-a))}{\theta} - \frac{\sin(\theta(x-b))}{\theta} \right) \, d\theta \]

\[ = \frac{1}{\pi} \left( \text{sgn}(x-a)S(T|x-a|) - \text{sgn}(x-b)S(T|x-b|) \right) \]

where \( S(U) := \int_{0}^{U} \sin(x) \, dx \), \( \forall U \in [0, \infty) \subset \mathbb{R} \). It is standard result that \( \lim_{U \to \infty} S(U) = \frac{\pi}{2} \) (see Example 33.9 of Bartle (1976) and Proposition 12.100). Then, \( \lim_{T \to \infty} f(x, T) = f_0(x) = \begin{cases} 0 & \text{if } x < a \text{ or } x > b \\ 1 & \text{if } a < x < b \\ \frac{1}{2} & \text{if } x = a \text{ or } x = b \end{cases} \)

\( \forall x \in \mathbb{R} \). By Bounded Convergence Theorem 11.77, we have

\[ \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{\exp(-i\theta a) - \exp(-i\theta b)}{i\theta} \Psi_X(\theta) \, d\theta = \lim_{T \to \infty} \int_{\mathbb{R}} f(x, T) \, dL(x) \]

\[ = \int_{\mathbb{R}} f_0(x) \, dL(x) = \frac{1}{2} L\{\{a\}\} + L(r_{a,b}) + \frac{1}{2} L\{\{b\}\} \]

\[ = \frac{1}{2}(F(b) + \lim_{t \to -b} F(t)) - \frac{1}{2}(F(a) + \lim_{t \to a} F(t)) \]

Thus, statement (i) is established.

(ii) Under the assumption that \( \Psi_X \in \mathcal{L}_1(\mathbb{R}, \mathbb{C}) \), let \( c := \| \Psi_X \|_{\mathcal{L}_1(\mathbb{R}, \mathbb{C})} \). By Claim 14.48.1, we have

\[ \left| \frac{\exp(-i\theta a) - \exp(-i\theta b)}{i\theta} \Psi_X(\theta) \right| \leq (b-a) |\Psi_X(\theta)| \]

which is then absolutely integrable over \( \mathbb{R} \). Then, by (i) and Lebesgue Dominated Convergence Theorem 11.91, we have, \( \forall a, b \in \mathbb{R} \) with \( a < b \),

\[ \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\exp(-i\theta a) - \exp(-i\theta b)}{i\theta} \Psi_X(\theta) \, d\mu_B(\theta) = \frac{1}{2} L\{\{a\}\} + L(r_{a,b}) \]

\[ + \frac{1}{2} L\{\{b\}\} = \frac{1}{2}(F(b) + \lim_{t \to -b} F(t)) - \frac{1}{2}(F(a) + \lim_{t \to a} F(t)) =: K(a, b) \]

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), let \( \delta := \frac{\pi (b-a)}{1 + \epsilon} > 0 \), \( \forall (a_1, b_1) \in \mathcal{B}_{\mathbb{R}^2}(\{a,b\}, \delta) \), we have

\[ |K(a_1, b_1) - K(a, b)| \]

\[ = \frac{1}{2\pi} \left| \int_{\mathbb{R}} \frac{\exp(-i\theta a_1) - \exp(-i\theta a) + \exp(-i\theta b) - \exp(-i\theta b_1)}{i\theta} \Psi_X(\theta) \, d\mu_B(\theta) \right| \]

\[ \leq \frac{1}{2\pi} \int_{\mathbb{R}} \left| \frac{\exp(-i\theta a_1) - \exp(-i\theta a) + \exp(-i\theta b) - \exp(-i\theta b_1)}{i\theta} \right| \]
\[
\begin{align*}
&\quad \cdot |\Psi_X(\theta)| \, d\mu_B(\theta) \\
&\leq \frac{1}{2\pi} \int_{\mathbb{R}} \left( |\exp(-i\theta a_1) - \exp(-i\theta a)| \right. \\
&\quad + \left. |\exp(-i\theta b) - \exp(-i\theta b_1)| \right) \\
&\quad \cdot |\Psi_X(\theta)| \, d\mu_B(\theta) \\
&\leq \frac{1}{2\pi} \left( |a - a_1| + |b - b_1| \right) |\Psi_X(\theta)| \, d\mu_B(\theta) \\
&= \frac{1}{2\pi} \left( |a - a_1| + |b - b_1| \right) \lim_{\theta \to 0} |\Psi_X(\theta)| \, d\mu_B(\theta) \\
&\leq \frac{1}{\sqrt{2\pi}} \sqrt{2((a - a_1)^2 + (b - b_1)^2)} \epsilon < \frac{1}{\sqrt{2\pi}} \delta c < \epsilon
\end{align*}
\]

where the first inequality follows from Proposition 11.92; and the third inequality follows from Claim 14.48.1. This shows that \( K(a, b) \) is continuous at \((a, b), \forall a, b \in \mathbb{R} \) with \( a < b \). Thus, we may conclude that \( F \) is continuous. Thus, we have

\[
K(a, b) = F(b) - F(a) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta a) - \exp(-i\theta b) \Psi_X(\theta) \, d\mu_B(\theta)
\]

Fix any \( a_0 \in \mathbb{R} \), we have

\[
\lim_{b \to a_0^+} \frac{F(b) - F(a_0)}{b - a_0} = \lim_{b \to a_0^-} \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta a_0) - \exp(-i\theta b) \Psi_X(\theta) \, d\mu_B(\theta)
\]

where the second equality follows by the Lebesgue Dominated Convergence Theorem 11.91. Similarly,

\[
\lim_{a \to a_0^-} \frac{F(a_0) - F(a)}{a_0 - a} = \lim_{a \to a_0^-} \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta a) - \exp(-i\theta a_0) \Psi_X(\theta) \, d\mu_B(\theta)
\]

Therefore, we have \( D F(a_0) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta a_0) \Psi_X(\theta) \, d\mu_B(\theta) \). By the arbitrariness of \( a_0 \), we have \( D F(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-i\theta x) \Psi_X(\theta) \, d\mu_B(\theta) = f(x) \), \( \forall x \in \mathbb{R} \). Again, by Lebesgue Dominated Convergence Theorem 11.91, we have \( f \) is continuous. Hence, \( F \) is \( C_1 \).

(iii) By (ii), we have (14.6b). By Definition 14.45, we have \( \Psi_X(\theta) = \int_{\mathbb{R}} \exp(i\theta x) \, dF(x) = \int_{\mathbb{R}} \exp(i\theta x) f(x) \, d\mu_B(x) \), where the second equality follows from Propositions 12.102 and 11.168.

This completes the proof of the theorem. \( \square \)

By Lévy Inversion Formula, Theorem 14.48, if the characteristic function \( g : \mathbb{R} \to \mathbb{C} \) for a random variable \( X \) with law \( \mu \) and cumulative distribution function \( F \) (as defined in Definition 14.14) is known, then (i)
of Theorem 14.48 implies that the value \( \frac{1}{n} \mu(\{a\}) + \mu((a, b)) + \frac{1}{n} \mu(\{b\}) \) is uniquely determined, \( \forall a, b \in \mathbb{R} \) with \( a < b \). Then, \( g \) uniquely determines \( \mu((a, b)) = F(b) - F(a), \forall a, b \in \mathbb{R} \) with \( a < b \) and \( F \) being continuous at \( a \) and \( b \). Since \( \mu \) is a probability measure, then \( g \) uniquely determines \( F(b), \forall b \in \mathbb{R} \) with \( F \) being continuous at \( b \). By Proposition 12.51, \( F \) is a function of bounded variation. Then, it is right continuous and monotonically nondecreasing. Then, \( F \) has countably many points of discontinuity (the total number of discontinuity with a jump of at least \( 1/n \) is at most \( n \)). Which means that the set of points in \( \mathbb{R} \) where \( F \) is continuous is dense in \( \mathbb{R} \). Then, \( F \) is uniquely determined by \( g \). This further implies that \( \mu \) is uniquely determined by \( g \) by Theorem 12.50.

### 14.7. Convergence in Distribution

**Definition 14.49** Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure space, \( m \in \mathbb{N}, (x_n)_{n=1}^{\infty} \) be a sequence of \( \mathbb{R}^m \)-valued random variables, where \( x_n : \Omega \to \mathbb{R}^m \) and admits law \( L_n, \forall n \in \mathbb{N} \), \( x_0 \) be an \( \mathbb{R}^m \)-valued random variable with law \( L_0 \). We will say that \((x_n)_{n=1}^{\infty}\) converges to \( x_0 \) in distribution if \( \lim_{n \in \mathbb{N}} L_n = L_0 \) weak*. In this case, we will write \( x_n \xrightarrow{\ast} x_0 \) as \( n \to \infty \), or \( \lim_{n \in \mathbb{N}} x_n \xrightarrow{\ast} x_0 \). Let \( F_n : \mathbb{R}^m \to [0, 1] \subset \mathbb{R} \) be the cumulative distribution functions of \( x_n \) as defined in Definition 14.14, \( \forall n \in \mathbb{Z}_+ \), we will write \( \lim_{n \in \mathbb{N}} F_n \xrightarrow{\ast} F_0 \) or \( F_n \xrightarrow{\ast} F_0 \) as \( n \to \infty \).

In the above definition, \( L_n \) is a probability measures on the measurable space \((\mathbb{R}^m, \mathcal{B}_B(\mathbb{R}^m))\), \( \forall n \in \mathbb{Z}_+ \). Then, \( L_n \in \text{MFt}(\mathbb{R}^m, \mathbb{R}) = \text{MF}(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m), \mathbb{R}) = (\mathcal{C}_0(\mathbb{R}^m, \mathbb{R}))^* \) by Riesz Representation Theorem 11.209, \( \forall n \in \mathbb{Z}_+ \). Thus, \( \lim_{n \in \mathbb{N}} L_n = L_0 \) weak* makes sense. Therefore, the definition makes sense.

We will note that the convergence in distribution defined above is actually weaker than the its classical definition (Williams, 1991, See Definition 17.1). The classical definition implies our definition, on the other hand, our definition when coupled with the uniform growth boundedness assumption yields the classical definition.

**Theorem 14.50 (Modes of Convergence)** Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure space, \( m \in \mathbb{N}, (x_n)_{n=0}^{\infty} \) be a sequence of \( \mathbb{R}^m \)-valued random variables, where \( x_n : \Omega \to \mathbb{R}^m \) and admits law \( L_n, \forall n \in \mathbb{N} \). Then, the following statements hold.

(i) If \( \lim_{n \in \mathbb{N}} x_n = x_0 \) a.e. in \( \Omega \), then \( \lim_{n \in \mathbb{N}} x_n = x_0 \) in measure in \( \Omega \).

(ii) If \( \lim_{n \in \mathbb{N}} x_n = x_0 \) in measure in \( \Omega \), then \( \lim_{n \in \mathbb{N}} x_n \xrightarrow{\ast} x_0 \).

**Proof**

(i) Let \( \lim_{n \in \mathbb{N}} x_n = x_0 \) a.e. in \( \Omega \). Then, by Definition 11.42, there exists \( E \in \mathcal{B} \) with \( \mathbb{P}(E) = 0 \) such that \( \lim_{n \in \mathbb{N}} x_n(\omega) = x_0(\omega), \forall \omega \in \Omega \setminus E \). Fix any \( \epsilon \in \mathbb{R}_+ \). Let \( E_n := \{ \omega \in \Omega \mid |x_n(\omega) - x_0(\omega)| > \epsilon \}, \forall n \in \mathbb{N} \).
Clearly, \( \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} E_j \subseteq E \). Hence, we have \( 0 \leq P(\bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} E_j) \leq P(E) = 0 \). This implies that, by Proposition 11.5 and \( P(\Omega) = 1 < \infty \), \( \lim_{n \to \infty} P(\bigcup_{j=n}^{\infty} E_j) = 0 \), which further implies that \( \lim_{n \in \mathbb{N}} P(E_n) = 0 \). Hence, \( \exists N \in \mathbb{N} \), \( \forall n \in \mathbb{N} \) with \( n \geq N \), we have \( P(E_n) < \epsilon \). Therefore, \( \lim_{n \in \mathbb{N}} x_n = x_0 \) in measure in \( \Omega \).

(ii) Let \( \lim_{n \in \mathbb{N}} x_n = x_0 \) in measure in \( \Omega \). Fix any \( z \in \mathcal{C}(\mathbb{R}^m, \mathbb{R}) \). Let \( \| z \|_{C(\mathbb{R}^m, \mathbb{R})} =: c \in \mathbb{R}_+ \). Fix any \( \epsilon \in \mathbb{R}_+ \). Since \( z \in \mathcal{C}(\mathbb{R}^m, \mathbb{R}) \), then \( \exists M \in \mathbb{R}_+ \) such that \( \forall s \in \mathbb{R}^m \) with \( |s| > M \), we have \( |z(s)| < \frac{\epsilon}{4} \). By Proposition 5.39, \( z_k \) is uniformly continuous, where \( K := \overline{B}_m(0, m + 1) \subseteq \mathbb{R}^m \) is a compact set by Proposition 5.40. Then, \( \exists \delta \in (0, \frac{\epsilon}{4} \land 1) \subseteq \mathbb{R} \), such that \( |z(s_1) - z(s_2)| < \frac{\epsilon}{4} \), \( \forall s_1, s_2 \in K \) with \( |s_1 - s_2| < \delta \). By Definition 11.56, there exists \( N \in \mathbb{N} \), \( \forall n \in \mathbb{N} \) with \( n \geq N \), we have \( P(\{ \omega \in \Omega \mid |x_n(\omega) - x_0(\omega)| > \frac{\delta}{4(1+c)} \}) < \frac{\delta}{4(1+c)} \). Fix any \( z \in \mathbb{N} \) with \( n \geq N \). Then, we have \( \langle L_n, z \rangle = \langle L_{0}, z \rangle \rangle = \int_{\mathbb{R}^m} |\{(dL_n(s), z(s))\} - \int_{\mathbb{R}^m} |\{(dL_0(s), z(s))\} | = \int_{\mathbb{R}^m} |\{z(x_n(\omega)) dP(\omega) - \int_{\mathbb{R}^m} z(x_n(\omega)) dP(\omega) - \int_{\mathbb{R}^m} z(z(x_n(\omega)) - z(x_0(\omega))) dP(\omega)\} - \sum_{\omega} \omega \rangle = \langle L_0, z \rangle \rangle < \epsilon \). Hence, \( \lim_{n \in \mathbb{N}} \langle L_n, z \rangle = \langle L_0, z \rangle \rangle \). By the arbitrariness of \( z \), we have \( \lim_{n \in \mathbb{N}} L_n = L_0 \) weak*. This proves that \( \lim_{n \in \mathbb{N}} x_n \Rightarrow x_0 \).

This completes the proof of the proposition. \( \square \)

**Definition 14.51** Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure space, \( m \in \mathbb{N} \), and \( M \) be a set of \( \mathbb{R}^m \)-valued random variables. \( M \) is said to be uniformly growth bounded if \( \forall \epsilon \in \mathbb{R}_+ \), \( \exists K \in \mathbb{R}_+ \), \( \forall x \in M \), we have \( x : \Omega \to \mathbb{R}^m \) admits law \( L_x : \mathbb{B}_m(\mathbb{R}^m) \to [0,1] \subseteq \mathbb{R} \) is such that \( \mathbb{P}(\{ \omega \in \Omega \mid |x(\omega)| > K \}) = L_x(\mathbb{R}^m \setminus \overline{B}_m(0, K)) < \epsilon \).

**Theorem 14.52 (Skorokhod Representation)** Let \( \mu \in \mathbb{P} \) be a probability measure on \( (\mathbb{R}, \mathbb{B}_m(\mathbb{R})) \) with cumulative distribution function \( F_n : \mathbb{R} \to [0,1] \subseteq \mathbb{R} \) as defined in Definition 14.14, \( \forall n \in \mathbb{Z}_+ \). Assume that \( \forall b \in \mathbb{R} \) with \( F_0 \) being continuous at \( b \), we have \( \lim_{n \in \mathbb{N}} F_n(b) = F_0(b) \). Then, there
exists a probability measure space \( \Omega := (\Omega, \mathcal{B}, P) \) and a sequence of random variables \( \{x_n\}_{n=0}^{\infty} \), such that \( x_n : \Omega \to \mathbb{R} \) admits law \( \mu_n, \forall n \in \mathbb{Z}_+ \), and \( \lim_{n \to 0} x_n(\omega) = x_0(\omega) \) a.e. \( \omega \in \Omega \).

**Proof**

Define \( E := \{ b \in \mathbb{R} \mid F_0 \) is continuous at \( b \} \subseteq \mathbb{R} \). Note that \( F_0 \) is monotonically nondecreasing and is of bounded variation. Then, \( F_0 \) has countably many points of discontinuity (the total number of discontinuity with a jump of at least \( 1/n \) is at most \( n \)).

Take \( \Omega := ((0,1), \mathcal{B}, P) \) to be the finite measure subspace of \( \mathbb{R} \). Fix any \( n \in \mathbb{Z}_+ \), define \( x_n^+(\omega) := \inf \{ z \in \mathbb{R} \mid F_n(z) > \omega \} \) and \( x_n^-(\omega) := \inf \{ z \in \mathbb{R} \mid F_n(z) \geq \omega \} \), \( \forall \omega \in (0,1) \subseteq \mathbb{R} \). By Proposition 12.54, \( F_n \) is \( \mathcal{B}_B(\mathbb{R}) \)-measurable. Since \( \mu_n \) is a probability measure, then \( \lim_{n \to \infty} F_n(b) = 0 \) and \( \lim_{n \to \infty} F_n(b) = 1 \). This implies that \( x_n^- \) is \( \mathbb{R} \)-valued and \( x_n^+ \) is \( \mathbb{R} \)-valued, \( \forall \omega \in (0,1) \subseteq \Omega \). It is easy to show that \( \omega \in \Omega \) if and only if \( x_n^-(\omega) \leq z \). Then, by Proposition 11.35, \( x_n^- \) is \( \mathcal{B} \)-measurable. Thus, \( x_n^- \) is a random variable on \( \Omega \). The preceding argument further implies that \( x_n^- \) admits law \( \mu_n \) since its cumulative distribution function as defined in Definition 14.14 is \( F_n \). It is easy to show that \( \{ \omega \in \Omega \mid x_n^-(\omega) \leq z \} \supseteq \{ \omega \in \Omega \mid \omega \leq F_n(z) \} \). Note that \( x_n^- \) is \( \mathbb{R} \)-valued and \( x_n^+ \) is \( \mathbb{R} \)-valued, \( \forall \omega \in \Omega \). Then, \( \{ \omega \in \Omega \mid x_n^+(\omega) \leq z \} \subseteq \{ \omega \in \Omega \mid \omega \leq F_n(z) \} \), \( \forall z \in \mathbb{R} \). Thus, we have \( (0,F(z)) \subseteq \{ \omega \in \Omega \mid x_n^+(\omega) \leq z \} \subseteq (0,F(z)) \). Then, \( \{ \omega \in \Omega \mid x_n^+(\omega) \leq z \} \) equals either \( (0,F(z)) \) or \( (0,F(z)) \), which is \( \mathcal{B} \)-measurable in either case. By the arbitrariness of \( z \) and Proposition 11.35, \( x_n^+ \) is \( \mathcal{B} \)-measurable. Hence, \( x_n^+ \) is a random variable on \( \Omega \). The preceding argument further implies that \( x_n^- \) admits law \( \mu_n \) and \( x_n^+ = x_n^- \) a.e. in \( \Omega \).

Fix \( \omega \in \Omega \). Let \( z \in \mathbb{R} \) with \( \omega < F_0(z) \). By the assumption, we have \( \lim_{n \to \infty} F_n(z) = F_0(z) > \omega \). Then, for sufficiently large \( n \), \( F_n(z) > \omega \), which implies that \( x_n^-(\omega) \leq z \). Combining the result of the above two paragraphs, we have \( x_n^- \leq \liminf_{n \to 0} x_n^-(\omega) \leq \limsup_{n \to 0} x_n^-(\omega) \leq \limsup_{n \to 0} x_n^+(\omega) \leq \omega \), \( \forall \omega \in \Omega \). When \( x_0^- = x_0^+ \), we have \( x_0^-(\omega) = \lim_{n \to 0} x_n^-(\omega) \). Since \( x_0^- = x_0^+ \) a.e. \( \omega \in \Omega \), we have \( x_0^- = \lim_{n \to 0} x_n^-(\omega) \) a.e. \( \omega \in \Omega \).
the proof of the theorem.

**Proposition 14.53** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, and $(x_n)_{n=0}^\infty$ be a sequence of $(\mathbb{R}\text{-}valued) random variables with $x_n : \Omega \to \mathbb{R}$ with law $\mu_n$, $\forall n \in \mathbb{Z}_+$, and $F_n : \mathbb{R} \to [0, 1] \subset \mathbb{R}$ be the cumulative distribution function for $\mu_n$, as defined in Definition 14.14, $\forall n \in \mathbb{Z}_+$. Then, the following statements hold.

(i) If $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle \in \mathbb{R}$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$, then $\lim_{n \to \infty} \mu_n \Rightarrow \mu_0$.

(ii) $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle \in \mathbb{R}$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$, if, and only if, $\lim_{n \to \infty} F_n(b) = F_0(b)$, $\forall b \in \mathbb{R}$ with $F_0$ being continuous at $b$.

(iii) If $\lim_{n \to \infty} \mu_n \Rightarrow \mu_0$ and $(x_n)_{n=1}^\infty$ is uniformly growth bounded, then $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle \in \mathbb{R}$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$.

**Proof**

(i) Note that $\lim_{n \to \infty} \mu_n \Rightarrow \mu_0$ means that $\lim_{n \to \infty} \mu_n = \mu_0$ weak*, that is $\lim_{n \to \infty} \langle \langle \mu_n, z \rangle \rangle = \langle \langle \mu_0, z \rangle \rangle$, $\forall z \in C_c(\mathbb{R}, \mathbb{R})$. By Riesz Representation Theorem 11.209, this is equivalent to $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle$, $\forall z \in C_c(\mathbb{R}, \mathbb{R}) \subseteq C_b(\mathbb{R}, \mathbb{R})$. Thus, (i) follows readily.

(ii) “Necessity” Let $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle \in \mathbb{R}$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$. Fix any $b \in \mathbb{R}$ and $\delta \in \mathbb{R}_+$. Define $h_{b, \delta} : \mathbb{R} \to [0, 1] \subset C_b(\mathbb{R}, \mathbb{R})$. Then, $\lim_{n \to \infty} \langle \langle \mu_n, h_{b, \delta} \rangle \rangle = \langle \langle \mu_0, h_{b, \delta} \rangle \rangle$, $F_n(b) \leq \langle \langle \mu_n, h_{b, \delta} \rangle \rangle$, and $\langle \langle \mu_0, h_{b, \delta} \rangle \rangle \leq F_0(b + \delta)$. Then, taking limit as $n \to \infty$ in the above three relations, we have $\limsup_{n \to \infty} F_n(b) \leq F_0(b + \delta)$. Since $F_0$ is a cumulative distribution function and therefore right continuous. Taking limit as $\delta \to 0^+$, we have $\limsup_{n \to \infty} F_n(b) \leq F_0(b)$. Similarly, we must have $\liminf_{n \to \infty} \langle \langle \mu_n, h_{b, \delta} \rangle \rangle = \langle \langle \mu_0, h_{b, \delta} \rangle \rangle$, $F_n(b + \delta) \geq \langle \langle \mu_n, h_{b, \delta} \rangle \rangle$, and $\langle \langle \mu_0, h_{b, \delta} \rangle \rangle \geq F_0(b)$. Taking limit as $n \to \infty$ in the above three relations, we have $\liminf_{n \to \infty} F_n(b + \delta) \geq F_0(b)$. Let $b := b + \delta$, we have $\liminf_{n \to \infty} F_n(b) \geq F_0(b - \delta)$, $\forall b \in \mathbb{R}$, $\forall \delta \in \mathbb{R}_+$. Now, fix any $b \in \mathbb{R}$ with $F_0$ being continuous at $b$. We have $F_0(b - \delta) \leq \liminf_{n \to \infty} F_n(b) \leq \limsup_{n \to \infty} F_n(b) \leq F_0(b)$, $\forall b \in \mathbb{R}$, $\forall \delta \in \mathbb{R}_+$. Now, take $\delta \to 0^+$, we have $F_0(b) = \liminf_{n \to \infty} F_n(b) \limsup_{n \to \infty} F_n(b) \leq F_0(b)$ and hence $\lim_{n \to \infty} F_n(b) = F_0(b)$. This completes the “Necessity” part of the proof.

“Sufficiency” Let $\lim_{n \to \infty} F_n(b) = F_0(b)$, $\forall b \in \mathbb{R}$ with $F_0$ being continuous at $b$. By Skorokhod Representation Theorem 14.52, there exists a probability measure space $\Omega := (\Omega, \mathcal{B}, P)$ and sequence of random variables $(Y_n)_{n=0}^\infty$ on the probability measure space $\Omega$ such that $Y_n$ admits law $\mu_n$ and cumulative distribution function $F_n$, $\forall n \in \mathbb{Z}_+$, and $\lim_{n \to \infty} Y_n = Y_0 \Rightarrow \mu_0$. Then, $\lim_{n \to \infty} \int_{\mathbb{R}} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}} \langle \langle d\mu_0(x), z(x) \rangle \rangle \in \mathbb{R}$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$. Thus, (ii) follows readily.
convergence is uniform with respect to
∀a < b
∀ This implies that,
arguments, that
and the inequalities follows from (14.7). By the assumption that (14.7), CONVERGENCE IN DISTRIBUTION
767
the first and the last equalities follow from Definition 14.14. This completes the “Sufficiency” part of the proof.
Hence, (ii) is proved.
(iii) Let limₙ→N µₙ ↠ µ₀. Fix any z ∈ C_b(ℝ, ℝ). By Bounded Convergence Theorem 11.77, ∫Ω ⟨⟨(dµₙ(y), z(y))⟩⟩ = ∫Ω z(Yₙ(ω)) dP₁(ω) = ∫Ω z(Yₙ(ω)) dP₂(ω) = ∫Ω ⟨⟨(dµ₀(y), z(y))⟩⟩ ∈ ℝ, where
the first and the last equalities follow from Definition 14.14. This completes the proof.

where the first and last equalities follow from Definition 14.14; and the first
and last inequalities follow from Proposition 11.92. Taking limit as n → ∞,

\[
\limsup_{n \to \infty} (Fₙ(b) - Fₙ(a + \frac{1}{k})) \leq \lim_{n \to \infty} \langle \langle \muₙ, hₖ,a,b \rangle \rangle = \langle \langle \mu₀, hₖ,a,b \rangle \rangle
\]

where the equality follows from \( \lim_{n \to \infty} \muₙ \Rightarrow \mu₀ \) and \( hₖ,a,b \in C_c(ℝ, ℝ) \);
and the inequalities follows from (14.7). By the assumption that \( (xₙ)_{n=1}^{∞} \)
is uniformly growth bounded, we have \( \lim_{n \to \infty} Fₙ(a) = 0 \), and the
convergence is uniform with respect to \( n \in \mathbb{N} \). It is easy to prove, using \( \epsilon - \delta \)
arguments, that

\[
0 \leq \lim_{n \to \infty} \liminf_{n \to \infty} Fₙ(a) \leq \lim_{n \to \infty} \limsup_{n \to \infty} Fₙ(a)
\]
\[
= \lim_{n \to \infty} \limsup_{n \to \infty} Fₙ(a) = 0
\]

This implies that, \( \forall b \in \mathbb{R} \) and \( \forall k \in \mathbb{N} \),

\[
\limsup_{n \to \infty} Fₙ(b) = \limsup_{n \to \infty} Fₙ(b) - \lim_{n \to \infty} \limsup_{n \to \infty} Fₙ(a + \frac{1}{k})
\]
\[
= \lim_{n \to \infty} \left( \limsup_{n \to \infty} Fₙ(b) - \limsup_{n \to \infty} Fₙ(a + \frac{1}{k}) \right)
\]
\[
\leq \liminf_{n \to \infty} \limsup_{n \to \infty} Fₙ(b) - Fₙ(a + \frac{1}{k}) \leq \liminf_{n \to \infty} \langle \langle \mu₀, hₖ,a,b \rangle \rangle
\]
\[
\lim_{a \to -\infty} (F_0(b + \frac{1}{k}) - F_0(a)) = F_0(b + \frac{1}{k})
\]
where the first equality follows from (14.9); the second equality follows from Proposition 3.83; the first inequality follows from Proposition 3.83 and the fact that all limits are real numbers in the closed interval \([0, 1] \subseteq \mathbb{R}\); the second inequality follows from (14.8); the third inequality follows from (14.7) with \(n = 0\); and the last equality follows from the fact that \(F_0\) is the cumulative distribution function for \(\mu_0\) as defined in Definition 14.14. We also have, \(\forall b \in \mathbb{R}\) and \(\forall k \in \mathbb{N}\),

\[
F_0(b) = \lim_{a \to -\infty} (F_0(b) - F_0(a + \frac{1}{k})) \leq \liminf_{a \to -\infty} (\langle \mu_0, h_{k,a,b} \rangle)
\]

\[
\leq \liminf_{a \to -\infty} \liminf_{n \in \mathbb{N}} (F_n(b + \frac{1}{k}) - F_n(a))
\]

\[
\leq \liminf_{n \in \mathbb{N}} \left( \liminf_{n \in \mathbb{N}} (F_n(b + \frac{1}{k}) - \liminf_{n \in \mathbb{N}} F_n(a)) \right)
\]

\[
= \liminf_{n \in \mathbb{N}} F_n(b + \frac{1}{k}) - \liminf_{n \in \mathbb{N}} F_n(a) = \liminf_{n \in \mathbb{N}} F_n(b + \frac{1}{k})
\]

(14.11)

where the first equality follows from the fact that \(F_0\) is the cumulative distribution function of \(\mu_0\) as defined in Definition 14.14; the first inequality follows from (14.7) with \(n = 0\); the second inequality follows from (14.8); the third inequality follows from Proposition 3.83 and the fact all limits are real numbers in the closed interval \([0, 1] \subseteq \mathbb{R}\); the second equality follows from Proposition 3.83; and the last equality follows from (14.9). Then, let \(b \in \mathbb{R}\) be any point such that \(F_0\) is continuous at \(b\), we have \(\limsup_{n \in \mathbb{N}} F_n(b) \leq \liminf_{k \in \mathbb{N}} F_0(b + \frac{1}{k}) = F_0(b) = \liminf_{k \in \mathbb{N}} \limsup_{n \in \mathbb{N}} F_n(b) = \liminf_{n \in \mathbb{N}} F_n(b) \leq \limsup_{n \in \mathbb{N}} F_n(b)\), where the first inequality follows from (14.10); the first and second equalities follow from the fact that \(F_0\) is continuous at \(b\); the second inequality follows from (14.11) with \(b\) substituted by \(b - \frac{1}{k}\); and the last equality follows from Proposition 3.83. Thus, we have \(\lim_{n \in \mathbb{N}} F_n(b) = F_0(b)\). By the sufficiency result of (ii), (iii) holds.

This completes the proof of the proposition. \(\square\)

14.8 Central Limit Theorem

**Lemma 14.54 (Helly)** Let \(m \in \mathbb{N}\), \(\mathcal{X}\) be an \(m\)-dimensional Banach space over \(\mathbb{K}_1\), \(k \in \mathbb{Z}_+\), \(\mathcal{Y}\) be a \(k\)-dimensional Banach space over \(\mathbb{K}_2\), \(\mathcal{X}_n := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu_n)\) be a \(\mathcal{Y}\)-valued measure space with \(\mathcal{P} \circ \mu_n(\mathcal{X}) \leq 1\), \(\forall n \in \mathbb{N}\). Then, there exists an finite \(\mathcal{Y}\)-valued measure space \((\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)\) with \(\mathcal{P} \circ \mu(\mathcal{X}) \leq 1\), and a subsequence \((\mu_{n_i})_{i=1}^\infty\), such that \(\lim_{i \in \mathbb{N}} \mu_{n_i} \Rightarrow \mu\).

**Proof** Fix any \(n \in \mathbb{N}\). Clearly, \(\mu_n \in M_{ft}(\mathcal{X}, \mathcal{Y})\) with \(\|\mu_n\|_{M_{ft}(\mathcal{X}, \mathcal{Y})} = \mathcal{P} \circ \mu_n(\mathcal{X}) \leq 1\). By Propositions 7.90 and 7.92, \(\mathcal{Y}\) and \(\mathcal{Y}^\ast\) are reflexive
Banach spaces. By Alaoglu Theorem 7.122 and Riesz Representation Theorem 11.209, $K := \overline{B}_{M_f(X,Y)}(\theta M_{f_l}(X,Y), 1) \subseteq M_{f_l}(X,Y) = (C_c(X,Y^*)^*)^*$ is weak* compact. Clearly, $(\mu_n)_{n=1}^\infty \subseteq K$. By Propositions 5.22 and 5.26, $K$ has the Bolzano-Weierstrass property. Then, exists $\mu \in K$ such that $(\mu_n)_{n=1}^\infty$ admits a weak* cluster point $\mu$. Fix any $z \in C_c(X,Y^*)$. Then, $\langle \langle \langle \mu_n, z \rangle \rangle \rangle_{n=1}^\infty \subseteq \overline{B}_{K_2} \left( \left\{ 0, \|z\|_{C_c(X,Y)} \right\} \right) \subset K_2$ has a cluster point $\langle \langle \mu, z \rangle \rangle$, by Proposition 3.66. Then, exists a subsequence $(\mu_{n_i})_{i=1}^\infty$ such that $\lim_{i \rightarrow \infty} \langle \langle \mu_{n_i}, z \rangle \rangle = \langle \langle \mu, z \rangle \rangle$. By Riesz Representation Theorem 11.209, $C_c(X,Y^*)$ is separable. Let $D := \{ z_1, z_2, \ldots \} \subseteq C_c(X,Y^*)$ be a countable dense set. Then, there exists a subsequence $(\mu_{n_{i_j}})_{j=1}^\infty$ such that $\lim_{j \rightarrow \infty} \langle \langle \mu_{n_{i_j}}, z \rangle \rangle = \langle \langle \mu, z \rangle \rangle$, and so on and so forth.

Take the diagonal sequence in the countably many subsequences, there exists a subsequence $(\mu_{n_{i_{j_k}}})_{k=1}^\infty$ such that $\lim_{k \rightarrow \infty} \langle \langle \mu_{n_{i_{j_k}}}, z_i \rangle \rangle = \langle \langle \mu, z_i \rangle \rangle, \forall z_i \in D$. Fix any $z \in C_c(X,Y^*)$. $\forall \epsilon \in \mathbb{R}_+$, there exists $z_i \in D$ such that $\|z - z_i\|_{C_c(X,Y^*)} < \frac{\epsilon}{4}$. For that $z_i$, $\exists L \in \mathbb{N}$, $\forall l \in \mathbb{N}$ with $l \geq L$, we have $|\langle \langle \mu_{n_{i_j}}, z \rangle \rangle - \langle \langle \mu, z \rangle \rangle| < \frac{\epsilon}{4}$. Then, $|\langle \langle \mu_{n_{i_j}}, z \rangle \rangle - \langle \langle \mu, z \rangle \rangle| \leq |\langle \langle \mu_{n_{i_j}}, z \rangle \rangle - \langle \langle \mu_{n_{i_j}}, z_i \rangle \rangle| + |\langle \langle \mu_{n_{i_j}}, z_i \rangle \rangle - \langle \langle \mu, z_i \rangle \rangle| + |\langle \langle \mu, z_i \rangle \rangle - \langle \langle \mu, z \rangle \rangle| < \|\mu_{n_{i_j}}\|_{M_{f_l}(X,Y)} \|z - z_i\|_{C_c(X,Y^*)} + \frac{\epsilon}{4} + \|\mu\|_{M_{f_l}(X,Y)} \|z - z_i\|_{C_c(X,Y^*)} \leq \epsilon$, where the second inequality follows from Proposition 7.72; and the last inequality follows from the preceding discussion of this paragraph. Hence, we have $\lim_{k \rightarrow \infty} \langle \langle \mu_{n_{i_k}}, z \rangle \rangle = \langle \langle \mu, z \rangle \rangle$. By the arbitrariness of $z \in C_c(X,Y^*)$, we have that $\lim_{k \rightarrow \infty} \mu_{n_{i_k}} \Rightarrow \mu$.

This completes the proof of the lemma.

\[ \]  

**Theorem 14.55 (Lévy’s Convergence Theorem)** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure, $(X_n)_{n=1}^\infty$ be a sequence of $(\mathbb{R}$-valued) random variables with characteristic function $\Psi_{X_n} : \mathbb{R} \rightarrow \mathbb{C}$, and $g : \mathbb{R} \rightarrow \mathbb{C}$ be a function that is continuous at 0. Assume that $\sup_{\theta \in \mathbb{R}}|\Psi_{X_n}(\theta)| = g(\theta), \forall \theta \in \mathbb{R}$. Then, $(X_n)_{n=1}^\infty$ is uniformly growth bounded; and there exists a probability measure space $\Omega := (\Omega, \mathcal{B}, P)$ and a random variable $X : \Omega \rightarrow \mathbb{R}$ with characteristic function $\Psi_X = g$ such that $\lim_{n \rightarrow \infty} X_n \Rightarrow X$.

**Proof** Fix any $n \in \mathbb{N}$. Let $X_n$ admits law $\mu_n$ and cumulative distribution function $F_n : \mathbb{R} \rightarrow [0,1] \subseteq \mathbb{R}$ as defined in Definition 14.14. Then, $\mu_n \in M_{f_l}(\mathbb{R}, \mathbb{R})$ with $\|\mu_n\|_{M_{f_l}(\mathbb{R}, \mathbb{R})} = P \circ \mu_n(\mathbb{R}) = \mu_n(\mathbb{R}) = 1$. By Helly’s Lemma 14.54, there exists $\mu \in M_{f_l}(\mathbb{R}, \mathbb{R})$ with $P \circ \mu(\mathbb{R}) \leq 1$ and a subsequence $(\mu_{n_{i_k}})_{k=1}^\infty$ such that $\lim_{k \rightarrow \infty} \mu_{n_{i_k}} \Rightarrow \mu$.

We will next show that $(X_n)_{n=1}^\infty$ is uniformly growth bounded. Fix any $\theta \in \mathbb{R}_+$. By Definition 14.45, we have $\Psi_{X_n}(\theta) = \int_{\mathbb{R}} \exp(i\theta x) \exp(n F_n(x), \forall \theta \in \mathbb{R}, \forall n \in \mathbb{N}$. Then, $\Psi_{X_n}(\theta) + \Psi_{X_n}(-\theta) = \int_{\mathbb{R}} 2 \cos(\theta x) \exp(n x) \in \mathbb{R}$. This implies that $g(\theta) + g(-\theta) = \lim_{n \rightarrow \infty} (\Psi_{X_n}(\theta) + \Psi_{X_n}(-\theta)) \in \mathbb{R}$, $\forall \theta \in \mathbb{R}$. Clearly, $g(0) = \lim_{n \rightarrow \infty} \Psi_{X_n}(0) = 1$, by Proposition 14.46.
Since $g$ is continuous at 0, then $\exists \delta \in \mathbb{R}^+$ such that $|1 - g(\theta)| < \frac{\epsilon}{\delta}$, $\forall \theta \in (-\delta, \delta) \subset \mathbb{R}$. This implies that $0 \leq \frac{1}{\delta} \int_{0}^{\delta} (2 - g(\theta) - g(-\theta)) \, d\theta \leq \frac{\epsilon}{\delta}$, where the first inequality follows from the fact that $g(\theta) + g(-\theta) \in \mathbb{R}$ and (ii) of Proposition 14.46. By Bounded Convergence Theorem 11.77, we have $\lim_{n \to \infty} \frac{1}{\delta} \int_{0}^{\delta} (2 - \Psi_{X_n}(\theta) - \Psi_{X_n}(-\theta)) \, d\theta \in [0, \frac{\epsilon}{\delta}] \subset \mathbb{R}$. Then, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, we have $0 \leq \frac{1}{\delta} \int_{0}^{\delta} (2 - \Psi_{X_n}(\theta) - \Psi_{X_n}(-\theta)) \, d\theta < \epsilon$, where the first inequality follows from (ii) of Proposition 14.46. Note that

$$
\epsilon > \frac{1}{\delta} \int_{0}^{\delta} (2 - \Psi_{X_n}(\theta) - \Psi_{X_n}(-\theta)) \, d\theta = \frac{1}{\delta} \int_{0}^{\delta} (2 - \Psi_{X_n}(\theta) - \Psi_{-X_n}(\theta)) \, d\theta
$$

$$
= \frac{1}{\delta} \int_{0}^{\delta} \int_{\mathbb{R}} (2 - \exp(i\theta x) - \exp(-i\theta x)) \, dF_n(x) \, d\theta
$$

$$
= \frac{1}{\delta} \int_{0}^{\delta} \int_{\mathbb{R}} (2 - 2\cos(\theta x)) \, dF_n(x) \, d\theta
$$

$$
= \int_{\mathbb{R}} \frac{1}{\delta} \int_{0}^{\delta} (2 - 2\cos(\theta x)) \, d\theta \, dF_n(x) = 2 \int_{\mathbb{R}} \left(1 - \frac{\sin(\delta x)}{\delta x}\right) \, dF_n(x)
$$

$$
\geq 2 \int_{E_\delta} \left(1 - \frac{1}{|\delta x|}\right) \, dF_n(x) \geq \int_{E_\delta} dF_n(x) = \mu_n(E_\delta) \geq 0
$$

where the first equality follows from Proposition 14.46; the second equality follows from Definition 14.45; the fourth equality follows from Fubini Theorem 12.31 (note that the integrand is absolutely integrable in the product measure space by Tonelli Theorem 12.29); the fifth equality follows from Theorem 12.83; and the second and third inequalities follow from our definition of $E_\delta := \{ x \in \mathbb{R} \mid |x| > \frac{\delta}{\epsilon} \}$. Hence, $(X_n)_{n=1}^\infty$ is uniformly growth bounded.

By Proposition 14.53, we have $\lim_{n \to \mathbb{N}} \langle \langle \mu_n, z \rangle \rangle = \langle \langle \mu, z \rangle \rangle$, $\forall z \in C_b(\mathbb{R}, \mathbb{R})$. Let $z = \chi_{\Omega, \mathbb{R}}$, we have $\mu(\mathbb{R}) = \langle \langle \mu, z \rangle \rangle = \lim_{n \to \mathbb{N}} \langle \langle \mu_n, z \rangle \rangle = \lim_{n \to \mathbb{N}} \mu_n(\mathbb{R}) = 1$. Combined with $\mathcal{P} \circ \mu(\mathbb{R}) \leq 1$, we conclude that $\mathcal{P} \circ \mu(\mathbb{R}) = 1$ and $\mu$ is a probability measure on $\mathcal{B}(\mathbb{R})$. Therefore, by Example 14.15, there exists a probability measure space $\Omega := (\Omega, \mathcal{F}, \mathcal{P})$ and a random variable $X : \Omega \to \mathbb{R}$ with law $\mu$. Let $F : \mathbb{R} \to [0, 1] \subset \mathbb{R}$ be the cumulative distribution function of $\mu$ as defined in Definition 14.14. Then, we have $g(\theta) = \lim_{n \to \mathbb{N}} \Psi_{X_n}(\theta) = \lim_{n \to \mathbb{N}} \int_{\mathbb{R}} \exp(i\theta x) \, d\mu_n(x) = \lim_{n \to \mathbb{N}} \langle \langle \mu_n, z_0(\theta) \rangle \rangle = \lim_{n \to \mathbb{N}} \langle \langle \mu_n, z_0(\theta) \rangle \rangle = \langle \langle \mu, z_0(\theta) \rangle \rangle = \int_{\mathbb{R}} \exp(i\theta x) \, d\mu(x) = \Psi_X(\theta), \forall \theta \in \mathbb{R}$, where the first equality follows from the assumption of the theorem; the second equality follows from Definition 14.45; the third equality follows from the definition $z_0(\theta) : \mathbb{R} \to \mathbb{C}$ is defined by $z_0(\theta)(x) = \exp(i\theta x), \forall x \in \mathbb{R}$; the fourth equality follows from Proposition 3.70; the fifth equality follows from the preceding discussion and $z_0(\theta) \in C_b(\mathbb{R}, \mathbb{R})$; the last equality follows from Definition 14.45. By Proposition 14.53, $\lim_{n \to \mathbb{N}} F_n(b) = F(b), \forall b \in \mathbb{R}$ with $F$ being continuous at $b$. 
By Lévy Inversion Formula, Theorem 14.48, and the discussion immediately after the theorem, \( g \) uniquely determines \( \mu \). Thus, we have shown that for the sequence \((\mu_n)_{n=1}^{\infty}\) there is a subsequence \((\mu_{n_i})_{i=1}^{\infty}\) such that \( \lim_{i \to \infty} \mu_{n_i} \Rightarrow \mu \), where \( \mu \) is the probability measure on \((\mathbb{R}, \mathcal{B}(\mathbb{R}))\) that is the law of a random variable \( X \) whose characteristic function is given by \( g \). Note that \( g \) is independent of the choice of the subsequence \((\mu_{n_i})_{i=1}^{\infty}\).

This clearly holds for any subsequence of \((\mu_n)_{n=1}^{\infty}\). Thus, \( \forall \) subsequence \((\mu_{n_i})_{i=1}^{\infty}\) of \((\mu_n)_{n=1}^{\infty}\), there exists a further subsequence \((\mu_{n_{i,j}})_{j=1}^{\infty}\) such that \( \lim_{n \to \infty} \mu_{n_{i,j}} \Rightarrow \mu \). By the uniqueness of the weak* limit (Proposition 7.120) and Helly’s Lemma 14.54, we have \( \lim_{n \to \infty} \mu_n \Rightarrow \mu \). Hence, \( \lim_{n \to \infty} X_n \Rightarrow X \).

This completes the proof of the theorem. \( \square \)

**Theorem 14.56 (Central Limit Theorem)** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure, \( b \in \mathbb{R}_+ \), \((X_n)_{n=1}^{\infty}\) be a sequence of independent \((\mathbb{R}\)-valued) random variables, \( X_n : \Omega \to \mathbb{R} \), with \( E(X_n) = 0 \), \( E(X_n^2) \leq b \), and \( E(|X_n|^3) \leq b \), \( \forall n \in \mathbb{N} \), and \( S_n := \sum_{i=1}^{n} X_i \). Assume that \( \lim_{n \to \infty} E(S_n^3)/n^3 = \infty \). Then, \( (S_n/\sqrt{E(S_n^2)})_{n=1}^{\infty} \) is uniformly growth bounded and \( \lim_{n \to \infty} S_n/\sqrt{E(S_n^2)} \Rightarrow X \), where \( X \) is a random variable with cumulative distribution function \( N(0,1) \).

**Proof** Clearly, \( E(S_n) = 0 \) and \( b_n := E(S_n^2) = \sum_{i=1}^{n} E(X_i^2) \leq nb \), \( \forall n \in \mathbb{N} \). Fix any \( n \in \mathbb{N} \). Let \( \Psi_{X_n} : \mathbb{R} \to \mathbb{C} \) be the characteristic function of \( X_n \) and \( \Psi_{V_n} : \mathbb{R} \to \mathbb{C} \) be the characteristic function of \( V_n := S_n/\sqrt{E(S_n^2)} \). Then, by Propositions 14.46 and 14.47, \( \Psi_{V_n}(\theta) = \prod_{i=1}^{n} \Psi_{X_i}(\theta/\sqrt{b_n}) \), \( \forall \theta \in \mathbb{R} \). By Proposition 14.46, we have \( \Psi_{X_n}(0) = 1, \Psi_{X_n}^{(1)}(0) = E(X_n) = 0, \Psi_{X_n}^{(2)}(0) = -E(X_n^2), \) and \( |\Psi_{X_n}^{(3)}(\theta)| = |E((iX_n)^3 \exp(i\theta X_n))| \leq E(|X_n|^3) \leq b \). By Taylor’s Theorem 9.48, we have \( R_n(\theta) := \Psi_{X_n}(\theta) - (1 - E(X_n^2)\theta^2) \) and

\[
|R_n(\theta)| \leq \frac{1}{6} E(|X_n|^3) |\theta|^3, \quad \forall \theta \in \mathbb{R}
\]

Let \( h_{n,i}(\theta) := -E(X_n^2)\theta^2/2 + R_i(\theta/\sqrt{b_n}), \forall n \in \mathbb{N}, i \in \mathbb{N} \). Fix any \( \theta \in \mathbb{R} \). Clearly, we have \( \lim_{n \to \infty} h_{n,i}(\theta) = 0 \) and the convergence is uniform in \( i \in \mathbb{N} \). For sufficiently large \( n \), we have \( |h_{n,i}(\theta)| < \frac{1}{2i}, \forall i \in \mathbb{N} \). By Taylor’s Theorem 9.48, we have \( |\ln(1 + x)| \leq 2x^2, \forall x \in [-\frac{1}{2}, \frac{1}{2}] \subseteq \mathbb{R} \). Then, we have

\[
\Psi_{V_n}(\theta) = \prod_{i=1}^{n} \Psi_{X_i}(\theta/\sqrt{b_n}) = \prod_{i=1}^{n} \left( 1 - \frac{E(X_n^2)}{b_n} \theta^2 + R_i(\theta/\sqrt{b_n}) \right)
\]

\[
= \prod_{i=1}^{n} \left( 1 + h_{n,i}(\theta) \right) = \exp \left( \sum_{i=1}^{n} \ln \left( 1 + h_{n,i}(\theta) \right) \right)
\]
\[ \exp \left( \sum_{i=1}^{n} \left( h_{n,i}(\theta) + 2(h_{n,i}(\theta))^2 \right) \right) \]
\[ \leq \exp \left( -\frac{\theta^2}{2} + \sum_{i=1}^{n} \left( \frac{b|\theta|^3}{6} b_{n}^{-3/2} + \frac{\theta^4 b^2}{b_n^2} + \frac{1}{9b_n^3} b^2 \theta^6 \right) \right) \]
\[ = \exp \left( -\frac{\theta^2}{2} + \sum_{i=1}^{n} \left( -\frac{b|\theta|^3}{6} b_{n}^{-3/2} - \frac{\theta^4 b^2}{b_n^2} + \frac{1}{9b_n^3} b^2 \theta^6 \right) \right) \]
and
\[ \Psi_{V_n}(\theta) = \exp \left( \sum_{i=1}^{n} \ln(1 + h_{n,i}(\theta)) \right) \geq \exp \left( \sum_{i=1}^{n} \left( h_{n,i}(\theta) - 2(h_{n,i}(\theta))^2 \right) \right) \]
\[ \geq \exp \left( -\frac{\theta^2}{2} + \sum_{i=1}^{n} \left( -\frac{b|\theta|^3}{6} b_{n}^{-3/2} - \frac{\theta^4 b^2}{b_n^2} + \frac{n\theta^4 b^2}{b_n^2} + \frac{n}{9b_n^3} b^2 \theta^6 \right) \right) \]

Hence, \( \lim_{n \to \infty} \Psi_{V_n}(\theta) = \exp(-\theta^2/2) =: g(\theta) \), if \( \lim_{n \to \infty} \frac{b_n^3}{n^2} = \infty \).

Note that \( g(\theta) \) is the characteristic function of a random variable \( X \sim N(0, 1) \). Then, the result follows from Lévy Convergence Theorem 14.55.

This completes the proof of the theorem. \( \square \)

### 14.9 Uniform Integrability and Martingales

**Definition 14.57** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( Y \) be a separable Banach space, \( \mathcal{F} \) be a collection of \( Y \)-valued random variables. \( \mathcal{F} \) is said to be uniformly integrable if \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists K \in [0, \infty) \subset \mathbb{R} \) such that, \( \forall X \in \mathcal{F}, \) we have \( \int_{\{\omega \in \Omega \mid X(\omega) > K\}} P \circ X(\omega) \, dP(\omega) < \epsilon \).

Clearly, for a uniformly integrable (UI) family \( \mathcal{F} \) of \( Y \)-valued random variables, let \( \epsilon = 1 \) and \( K_1 \in [0, \infty) \subset \mathbb{R} \) be as in Definition 14.57. Then, \( \forall X \in \mathcal{F}, \) we have

\[
E(P \circ X) = \int_{\Omega} P \circ X(\omega) \, dP(\omega) = \int_{\{\omega \in \Omega \mid P \circ X(\omega) > K_1\}} P \circ X(\omega) \, dP(\omega) + \int_{\{\omega \in \Omega \mid P \circ X(\omega) \leq K_1\}} P \circ X(\omega) \, dP(\omega) < 1 + \int_{\{\omega \in \Omega \mid P \circ X(\omega) \leq K_1\}} K_1 \, dP(\omega) \leq 1 + K_1
\]

where the first equality follows from Definition 14.1; the second equality follows from Proposition 11.89 and Bounded Convergence Theorem 11.77; and the first inequality follows from Proposition 11.89 and Definition 14.57.
Hence, \( E(P \circ X) < K_1 + 1 \), \( \forall X \in \mathcal{F} \), and \( \mathcal{F} \subseteq \mathcal{B}_{L_1(\Omega, Y)}(\vartheta_{L_1(\Omega, Y)}, 1 + K_1) \subseteq L_1(\Omega, Y) \).

**Proposition 14.58** Let \( \Omega := (\Omega, \mathcal{B}, \mathcal{P}) \) be a probability measure space, \( Y \) be a separable Banach space, \( \mathcal{F} \) be a collection of \( Y \)-valued random variables. Then, the following statements hold.

(i) If \( \mathcal{F} \) is bounded in \( L^p(\Omega, Y) \) with \( p \in (1, \infty) \subseteq \mathbb{R} \), that is \( \exists A \subseteq [0, \infty) \subseteq \mathbb{R} \) such that \( E(P \circ X) \leq A \), \( \forall X \in \mathcal{F} \). Then, \( \mathcal{F} \) is uniformly integrable.

(ii) If there exists \( g : \Omega \to [0, \infty) \subseteq \mathbb{R} \) such that \( g \) is \( \mathcal{B} \)-measurable, \( E(g) < \infty \), and \( P \circ X(\omega) \leq g(\omega) \) a.e. \( \omega \in \Omega \), \( \forall X \in \mathcal{F} \). Then, \( \mathcal{F} \) is uniformly integrable.

**Proof**

(i) Fix any \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \). Let \( K := (\frac{\epsilon}{A})^{\frac{1}{p-1}} \). \( \forall X \in \mathcal{F} \), by Propositions 11.38 and 7.21, \( P \circ X \) is \( \mathcal{B} \)-measurable, \( \forall q \in [1, \infty) \subseteq \mathbb{R} \). Then,

\[
\int_{\{ \omega \in \Omega \mid P \circ X(\omega) > K \}} P \circ X(\omega) \, dP(\omega) \\
\leq \int_{\{ \omega \in \Omega \mid P \circ X(\omega) > K \}} P_p \circ X(\omega) K^{1-p} \, dP(\omega) \\
\leq K^{1-p} \int_{\{ \omega \in \Omega \mid P \circ X(\omega) > K \}} P_p \circ X(\omega) \, dP(\omega) \\
\leq K^{1-p} \int_{\Omega} P_p \circ X(\omega) \, dP(\omega) \leq \frac{\epsilon}{A} A = \epsilon
\]

where the inequalities follow from Proposition 11.83. Hence, \( \mathcal{F} \) is uniformly integrable.

(ii) Fix any \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \). By Proposition 11.84, \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) such that \( \forall E \in \mathcal{B} \) with \( P(E) < \delta \), we have \( \int_E g \, dP < \epsilon \), since \( E(g) < \infty \). Note that \( P \left( \bigcap_{K \in \mathbb{N}} \{ \omega \in \Omega \mid g(\omega) > K \} \right) = P(\emptyset) = 0 \). By Proposition 11.5, we have \( 0 = \lim_{K \in \mathbb{N}} P \left( \{ \omega \in \Omega \mid g(\omega) > K \} \right) \). Then, \( \exists K \in \mathbb{N} \) such that \( E := \{ \omega \in \Omega \mid g(\omega) > K \} \in \mathcal{B} \) and \( P(E) < \delta \). \( \forall X \in \mathcal{F} \),

\[
\int_{\{ \omega \in \Omega \mid P \circ X(\omega) > K \}} P \circ X(\omega) \, dP(\omega) \\
\leq \int_{\mathcal{E}} P \circ X \, dP \leq \int_E g \, dP < \epsilon
\]

where the first inequality follows Proposition 11.83 and the fact that \( \exists E_0 \in \mathcal{B} \) with \( P(E_0) = 0 \) such that \( E_0 \cup E \geq \{ \omega \in \Omega \mid P \circ X(\omega) > K \} \); the second inequality follows from Proposition 11.83; and the last inequality follows from \( P(E) < \delta \) and the preceding discussion. Hence, \( \mathcal{F} \) is uniformly integrable.

This completes the proof of the proposition. \( \square \)
Theorem 14.59 Let $\Omega := (\Omega, B, P)$ be a probability measure space, $\gamma$ be a separable Banach space, $X \in L_1(\Omega, \gamma)$ be a $\gamma$-valued random variable. Then, the set $F := \{ f \in L_1(\Omega, \gamma) \mid \exists \mathcal{B} \subseteq B \text{ which is a sub-}\sigma\text{-algebra on } \Omega, f \in E(X|\mathcal{B}) \}$ is uniformly integrable.

Proof Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By Proposition 11.84, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall E \in B$ with $P(E) < \delta$, we have $\int_E P \circ X dP < \epsilon$. Let $K \in [0, \infty) \subset \mathbb{R}$ be such that $K^{-1} E(P \circ X) < \delta$. Fix any $\mathcal{B} \subseteq B$ with $\mathcal{B}$ being a $\sigma$-algebra on $\Omega$. Let $\Omega := (\Omega, \mathcal{B}, P := P|_\mathcal{B})$. Fix any $f \in E(X|\mathcal{B})$. By (i) of Proposition 14.11 and Proposition 8.25, we have $P \circ f(\omega) \leq f(\omega)$ a.e. $\omega \in \Omega$, where $f \in E(P \circ X|\mathcal{B})$. Let $E := \{ \omega \in \Omega \mid P \circ f(\omega) > K \} \in \mathcal{B}$, since $f$ is $\mathcal{B}$-measurable. Then, $K P(E) = K \int_E 1 dP \leq \int_E P \circ f dP \leq \int_\Omega P \circ f dP = \int_\Omega P \circ X dP = E(P \circ X)$, where the first equality follows from Proposition 11.75; the first, second, and third inequality follows from Proposition 11.83; and the second equality follows from Definition 14.7. This implies that $P(E) = P(E) \leq E(P \circ X) < \delta$. Therefore, $\int_E P \circ f dP = \int_E P \circ f dP \leq \int_\Omega P \circ X dP = E(P \circ X) < \epsilon$, where the first equality follows from the fact $f$ is $\mathcal{B}$-measurable and Proposition 11.72; the first inequality follows from Proposition 11.83; and the second equality follows from Definition 14.7. Hence, $F$ is uniformly integrable. This completes the proof of the theorem.

Proposition 14.60 (Chebyshev Inequality) Let $\Omega := (\Omega, B, P)$ be a probability measure space, $X \in L_1(\Omega, \mathbb{R})$ be a $\mathbb{R}$-valued random variable. Assume that $X(\omega) \geq 0$ a.e. in $\Omega$. Then, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, we have $P(\{ \omega \in \Omega \mid X(\omega) \geq \epsilon \}) \leq c^{-1} E(X)$.

Proof We have the following derivations.

\[
P(\{ \omega \in \Omega \mid X(\omega) \geq \epsilon \}) = \int_{\{ \omega \in \Omega \mid X(\omega) \geq \epsilon \}} 1 dP \\
\leq \int_{\{ \omega \in \Omega \mid X(\omega) \geq \epsilon \}} c^{-1} X(\omega) dP(\omega) = c^{-1} \int_{\{ \omega \in \Omega \mid X(\omega) \geq \epsilon \}} X(\omega) dP(\omega) \\
\leq c^{-1} \int_\Omega X(\omega) dP(\omega) = c^{-1} E(X)
\]

where the first equality follows from Proposition 11.75; the first inequality follows from Proposition 11.89; the second equality follows from Proposition 11.89; and the second inequality follows from the fact that $X(\omega) \geq 0$ a.e. in $\Omega$.

Theorem 14.61 Let $\Omega := (\Omega, B, P)$ be a probability measure space, $\gamma$ be a separable Hilbert space, $(X_n)_{n=1}^\infty \subseteq L_1(\Omega, \gamma) := Z$ be a sequence of $\gamma$-valued random variables and $X \in L_1(\Omega, \gamma)$ be a $\gamma$-valued random variable. Then, $\lim_{n \to \infty} X_n = X$ in $Z$, or equivalently $\lim_{n \to \infty} E(P \circ (X_n - X)) = 0$ implies that the following two conditions are satisfied.


14.9. DECOMPOSITION OF MEASURES

(i) \( \lim_{n \to \infty} X_n = X \) in measure in \( \Omega \).

(ii) The sequence \( (X_n)_{n=1}^{\infty} \) is uniformly integrable.

The converse of the above holds if \( Y \) is a separable real Hilbert space.

Proof

"Necessity" Let \( \lim_{n \to \infty} E(P \circ (X_n - X)) = 0 \). Fix any \( \epsilon \in (0, 1) \subset \mathbb{R} \). Then, \( \exists n_0 \in \mathbb{N} \), \( \forall n \in \mathbb{N} \) with \( n_0 \leq n \), we have \( E(P \circ (X_n - X)) < \epsilon^2 < \epsilon \). By Chebyshev Inequality Proposition 14.60, we have \( P(\{ \omega \in \Omega \mid P \circ (X_n - X)(\omega) > \epsilon \}) \leq \epsilon^{-1} E(P \circ (X_n - X)) < \epsilon \), \( \forall n \in \mathbb{N} \) with \( n_0 \leq n \). Hence, \( \lim_{n \to \infty} P(\{ \omega \in \Omega \mid P \circ (X_n - X)(\omega) > \epsilon \}) = 0 \). By the arbitrariness of \( \epsilon \), we have (i).

By Proposition 11.84, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \forall E \in \mathcal{B} \) with \( P(E) < \delta \), we have \( \int_E P \circ X_n dP < \epsilon \), \( n = 1, \ldots, n_0 \); and \( \int_E P \circ X dP < \epsilon \). Clearly, \( \sup_{n \in \mathbb{N}} E(P \circ X_n) =: M \in (0, \infty) \subset \mathbb{R} \). Choose \( K > 0 \) such that \( K^{-1}M < \delta \). Then, \( \forall n, K \in \mathbb{N} \), by Chebyshev Inequality Proposition 14.60, we have \( P(\{ \omega \in \Omega \mid P \circ X_n(\omega) > K \}) \leq K^{-1}E(P \circ X_n) \leq K^{-1}M < \delta \). If \( n \geq n_0 \), then \( \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} P \circ X dP \leq \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} P \circ X dP + \int_{\omega \in \Omega \mid P \circ X_n(\omega) \leq K} P \circ X dP < \epsilon + \int_{\omega \in \Omega} P \circ (X_n - X) dP < 2\epsilon, \)
where the first inequality follows from Definition 7.1 and Proposition 11.83; the second inequality follows from the preceding discussion and Proposition 11.83; and the last inequality follows from the preceding discussion. On the other hand, if \( n < n_0 \), we have \( \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} P \circ X dP < \epsilon \) by the preceding discussion. Thus, \( (X_n)_{n=1}^{\infty} \) is uniformly integrable.

"Sufficiency" Let (i) and (ii) hold and \( Y \) be a separable real Hilbert space. \( \forall K \in (0, \infty) \subset \mathbb{R} \), define \( \varphi_K : y \to \mathcal{B}_y(\vartheta_y, K) \) by \( \varphi_K(y) = \begin{cases} K \varepsilon(\vartheta_y) y & \text{if } \|y\| > K \\ y & \text{if } \|y\| \leq K \end{cases} \), \( \forall y \in Y \). Clearly, \( \varphi_K \) is continuous. Fix any \( \epsilon \in (0, \infty) \subset \mathbb{R} \). By (ii) and the fact that \( X \in L_1(\Omega, Y) \), we choose \( K > 0 \) such that \( \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} P \circ X_n dP < \frac{\epsilon}{2\delta} \), \( \forall n \in \mathbb{N} \); and \( \int_{\omega \in \Omega \mid P \circ X(\omega) > K} P \circ X dP < \frac{\epsilon}{4\delta} \), \( \forall n \in \mathbb{N} \), we have \( E(P \circ (\varphi_K \circ X_n - X_n)) = \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} \left\| X_n(\omega) - \frac{K}{\|X_n(\omega)\|} Y \right\| dP(\omega) = \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} \left(1 - \frac{K}{\|X_n(\omega)\|}\right) \|X_n(\omega)\| dP(\omega) \leq \int_{\omega \in \Omega \mid P \circ X_n(\omega) > K} P \circ X_n dP < \frac{\epsilon}{4\delta} \), where the first inequality follows from Proposition 11.83. Similarly, we have \( E(P \circ (\varphi_K \circ X - X)) < \frac{\epsilon}{4\delta} \). We need the following intermediate result.

Claim 14.61.1 \( \|\varphi_K(x) - \varphi(y)\|_Y \leq \|x - y\|_Y \), \( \forall x, y \in Y \).

Proof of claim: We will distinguish four exhaustive and mutually exclusive cases: Case 1: \( x, y \in \mathcal{B}_y(\vartheta_y, K) \); Case 2: \( \|x\|_Y > K \) and \( \|y\|_Y > K \); Case 3: \( \|x\|_Y > K \) and \( \|y\|_Y \leq K \); Case 4: \( \|x\|_Y \leq K \) and \( \|y\|_Y > K \).

Case 1: \( x, y \in \mathcal{B}_y(\vartheta_y, K) \). Then, we have \( \|\varphi_K(x) - \varphi_K(y)\|_Y = \|x - y\|_Y \). This case is proved.
14.10 Existence of the Wiener Process

Proposition 14.62 Let \( \Omega := (\Omega, \mathcal{B}, \mathbb{P}) \) be a probability measure space, \( \mathbb{Y} \) be a separable reflexive Banach space with \( \mathbb{Y}^* \) being separable, \( \mathcal{B} \subseteq \mathcal{B} \) be a sub-\( \sigma \)-algebra, \( \hat{\Omega} := (\hat{\Omega}, \hat{\mathcal{B}}, \hat{\mathbb{P}} := \mathbb{P}|_{\hat{\mathcal{B}}} ) \) be probability measure space, \( p \in [1, \infty) \subset \mathbb{R}, f \in L_p(\hat{\Omega}, \hat{\mathbb{Y}}) \) be a \( \hat{\mathbb{Y}} \)-valued random variable, \( g \in \mathbb{E}(\mathbb{P}_p \circ \hat{\mathbb{P}}|_{\hat{\mathcal{B}}}|_{\hat{\mathcal{B}}} \circ f) \)}
14.10. EXISTENCE OF THE WIENER PROCESS

777

Proof. Then, by (i) of Proposition 14.11 and Proposition 8.25, we have

$$f|\mathcal{B}) \in L_1(\Omega, \mathbb{R}), \text{ and } h \in E(f|\mathcal{B}) \in L_1(\Omega, y).$$

Then, $$h \in L_p(\Omega, y)$$ and $$P_p \circ h \leq g \text{ a.e. in } \Omega.$$ 

Proof. By Definition 14.7, $$g \in L_1(\Omega, [0, \infty))$$ and $$h \in L_1(\Omega, y)$$ exists. Then, by (i) of Proposition 14.11 and Proposition 8.25, we have $$P_p \circ h \leq l \text{ a.e. in } \Omega,$$ where $$l \in E(P \circ f|\mathcal{B}) \in L_1(\Omega, [0, \infty)),$$ and $$P_p \circ h = (P \circ h)^p \leq l^p \leq g \text{ a.e. in } \Omega.$$ Then, $$\|h\|_{L_p(\Omega, y)}^p = E(P_p \circ h) = E(P_p \circ f) \leq E(g) = E(P_p \circ f) < \infty,$$ where the first equality follows from Definition 14.1; the second equality follows from Proposition 11.72; the first inequality follows from Proposition 11.83; and the third equality follows from (n) of Proposition 14.11; and the last inequality follows from $$f \in L_p(\Omega, y).$$ Hence, $$h \in L_p(\Omega, y).$$ This completes the proof of the proposition.

Theorem 14.63 Let $$\Omega := (\Omega, \mathcal{B}, \mathbb{P})$$ be a probability measure space, $$\mathcal{Y}$$ be a separable real Hilbert space, $$\mathcal{B} \subseteq \mathbb{B}$$ be a sub-$$\sigma$$-algebra, $$\hat{\Omega} := (\hat{\Omega}, \hat{\mathcal{B}}, \hat{\mathbb{P}} := \mathbb{P}|_{\hat{\mathcal{B}}})$$ be probability measure space, $$f \in L_2(\hat{\Omega}, \mathcal{Y})$$ be a $$\mathcal{Y}$$-valued random variable, and $$h \in E(f|\mathcal{B}) \in L_1(\hat{\Omega}, \mathcal{Y}).$$ Then, $$h \in L_2(\hat{\Omega}, \mathcal{Y})$$ and $$\|f - h\|_{L_2(\hat{\Omega}, \mathcal{Y})}^2 = E(P_2 \circ (f - h)) = E(P_2 \circ f) - E(P_2 \circ h) \leq \|f - g\|_{L_2(\Omega, \mathcal{Y})}^2, \forall g : \Omega \to \mathcal{Y}$$ that is $$\mathcal{B}$$-measurable, with equality holding if, and only if, $$g = h \text{ a.e. in } \hat{\Omega}.$$ 

Proof. By Definition 14.7, $$E(f|\mathcal{B}) \in L_1(\Omega, \mathcal{Y})$$ exists. Hence, $$h \in L_1(\Omega, \mathcal{Y})$$ exists. By Proposition 14.62, $$h \in L_2(\hat{\Omega}, \mathcal{Y}).$$ Then, $$\lambda := \inf_{g : \Omega \to \mathcal{Y} \text{being } \mathcal{B} \text{-measurable}} \|f - g\|_{L_2(\hat{\Omega}, \mathcal{Y})}^2 \in [0, \infty) \subset \mathbb{R}.$$ Thus, the infimization problem can only consider those $$g \in L_2(\hat{\Omega}, \mathcal{Y}).$$ Then, we have

$$\lambda = \inf_{g \in L_2(\mathcal{Y})} \|f - g\|_{L_2(\hat{\Omega}, \mathcal{Y})}^2 = \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E(P_2 \circ (f - g))$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E(E((f, f)_y - 2 \langle f, g \rangle_y + \langle g, g \rangle_y |\mathcal{B}))$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E(E((f, f)_y |\mathcal{B}) - 2E((f, g)_y |\mathcal{B}) + E((g, g)_y |\mathcal{B}))$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E(E((f, f)_y |\mathcal{B}) - 2E((f, f)_y |\mathcal{B})) + E((g, g)_y |\mathcal{B}))$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E((f, f)_y |\mathcal{B}) - 2\langle g, E(f|\mathcal{B}) \rangle + (g, g)_y$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E((f, f)_y |\mathcal{B}) - 2\langle g, E(f|\mathcal{B}) \rangle + (g, g)_y$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E((f, f)_y |\mathcal{B}) - 2\langle g, E(f|\mathcal{B}) \rangle + (g, g)_y$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} (E((f, f)_y |\mathcal{B}) - 2\langle (g, h) \rangle + (g, g)_y)$$

$$= \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E((f, f)_y) - 2\langle (g, h) \rangle + (g, g)_y$$

$$= E((f, f)_y) + \inf_{g \in L_2(\hat{\Omega}, \mathcal{Y})} E(-2\langle h, g \rangle_y + (g, g)_y)$$
\begin{align*}
= & \mathbb{E}(\langle f, f \rangle_Y) + \inf_{g \in L^2(\Omega, Y)} \mathbb{E}(\langle (h, h) \rangle_Y + \langle g - h, g - h \rangle_Y) \\
= & \mathbb{E}(\langle f, f \rangle_Y) - \mathbb{E}(\langle (h, h) \rangle_Y) + \inf_{g \in L^2(\Omega, Y)} \mathbb{E}(\langle g - h, g - h \rangle_Y) \\
= & \|f\|_{L^2(\Omega, Y)}^2 - \|h\|_{L^2(\Omega, Y)}^2 + \inf_{g \in L^2(\Omega, Y)} \|g - h\|_{L^2(\Omega, Y)}^2 \\
= & \|f\|_{L^2(\Omega, Y)}^2 - \|h\|_{L^2(\Omega, Y)}^2 = \|f - h\|_{L^2(\Omega, Y)}^2
\end{align*}

where the first equality follows from the preceding discussion; the second equality follows from Definition 14.1; the third equality follows from (a) of Proposition 14.11; the fourth equality follows from Proposition 13.2 and Definition 13.1; the fifth equality follows from (c) of Proposition 14.11; the sixth equality follows from (iv) of Theorem 13.15 and (b) of Proposition 14.11; the seventh equality follows from (h) of Proposition 14.11; the eighth and ninth equality follows from Proposition 11.9; the eleventh equality follows from Definition 13.1; the twelfth equality follows from Definition 14.1 and Proposition 13.2; the thirteenth equality follows whenever \( g = h \) a.e. in \( \Omega \); and the last equality follows from the fact we substitute \( g = h \) a.e. in \( \Omega \) into the original infimization problem.

This completes the proof of the theorem. \( \square \)

The above theorem says that minimization problem

\[ \min_{g: \Omega \to Y \text{ being } \mathcal{B}-\text{measurable}} \|f - g\|_{L^2(\Omega, Y)}^2 \]

admits a unique minimizer (modulo a set of measure 0) given by a version of \( \mathbb{E}(f|\mathcal{B}) \).

**Proposition 14.64** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( m \in \mathbb{N} \), and \( (x_n)_{n=0}^{\infty} \) be a sequence of \( \mathbb{R}^m \)-valued random variables with \( x_n: \Omega \to \mathbb{R}^m \) with law \( \mu_n, \forall n \in \mathbb{Z}_+ \), and \( F_n: \mathbb{R}^m \to [0, 1] \subset \mathbb{R} \) be the cumulative distribution function for \( \mu_n \) as defined in Definition 14.14, \( \forall n \in \mathbb{Z}_+ \). Then, the following statements hold.

(i) If \( \lim_{n \in \mathbb{N}} \int_{\mathbb{R}^m} \langle (d\mu_n(x), z(x)) \rangle = \int_{\mathbb{R}^m} \langle (d\mu_0(x), z(x)) \rangle \in \mathbb{R}, \forall z \in \mathfrak{C}_b(\mathbb{R}^m, \mathbb{R}) \), then \( \lim_{n \in \mathbb{N}} \mu_n \Rightarrow \mu_0 \).

(ii) \( \lim_{n \in \mathbb{N}} \int_{\mathbb{R}^m} \langle (d\mu_n(x), z(x)) \rangle = \int_{\mathbb{R}^m} \langle (d\mu_0(x), z(x)) \rangle \in \mathbb{R}, \forall z \in \mathfrak{C}_b(\mathbb{R}^m, \mathbb{R}) \) implies that \( \lim_{n \in \mathbb{N}} F_n(b) = F_0(b), \forall b \in \mathbb{R}^m \) with \( F_0 \) being continuous at \( b \).

(iii) If \( \lim_{n \in \mathbb{N}} \mu_n \Rightarrow \mu_0 \) and \( (x_n)_{n=1}^{\infty} \) is uniformly growth bounded, then \( \lim_{n \in \mathbb{N}} F_n(b) = F_0(b), \forall b \in \mathbb{R}^m \) with \( F_0 \) being continuous at \( b \).

(iv) If \( \lim_{n \in \mathbb{N}} \mu_n \Rightarrow \mu_0 \), \( (x_n)_{n=1}^{\infty} \) is uniformly growth bounded, and \( \lim_{n \in \mathbb{N}} F_n(b) =: F(b) \in [0, 1] \subset \mathbb{R}, \forall b \in \mathbb{R}^m \), and \( F: \mathbb{R}^m \to [0, 1] \subset \mathbb{R} \) is a valid cumulative distribution function for some \( \mathbb{R}^m \)-valued random variable, then \( F = F_0 \).
14.10. EXISTENCE OF THE WIENER PROCESS

\textbf{Proof}  
(i) Note that \( \lim_{n \to N} \mu_n \to \mu_0 \) means that \( \lim_{n \to N} \mu_n = \mu_0 \) weak*, that is \( \lim_{n \to N} \langle \langle \mu_n, z \rangle \rangle = \langle \langle \mu_0, z \rangle \rangle, \forall z \in \mathcal{C}_c(\mathbb{R}^m, \mathbb{R}) \). By Riesz Representation Theorem 11.209, this is equivalent to \( \lim_{n \to N} \int_{\mathbb{R}^m} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}^m} \langle \langle d\mu_0(x), z(x) \rangle \rangle, \forall z \in \mathcal{C}_c(\mathbb{R}^m, \mathbb{R}) \subseteq \mathcal{C}_b(\mathbb{R}^m, \mathbb{R}) \). Thus, (i) follows readily.

(ii) Let \( \lim_{n \to N} \int_{\mathbb{R}^m} \langle \langle d\mu_n(x), z(x) \rangle \rangle = \int_{\mathbb{R}^m} \langle \langle d\mu_0(x), z(x) \rangle \rangle, \forall z \in \mathcal{C}_b(\mathbb{R}^m, \mathbb{R}) \). Fix any \( b \in \mathbb{R}^m \) and \( \delta \in \mathbb{R}_+ \). Define \( h_{b,\delta} : \mathbb{R}^m \to [0, 1] \) by \( h_{b,\delta}(x) = \begin{cases} 
1 & x \leq b \\
1 - \frac{\max(x-b)}{\delta} & x \leq b + \delta 1_m \text{ and } x \neq b, \forall x \in \mathbb{R}^m. \\
0 & x \neq b + \delta 1_m 
\end{cases} \)

Clearly, \( h_{b,\delta} \in \mathcal{C}_b(\mathbb{R}^m, \mathbb{R}) \). Then, \( \lim_{n \to N} \langle \langle \mu_n, h_{b,\delta} \rangle \rangle = \langle \langle \mu_0, h_{b,\delta} \rangle \rangle \), \( F_n(b) \leq \langle \langle \mu_0, h_{b,\delta} \rangle \rangle \), and \( \langle \langle \mu_0, h_{b,\delta} \rangle \rangle \leq F_0(b + \delta) \). Then, taking limit as \( n \to \infty \) in the above three relations, we have \( \limsup_{n \to N} F_n(b) \leq F_0(b + \delta) \).

Since \( F_0 \) is a cumulative distribution function and therefore right continuous. Taking limit as \( \delta \to 0^+ \), we have \( \limsup_{n \to N} F_n(b) \leq F_0(b) \). Similarly, we must have \( \lim_{n \to N} \langle \langle \mu_n, h_{b,\delta} \rangle \rangle = \langle \langle \mu_0, h_{b,\delta} \rangle \rangle \), \( F_n(b + \delta) \geq \langle \langle \mu_n, h_{b,\delta} \rangle \rangle \), and \( \langle \langle \mu_0, h_{b,\delta} \rangle \rangle = F_0(b) \). Taking limit as \( n \to \infty \) in the above three relations, we have \( \liminf_{n \to N} F_n(b + \delta) \geq F_0(b) \). Let \( b : = b + \delta 1_m \), we have \( \liminf_{n \to N} F_n(b) \geq F_0(b - \delta 1_m), \forall b \in \mathbb{R}^m, \forall \delta \in \mathbb{R}_+ \). Now, fix any \( b \in \mathbb{R}^m \) with \( F_0 \) being continuous at \( b \). We have \( F_0(b - \delta 1_m) \leq \liminf_{n \to N} F_n(b) \leq \limsup_{n \to N} F_n(b) \leq F_0(b) \). Now, take \( \delta \to 0^+ \), we have \( F_0(b) \leq \liminf_{n \to N} F_n(b) \leq \limsup_{n \to N} F_n(b) \leq F_0(b) \) and hence \( \lim_{n \to N} F_n(b) = F_0(b) \). Hence, (ii) is proved.

(iii) Let \( \lim_{n \to N} \mu_n \to \mu_0 \). Fix any \( z \in \mathcal{C}_b(\mathbb{R}^m, \mathbb{R}) \). By Bounded Convergence Theorem 11.77, \( \int_{\mathbb{R}^m} \langle \langle d\mu_n(x), z(x) \rangle \rangle \in \mathbb{R}, \forall n \in \mathbb{Z}_+ \), \( \forall a, b \in \mathbb{R}^m \) with \( a < b, \forall k \in \mathbb{N} \) with \( \min(b-a) > \frac{1}{k} \), by Proposition 5.62, there exists a continuous function \( h_{k,a,b} : \mathbb{R}^m \to [0, 1] \subseteq \mathbb{R} \) such that \( h_{k,a,b}(x) = 1 \), \( \forall x \in E_{k,1} := \{ \bar{x} \in \mathbb{R}^m | a + \frac{1}{k} 1_m \leq x \leq \bar{x} \} \subseteq \mathbb{R}^m \) and \( h_{k,a,b}(x) = 0, \forall x \in E_{k,2} := \mathbb{R}^m \setminus \{ \bar{x} \in \mathbb{R}^m | a < x \leq b + \frac{1}{k} 1_m \} \). Clearly, \( h_{k,a,b} \in \mathcal{C}_c(\mathbb{R}^m, \mathbb{R}) \).

Then,

\[
\Delta F_n(r_{a+b}1_m, b) = \int_{\mathbb{R}^m} \chi_{r_{a+b}1_m} \mathbb{R}^m \, d\mu_n \leq \int_{\mathbb{R}^m} h_{k,a,b} \, d\mu_n = \langle \langle \mu_n, h_{k,a,b} \rangle \rangle
\]

\[
\leq \int_{\mathbb{R}^m} \chi_{r_{a+b}1_m} \mathbb{R}^m \, d\mu_n = \Delta F_n(r_{a+b}1_m, b); \quad \forall n \in \mathbb{Z}_+, \forall a, b \in \mathbb{R}^m, \forall k \in \mathbb{N} \text{ with } a + \frac{1}{k} \leq b \quad (14.12)
\]

where the first and last equalities follow from Definition 14.14; and the first and last inequalities follow from Proposition 11.92. Taking limit as \( n \to \infty \),

\[
\limsup_{n \to N} \Delta F_n(r_{a+b}1_m, b) \leq \lim_{n \to N} \langle \langle \mu_n, h_{k,a,b} \rangle \rangle = \langle \langle \mu_0, h_{k,a,b} \rangle \rangle \quad (14.13)
\]

\[
\leq \liminf_{n \to N} \Delta F_n(r_{a+b}1_m, b); \quad \forall a, b \in \mathbb{R}^m, \forall k \in \mathbb{N} \text{ with } a + \frac{1}{k} \leq b
\]
where the equality follows from \( \lim_{n \in \mathbb{N}} \mu_n \xrightarrow{} \mu_0 \) and \( h_{k,a,b} \in C_c(\mathbb{R}^m, \mathbb{R}) \); and the inequalities follow from (14.12). By the assumption that \( (x_n)_{n=1}^\infty \) is uniformly growth bounded, we have \( \lim_{\min d \to -\infty} \max_{n \in \mathbb{N}} F_n(d) = 0 \), and the convergence is uniform with respect to \( n \in \mathbb{N} \). It is easy to prove, using \( \epsilon - \delta \) arguments, that

\[
0 \leq \lim_{\min d \to -\infty} \lim_{n \in \mathbb{N}} \inf \sum F_n(d) \leq \lim_{\min d \to -\infty} \sum \sup F_n(d) = \lim_{n \in \mathbb{N}} \lim_{\min d \to -\infty} F_n(d) = 0 \quad (14.14)
\]

This implies that, \( \forall b \in \mathbb{R}^m \) and \( \forall k \in \mathbb{N}, \)

\[
\lim_{\sup_{n \in \mathbb{N}} F_n(b)} = \lim_{\sup_{n \in \mathbb{N}}} \lim_{\max a \to -\infty} \lim_{\min n \in \mathbb{N}} F_n(r_{a,b} + \frac{1}{k} 1_m) = \lim_{\max a \to -\infty} \lim_{\min n \in \mathbb{N}} F_n(r_{a,b} + \frac{1}{k} 1_m) = \lim_{\max a \to -\infty} F_0(r_{a,b} + \frac{1}{k} 1_m) = 0 \quad (14.15)
\]

where the first equality follows from (14.14); the second equality follows from (14.14); the first inequality follows from (14.13) and the fact that all limits are real numbers in the closed interval \([0, 1] \subset \mathbb{R} \); the second inequality follows from (14.12) with \( n = 0 \); and the last equality follows from (14.14) and Definition 12.41. We also have, \( \forall b \in \mathbb{R}^m \) and \( \forall k \in \mathbb{N}, \)

\[
F_0(b) = \lim_{\max a \to -\infty} \Delta F_0(r_{a,b} + \frac{1}{k} 1_m) \leq \lim_{\max a \to -\infty} \lim_{\min n \in \mathbb{N}} F_n(r_{a,b} + \frac{1}{k} 1_m) = \lim_{\max a \to -\infty} \inf_{n \in \mathbb{N}} F_n(r_{a,b} + \frac{1}{k} 1_m) = \inf_{n \in \mathbb{N}} \sup_{n \in \mathbb{N}} F_n(b) = \inf_{n \in \mathbb{N}} \sup_{n \in \mathbb{N}} F_n(b) \leq \lim_{\max a \to -\infty} \inf_{n \in \mathbb{N}} F_n(b) \quad (14.16)
\]

where the first equality follows from (14.14) and Definition 12.41; the first inequality follows from (14.12) with \( n = 0 \); the second inequality follows from (14.13); the second equality follows from (14.14); and the last equality follows from (14.14). Then, let \( b \in \mathbb{R}^m \) be any point such that \( F_0 \) is continuous at \( b \), we have \( \limsup_{n \in \mathbb{N}} F_n(b) \leq \lim_{k \in \mathbb{N}} F_0(b + \frac{1}{k} 1_m) \leq \lim_{k \in \mathbb{N}} \inf_{n \in \mathbb{N}} F_n(b) \leq \lim_{k \in \mathbb{N}} \sup_{n \in \mathbb{N}} F_n(b) \), where the first inequality follows from (14.15); the first and second equalities follow from the fact that \( F_0 \) is continuous at \( b \); the second inequality follows from (14.16) with \( b \) substituted by \( b - \frac{1}{k} \); and the last equality follows from Proposition 3.83. Thus, we have \( \lim_{n \in \mathbb{N}} F_n(b) = F_0(b) \). This establishes (iii).

(iv) \( \forall a, b \in \mathbb{R}^m \) with \( a < b \), \( \forall k \in \mathbb{N} \) with \( \min(b - a) > \frac{1}{k} \), let \( h_{k,a,b} : \mathbb{R}^m \to [0, 1] \subset \mathbb{R} \) be as defined in (iii). Clearly, \( h_{k,a,b} \in C_c(\mathbb{R}^m, \mathbb{R}) \). Then, (14.12)–(14.16) holds. Taking limit as \( n \to \infty \) in (14.12),

\[
\Delta F(r_{a,b} + \frac{1}{k} 1_m) = \lim_{n \in \mathbb{N}} \Delta F_n(r_{a,b} + \frac{1}{k} 1_m) \leq \lim_{n \in \mathbb{N}} \langle \mu_n, h_{k,a,b} \rangle = \langle \mu_0, h_{k,a,b} \rangle
\]
where the first equality follows from the assumption; the second equality follows from (14.14); the third equality follows from (14.14); the first inequality follows from (14.12); and the last equality follows from the assumption. This further implies that, \( \forall k \in \mathbb{N} \),

\[
F(b) = \lim_{n \to \infty} F_n(b) = \lim_{n \to \infty} \sum_{j=1}^{m} \Delta F_n r_{a,b}^{k,1_m} \leq \lim_{n \to \infty} \langle \mu_0, h_{k,a,b} \rangle
\]

where the first equality follows from (14.14) and Definition 12.41; the first inequality follows from (14.17); the second inequality follows from (14.12) with \( n = 0 \); and the last equality follows from (14.14) and Definition 12.41. Taking limit \( k \to \infty \), we have \( \tilde{F}(b) \leq F_0(b) \), \( \forall b \in \mathbb{R}^m \), since \( F_0 \) is continuous on the right. We also have, \( \forall b \in \mathbb{R}^m \) and \( \forall k \in \mathbb{N} \),

\[
F_0(b) = \lim_{n \to \infty} \Delta F_n r_{a,b}^{k,1_m} \leq \lim_{n \to \infty} \langle \mu_0, h_{k,a,b} \rangle
\]

where the first equality follows from (14.14) and Definition 12.41; the first inequality follows from (14.12) with \( n = 0 \); the second inequality follows from (14.13); the second and third equalities follow from (14.14); and the last equality follows from the assumption. Then, again taking limit \( k \to \infty \), we have \( \tilde{F}(b) \leq F_0(b), \forall b \in \mathbb{R}^m \), since \( \tilde{F} \) is continuous on the right. Thus, we have \( \tilde{F}(b) = F_0(b) \), \( \forall b \in \mathbb{R}^m \). This establishes (iv).

This completes the proof of the proposition. \( \square \)

**Proposition 14.65** Let \( \Omega := (\Omega, \mathcal{B}, \mathbf{P}) \) be a probability measure space, \( m \in \mathbb{N} \), \( (X_i)_{i=1}^{\infty} \) be a sequence of independent identically distributed random variable with \( X_i \sim N(0,1) \), \( \forall i \in \mathbb{N} \), \( \alpha_1, \ldots, \alpha_m \in L^1(\mathbb{R}) \) be linearly independent, and \( (Y_j)_{j=1}^{m} \) be random variables defined by

\[
Y_j(\omega) := \begin{cases} \sum_{i=1}^{\infty} \alpha_{j,i} X_i(\omega) & \forall \omega \in F_j \\ 0 & \forall \omega \in \Omega \setminus F_j \end{cases}, \forall \omega \in \Omega
\]

where \( F_j := \{ \omega \in \Omega | \sum_{i=1}^{\infty} \alpha_{j,i} X_i(\omega) \in \mathbb{R} \}, j = 1, \ldots, m \); \( \alpha_j = (\alpha_{j,1}, \alpha_{j,2}, \ldots) \), \( j = 1, \ldots, m \). Then, \( Y := (Y_1, \ldots, Y_m) \) is an \( \mathbb{R}^m \)-valued Gaussian random variable with distribution \( N(0_m, K) \), where \( K = (\sum_{i=1}^{\infty} \alpha_{i,j} \alpha_{j,l})_{m \times m} \), and \( \mathbf{P}(F_j) = 0, j = 1, \ldots, m \), if \( K \in S_{+} \mathbb{R}^m \).
CHAPTER 14. PROBABILITY THEORY

Proof  Let \( K \in S_+ \mathbb{R}^m \). Then, \( \alpha_j \in l_2(\mathbb{R}) =: \mathcal{X}, j = 1, \ldots, m \). By Proposition 13.42, \( S_+ \mathbb{R}^m \) is an open subset of \( S_{\mathbb{R}^m} \), then, \( \exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n \geq n_1 \), we have \( K_n := (\sum_{i=1}^n \alpha_{i,j})_{m \times m} \in S_+ \mathbb{R}^m \). Then, the matrix \( V_n := (\alpha_{i,j})_{m \times n} \) is of full row rank, \( \forall n \in \mathbb{N} \) with \( n \geq n_1 \), since \( K_n = V_n V_n' \). By Proposition 14.22, we have \( Y^{(n)} := V_n X^{(n)} \) is a \( \mathbb{R}^m \)-valued Gaussian random variable with distribution \( N(0_n, K_n) \), \( \forall n \in \mathbb{N} \) with \( n \geq n_1 \), where \( X^{(n)} := (X_1, \ldots, X_n) \) is an \( \mathbb{R}^n \)-valued Gaussian random variable with distribution \( N(0_n, I_{\mathbb{R}^n}) \). It is easy to see that \[ \| Y^{(n)} - Y^{(v)} \|_{L^2(\Omega, \mathbb{R}^m)}^2 = \int_{\Omega} P_2 \circ (Y^{(n)} - Y^{(v)}) \, dP = E(\mathcal{P}_2 \circ (Y^{(n)} - Y^{(v)})) = E(\sum_{i=1}^m \sum_{v=1}^m \alpha_{i,v} X_i (\sum_{k=1}^m \alpha_{v,k} X_k)) = \sum_{i=1}^m \sum_{v=1}^m E(\alpha_{i,v}^2 X_i^2) = \sum_{i=1}^m \sum_{v=1}^m \alpha_{i,v}^2 \rightarrow 0 \] as \( n_2 \rightarrow \infty \), \( \forall u, v \in \mathbb{N} \) with \( n_2 \leq u \leq v \). Thus, we have \( Y^{(n)} \) is a Cauchy sequence in \( L_2(\Omega, \mathbb{R}^m) \), and converges to \( Y \) almost everywhere in \( \Omega \). Hence, \( P(F_j) = 0, j = 1, \ldots, m \). By Mode of Convergence Theorem 14.50, we have \( Y^{(n)} \rightarrow Y \). By Chebyshev Inequality Proposition 14.60, we have that \( (Y^{(n)})_{n=1}^\infty \) is uniform growth bounded. By Lebesgue Dominated Convergence Theorem 11.91 and Definition 14.19, we have \( \lim_{n \in \mathbb{N}} F_{Y^{(n)}}(z) = \tilde{F}(z), \forall z \in \mathbb{R}^m \), where \( F_{Y^{(n)}} : \mathbb{R}^m \rightarrow [0, 1] \subset \mathbb{R} \) is the cumulative distribution function of \( Y^{(n)} \), \( \forall n \in \mathbb{N} \), and \( \tilde{F} : \mathbb{R}^m \rightarrow [0, 1] \subset \mathbb{R} \) is the cumulative distribution function of some \( \mathbb{R}^m \)-valued Gaussian random variable with distribution \( N(0_n, K) \). By Proposition 14.64, \( Y \sim N(0_n, K) \).

This completes the proof of the proposition. \( \square \)

Example 14.66  Let \( I := [0, 1] \subset \mathbb{R} \), and \( I := ((P(I) = (0, 1], [\cdot, \cdot]), B, P) \) be the finite metric measure subspace of \( \mathbb{R} \), which is a probability measure space. Define a sequence of simple functions \((h_n)_{n=1}^\infty, h_n : I \rightarrow [-1, 1] \subset \mathbb{R} \)

by \( h_n(x) = \begin{cases} 1 & 2^{-m}2i < x \leq 2^{-m}(2i + 1) \\ -1 & 2^{-m}(2i + 1) < x \leq 2^{-m}(2i + 2) \\ 0 & \text{otherwise} \end{cases} \) \( \forall x \in I, \forall n \in \mathbb{N} \), \( m := \lceil \log_2 n \rceil \in \mathbb{Z}_+, i := \lceil n - 1 - 2^{m-1} \rceil \) \( \forall 0 \in \mathbb{Z}_+ \). This sequence of functions is said to be the Haar functions. See Figure 14.1 for the plot of the first eight of Haar functions. Clearly, \((h_n)_{n=1}^\infty\) is an orthogonal sequence in \( Z \subset L_2(I, \mathbb{R}) \). We will further define a sequence of simple functions \((H_n)_{n=1}^\infty\) by \( H_n := \frac{1}{\|h_n\|_2} h_n, \forall n \in \mathbb{N} \), which is then an orthonormal sequence in \( Z \). We will show that \((H_n)_{n=1}^\infty\) is a complete orthonormal sequence in \( Z \subset L_2(I, \mathbb{R}) \).

View each \( H_n \) as a random variable, \( \forall n \in \mathbb{N} \). Let \( B_m := \sigma((H_n)_{n=1}^{2^m}), \forall m \in \mathbb{Z}_+ \). We observe that \( B_m = \sigma(\mathcal{H}_m) \), where \( \mathcal{H}_0 := \{ \emptyset, (0, 1] \} \), and \( \mathcal{H}_{m+1} := \mathcal{H}_m \cup \{ (0, 2^{-m-1}], (2^{-m-1}, 2^{-m}], \ldots, ((2^{m+1}-1)2^{-m-1}, 1] \} \), \( \forall m \in \mathbb{Z}_+ \). It should be clear that \( \mathcal{H}_m \) is a \( \pi \)-system, \( \forall m \in \mathbb{Z}_+ \). By Definition 12.17. We need the following intermediate result.

Claim 14.66.1 \( \forall E \in B_m, \exists a_1, \ldots, a_{2^m} \in \mathbb{R} \) such that \( \chi_{E,i} = \sum_{i=1}^{2^m} a_i h_i \).
Figure 14.1: The first eight Haar functions.
Definition 14.67 Let $T \in \mathbb{R}_+$, $[0, T] \subset \mathbb{R}$, and $\mathcal{I} := \{([0, T], | \cdot |), \mathbb{B}, \mu\}$ be the finite metric measure subspace of $\mathbb{R}$ with partial ordering $\leq$, $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, and $\mathcal{I} \times \Omega := ([0, T] \times \Omega, \mathcal{B} \times \mathcal{P}, \nu)$ be the finite product measure space. A Wiener process on $\mathcal{I} \times \Omega$ is a mapping $w : \mathcal{I} \times \Omega \to \mathbb{R}$ that satisfies

(a) $w$ is $\mathbb{B}$-measurable;

Proof of claim: We will prove the claim using mathematical induction on $m \in \mathbb{Z}_+$. 

1° $m = 0$. This case is straightforward since $\mathcal{B}_0 = \mathcal{H}_0$.

2° Assume that the result holds for $0 \leq m \leq k \in \mathbb{Z}_+$. 

3° Consider the case $m = k + 1 \in \mathbb{N}$. Since $\mathcal{B}_{k+1} = \sigma(\mathcal{H}_{k+1})$, we just need to show the result holds for $\forall E \in \mathcal{H}_{k+1}$. Clearly, $\mathcal{B}_0 \subseteq \mathcal{B}_1 \subseteq \cdots \subseteq \mathcal{B}_{k+1} \subseteq \cdots$, since $\mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \cdots \subseteq \mathcal{H}_{k+1} \subseteq \cdots$. Then, if $E \in \mathcal{H}_k$, the result holds by 2°. On the other hand, if $E \in \mathcal{H}_{k+1} \setminus \mathcal{H}_k$, then $E = (2^{-k-1}i, 2^{-k-1}(i + 1))$ for some $i \in \{0, 1, \ldots, 2^{k+1} - 1\}$. Let $j = \lfloor \frac{i}{2} \rfloor \in \{0, 1, \ldots, 2^k - 1\}$ and $E := (2^{-k}j, 2^{-k}(j + 1)) \in \mathcal{H}_k$. Then, $\chi_{E, 1} = \frac{1}{2}(\chi_{E, 1} + (-1)^jh_{2^{k+1}+j})$. By 2°, $\chi_{E, 1} = \sum_{i=1}^{2^k} \alpha_i h_i$ for some $\alpha_1, \ldots, \alpha_{2^k} \in \mathbb{R}$. Then, we have $\chi_{E, 1} = \sum_{i=1}^{2^{k+1}} \alpha_i h_i$ for some $\alpha_1, \ldots, \alpha_{2^{k+1}} \in \mathbb{R}$. This completes the proof for $m = k + 1$, and completes the inductive process.

This completes the proof of the claim. □

Fix any $f \in \mathcal{Z}$. Fix $m \in \mathbb{Z}_+$. By Theorem 14.63, we have $\forall g_m \in E(f|\mathcal{B}_m)$, $g_m = \arg\min_{g : [0, T] \to \mathbb{R}} \| f - g \|_2$. By $\mathcal{B}_m = \sigma(\mathcal{H}_m) = \sigma((\mathcal{H}_n)_{n=1}^m)$, we have $g_m$ is a real-valued simple function on the measurable space $([0, 1], \mathcal{B}_m)$. By Claim 14.66.1, we have $g_m$ is a linear combination of $(h_n)_{n=1}^m$. Therefore, $g_m = \arg\min_{g \in \text{span}((H_n)_{n=1}^m)} \| f - g \|_2$. By Proposition 13.31, we have $g_m = \sum_{i=1}^{2^m} (f, H_i) H_i$.

Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$, by Proposition 11.182, $\exists \tilde{g} \in C(\mathbb{I}, \mathbb{R}) \cap \mathcal{Z}$ such that $\| f - \tilde{g} \|_2 < \frac{\epsilon}{2}$. Since $\tilde{g} \in \mathcal{Z}$, $\exists m_1 \in \mathbb{Z}_+$, we have $\int_0^{2^{-m_1}} \mathcal{P}_2 \circ \tilde{g}_i \, d\mu < \frac{\epsilon^2}{8}$. $\tilde{g}_i |_{[2^{-m_1}, 1]}$ is continuous and hence uniformly continuous, by Proposition 5.39. Then, $\exists m_2 \in \mathbb{Z}_+ \text{ with } m_2 \geq m_1$, we have $|\tilde{g}(t_1) - \tilde{g}(t_2)| < \frac{\epsilon}{2}$, $\forall t_1, t_2 \in [2^{-m_1}, 1] \subset \mathbb{R}$ with $|t_1 - t_2| < 2^{-m_2}$. Then, $\forall m \in \mathbb{Z}_+$ with $m \geq m_2$, there exists a function $\tilde{g}_m \in \text{span}((H_n)_{n=1}^{2^m})$, defined by $\tilde{g}_m(t) = \begin{cases} 0 & 0 \leq t < 2^{-m_1} \\ \tilde{g}(2^{-m_2}i) - \tilde{g}(2^{-m_2}i + 1), & 2^{-m_2}i \leq t < 2^{-m_2}(i + 1), i = 2^{m_2-1}, \ldots, 2^{m_2} - 1 \end{cases}$, $\forall t \in \mathbb{I}$, such that $\| \tilde{g}_m - \tilde{g} \|_2 < \frac{\epsilon}{2}$. (Note that $\tilde{g}_m \in \text{span}((H_n)_{n=1}^{2^m})$.) Then, we have $\| f - \tilde{g}_m \|_2 < \epsilon$. By the previous paragraph, we have $\| f - g_m \|_2 \leq \| f - \tilde{g}_m \|_2 < \epsilon$. Hence, $\lim_{m \in \mathbb{N}} g_m$ is in $\mathcal{Z}$, and $[f] \in \text{span}((H_n)_{n=1}^{\infty})$ and $(H_n)_{n=1}^{\infty} \subset \mathcal{Z}$ is a complete orthonormal sequence in $\mathcal{Z}$. □
(b) $w(0, \omega) = 0$, $\forall \omega \in \Omega$;

(c) $\forall \omega \in \Omega$, $w(\cdot, \omega) : \mathcal{I} \to \mathbb{R}$ is continuous; (i.e., the sample paths are continuous)

(d) the random process is of independent increment, i.e., $\forall n \in \mathbb{N}$, $\forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n \leq T$, we have $w_{t_i} - w_{s_i}$ is a Gaussian random variable with mean 0 and covariance $t_i - s_i$, $i = 1, \ldots, n$; and the random variables $w_{t_i} - w_{s_i}$, $i = 1, \ldots, n$, are independent, where $(w_t)_{t \in \mathcal{I}}$ is defined by $w_t(\omega) = w(t, \omega)$, $\forall \omega \in \Omega$, $\forall t \in \mathcal{I}$.

**Example 14.68** Let $\mathcal{I} := ((0, 1], \cdot | \cdot), \mathcal{B}, \mu)$ be the finite compact metric measure subspace of $\mathbb{R}$ with partial ordering $\leq$, $\mathcal{I} := ((\mathbb{P}(\mathcal{I}), \cdot | \cdot), \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$, $(H_n)_{n=1}^{\infty}$ be the Haar functions as defined in Example 14.66, $\mathcal{Z} := L_2(\mathbb{P}, \mathbb{R})$, and $\mathcal{Z} := L_2(\mathcal{I}, \mathbb{R})$. Define $(a_n)_{n=1}^{\infty}$ be a sequence of functions: $a_n : \mathcal{I} \to \mathbb{R}$ with $a_n(t) := \int_{\mathcal{I}} dH_n(s)\,ds$, $\forall t \in \mathcal{I}$ and $\forall n \in \mathbb{N}$. Clearly, $a_n(t) = \langle \chi_{(a_n \cdot 1]} \cdot [H_n] \rangle_{\mathcal{Z}}$, $\forall t \in \mathcal{I}$, $\forall n \in \mathbb{N}$. By Fundamental Theorem on Modeling 14.13, and Example 14.14, there exists a probability measure space $\Omega := (\Omega, \mathcal{B}, \mu)$ and random variables $(X_t)_{t \in \mathcal{I}}$, $X_t : \Omega \to \mathbb{R}$, such that $X_t \sim N(0, 1)$, and the random variables $(X_t)_{t \in \mathcal{I}}$ are independent. Let $\mathcal{I} \times \Omega := ([0, 1] \times \Omega, \mathcal{B}, \nu)$ be the finite product measure space, which happens to be a probability measure space. Define the partial sum $w_n : \mathcal{I} \times \Omega \to \mathbb{R}$ by $w_n(t, \omega) = \sum_{i=1}^{n} X_i(\cdot)\,a_i(t)$, $\forall t \in \mathcal{I}$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. Clearly, $w_n$ is $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, by Propositions 11.38 and 11.39; and $w_n(\cdot, \omega) : \mathcal{I} \to \mathbb{R}$ is continuous, $\forall \omega \in \Omega$.

Note that $a_k(t) = \frac{1}{\|h_k\|_2} \langle \chi_{(a_k \cdot 1]} \cdot [h_k] \rangle_2$, $\forall k \in \mathbb{N}$, and $m := \|\log_2 k\|_2$; and the function $\hat{a}_k : \mathcal{I} \to [0, \infty) \subset \mathbb{R}$ defined by $\hat{a}_k(t) = \langle \chi_{(a_k \cdot 1]} \cdot [h_k] \rangle_2$, $\forall t \in \mathcal{I}$, has the same support as $h_k$, and has maximum absolute value $2^{-m}$. Therefore, $a_k$ has the same support as $h_k$, is nonnegative real-valued, and has maximum $\hat{a}_k = 2^{-m}2^{-\frac{m}{2}+1} = 2^{-\frac{m+1}{2}}$, where $m := \log_2 k$. Therefore, $w_n(t, \omega) = X_t(\cdot)\,a_1(t) + \sum_{j=1}^{n} X_i(\cdot)\,a_i(t) =: X_t(\cdot)\,a_1(t) + \sum_{j=1}^{n} b_j(t, \omega)$, $\forall t \in \mathcal{I}$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{N}$. $w_n(\cdot, \omega)$ will converge in $\mathcal{C}(\mathcal{I}, \mathbb{R})$ if $\sum_{j=1}^{n} \max_{t \in \mathcal{I}} |b_j(t, \omega)| < \infty$, by Proposition 7.27 and Example 7.30. Note that $\max_{t \in \mathcal{I}} |b_j(t, \omega)| = \max_{t \in \mathcal{I}} \left| \sum_{i=2^{i-1}+1}^{2^{i+1}} X_i(\cdot)\,a_i(t) \right| = 2^{-\frac{m+1}{2}} \max_{i \in \{2^{i-1}+1, \ldots, 2^i\}} |X_i(\cdot)| =: b_jM_j(\omega)$, $\forall j \in \mathbb{N}$, $\forall \omega \in \Omega$.

Define $w : \mathcal{I} \times \Omega \to \mathbb{R}$ by $w(t, \omega) =$ \begin{align*} &\lim_{n \to \infty} w_n(t, \omega) &\omega \in F \setminus \omega \in F \setminus F, \\
&0 &\forall \omega \in \Omega, \\
\end{align*} 
$\forall t \in \mathcal{I}$, $\forall \omega \in \Omega$, where $F := \{ \omega \in \Omega \mid \lim_{n \to \infty} \sum_{j=1}^{n} b_jM_j(\omega) < \infty \}$. We will show that $w$ is a Wiener process on $\mathcal{I} \times \Omega$.

First, we show that $P(F) = 1$ by proving $E(\sum_{j=1}^{\infty} b_jM_j) < \infty$. Note that $M_j(\omega) = \max_{i \in \{2^{i-1}+1, \ldots, 2^i\}} |X_i(\cdot)| \leq \left( \sum_{i=2^{i-1}+1}^{2^i}(X_i(\cdot))^4 \right)^{1/4} =:
\[ Y_j(\omega), \forall \omega \in \Omega, \forall j \in \mathbb{N}. \text{ Then, } E(M_j) \leq E(Y_j) \leq \left( \sum_{j=2^j-1+1}^{2^j} E(X_j^2) \right)^{1/4} = (3 \cdot 2^{j-1})^{1/4} = \sqrt{3} \cdot 2^{j/4}, \text{ where the second inequality follows from Jensen’s Inequality Theorem 11.98. This implies that } E\left(\sum_{j=1}^{\infty} b_j M_j\right) = \sum_{j=1}^{\infty} b_j E(M_j) \leq \sum_{j=1}^{\infty} 2^{-j/4} \cdot \sqrt{3} \cdot 2^{j/4} = \sum_{j=1}^{\infty} \sqrt{3} \cdot 2^{-j/4} = \sqrt{24}. \text{ Hence, we have } P(F) = 1 \text{ and } w(t, \omega) = \lim_{n \to \infty} w_n(t, \omega) \text{ a.e. } (t, \omega) \in \mathbb{I} \times \Omega. \text{ By Propositions 11.40 and 11.41, } w \text{ is } \bar{B}\text{-measurable. Hence, } w \text{ satisfies (a) of Definition 14.67.}

We will check the rest of the conditions for } w \text{ to be a Wiener process on } \mathbb{I} \times \Omega.

(b) Clearly, } w_n(0, \omega) = 0, \forall n \in \mathbb{N}, \forall \omega \in \Omega. \text{ Then, } w(0, \omega) = \left\{ \begin{array}{ll}
\lim_{n \to \infty} w_n(0, \omega) & = 0, \omega \in F \\
0 & \omega \in \Omega \setminus F = 0, \forall \omega \in \Omega.
\end{array} \right.

(c) Clearly, } w(\cdot, \omega) : \mathbb{I} \to \mathbb{R} \text{ is continuous, } \forall \omega \in \Omega.

(d) Since the Gaussian random variables } (X_t)_{t \in \mathbb{I}} \text{ are independent, then } \forall n \in \mathbb{N}, \forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n \leq 1, \text{ we will prove that } (w_{t_1} - w_{s_1}, \ldots, w_{t_n} - w_{s_n}) \text{ is } \mathbb{R}^n\text{-valued Gaussian random variable with distribution } N(0_n, Q) \text{ by Proposition 14.65, where } Q := \text{block diagonal } (t_1 - s_1, \ldots, t_n - s_n) \in \mathcal{S}_{+} \mathbb{R}^n, \text{ and } (w_t)_{t \in \mathbb{I}} \text{ is defined by } w_t(\omega) = w(t, \omega), \forall \omega \in \Omega, \forall t \in \mathbb{I}. \text{ Clearly, } E(\sum_{j=1}^{\infty} (s_j - t_j)) = 0, j = 1, \ldots, n. \text{ Furthermore,}

\[ E((w_j - w_{s_j})^2) = E\left( \sum_{i=1}^{\infty} (a_i(t_j) - a_i(s_j))^2 \right) = \sum_{i=1}^{\infty} (a_i(t_j) - a_i(s_j))^2
\]

where the second equality follows from the independence of the random variables } (X_t)_{t \in \mathbb{I}}; \text{ and the fourth and fifth equalities follow from the fact that } ([H_i])_{i=1}^{\infty} \text{ is a complete orthonormal sequence in } \mathbb{Z}. \text{ It is easy to check that}

\[ E((w_k - w_{s_k})(w_l - w_{s_l})) = E\left( \sum_{i=1}^{\infty} X_i (a_i(t_k) - a_i(s_k)) (\sum_{j=1}^{\infty} X_j (a_j(t_l) - a_j(s_l))) \right)
\]

\[ = \sum_{i=1}^{\infty} E(X_i^2 (a_i(t_k) - a_i(s_k)) (a_i(t_l) - a_i(s_l)))\]
14.10. EXISTENCE OF THE WIENER PROCESS

\[ = \sum_{i=1}^{\infty} (a_i(t_k) - a_i(s_k))(a_i(t_l) - a_i(s_l)) \]

\[ = \sum_{i=1}^{\infty} \left( \left\langle x(s_i,t_k), x(s_i,t_l), t \right\rangle_z \cdot \left\langle [H_i] \right\rangle_z \right) \]

\[ = \left\langle \sum_{i=1}^{\infty} \left( \left\langle x(s_i,t_k), x(s_i,t_l), t \right\rangle_z \cdot \left\langle [H_i] \right\rangle_z \right) \right\rangle_z \]

\[ = \left\langle \left\langle x(s_i,t_k), x(s_i,t_l), t \right\rangle_z \cdot \left\langle [H_i] \right\rangle_z \right\rangle_z = 0, \quad \forall 1 \leq k, l \leq n \text{ with } l \neq k \]

where the second equality follows from the independence of the random variables \((X_i)_{i=1}^{\infty}\); the fourth equality follows from Definition 13.1; and the fifth and sixth equalities follow from the fact that \(([H_i])_{i=1}^{\infty}\) is a complete orthonormal sequence in \(Z\). Hence, \((w_{t_0} - w_{s_1}, \ldots, w_{t_n} - w_{s_n})\) is \(\mathbb{R}^n\)-valued Gaussian random variable with distribution \(N(0, Q)\). Then, \(w_{t_1} - w_{s_1}, \ldots, w_{t_n} - w_{s_n}\) are independent by Proposition 14.21.

This completes the proof of existence of the Wiener process on \(\bar{I} \times \bar{\Omega}\). ⊗

**Definition 14.69** Let \(\bar{\mathbb{I}} := \{([0,1], | \cdot |), \bar{\mathbb{B}}, \bar{\mu}\}\) be the \(\sigma\)-finite metric measure subspace of \(\mathbb{R}\) with partial ordering \(\leq\), \(\bar{\Omega} := (\Omega, \mathcal{B}, \mu)\) be a probability measure space, and \(\bar{\mathbb{I}} \times \bar{\Omega} := ([0, \infty) \times \bar{\Omega}, \bar{\mathcal{B}}, \bar{\nu})\) be the \(\sigma\)-finite product measure space. A standard Wiener process on \(\bar{\mathbb{I}} \times \bar{\Omega}\) is a mapping \(w : \bar{\mathbb{I}} \times \bar{\Omega} \to \mathbb{R}\) that satisfies

(a) \(w\) is \(\bar{\mathcal{B}}\)-measurable;

(b) \(w(0, \omega) = 0, \forall \omega \in \Omega\);

(c) \(\forall \omega \in \Omega, w(\cdot, \omega) : \bar{\mathbb{I}} \to \mathbb{R}\) is continuous; (i.e., the sample paths are continuous)

(d) the random process is of independent increment, i.e., \(\forall n \in \mathbb{N}, \forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n < \infty, \) we have \(w_{t_i} - w_{s_i}\) is a Gaussian random variable with mean 0 and covariance \(t_i - s_i, i = 1, \ldots, n;\) and the random variables \(w_{t_i} - w_{s_i}, i = 1, \ldots, n,\) are independent, where \((w_t)_{t \in \bar{\mathbb{I}}}\) is defined by \(w_t(\omega) = w(t, \omega), \forall \omega \in \Omega, \forall t \in \bar{\mathbb{I}}\).

**Example 14.70** Let \(\bar{\mathbb{I}} := \{([0,1], | \cdot |), \bar{\mathbb{B}}, \mu\}\) be the finite compact metric measure subspace of \(\mathbb{R}\) with partial ordering \(\leq\). By Example 14.68, there exists probability measure spaces \(\Omega_n := (\Omega_n, \mathcal{B}_n, \mathbb{P}_n), \forall n \in \mathbb{N},\) each with a Wiener process \(w_n : \bar{\mathbb{I}} \times \Omega_n \to \mathbb{R}\). By Fundamental Theorem on Modeling 14.13, there exists a unique probability measure space \(\mathbb{I} := (\Omega, \mathcal{B}, \mathbb{P}) = (\prod_{n \in \mathbb{N}} \Omega_n, \mathcal{B}, \mathbb{P} := \prod_{n \in \mathbb{N}} \mathcal{B}_n, \mathcal{P} := \prod_{n \in \mathbb{N}} \mathbb{P}_n),\) on which \(\mathcal{B}_n := \{ \prod_{i \in \mathbb{N}} B_i \in \mathcal{B} \mid B_i \in \mathcal{B}_i, \forall i \in \mathbb{N}; \text{ and } B_i = \Omega_i, \forall i \in \mathbb{N} \text{ with } i \neq n \} \subseteq \mathcal{B}, n \in \mathbb{N}\) are independent \(\sigma\)-algebras. Let \(\mathbb{I} := \{([0,1], | \cdot |), \mathcal{B}, \mu\}\) be the \(\sigma\)-finite metric measure subspace of \(\mathbb{R}\) with partial ordering \(\leq\), and \(\bar{\mathbb{I}} \times \bar{\Omega} := ([0, \infty) \times \bar{\Omega}, \bar{\mathcal{B}}, \bar{\nu})\) be
the $\sigma$-finite product measure space. Define $w: \mathbb{I} \times \Omega \rightarrow \mathbb{R}$ by $w(t, \omega) = \sum_{i=1}^{\infty} w_i(t, \omega) + w(t - i + 1, \omega), \forall t \in \mathbb{I}$, and $i := \lfloor t \rfloor + 1 \in \mathbb{N}, \forall \omega \in \Omega$. It is easy to check that $w$ is a standard Wiener process on $\mathbb{I} \times \Omega$.

**Definition 14.71** Let $\mathbb{I} := (\{0, \infty\}, | \cdot |, \mathbb{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$, $\Omega := (\Omega, \mathcal{B}, \mathbb{P})$ be a probability measure space, $\mathbb{I} \times \Omega := (\{0, \infty\} \times \Omega, \mathbb{B}, \nu)$ be the $\sigma$-finite product measure space. A $\mathbb{R}^m$-valued standard Wiener process on $\mathbb{I} \times \Omega$ is a mapping $w: \mathbb{I} \times \Omega \rightarrow \mathbb{R}^m$ that satisfies

(a) $w$ is $\mathbb{B}$-measurable;

(b) $w(0, \omega) = 0_m, \forall \omega \in \Omega$;

(c) $\forall \omega \in \Omega, w(. , \omega): \mathbb{I} \rightarrow \mathbb{R}^m$ is continuous; (i. e., the sample paths are continuous)

(d) the random process is of independent increment, i. e., $\forall n \in \mathbb{N}, \forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n \leq T$, we have $w_{t_i} - w_{s_i}$ is a $\mathbb{R}^m$-valued Gaussian random vector with mean $0_m$ and covariance $(t_i - s_i) \text{id}_{\mathbb{R}^m}, i = 1, \ldots, n$; and the random vectors $w_{t_i} - w_{s_i}, i = 1, \ldots, n$, are independent, where $(w_t)_{t \in \mathbb{I}}$ is defined by $w(t, \omega) = w(t, \omega), \forall \omega \in \Omega, \forall t \in \mathbb{I}$.

**Example 14.72** Let $\mathbb{I} := (\{0, \infty\}, | \cdot |, \mathbb{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$ with partial ordering $\leq$, $m \in \mathbb{N}$, and $J := \{1, \ldots, m\}$. By Example 14.70, there exists probability measure spaces $\Omega_i := (\Omega_i, \mathcal{B}_i, \mathbb{P}_i), \forall i \in J$, each with a standard Wiener process $w_i$ defined on $\mathbb{I} \times \Omega_i$. By Fundamental Theorem on Modeling 14.13, there exists a unique probability measure space $\Omega := (\Omega, \mathcal{B}, \mathbb{P}) := (\prod_{i=1}^{m} \Omega_i, \mathcal{B}, \mathbb{P} := \prod_{i=1}^{m} \mathcal{B}_i)$, on which $\mathcal{B} := \{ \bigcap_{j=1}^{m} B_j \in \mathcal{B} \mid B_j \in \mathcal{B}_j, \forall j \in J; \text{ and } B_j = \Omega_j, \forall j \in J \text{ with } j \neq i \} \subset \mathcal{B}, i \in J$ are independent $\sigma$-algebras. Let $\mathbb{I} \times \Omega := (\mathbb{I} \times \Omega, \mathcal{B}, \nu)$ be the $\sigma$-finite product measure space, and define $w: \mathbb{I} \times \Omega \rightarrow \mathbb{R}^m$ by $w(t, \omega) = (w_1(t, \omega), \ldots, w_m(t, \omega)) \in \mathbb{R}^m$, $\forall t \in \mathbb{I}, \forall \omega \in \Omega$. It is easy to check that $w$ is a $\mathbb{R}^m$-valued standard Wiener process on $\mathbb{I} \times \Omega$.

**Proposition 14.73** Let $m \in \mathbb{N}$, $\mathbb{I} := (\{0, \infty\}, | \cdot |, \mathbb{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$ with partial ordering $\leq$, $\Omega := (\Omega, \mathcal{B}, \mathbb{P})$ be a probability measure space, $\mathbb{I} \times \Omega := (\{0, \infty\} \times \Omega, \mathcal{B}, \nu)$ be the $\sigma$-finite product measure space, $w: \mathbb{I} \times \Omega \rightarrow \mathbb{R}^m$ be an $\mathbb{R}^m$-valued standard Wiener process (as defined in Example 14.72), and $T \in \mathbb{B}(\mathbb{R}^m, \mathbb{R}^m)$ is invertible. Then, $\tilde{w} := T w: \mathbb{I} \times \Omega \rightarrow \mathbb{R}^m$ is an $\mathbb{R}^m$-valued stochastic process on $\mathbb{I} \times \Omega$ satisfying

(i) $\tilde{w}$ is $\mathbb{B}$-measurable;

(ii) $\tilde{w}(0, \omega) = 0_m, \forall \omega \in \Omega$;
(iii) \( \forall \omega \in \Omega, \tilde{w}(\cdot, \omega): I \rightarrow \mathbb{R}^m \) is continuous; (i.e., the sample paths are continuous)

(iv) the \( \mathbb{R}^m \)-valued random variable \( \tilde{w}_t: \Omega \rightarrow \mathbb{R}^m \), defined by \( \tilde{w}_t(\omega) = \tilde{w}(t, \omega) = Tw(t, \omega), \forall \omega \in \Omega, \forall t \in I, \) is of independent increment: i.e., \( \forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n < \infty, \tilde{w}_{t_1} - \tilde{w}_{s_1}, \ldots, \tilde{w}_{t_n} - \tilde{w}_{s_n} \) are independent Gaussian random variables with mean \( 0_m \) and covariance \( (t_i - s_i)TT', i = 1, \ldots, n; \)

(v) the auto-covariance function \( E(\tilde{w}_{t_1} \otimes \tilde{w}_{t_2}) = \min \{ t_1, t_2 \} TT', \forall t_1, t_2 \in I. \)

(vi) if in addition, \( T \) is a unitary matrix, i.e., \( TT' = \text{id}_{\mathbb{R}^m}, \) then \( \tilde{w} \) is an \( \mathbb{R}^m \)-valued standard Wiener process on \( I \times \Omega. \)

**Proof**

(i) This is immediate from Definition 14.71 and Proposition 11.38.

(ii) \( \tilde{w}(0, \omega) = Tw(0, \omega) = T0_m = 0_m, \forall \omega \in \Omega. \)

(iii) This is immediate from Definition 14.71 and Proposition 3.12.

(iv) By \( w \) being a \( \mathbb{R}^m \)-valued standard Wiener process, \( \forall 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n < \infty, \tilde{w}_{t_1} - \tilde{w}_{s_1}, \ldots, \tilde{w}_{t_n} - \tilde{w}_{s_n} \) are independent \( \mathbb{R}^m \)-valued Gaussian random variables with mean \( 0_m \) and covariance \( (t_i - s_i)\text{id}_{\mathbb{R}^m}. \) Then, \( W := (w_{t_1} - w_{s_1}, \ldots, w_{t_n} - w_{s_n}) \sim N(0_m, Q), \) where \( Q := \text{block diagonal } \left( (t_1 - s_1)\text{id}_{\mathbb{R}^m}, \ldots, (t_n - s_n)\text{id}_{\mathbb{R}^m} \right). \) Then, we have \( \tilde{W} := (\tilde{w}_{t_1} - \tilde{w}_{s_1}, \ldots, \tilde{w}_{t_n} - \tilde{w}_{s_n}) = (T(w_{t_1} - w_{s_1}), \ldots, T(w_{t_n} - w_{s_n})) = \text{block diagonal } (T, \ldots, T)\tilde{W} =: T_{\tilde{W}}. \) By the invertibility of \( T \), then \( T_{\tilde{W}} \) is invertible. By Proposition 14.22, we have \( \tilde{W} \sim N(0_m, T_{\tilde{W}}T_{\tilde{W}}' =: \tilde{Q}). \) Clearly, \( \tilde{Q} = \text{block diagonal } ((t_1 - s_1)TT', \ldots, (t_n - s_n)TT') \). Hence, \( \tilde{w}_{t_1} - \tilde{w}_{s_1}, \ldots, \tilde{w}_{t_n} - \tilde{w}_{s_n} \) are independent \( \mathbb{R}^m \)-valued Gaussian random variables with mean \( 0_m \) and covariance \( (t_i - s_i)TT', i = 1, \ldots, n, \) by Proposition 14.21.

(v) Without loss of generality, assume \( 0 \leq t_1 \leq t_2. \) The result is trivial if \( t_1 = 0. \) If \( t_1 > 0, \) then \( \tilde{E}(\tilde{w}_{t_1} \otimes \tilde{w}_{t_2}) = \tilde{E}(\tilde{w}_{t_1} \otimes (\tilde{w}_{t_2} - \tilde{w}_{t_1})) + \tilde{E}(\tilde{w}_{t_1} \otimes \tilde{w}_{t_1}) = 0_{m \times m} + t_1 TT' = \min \{ t_1, t_2 \} TT', \) where the first equality follows from Proposition 11.92; and the second equality follows from (iv).

(vi) The result follows directly from (i) – (iv), and Definition 14.71.

This completes the proof of the proposition.

**Proposition 14.74** Let \( m \in \mathbb{N}, I := ([0, \infty), | \cdot |), \mathcal{B}, \mu \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \) with partial ordering \( \leq, \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( I \times \Omega := ([0, \infty) \times \Omega, \mathcal{B}, \nu) \) be the \( \sigma \)-finite product measure space, \( w: I \times \Omega \rightarrow \mathbb{R}^m \) be an \( \mathbb{R}^m \)-valued standard Wiener process (as defined in Example 14.72), and \( \alpha \in (0, \infty) \subset \mathbb{R}. \) Then, \( \tilde{w}: I \times \Omega \rightarrow \mathbb{R}^m, \) defined by \( \tilde{w}(t, \omega) = \frac{1}{\alpha}w(\alpha^2 t, \omega), \forall \omega \in \Omega, \forall t \in I, \) is an \( \mathbb{R}^m \)-valued standard Wiener process.

**Proof** This is immediate from Proposition 14.73 and Definition 14.71.
Proposition 14.75 Let \( I := ([0, \infty), \cdot, | \cdot|, \hat{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \) with partial ordering \( \leq \). \( \Omega := (\Omega, B, P) \) be a probability measure space, \( I \times \Omega := ([0, \infty) \times \Omega, B, \nu) \) be the \( \sigma \)-finite product measure space, \( w : I \times \Omega \to \mathbb{R} \) be an standard Wiener process (as defined in Example 14.72). Then, there exists \( F \in B \) with \( P(F) = 1 \) such that \( \forall \omega \in F, \lim_{h \to 0^+} \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| = \infty, \forall t \in I \).

**Proof** Fix any \( m \in \mathbb{Z}^+ \), we consider the interval \( t \in [m, m + 1) \). Let \( F_m := \{ \omega \in \Omega \mid \lim_{h \to 0^+} \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| = \infty, \forall t \in [m, m + 1) \} \). Note that \( F_m \) may not be \( B \)-measurable. Define \( \hat{F} := \bigcap_{m=0}^\infty F_m \).

Then, \( \forall \omega \in \hat{F}, \lim_{h \to 0^+} \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| = \infty, \forall t \in I \).

Let \( A_{m,j,k} := \bigcup_{m \leq t < m+1} \bigcup_{0 < h \leq \frac{1}{k}} \{ \omega \in \Omega \mid \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| \leq j \} \), \( \forall j, k \in \mathbb{N} \). Clearly, \( A_{m,j,k} \subseteq \Omega \setminus F_m, \forall j, k \in \mathbb{N} \), and \( F_m \subseteq \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus A_{m,j,k}) \).

\( \forall \omega \in \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus A_{m,j,k}), \) we have \( \forall j \in \mathbb{N}, \forall k \in \mathbb{N}, \forall t \in [m, m + 1), \exists h \in (0, \frac{1}{k}], \) such that \( \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| > j \). Then, \( \omega \in F_m \).

Hence, \( \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus A_{m,j,k}) \subseteq F_m \). Thus, \( F_m = \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus A_{m,j,k}) \). Note that \( A_{m,j,k} \)’s may not be \( B \)-measurable.

Let \( C_{m,j,k} := \bigcap_{h=4k}^\infty \bigcup_{m=1}^n \bigcup_{j=1}^3 \{ \omega \in \Omega \mid |w(m + \frac{i+n-1}{n}, \omega) - w(m + \frac{i+n}{n}, \omega)| \leq 1/n \} \in B, \forall j, k \in \mathbb{N} \).

We will show that \( A_{m,j,k} \subseteq C_{m,j,k}, \forall j, k \in \mathbb{N}, \) and \( P(C_{m,j,k}) = 0, \forall j, k \in \mathbb{N} \). This yields that \( F_m \supseteq \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus C_{m,j,k}), \forall j, k \in \mathbb{N}, \) and \( F \supseteq F := \bigcap_{m=0}^\infty \bigcap_{j,k \in \mathbb{N}} (\Omega \setminus C_{m,j,k}) \in B, \) and \( P(F) = 1 \). This set \( F \) is what we seek.

Fix any \( j, k \in \mathbb{N} \). Fix any \( \omega \in A_{m,j,k} \). Then, \( \exists t \in [m, m + 1) \) such that \( \omega \in \bigcap_{h \leq \frac{1}{k}} \{ \omega \in \Omega \mid \frac{1}{h} |w(t + h, \omega)) - w(t, \omega)| \leq j \} \). Fix any \( n \geq 4k \). We have \( \frac{1}{n} \leq \frac{1}{4k} \). This implies that \( \exists i \in \{1, \ldots, n\} \) such that \( m + \frac{i+n-1}{n} \leq t < m + \frac{i+n}{n} \).

Fix any \( v \in \{1, 2, 3\} \). Then, we have \( m + \frac{i+n-1}{n} \leq t + \frac{v}{n} < m + \frac{i+n}{n}, 0 < m + \frac{i+n-1}{n} - t \leq \frac{v}{n} < \frac{v}{n} = \frac{1}{4k}, \) and \( 0 < m + \frac{i+n}{n} - t \leq \frac{v}{n} \leq \frac{1}{4k} \).

This leads to \( |w(m + \frac{i+n-1}{n}, \omega) - w(m + \frac{i+n}{n}, \omega)| \leq |w(m + \frac{i+n-1}{n} + j + (m + \frac{i+n}{n} - t)j) - (m + \frac{i+n-1}{n} - t)j| \leq \frac{1}{4k} + \frac{1}{4k} = \frac{1}{2k} \).

Hence, \( \omega \in C_{m,j,k} \). Thus, \( A_{m,j,k} \subseteq C_{m,j,k}, \forall j, k \in \mathbb{N} \).

Now, \( P(\{ \omega \in \Omega \mid |w(m + \frac{i+n-1}{n}, \omega) - w(m + \frac{i+n}{n}, \omega)| \leq \frac{1}{2k} \}) = \int_{-\frac{1}{2k}}^{\frac{1}{2k}} \sqrt{2\pi n} \cdot \exp\left(-\frac{n x^2}{2}\right) dx \leq \frac{1}{\sqrt{2\pi n}}, \) where the equality follows from the fact \( w_{m + \frac{i+n-1}{n}} - w_{m + \frac{i+n}{n}} \sim \mathcal{N}(0, \frac{1}{n}) \).

Then, \( P(\bigcap_{i=1}^{n} \{ \omega \in \Omega \mid |w(m + \frac{i+n-1}{n}, \omega) - w(m + \frac{i+n}{n}, \omega)| \leq \frac{1}{2k} \}) \leq \left(\frac{14j}{\sqrt{2\pi n}}\right)^3 \) where the inequality follows from the fact that the three events in the intersection are independent. This implies that \( P(\bigcup_{i=1}^{n} \{ \omega \in \Omega \mid |w(m + \frac{i+n-1}{n}, \omega) - w(m + \frac{i+n-1}{n}, \omega)| \leq \frac{1}{2k} \}) \leq n \left(\frac{14j}{\sqrt{2\pi n}}\right)^3 = \frac{1}{\sqrt{n}} \left(\frac{14j}{\sqrt{2\pi}}\right)^3 \to 0 \) as \( n \to \infty \).
14.11. Martingales With General Index Set

Definition 14.76 Let $X := (X, \mathcal{B}(X), \mu)$ be a nonempty $\sigma$-finite topological measure space with the partial ordering $\preceq$ defined on it, $\Omega := (\Omega, \mathcal{B}, (\mathcal{B}_t)_{t \in X}, P)$ be a filtered probability measure space, $X \times \Omega := (X \times \Omega, \mathcal{B} \times \mathcal{B}, \nu)$ be the $\sigma$-finite product measure space, $\mathcal{Y}$ be a Banach space, $M : X \times \Omega \to \mathcal{Y}$, or equivalently $(M_t)_{t \in X}$, be a $\mathcal{Y}$-valued stochastic process adapted to $(\mathcal{B}_t)_{t \in X}$. Then, the stochastic process $(M_t)_{t \in X}$ is said to be a Martingale if

(i) $M_t \in \mathcal{L}_1(\Omega, Y_t), \forall t \in X$, where $Y_t$ is a separable Banach subspace of $\mathcal{Y}$;

(ii) $M_s \in E(M_t|\mathcal{B}_s), \forall s, t \in X$ with $s \leq t$.

Let $\mathcal{Y} = \mathbb{R}$. The stochastic process $(M_t)_{t \in X}$ is said to be a super Martingale if it satisfies (i) and $M_s \geq E(M_t|\mathcal{B}_s)$ a.e. in $\Omega$, $\forall s, t \in X$ with $s \leq t$.

Let $\mathcal{Y} = \mathbb{R}$. The stochastic process $(M_t)_{t \in X}$ is said to be a sub Martingale if it satisfies (i) and $M_s \leq E(M_t|\mathcal{B}_s)$ a.e. in $\Omega$, $\forall s, t \in X$ with $s \leq t$.

Proposition 14.77 Let $m \in \mathbb{N}$, $I := ((0, \infty), | \cdot |, \mathcal{B}, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$ with the partial ordering $\leq$, $\Omega := (\Omega, \mathcal{B}, (\mathcal{B}_t)_{t \in I}, P)$ be a filtered probability measure space, $\alpha \in \mathbb{R}^m$, $I \times \Omega := ([0, \infty) \times \Omega, \mathcal{B}, \nu)$ be the $\sigma$-finite product measure space, $w : I \times \Omega \to \mathbb{R}^m$ be an $\mathbb{R}^m$-valued standard Wiener process (as defined in Example 14.72); $(\mathcal{B}_t)_{t \in I}$ be the natural filtration of the process $(w_t)_{t \in I}$, which is defined as $w_t(\omega) = w(t, \omega), \forall \omega \in \Omega, \forall t \in I$; and $M : I \times \Omega \to \mathbb{R}$, defined by $M(t, \omega) := \exp((w(t, \omega), \alpha)_{\mathbb{R}^m} - \frac{(\alpha, \alpha)_{\mathbb{R}^m}}{2}t), \forall \omega \in \Omega, \forall t \in I$. Then, $(M_t)_{t \in I}$, defined by $M_t(\omega) := M(t, \omega), \forall \omega \in \Omega, \forall t \in I$, is a Martingale adapted to $(\mathcal{B}_t)_{t \in I}$.

Proof By Propositions 11.38 and 11.39, $M$ is $\mathcal{B}$-measurable. By Proposition 11.38, $M_t$ is $\mathcal{B}_t$-measurable, $\forall t \in I$. Hence, $(M_t)_{t \in I}$ is a stochastic process adapted to the filtration $(\mathcal{B}_t)_{t \in I}$.

Clearly, $M_0(\omega) = 1, \forall \omega \in \Omega$. Then, $E(M_0) = 1, \forall t \in I$ with $t > 0$, we have $w_t \sim N(0_m, t\text{id}_{\mathbb{R}^m})$. Then,

$E(M_t) = E(\exp((w_t, \alpha)_{\mathbb{R}^m} - \frac{(\alpha, \alpha)_{\mathbb{R}^m}}{2}t))$

$= E(\exp((w_t, \alpha)_{\mathbb{R}^m})) \exp(-\frac{(\alpha, \alpha)_{\mathbb{R}^m}}{2}t)$

Thus, we have $0 \leq P(C_{m,j,k}) \leq \min_{n \geq 4k} \frac{1}{\sqrt{n}} \frac{144}{\sqrt{2\pi n}} = 0, \forall j, k \in \mathbb{N}$.

This completes the proof of the proposition. $\square$

This proposition shows that the Wiener process is nowhere differentiable on $[0, \infty)$ with probability 1.
\[= \exp\left(-\frac{\langle \alpha, \alpha \rangle_{\mathbb{R}^m}}{2} - t\right) \int_{\mathbb{R}^m} \exp\left(\frac{1}{2t} \langle x, x \rangle_{\mathbb{R}^m}\right) d\mu_{\mathbb{R}^m}(x)\]
\[= \int_{\mathbb{R}^m} \exp\left(\frac{1}{2t} \langle x, x \rangle_{\mathbb{R}^m}\right) d\mu_{\mathbb{R}^m}(x)\]
\[= \int_{\mathbb{R}^m} \exp\left(-\frac{1}{2t} \langle x, x \rangle_{\mathbb{R}^m}\right) d\mu_{\mathbb{R}^m}(x) = 1\]

where the second equality follows from Proposition 11.92; the third equality follows from the Definitions 14.14 and 14.19; and the fourth equality follows from simple algebra; the fifth equality follows from Change of Variable, Theorem 12.91, with \(\bar{x} := x - ta\); and the last equality is well-known.

Hence, \(M_t \in L_1(\Omega, \mathcal{F}, \mathbb{P})\), \(\forall t \in \mathbb{R}\).

Fix any \(s, t \in \mathbb{R}\) with \(s \leq t\). If \(t = s\), then \(M_s \in \mathbb{B}(\mathcal{M}_t)\). Consider the case \(t > s\), we have \(\mathbb{E}(M_s \mid \mathcal{B}_s) = \mathbb{E}(M_t \exp(\langle (w_t - w_s), \alpha \rangle_{\mathbb{R}^m} - \frac{\langle \alpha, \alpha \rangle_{\mathbb{R}^m}}{2} (t - s)))|\mathcal{B}_s) = M_s \mathbb{E}(\exp(\langle (w_t - w_s), \alpha \rangle - \frac{\langle \alpha, \alpha \rangle_{\mathbb{R}^m}}{2} (t - s)))|\mathcal{B}_s) = M_s \text{ a.e. in } \Omega_s\),

where the second equality follows from (h) of Proposition 14.11; the last equality follows from the independence of \(\mathcal{B}_s\) and \(\sigma(w_t - w_s)\), the derivation of the previous paragraph, and (k) of Proposition 14.11; and \(\Omega_s := (\Omega, \mathcal{B}_s, \mathbb{P}_s := \mathbb{P}|_{\mathcal{B}_s})\). Hence, \(M_s \in \mathbb{E}(M_t | \mathcal{B}_s), \forall s, t \in \mathbb{R}\) with \(s \leq t\).

This completes the proof of the proposition.

\[\square\]

**Lemma 14.78** Let \(\mathbb{I} := ((0, \infty), \cdot), \mathcal{B}, \mu\) be the \(\sigma\)-finite metric measure space of \(\mathbb{R}\) with partial ordering \(\leq\), \(\Omega := (\Omega, \mathcal{B}, \mathbb{P})\) be a probability measure space, \(\mathbb{I} \times \Omega := ((0, \infty) \times \Omega, \mathcal{B}, \nu)\) be the \(\sigma\)-finite product measure space, \(w : \mathbb{I} \times \Omega \rightarrow \mathbb{R}\) be an standard Wiener process (as defined in Example 14.70), \((\mathcal{B}_t)_{t \in \mathbb{I}}\) be the natural filtration of the process \((w_t)_{t \in \mathbb{I}}\), where \(w_t(\omega) = w(\omega)(t), \forall \omega \in \Omega, \forall t \in \mathbb{I}\), and \(\alpha, \beta \in (0, \infty) \subset \mathbb{R}\). Then, \(\mathbb{P}(\{\omega \in \Omega \mid w(t, \omega) > \frac{\alpha}{2} + \beta, \text{ for some } t \in \mathbb{I}\}) = \exp(-\alpha \beta)\).

**Proof** Let \(A := \{\omega \in \Omega \mid w(t, \omega) > \frac{\alpha}{2} + \beta, \text{ for some } t \in \mathbb{I}\}\) and \(M : \mathbb{I} \times \Omega \rightarrow \mathbb{R}\) be defined by \(M(t, \omega) = \exp(\alpha w(\omega)(t) - \frac{\alpha^2}{2} t), \forall \omega \in \Omega, \forall t \in \mathbb{I}\). By Proposition 14.77, the stochastic process \((M_t)_{t \in \mathbb{I}}\), defined by \(M_t(\omega) = M(t, \omega), \forall \omega \in \Omega, \forall t \in \mathbb{I}\), is a martingale adapted to the filtration \((\mathcal{B}_t)_{t \in \mathbb{I}}\) and \(\mathbb{E}(M_t) = 1, \forall t \in \mathbb{I}\). Then, \(A = \{\omega \in \Omega \mid \sup_{t \in \mathbb{I}} M_t(\omega) > \exp(\alpha \beta)\}\). Let \(M^*(\omega) := \sup_{t \in \mathbb{I}} M_t(\omega), \forall \omega \in \Omega\). Then, \(A = \{\omega \in \Omega \mid M^*(\omega) > \exp(\alpha \beta)\}\). Fix any \(N \in \mathbb{N}\), define \((f^n(N))_{n \in \mathbb{Z}^+}\) by \(f^n(N)(\omega) = M(\frac{\alpha}{2n}, \omega), \forall n \in \mathbb{Z}^+\) and \(\forall \omega \in \Omega\). Then, \((f^n(N))_{n \in \mathbb{Z}^+}\) is a martingale indexed by \(\mathbb{Z}^+\). By the continuity of \(M(\cdot, \omega) : \mathbb{I} \rightarrow \mathbb{R}, \forall \omega \in \Omega\), we have \(A = \{\omega \in \Omega \mid M^*(\omega) > \exp(\alpha \beta)\} = \bigcup_{N \in \mathbb{N}} \{\omega \in \Omega \mid \sup_{n \in \mathbb{Z}^+} f^n(N)(\omega) > \exp(\alpha \beta)\}\). The last element of follows from Proposition 11.40. Fix any \(N \in \mathbb{N}\), consider the martingale \((f^n(N))_{n \in \mathbb{Z}^+}\). Let \(T^{(N)} : \Omega \rightarrow \mathbb{Z}^+ \cup \{\infty\}\) be defined by \(T^{(N)}(\omega) := \inf \{n \in \mathbb{Z}^+ \mid f^n(N)(\omega) > \exp(\alpha \beta)\}\), \(\forall \omega \in \Omega\). By
14.11. MARTINGALES WITH GENERAL INDEX SET

Example 14.28, $T^{(N)}$ is a stopping time for the Martingale $(f_n^{(N)})_{n \in \mathbb{Z}_+}$.

By Theorem 14.31, the stopped stochastic process $(Y_n^{(N)})_{n \in \mathbb{Z}_+}$, defined by $Y_n^{(N)}(\omega) = f_{T^{(N)}(\omega) \wedge n}(\omega)$, $\forall \omega \in \Omega$, $\forall n \in \mathbb{Z}_+$, is a Martingale. Note that $Y_0^{(N)}(\omega) = M_0(\omega) = 1$, $\forall \omega \in \Omega$. Then, $E(Y_n^{(N)}) = 1$, $\forall n \in \mathbb{Z}_+$.

By Doob’s Forward Convergence Theorem 14.37, we have $Y_\infty^{(N)}(\omega) := \lim_{t \to \infty} Y_t^{(N)}(\omega)$ satisfies $Y_\infty^{(N)}(\omega)$.

Lemma 14.79 Let $\alpha \in \mathbb{R}_+$, $V : (\alpha, \infty) \to \mathbb{R}$ be continuous, and $\tilde{V} : (0, \alpha) \to \mathbb{R}$ be continuous. Then,

$$\limsup_{t \to \infty} V(t) = \inf_{t \in \mathbb{R}_+, t > \alpha} \sup_{s \in \mathbb{R}_+, s > t} V(s) = \inf_{t \in \mathbb{Q}_+, t > \alpha} \sup_{s \in \mathbb{Q}_+, s > t} V(s) = \inf_{t \in (0, \alpha)} \sup_{s \in \mathbb{R}_+, s < t} V(s) = \inf_{t \in \mathbb{Q}_+, 0 < t < \alpha} \sup_{s \in \mathbb{Q}_+, 0 < s < t} V(s) \in \mathbb{R}_e$$

Proof First, consider the function $V$. By Definition 3.84, we have

$$\limsup_{t \to -\infty} V(t) = \inf_{t \in \mathbb{R}_+, t > \alpha} \sup_{s \in \mathbb{R}_+, s > t} V(s) \in \mathbb{R}_e.$$ Define $H : (\alpha, \infty) \to \mathbb{R}_e$ by $H(t) := \sup_{s \in \mathbb{R}_+, s > t} V(s) \in \mathbb{R}_e$, $\forall t \in \mathbb{R}$ with $t > \alpha$. Since $V$ is continuous, we have $H(t) = \sup_{s \in \mathbb{Q}_+, s > t} V(s) \in \mathbb{R}_e$, $\forall t \in \mathbb{R}$ with $t > \alpha$.

We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\limsup_{t \to -\infty} V(t) < +\infty$; Case 2: $\limsup_{t \to -\infty} V(t) = +\infty$. Case 1: $\limsup_{t \to -\infty} V(t) < +\infty$. Then, $\inf_{t \in \mathbb{R}_+, t > \alpha} H(t) < +\infty$. Clearly, $H$ is monotonically nonincreasing on $(\alpha, \infty)$. If there exists $t_0 \in (\alpha, \infty)$ with $H(t_0) = \infty$, we can show that $H(t) = \infty$, $\forall t \in (\alpha, \infty)$. Then, we have $H : (\alpha, \infty) \to \mathbb{R}$, since $V : (\alpha, \infty) \to \mathbb{R}$. We need the following intermediate result.

Claim 14.79.1 $H : (\alpha, \infty) \to \mathbb{R}$ is continuous.

Proof of claim: Fix any $t \in (\alpha, \infty)$. $H(t) = \sup_{s \in \mathbb{R}_+, s > t} V(s) \in \mathbb{R}$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. Clearly, $V(t) \leq H(t)$, since $V$ is continuous. We will distinguish two exhaustive and mutually exclusive cases: Case A: $H(t) > V(t)$; Case B: $H(t) = V(t)$.
Case A: $H(t) > V(t)$. Then, by continuity of $V$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $|V(t) - V(t)| < \frac{\delta}{2}(H(t) - V(t))$, $\forall t \in B_{\mathbb{R}}(t, \delta) \cap (\alpha, \infty) =: D$. This implies that $H(l) = H(t), \forall l \in D$. Then, $|H(l) - H(t)| = 0$, $\forall l \in D$.

Case B: $H(t) = V(t)$. By continuity of $V$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $|V(t) - V(t)| < \epsilon$, $\forall t \in B_{\mathbb{R}}(t, \delta) \cap (\alpha, \infty) =: D$. Fix any $l \in D$. If $l > t$, we have $H(t) - \epsilon = V(t) - \epsilon < V(l) \leq \sup_{s \geq t} V(s) = H(l) \leq H(t)$. Thus, we have $|H(l) - H(t)| \leq \epsilon$. On the other hand, if $l < t$, we have $H(t) \leq H(l) = \sup_{s \geq t} V(s) = \max \{ \sup_{s \in [l, t]} V(s), \sup_{s \in [t, \infty)} V(s) \} = H(t) + \epsilon$. This yields $|H(l) - H(t)| \leq \epsilon$.

Hence, we have $|H(l) - H(t)| \leq \epsilon$, $\forall l \in D$.

In both cases, we have found $\delta \in (0, \infty) \subset \mathbb{R}$ such that $|H(l) - H(t)| \leq \epsilon$, $\forall l \in B_{\mathbb{R}}(t, \delta) \cap (\alpha, \infty) =: D$. Hence, $H$ is continuous at $t$. By the arbitrariness of $t$, we conclude that $H$ is continuous. This completes the proof of the claim.

By Claim 14.79.1, we have $\inf_{t \in \mathbb{R}, t > \alpha} H(t) = \inf_{t \in \mathbb{R}, t > \alpha} H(t)$. This leads to $\limsup_{t \to \infty} V(t) = \inf_{t \in \mathbb{R}, t > \alpha} H(t) = \inf_{t \in \mathbb{R}, t > \alpha} H(t) = \inf_{t \in \mathbb{R}, t > \alpha} \sup_{s \in \mathbb{R}, s > t} V(s)$. Hence, the result holds in this case.

Case 2: $\limsup_{t \to \infty} V(t) = +\infty$. Then, $\inf_{t \in \mathbb{R}, t > \alpha} H(t) = \infty$. This implies that $H(t) = \infty$, $\forall t \in (\alpha, \infty) \subset \mathbb{R}$. This leads to $\limsup_{t \to \infty} V(t) = \inf_{t \in \mathbb{R}, t > \alpha} H(t) = \inf_{t \in \mathbb{R}, t > \alpha} \sup_{s \in \mathbb{R}, s > t} V(s)$. Hence, the result holds in this case.

In both cases, we have shown that the desired result holds.

Next, consider the function $\bar{V}$. By Definition 3.84,

$$
\lim_{t \to 0} \bar{V}(t) = \inf_{t \in \mathbb{R}, t < \alpha} \sup_{s \in \mathbb{R}, s < t} \bar{V}(s) = \inf_{t \in \mathbb{R}, t > \frac{1}{2}} \sup_{s \in \mathbb{R}, s > t} \bar{V}(s) = \frac{1}{s} \limsup_{t \to \infty} \bar{V}(s) = \frac{1}{s} \limsup_{t \to \infty} \bar{V}(t) = \frac{1}{s} \limsup_{t \to \infty} \bar{V}(t) = \frac{1}{s} \limsup_{t \to \infty} \bar{V}(t)
$$

where the first equality follows from Definition 3.84; and the third and fourth equalities follow from the result for function $V$ with $V(t) := \bar{V}(\frac{1}{s})$.

This completes the proof of the lemma. □

**Theorem 14.80 (Law of Iterated Logarithm)** Let $((0, \infty), | \cdot |, \mathcal{B}, \mu) =: \mathbb{I}$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$ with the partial ordering $\leq$, $\Omega := (\Omega, \mathcal{B}, \mathbb{P})$ be a probability measure space, $\mathbb{I} \times \Omega := ((0, \infty) \times \Omega, \mathcal{B}, \nu)$ be the $\sigma$-finite product measure space, $w : \mathbb{I} \times \Omega \to \mathbb{R}$ be an standard Wiener process, and the stochastic process $(w_t)_{t \in \mathbb{I}}$ defined by $w_t(\omega) = w(\omega)(t)$, $\forall \omega \in \Omega$, $\forall t \in \mathbb{I}$. Then, the following statements hold.

(i) $\limsup_{t \to \infty} \frac{w_t}{\sqrt{2t \ln(\ln(t))}} = 1$ a.e. in $\Omega$.

(ii) $\liminf_{t \to \infty} \frac{w_t}{\sqrt{2t \ln(\ln(t))}} = -1$ a.e. in $\Omega$. □
14.11. Martingales with General Index Set

(iii) \( \limsup_{t \to 0^+} \frac{w_t}{\sqrt{2t \ln(\ln(\frac{1}{t}))}} = 1 \) a.e. in \( \Omega \).

(iv) \( \liminf_{t \to 0^+} \frac{w_t}{\sqrt{2t \ln(\ln(\frac{1}{t}))}} = -1 \) a.e. in \( \Omega \).

(v) \( \tilde{w} : I \times \Omega \to \mathbb{R} \) defined by

\[
\tilde{w}(t, \omega) := \begin{cases} 
t w(\frac{1}{t}, \omega) & \text{if } (t, \omega) \in (0, \infty) \times E \\
0 & \text{otherwise}
\end{cases}
\]

\( \forall \omega \in \Omega, \forall t \in I, \) where \( E := \{ \omega \in \Omega \mid \limsup_{t \to \infty} \frac{w_t}{\sqrt{2t \ln(\ln(t))}} = 1 \) and \( \liminf_{t \to \infty} \frac{w_t}{\sqrt{2t \ln(\ln(t))}} = -1 \}, \) is also a standard Wiener process on \( \mathbb{I} \times \Omega \).

Proof (i) Let \( h : (e, \infty) \to (0, \infty) \subset \mathbb{R} \) be defined by \( h(t) = \sqrt{2t \ln(\ln(t))}, \forall t \in (e, \infty) \subset \mathbb{R} \). Clearly, \( \lim_{t \to \infty} h(t) = \infty \) and \( h \) is strictly increasing. Fix any \( \theta, \delta \in \mathbb{Q} \) with \( 0 < \delta < 1 < \theta \). Let \( \alpha_n := (1 + \delta) \theta^{-n} h(\theta^n) > 0 \) and \( \beta_n := \frac{1}{2} h(\theta^n) > 0, \forall n \in \mathbb{N} \) with \( n \geq n_{0, \theta} \), where \( \theta^{n_{0, \theta}} > e \). Then, \( \alpha_n \beta_n = \frac{1}{2} (1 + \delta) \theta^{-n} (h(\theta^n))^2 = (1 + \delta) \ln(n \ln(\theta^n)) = (1 + \delta) \ln(n \ln(\theta)) = \ln((n \ln(\theta))^{1+\delta}) \) and \( \exp(-\alpha_n \beta_n) = \frac{c(\theta)}{n^{1+\delta}} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{0, \theta} \), where \( c(\theta) := (\ln(\theta))^{1-\delta} \). Therefore, we have \( \sum_{n=n_{0, \theta}}^{\infty} \exp(-\alpha_n \beta_n) < +\infty \).

By Lemma 14.78, we have \( P(\{ \omega \in \Omega \mid w(t, \omega) > \frac{\alpha_n}{2} t + \beta_n \text{ for some } t \in I \}) \leq \exp(-\alpha_n \beta_n) \). This implies that \( \sum_{n=n_{0, \theta}}^{\infty} P(\{ \omega \in \Omega \mid w(t, \omega) > \frac{\alpha_n}{2} t + \beta_n \text{ for some } t \in I \}) < \infty \). By Borel-Cantelli Lemma 14.6, \( P(\bigcap_{n=n_{0, \theta}}^{\infty} \bigcap_{m=m_{0, \theta}}^{\infty} \{ \omega \in \Omega \mid w(t, \omega) > \frac{\alpha_m}{2} t + \beta_m \text{ for some } t \in I \}) = 0 = P(\{ \omega \in \Omega \mid w(t, \omega) > \frac{\alpha_m}{2} t + \beta_m \text{ for some } t \in I \text{ and for infinitely many } n \geq n_{0, \theta} \}) \). Then, \( P(\Omega_{\theta, \delta}) = 1 \). Fix any \( \omega \in \Omega_{\theta, \delta} \). \( \exists N_{\theta, \delta}(\omega) \in \mathbb{N} \) with \( N_{\theta, \delta}(\omega) \geq n_{0, \theta} \) such that \( w(t, \omega) \leq \frac{\alpha_n}{2} t + \beta_n \), \( \forall t \in I \) and \( \forall n > N_{\theta, \delta}(\omega) \). Fix any \( n > N_{\theta, \delta}(\omega) \). Let \( \forall t \in [\theta^n, \theta^{n+1}] \subset \mathbb{R} \), we have \( w(t, \omega) \leq \frac{\alpha_n}{2} t + \beta_n = \frac{\theta^{n+1} h(\theta^n)}{2} + \frac{1}{2} h(\theta^n) \leq (1 + \delta) \theta h(\theta^n) \leq (1 + \delta) h(t) \). Hence, we have \( \sup_{t \geq \theta^{N_{\theta, \delta}(\omega)+1}} \frac{w(t, \omega)}{h(t)} \leq (1 + \delta) h(t) \). This implies that \( \limsup_{t \to \infty} \frac{w(t, \omega)}{\theta(t)} \leq (1 + \delta) h(t) \), \( \forall \omega \in \Omega_{\theta, \delta} \). By Lemma 14.79, \( \limsup_{t \to \infty} \frac{w(t, \omega)}{\theta(t)} : \Omega \to \mathbb{R} \) is \( \mathcal{E} \)-measurable. Then, we have \( \limsup_{t \to \infty} \frac{w(t, \omega)}{\theta(t)} \leq (1 + \delta) h(t) \text{ a.e. in } \Omega \). By the arbitrariness of \( \theta \) and \( \delta \), we have \( \limsup_{t \to \infty} \frac{w(t, \omega)}{h(t)} \leq 1 \text{ a.e. in } \Omega \).

Next, we will show that \( \limsup_{t \to \infty} \frac{w(t, \omega)}{\theta(t)} \geq 1 \text{ a.e. in } \Omega \). Let \( A_n := \{ \omega \in \Omega \mid w(\theta^n, \omega) - w(\theta^{n-1}, \omega) > \sqrt{1 - \theta^{-1} h(\theta^n)} \} \in \mathcal{E}, \forall n \in \mathbb{N} \) with \( n \geq n_{0, \theta} \). Clearly, \( A_n, n \geq n_{0, \theta} \), are independent events, since \( w_t \) is of independent increments. We need the following inequality.

Claim 14.80.1 \( \frac{1}{1+x} \exp(-\frac{x^2}{2}) < \int_x^{\infty} \exp(-\frac{y^2}{2}) \, dy < \frac{1}{2} \exp(-\frac{x^2}{2}), \forall x \in (0, \infty) \subset \mathbb{R} \).
**Proof of claim:** Fix any \( x \in (0, \infty) \subset \mathbb{R} \). Then, we have

\[
\int_{x}^{\infty} \exp\left(-\frac{y^2}{2}\right) \, dy < \int_{x}^{\infty} \frac{y}{x} \exp\left(-\frac{y^2}{2}\right) \, dy
\]

\[
= \frac{1}{x} \int_{x}^{\infty} \exp\left(-\frac{y^2}{2}\right) \, dy \frac{y^2}{2} = \frac{1}{x} \int_{x}^{\infty} \exp(-z) \, dz = \frac{1}{x} \exp\left(-\frac{x^2}{2}\right)
\]

where the second equality follows from Change of Variable Theorem 12.91 with \( z = \frac{y^2}{2} \). On the other hand,

\[
\int_{x}^{\infty} \exp\left(-\frac{y^2}{2}\right) \, dy > \int_{x}^{\infty} \frac{1 + y^{-2}}{1 + x^{-2}} \exp\left(-\frac{y^2}{2}\right) \, dy
\]

\[
= \frac{x^2}{1 + x^2} \int_{x}^{\infty} (1 + y^{-2}) \exp\left(-\frac{y^2}{2}\right) \, dy = \frac{x^2}{1 + x^2} \left( -\frac{1}{y} \exp\left(-\frac{y^2}{2}\right) \right)_{x}^{\infty}
\]

\[
= \frac{x}{1 + x^2} \exp\left(-\frac{x^2}{2}\right)
\]

This completes the proof of the claim.

Then, \( P(A_n) = P\left( \left\{ \omega \in \Omega \mid \frac{u_{\theta n} - u_{\theta n+1}}{\sqrt{\theta n - \theta n-1}} > \frac{1 - \theta^{-1}}{\sqrt{\theta n - \theta n-1}} h(\theta^n) \right\} \right) = \int_{\Phi} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \, dy < \frac{1}{\sqrt{2\pi}} \frac{\exp\left(-\frac{x^2}{2}\right)}{\exp\left(-\frac{1}{2}\right)} = \frac{1}{\sqrt{2\pi} \ln(n)} \]

with \( n \geq n_{0, \theta} \), where the second equality follows from the fact that \( \frac{u_{\theta n} - u_{\theta n+1}}{\sqrt{\theta n - \theta n-1}} \sim N(0, 1) \); and the inequality follows from Claim 14.80.1. Again by Claim 14.80.1, we have \( P(A_n) > \frac{1}{\sqrt{2\pi} \ln(n)} = \frac{\sqrt{\theta n - \theta n-1}}{\sqrt{2\pi} \ln(n)} \exp\left(-\frac{1}{2}\right) \)

This implies that \( \exists n_{1, \theta} \in \mathbb{N} \) with \( n_{1, \theta} \geq n_{0, \theta} \) such that \( 1 < x < 2 \sqrt{\ln(n)} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{1, \theta} \). Then, we have \( P(A_n) > \frac{1}{\sqrt{2\pi} \ln(n)} \frac{\exp(-\frac{x^2}{2})}{\exp\left(-\frac{1}{2}\right)} = \frac{1}{\sqrt{2\pi} \ln(n)} \frac{1}{\sqrt{2\pi} \ln(n)} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{1, \theta} \).

This yields that \( \sum_{n=n_{0, \theta}}^{\infty} P(A_n) = \infty \), by Problem 34.M of Bartle (1976). By Borel-Cantelli Lemma 14.6, \( P(F) := P(\bigcap_{n=n_{0, \theta}}^{\infty} A_n) = 1 = P\left( \left\{ \omega \in \Omega \mid \frac{u_{\theta n} - u_{\theta n+1}}{\sqrt{\theta n - \theta n-1}} > \frac{1 - \theta^{-1}}{\sqrt{\theta n - \theta n-1}} h(\theta^n) \right\} \right) \).

Thus, we have equivalently that \( \sqrt{1 - \theta^{-1}} \leq \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n} - u_{\theta n+1}|}{h(\theta^n)} \leq \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n}|}{h(\theta^n)} + \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n+1}|}{h(\theta^n)} \leq \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n}|}{h(\theta^n)} + \frac{1}{\sqrt{\theta^n}} \), \( \forall \omega \in F \), where the second inequality follows from Proposition 3.83; and the last inequality follows from the facts that \( \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n}|}{h(\theta^n)} \leq \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n}|}{h(\theta^n)} \leq 1 \) (since \( u_{\theta} \) is a standard Wiener process by Proposition 14.73) and \( \lim_{n \in \mathbb{N}} \frac{h(\theta^n-1)}{h(\theta^n)} = \frac{1}{\sqrt{\theta^n}} \). This leads to \( \limsup_{n \rightarrow \infty} \frac{u_{\theta n}}{h(\theta^n)} \geq \limsup_{n \in \mathbb{N}} \frac{|u_{\theta n}|}{h(\theta^n)} \geq \sqrt{1 - \theta^{-1}} - \frac{1}{\sqrt{\theta^n}} \), \( \forall \omega \in F \).
Lemma 14.79, \( \limsup_{t \to \infty} \frac{w_t}{h_t(t)} : \Omega \to \mathbb{R} \) is \( \mathcal{B} \)-measurable. Then, we have \( \limsup_{t \to \infty} \frac{w_t}{h_t(t)} \geq \sqrt{1 - \theta^{-1}} - \sqrt{\frac{\theta}{\pi}} \) a.e. in \( \Omega \). By the arbitrariness of \( \theta \), we have \( \limsup_{t \to \infty} \frac{w_t}{h_t(t)} \geq 1 \) a.e. in \( \Omega \).

This proves (i).

(ii) By Proposition 14.73, \((-w_t)_{t \in I} \) is also a standard Wiener process. Then, by (i), we have \( \limsup_{t \to \infty} \frac{w_t}{h_t(t)} = -\liminf_{t \to \infty} \frac{w_t}{h_t(t)} = 1 \) a.e. in \( \Omega \), where the first equality follows from Proposition 3.85; and the second equality follows from (i). This proves (ii).

(iii) Let \( \bar{h} : (0, \frac{1}{\sqrt{\pi}}) \to (0, \infty) \subset \mathbb{R} \) be defined by \( \bar{h}(t) = \sqrt{2t \ln(\frac{1}{t})} \), \( \forall t \in (0, \frac{1}{\sqrt{\pi}}) \subset \mathbb{R} \). Clearly, \( \lim_{t \to 0} \bar{h}(t) = 0 \) and \( \bar{h} \) is strictly increasing on \((0, t_0) \subset \mathbb{R} \), for some \( t_0 \in (0, \frac{1}{\sqrt{\pi}}) \). Fix any \( \theta, \delta \in \mathbb{Q} \) with \( 0 < \theta, \delta < 1 \). Let \( n_0, \theta \in \mathbb{N} \) be such that \( \theta^{n_0, \theta} < t_0 \), and \( \alpha_n := (1 + \delta)\theta^{-n}\bar{h}(\theta^n) > 0 \) and \( \beta_n := \frac{1}{2}\bar{h}(\theta^n) > 0, \forall n \in \mathbb{N} \) with \( n \geq n_0, \theta \). Then, \( \alpha_n \beta_n = \frac{1}{2}(1 + \delta)\theta^{-n}(\bar{h}(\theta^n))^2 = (1 + \delta)\ln(\theta^{-n}) = (1 + \delta)\ln(n \ln(\frac{1}{\theta})) = \ln((n \ln(\frac{1}{\theta}))^{1+\delta}) \) and \( \exp(-\alpha_n \beta_n) = \frac{c(\theta)}{\theta^n}, \forall n \in \mathbb{N} \) with \( n \geq n_0, \theta \), where \( c(\theta) := \left(\ln(\theta)\right)^{-1-\delta} \). Therefore, we have \( \sum_{n=n_0, \theta}^{\infty} \exp(-\alpha_n \beta_n) < +\infty \).

By Lemma 14.78, we have \( P\{ \omega \in \Omega \mid w(t, \omega) > \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n) \text{ for some } t \in I \} \leq \exp(-\alpha_n \beta_n) \). This implies that \( \sum_{n=n_0, \theta}^{\infty} P\{ \omega \in \Omega \mid w(t, \omega) > \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n) \text{ for some } t \in I \} < \infty \). By Borel-Cantelli Lemma 14.6, \( P\{ \bigcap_{n=n_0, \theta}^{\infty} \bigcup_{t=0}^{t_0} \{ \omega \in \Omega \mid w(t, \omega) > \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n) \text{ for some } t \in I \} = 0 = P\{ \omega \in \Omega \mid w(t, \omega) > \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n) \text{ for some } t \in I \text{ and for infinitely many } n \geq n_0, \theta \} =: P(\Omega \setminus \Omega_{\theta, \delta}) \). Then, \( P(\Omega_{\theta, \delta}) = 1 \). Fix any \( \omega \in \Omega_{\theta, \delta}, \exists n_0, \theta, \delta(\omega) \in \mathbb{N} \) with \( N_{\theta, \delta}(\omega) \geq n_0, \theta \) such that \( w(t, \omega) \leq \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n), \forall t \in I \) and \( \forall n > N_{\theta, \delta}(\omega) \). Fix any \( n > N_{\theta, \delta}(\omega) \). \( \forall t \in [\theta^n, \theta^{n-1}] \subset \mathbb{R} \), we have \( w(t, \omega) \leq \frac{\theta^n + 1}{\theta^n} \bar{h}(\theta^n) + \frac{1}{2}\bar{h}(\theta^n) \leq \frac{1}{2}\bar{h}(\theta^n) \). Hence, we have \( \sup_{0 < t < \theta^n} \frac{w(t, \omega)}{h_t(t)} \leq \frac{1}{2}\bar{h}(\theta^n) \). This implies that \( \limsup_{t \to 0^+} \frac{w_t}{h_t(t)} = \frac{1}{2}\bar{h}(\theta^n) \), \( \forall \omega \in \Omega_{\theta, \delta} \). By Lemma 14.79, \( \limsup_{t \to 0^+} \frac{w_t}{h_t(t)} : \Omega \to \mathbb{R} \) is \( \mathcal{B} \)-measurable. Then, we have \( \limsup_{t \to 0^+} \frac{w_t}{h_t(t)} \geq 1 \) a.e. in \( \Omega \). By the arbitrariness of \( \theta \) and \( \delta \), we have \( \limsup_{t \to 0^+} \frac{w_t}{h_t(t)} \geq 1 \) a.e. in \( \Omega \).

Next, we will show that \( \limsup_{t \to 0^+} \frac{w_t}{h_t(t)} \geq 1 \) a.e. in \( \Omega \). Let \( A_n := \{ \omega \in \Omega \mid w(\theta^n, \omega) - w(\theta^{n+1}, \omega) > \sqrt{1 - \theta^{-1}} \bar{h}(\theta^n) \} \in \mathcal{B}, \ forall n \in \mathbb{N} \) with \( n \geq n_0, \theta \). Clearly, \( A_n, n \geq n_0, \theta \), are independent events, since \( w_t \) is of independent increments. Then, \( P(A_n) = P(\{ \omega \in \Omega \mid \frac{w(t, \omega) - w(t+1, \omega)}{\sqrt{\theta^n - \theta^{n+1}}} > \sqrt{1 - \theta^{-1}} \bar{h}(\theta^n) \}) = \int_{\sqrt{\theta^n - \theta^{n+1}} \bar{h}(\theta^n)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2}) dy \leq \frac{1}{\sqrt{2\pi}} \sqrt{\theta^n} \bar{h}(\theta^n) \).
We have \( \forall n \in \mathbb{N} \) with \( n \geq n_{0, \theta} \), where \( x := \frac{1}{\sqrt{2n}} \bar{h}(\theta^n) = \sqrt{2 \ln(n \ln(1/\theta))} = \sqrt{2 \ln(n/2) + 2 \ln(\ln(1/\theta))} \). This implies that \( \exists n_{1, \theta} \in \mathbb{N} \) with \( n_{1, \theta} \geq n_{0, \theta} \) such that \( 1 < x < 2 \sqrt{\ln(n)} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{1, \theta} \). Then, we have \( P(A_n) > \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) > \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{4\ln(n \ln(1/\theta))}) = \frac{1}{\sqrt{2\pi \ln(n \ln(1/\theta))}} n^{\frac{1}{2}} \exp(-\frac{1}{4\ln(n \ln(1/\theta))}) \), \( \forall n \in \mathbb{N} \) with \( n \geq n_{1, \theta} \). This yields that \( \sum_{n=n_{0, \theta}}^\infty P(A_n) = \infty \), by Problem 34.M of Bartle (1976). By Borel-Cantelli Lemma 14.66, 

\[
P(F) := P\left( \bigcap_{n=n_{0, \theta}}^\infty A_n \right) = 1 = P\left( \left\{ \omega \in \Omega \ \bigg| \ \frac{w_{\theta n} - w_{\theta n+1}}{h(\theta^n)} > \frac{1}{\sqrt{\ln(1/\theta)}}, \theta^n \right\} \right).
\]

Thus, we have equivalently that \( \sqrt{1 - \theta} \leq \limsup_{n \in \mathbb{N}} \frac{\theta_{\theta n}}{h(\theta^n)} \leq \sqrt{1 - \theta} \), \( \forall \omega \in F \), where the second inequality follows from Proposition 3.83; and the last inequality follows from the facts that \( \limsup_{n \in \mathbb{N}} \frac{\theta_{\theta n}}{h(\theta^n)} \leq \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} \leq 1 \) (since \( -w_\theta \) is a standard Wiener process by Proposition 14.73) and \( \lim_{n \in \mathbb{N}} \frac{\theta_{\theta n}}{h(\theta^n)} = \sqrt{1 - \theta} \).

This leads to \( \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} \geq \sqrt{1 - \theta} - \sqrt{\theta} \), \( \forall \omega \in F \).

By Lemma 14.79, \( \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} : \Omega \to \mathbb{R} \) is \( \mathcal{B} \)-measurable. Then, we have \( \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} \geq \sqrt{1 - \theta} - \sqrt{\theta} \) a.e. in \( \Omega \). By the arbitrariness of \( \theta \), we have \( \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} \geq 1 \) a.e. in \( \Omega \).

This proves (iii).

(iv) By Proposition 14.73, \( (w_\theta)_{t \in \mathbb{I}} \) is also a standard Wiener process. Then, by (iii), we have \( \limsup_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} = \liminf_{t \to 0^+} \frac{w_{\theta(t)}}{h(t)} = 1 \) a.e. in \( \Omega \), where the first equality follows from Proposition 3.85; and the second equality follows from (iii). This proves (iv).

(v) By (i) and (ii), \( E \in \mathcal{B} \) and \( P(E) = 1 \). Then, \( (0, \infty) \times E \in \mathcal{B} \).

Let \( F \) denote the measurable rectangles in the product space \((0, \infty) \times \Omega\), and define the mapping \( \Psi : (0, \infty) \times \Omega \to (0, \infty) \times \Omega \) by \( \Psi(t, \omega) = (\frac{1}{\theta(t)}, \omega) \), \( \forall (t, \omega) \in (0, \infty) \times \Omega \). For any set \( G_1 \times G_2 \in \mathcal{F} \), we have the mapping \( \Psi_{\text{inv}}(G_1 \times G_2) \in \mathcal{F} \subseteq \mathcal{B} \). Consider the set of all subsets \( G \subseteq (0, \infty) \times \Omega \) such that \( \Psi_{\text{inv}}(G) \in \mathcal{B} \), it is easy to show that this set is a \( \sigma \)-algebra containing \( \mathcal{F} \). Then, the mapping \( \bar{w} : (0, \infty) \times \Omega \to \mathbb{R} \) defined by \( \bar{w}(t, \omega) = w(\frac{1}{\theta(t)}, \omega) \), \( \forall (t, \omega) \in (0, \infty) \times \Omega \) is \( \mathcal{B} \)-measurable. By Propositions 11.38, 11.39, and 11.41, we have \( \bar{w} : \Omega \to \mathbb{R} \) is \( \mathcal{B} \)-measurable. This proves (a) in Definition 14.69. Clearly, \( \bar{w}(0, \omega) = 0 \), \( \forall \omega \in \Omega \). This establishes (b) in Definition 14.69. \( \forall \omega \in \Omega \), if \( \omega \in E \), then \( \bar{w}(\cdot, \omega) : \mathbb{I} \to \mathbb{R} \) is continuous at any \( t \in \mathbb{I} \); if \( \omega \notin \Omega \), then \( \bar{w}(\cdot, \omega) : \mathbb{I} \to \mathbb{R} \) is the identically zero function, and is continuous. Hence, (c) of Definition 14.69 is established. Fix any \( 0 = s_1 < t_1 = s_2 < t_2 = \cdots < s_n < t_n < \infty \). Then, we have \( \bar{w}_{t_1} - \bar{w}_{s_1} = t_1 w_{1/t_1}, \cdots, \bar{w}_{t_n} - \bar{w}_{s_n} = t_n w_{1/t_n} - s_n w_{1/s_n} \). Since \((w_t)_{t \in \mathbb{I}}\) is a standard Wiener process, we have \( E(t_1 w_{1/t_1}) = 0 \), \( E(t_1^2 w_{1/t_1}^2) = t_1 \), \( E(t_1 w_{1/t_1} - s_i w_{1/s_i}) = 0 \), \( E((t_1 w_{1/t_1} - s_i w_{1/s_i})^2) = t_1 + s_i - 2t_i s_i E(w_{1/t_1} w_{1/s_i}) = t_1 + s_i - 2t_i s_i \).
2t_is \frac{1}{s_i} = t_i - s_i, \; i = 2, \ldots, n. Furthermore, \( E(t_i w_1/t_t (t_i w_1/t_i - s_i w_1/s_i)) = t_i t_i E(w_1/t_t w_1/t_t) - t_i s_i E(w_1/t_t w_1/s_t) = t_i - t_i = 0, \; \; i = 2, \ldots, n; \) and 
\( E((t_i w_1/t_t - s_i w_1/s_t) (t_j w_1/t_t - s_j w_1/s_t)) = t_i t_j \frac{1}{t_j} - s_i t_j \frac{1}{s_j} - t_j - s_j = t_i - s_i - t_j + s_j = 0, \; \forall 2 \leq i < j \leq n. \) Hence, we have 
\( \{\tilde{w}_1 - \tilde{w}_s, \ldots, \tilde{w}_n - \tilde{w}_s\} \) 
admits mean \( \theta_n \) and covariance matrix \( K \) be block diagonal \( (t_i - s_i, t_i - s_i) \in S_+ \mathbb{R}_n. \) By Proposition 14.65 and the fact that \( (\mu_t)_{t \in I} \) is a standard Wiener process, 
\( (\tilde{w}_1 - \tilde{w}_s, \ldots, \tilde{w}_n - \tilde{w}_s) \sim N(\theta_n, K). \) By Proposition 14.21, 
\( \tilde{w}_t - \tilde{w}_x, \ldots, \tilde{w}_n - \tilde{w}_s \) are independent. This establishes (d) of Definition 14.69. Hence, the process \( (\tilde{w}_t)_{t \in I} \) defined by 
\( \tilde{w}_t(\omega) = \tilde{w}(t, \omega), \; \forall \omega \in \Omega, \; \forall t \in I, \) is a standard Wiener process.

This completes the proof of the theorem. \( \Box \)

**Example 14.81** Let \( \underline{I} := (|0, \infty|, |\cdot|), \underline{B}, \mu \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \) with the partial ordering \( \leq_\circ \) be a normed linear space, 
\( \Omega := (\Omega, \mathbb{B}, (\mathcal{B}_t)_{t \in I}, \mathbb{P}) \) be a filtered probability measure space, \( \Omega \times \Omega = ((0, \infty), \times, \mathbb{B}, (\nu_t)_{t \in I}) \) be the \( \sigma \)-finite product measure space, 
\( (X_t)_{t \in I} \) be an \( \mathbb{R} \)-valued adapted stochastic process on \( \Omega \) with continuous sample path 
\( \forall \omega \in \Omega, \; F \subseteq \mathbb{Y} \) be a closed set, and \( T : \Omega \rightarrow \mathbb{R} \cup \{ \infty \} \) be defined by 
\( T(\omega) := \inf \{ t \in I \mid X_t(\omega) \in F \}, \; \forall \omega \in \Omega. \) Then, \( T \) is a stopping time. 

This is because, \( \forall t \in I, \{ \omega \in \Omega \mid T(\omega) \leq t \} = \bigcup_{s \in [0, t]} \{ \omega \in \Omega \mid X_s(\omega) \in F \} = \bigcup_{s \in [0, t]} \{ \omega \in \Omega \mid \inf_{s \in [0, t]} \text{dist}(X_s(\omega), F) = 0 \} = \{ \omega \in \Omega \mid \inf_{s \in [0, t]} \text{dist}(X_s(\omega), F) = 0 \} \in \mathcal{B}_t, \) where the first 
equality follows from the fact that \( (X_t)_{t \in I} \) has continuous sample path, 
\( \forall \omega \in \Omega, \) which further implies that \( X_{T(\omega)}(\omega) \in F \) if \( T(\omega) < \infty; \) the second 
equality follows from the fact that \( (X_t)_{t \in I} \) is adapted to the filtration 
\( (\mathcal{B}_t)_{t \in I}; \) the third equality follows from the fact that \( (X_t)_{t \in I} \) has continuous sample path, 
\( \forall \omega \in \Omega; \) the fourth equality follows from the fact that 
\( \text{dist}(\cdot, F) : \mathbb{Y} \rightarrow [0, \infty] \subseteq \mathbb{R}_e \) is a continuous function; and the last 
element follows from the fact that \( \text{dist}(X_s(\cdot), F) : \Omega \rightarrow [0, \infty] \subseteq \mathbb{R}_e \) is 
\( \mathcal{B}_t \)-measurable by Proposition 11.38, \( \forall s \in [0, t] \cap \mathbb{Q}, \) and the set \( [0, t] \cap \mathbb{Q} \) is countable. \( \Box \)

**Proposition 14.82** Let \( \underline{I} := (|0, \infty|, |\cdot|), \underline{B}, \mu \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}, \) \( \Omega := (\Omega, \mathbb{B}, (\mathcal{B}_t)_{t \in I}, \mathbb{P}) \) be a filtered probability measure space, and \( \tau, \kappa \) are stopping times with values in \( I \cup \{ \infty \}. \) Then, \( \tau \lor \kappa, \) 
\( \tau \land \kappa \) and \( \tau + \kappa \) are stopping times.

**Proof** By Definition 14.27, \( \tau : \Omega \rightarrow \underline{I} \cup \{ \infty \} = [0, \infty] \subseteq \mathbb{R}_e \) is 
\( \mathcal{B} \)-measurable and \( \{ \omega \in \Omega \mid \tau(\omega) \leq t \} \in \mathcal{B}_t, \; \forall t \in I. \) The same 
thing holds for \( \kappa. \) By Propositions 11.36 and 11.40, \( \tau \lor \kappa, \tau \land \kappa \) and \( \tau + \kappa \) are functions of \( \Omega \) to \( [0, \infty] \subseteq \mathbb{R}_e \) and are \( \mathcal{B} \)-measurable. Fix any \( t \in I, \) we have 
\( \{ \omega \in \Omega \mid \tau \lor \kappa \leq t \} = \{ \omega \in \Omega \mid \tau \leq t \} \cap \{ \omega \in \Omega \mid \kappa \leq t \} \in \mathcal{B}_t. \) This implies 
that \( \tau \lor \kappa \) is a stopping time. In addition, \( \{ \omega \in \Omega \mid \tau \land \kappa \leq t \} = \{ \omega \in \Omega \mid \tau \leq t \} \cup \{ \omega \in \Omega \mid \kappa \leq t \} \in \mathcal{B}_t. \) This yields that 
\( \tau \land \kappa \) is a stopping time. Furthermore, 
\( \{ \omega \in \Omega \mid \tau + \kappa \leq t \} = \Omega \setminus \{ \omega \in \Omega \mid \tau + \kappa > t \} = \Omega \setminus \bigcup_{r \in \mathbb{Q}, r > 0} \{ \omega \in \Omega \mid \tau > r \} \cap \{ \omega \in \Omega \mid \kappa > t - r \} \cup \{ \omega \in \Omega \mid \tau + \kappa > t \} = \Omega \setminus \bigcup_{r \in \mathbb{Q}, r > 0} \{ \omega \in \Omega \mid \tau > r \} \cap \{ \omega \in \Omega \mid \kappa > t - r \} \cup \{ \omega \in \Omega \mid \tau + \kappa > t \} \).
Let \( \mathbb{I} := ((0, \infty), |\cdot|), \mathcal{S}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \) with the partial ordering \( \leq \), \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_t)_{t \in I}, P) \) be a filtered probability measure space, \( \mathbb{I} \times \Omega = ([0, \infty) \times \Omega, \mathcal{S}, \nu) \) be the \( \sigma \)-finite product measure space, \( w : \mathbb{I} \times \Omega \to \mathbb{R} \) be the standard Wiener process on \( \Omega \) (as defined in Example 14.70), and the stochastic process \((w_t)_{t \in I}\), where \( w_t(\omega) = w(\omega)(t), \forall \omega \in \Omega, \forall t \in I \), is adapted to the filtration \((\mathcal{B}_t)_{t \in I}\).

Expand \( \exp(ax - \frac{a^2}{2}t) \) as an analytic function of \( a \) to obtain

\[
\exp(ax - \frac{a^2}{2}t) = \sum_{n=0}^{\infty} H_n(x, t) \frac{a^n}{n!}, \quad \forall a \in \mathbb{R}
\]  

(14.20)

where \( H_n(x, t) : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \) is \( C_{\infty} \), \( \forall n \in \mathbb{N} \). Clearly, \( H_n(x, t) \) is a \( n \)th order polynomial in \( x \) and a \( \left\lfloor \frac{n}{2} \right\rfloor \)th order polynomial in \( t \). Note that \( \exp(|\alpha||x| + \frac{a^2}{2}|t|) \geq \sum_{n=0}^{\infty} |H_n(x, t)| \frac{|\alpha|^n}{n!}, \forall (\alpha, x, t) \in \mathbb{R}^3 \). Then, \((14.20)\) holds \( \forall (\alpha, x, t) \in \mathbb{R}^3 \) with the right-hand-side converges absolutely. The sequence \( (H_n(x, t))_{n=0}^{\infty} \) is called the Hermite polynomials. The leading Hermite polynomials are \( H_0(x, t) = 1, H_1(x, t) = x, H_2(x, t) = x^2 - t, H_3(x, t) = x^3 - 3xt, \ldots, \forall (x, t) \in \mathbb{R}^2 \).

Now, consider the adapted stochastic process \((M_t)_{t \in I}\), defined by \( M_t(\omega) = \exp(\alpha w_t(\omega) - \frac{\alpha^2}{2}t) = \sum_{n=0}^{\infty} H_n(w_t(\omega), t) \frac{\alpha^n}{n!}, \forall \omega \in \Omega, \forall t \in \mathbb{I} \). By Proposition 14.77, \((M_t)_{t \in I}\) is a Martingale. Then, we have \( \mathbb{E}(M_t|\mathcal{B}_s) = M_s \) a.e. in \( \Omega_s, \forall s, t \in \mathbb{I} \) with \( s \leq t \), where \( \Omega_s := (\Omega, \mathcal{B}_s, P_s := P|_{\mathcal{B}_s}) \) is a probability measure space. Since \( \exp(\alpha|w_t(\omega)| + \frac{\alpha^2}{2}|t|) \geq \sum_{n=0}^{\infty} |H_n(w_t(\omega), t)| \frac{|\alpha|^n}{n!} \), \( \forall (\alpha, t, \omega) \in \mathbb{R} \times \mathbb{I} \times \Omega \), then

\[
\mathbb{E}(\sum_{n=0}^{\infty} |H_n(w_t(\omega), t)| \frac{|\alpha|^n}{n!} \leq \mathbb{E}(M_t|\mathcal{B}_s) = \sum_{n=0}^{\infty} \mathbb{E}(H_n(w_t, t)|\mathcal{B}_s) \frac{\alpha^n}{n!} = M_s = \sum_{n=0}^{\infty} H_n(w_s, t) \frac{\alpha^n}{n!} \), a.e. in \( \Omega_s, \forall s, t \in \mathbb{I} \) with \( s \leq t \), \( \forall \alpha \in \mathbb{R} \), where the first equality follows from (g) of Proposition 14.11; the second equality follows from \((M_t)_{t \in I}\) being a Martingale; and the last equality follows from the definition of \((M_t)_{t \in I}\). By the arbitrariness of \( \alpha \), we have \( \mathbb{E}(H_n(w_t, t)|\mathcal{B}_s) = H_n(w_s, t) \), a.e. in \( \Omega_s, \forall t, s \in \mathbb{I} \) with \( s \leq t \), \( \forall n \in \mathbb{Z}_+ \).

Hence, \((H_n(w_t, t))_{t \in I}\) is a Martingale, \( \forall n \in \mathbb{Z}_+ \).

### 14.12 Stochastic Integral

**Definition 14.84** Let \( \mathbb{I} := ((0, \infty), |\cdot|), \mathcal{S}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R} \), \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_t)_{t \in I}, P) \) be a filtered probability measure space, \( w : \Omega \to C(\mathbb{I}, \mathbb{R}) \) be the standard Wiener process on \( \Omega \) (as defined in Example 14.70), \((\mathcal{B}_t)_{t \in I}\) be the natural filtration of the stochastic process \((w_t)_{t \in I}\), where \( w_t(\omega) = w(\omega)(t), \forall \omega \in \Omega, \forall t \in I \), and \( \phi : [a, b] \to \mathbb{R} \)
be of bounded variation, where \( a, b \in \mathbb{I} \) with \( a \leq b \). Define \( I : \Omega \to \mathbb{R} \) by
\[
I(\omega) := \phi(b)w(\omega)(b) - \phi(a)w(\omega)(a) - \int_{r_{a,b}} w(\omega)(s) \, d\phi(s), \quad \forall \omega \in \Omega,
\]
where the integral is the Lebesgue-Stieltjes integral as introduced in Definition 12.57.

We will denote \( I \) by \( \int_{r_{a,b}} \phi(s) \, dw_s \), which is called the (Wiener) integral of \( \phi \) with respect to \( w_t \) on the interval \( r_{a,b} \).

Since \( \phi : [a, b] \to \mathbb{R} \) is of bounded variation, then by Theorem 12.50, there exists a finite \( \mathbb{R} \)-valued measure \( \mu \) on the measurable space \((r_{a,b}, \mathcal{B}(r_{a,b}))\) such that \( \phi \) is a cumulative distribution function of \( \mu \). By Jordan Decomposition Theorem 11.162, there exists finite measures \( \mu_+ \) and \( \mu_- \) on \((r_{a,b}, \mathcal{B}(r_{a,b}))\), such that \( \mu = \mu_+ - \mu_- \) and \( \mathcal{P} \circ \mu = \mu_+ + \mu_- \). Let
\[
\phi_+ : [a, b] \to \mathbb{R} \quad \text{and} \quad \phi_- : [a, b] \to \mathbb{R}
\]
be the cumulative distribution function for \( \mu_+ \) and \( \mu_- \), respectively, according to Proposition 12.51. Clearly, \( \phi_+ \) and \( \phi_- \) are monotonically nondecreasing on \([a, b]\), and \( \phi = \phi_+ - \phi_- \).

By Theorems 29.7 and 30.2 of Bartle (1976), we have \( I_+ : \Omega \to \mathbb{R} \) and \( I_- : \Omega \to \mathbb{R} \), defined by
\[
I_+(\omega) := \int_{[a,b]} \phi_+(s) \, dw(\omega)(s) \quad \text{and} \quad I_-(\omega) := \int_{[a,b]} \phi_-(s) \, dw(\omega)(s), \quad \forall \omega \in \Omega,
\]
respectively, both exist, where the integrals in the definition of \( I_+ \) and \( I_- \) are Riemann-Stieltjes integrals, and we have
\[
I_+(\omega) = \phi_+(b)w(\omega)(b) - \phi_+(a)w(\omega)(a) - \int_{r_{a,b}} w(\omega)(s) \, d\phi_+(s) = \int_{r_{a,b}} \phi_+(s) \, dw(\omega)(s)
\]
and \( I_-(\omega) = \phi_-(b)w(\omega)(b) - \phi_-(a)w(\omega)(a) - \int_{r_{a,b}} w(\omega)(s) \, d\phi_-(s) = \int_{r_{a,b}} \phi_-(s) \, dw(\omega)(s) \), where the integrals in right-hand-sides of the preceding two equalities are Wiener integrals. By Theorem 29.5 of Bartle (1976), we have
\[
\int_{r_{a,b}} \phi(s) \, dw(\omega)(s) =: I(\omega) = I_+(\omega) - I_-(\omega) = \int_{[a,b]} (\phi_+(s) - \phi_-(s)) \, dw(\omega)(s) = \int_{[a,b]} \phi(s) \, dw(\omega)(s), \quad \forall \omega \in \Omega,
\]
where the integral in the left-hand-side of the above equation is the Wiener integral of Definition 14.84; the first equality follows from the definition of \( I(\omega) \) and the preceding discussion; the second equality follows from Theorem 29.5 of Bartle (1976); the third equality follows from the preceding discussion; and the integral in the right-hand-side of the equation is Riemann-Stieltjes integral. Thus, we can conclude that the Wiener integral on semi-open interval \( r_{a,b} \) is simply the Riemann-Stieltjes integral on the closed interval \([a, b]\).

Note that the integrand \( \phi \) must be of bounded variation according to Definition 12.41, which must be necessarily right continuous. Hence,\[
\int_{r_{a,b}} \phi(s) \, dw(\omega)(s) = \lim_{h \to 0^+} \int_{r_{a+h,b]} \phi(s) \, dw(\omega)(s) = \int_{[a,b]} \phi(s) \, dw(\omega)(s), \quad \forall \omega \in \Omega,
\]
where the first integral is the Wiener integral, the second integral is the Riemann-Stieltjes integral, and the last integral is the improper Riemann-Stieltjes integral.

**Theorem 14.85 (Itô’s Formula)**
Chapter 15

Numerical Methods

15.1 Newton’s Method

Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and $F : \mathcal{X} \to \mathcal{Y}$. To solve the equation $F(x) = \vartheta$, we may use Newton’s method. Assume that $F$ is twice Fréchet differentiable, we set $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1} F(x_n)$, $\forall n \in \mathbb{N}$. Then, when $x_1$ is sufficiently close to a solution, the sequence $(x_n)_{n=1}^{\infty}$ converges to a solution $x_{opt}$ with $F(x_{opt}) = \vartheta$. This result is formalized in the following proposition.

**Proposition 15.1** Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces over $\mathbb{K}$, $D_1 \subseteq \mathcal{X}$, $F : D_1 \to \mathcal{Y}$, and $x_1 \in \mathcal{X}$. Assume that $\exists \beta_1, \eta_1, h_1, K \in [0, \infty) \subset \mathbb{R}$ such that

(i) $h_1 := \beta_1 \eta_1 K \leq \frac{1}{2}$;

(ii) $\text{dom} \left( F^{(2)} \right) \supseteq B_{\mathcal{X}}(x_1, 2h_1 \eta_1) =: D \subseteq B_{\mathcal{X}}(x_1, 2\eta_1)$ and, $\forall x \in D$, $\|F^{(2)}(x)\| \leq K$;

(iii) $F^{(1)}(x_1)$ is bijective, $\|(F^{(1)}(x_1))^{-1}\| \leq \beta_1$, and $\|(F^{(1)}(x_1))^{-1} \cdot F(x_1)\| \leq \eta_1$;

Then, $\exists (x_n)_{n=1}^{\infty} \subseteq D$, defined by $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1} F(x_n)$, $\forall n \in \mathbb{N}$, that converges to $x_{opt} \in D$ with $F(x_{opt}) = \vartheta$.

**Proof** We will use mathematical induction to prove the following claim.

**Claim 15.1.1** $\forall n \in \mathbb{N}$, we have

(a) $F^{(1)}(x_n)$ is bijective and $\|(F^{(1)}(x_n))^{-1}\| \leq \beta_n := \frac{\beta_1 h_1}{1 - h_1 \eta_1} \in [0, \infty) \subset \mathbb{R}$;

(b) $\|(F^{(1)}(x_n))^{-1} F(x_n)\| \leq \eta_n := \frac{h_1 \eta_1 h_1}{2(1 - h_1 \eta_1)} \in [0, \infty) \subset \mathbb{R}$;
(c) \( x_{n+1} \in \mathcal{B}_y \left( x_1, \frac{2-2h_k}{2-3h_k} \left( 1 - \left( \frac{h_k}{2(1-h_k)} \right)^{n_1} \right) \right) ; \)

(d) \( 0 \leq h_n := \beta_n \eta_n K \leq h_{n-1} \leq \frac{1}{2}. \)

**Proof of claim:**  
1° Clearly, (a) – (d) are satisfied for \( n = 1. \)
2° Assume that (a) – (d) are satisfied for \( n \leq k \in \mathbb{N}. \)
3° Consider the case of \( n = k + 1. \) By Mean Value Theorem 9.23, \( \exists t_0 \in (0,1) \subset \mathbb{R} \) such that \( \| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \| \leq \| F^{(2)}(t_0 x_{k+1} + (1 - t_0)x_k) \| \| x_{k+1} - x_k \| \leq K \eta_k. \) Then, \( \| (F^{(1)}(x_k))^{-1} \| \| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \| \leq \beta k \eta K = h_k < 1. \) By Proposition 9.55, \( F^{(1)}(x_{k+1}) \) is bijective and

\[
\left\| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1} \right\| \leq \frac{\| (F^{(1)}(x_k))^{-1}\|^2 \| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \|}{1 - \| (F^{(1)}(x_k))^{-1}\| \| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \|} \leq \frac{\beta k^2 \eta K}{1 - \beta k \eta K} = \frac{h_k \beta_k}{1 - h_k}
\]

This leads to \( \| (F^{(1)}(x_{k+1}))^{-1}\| \leq \| (F^{(1)}(x_k))^{-1}\| + \| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1}\| \leq \beta_k + \frac{h_k \beta_k}{1 - h_k} = \beta_k / (1 - h_k) = \beta_{k+1}. \) Hence, (a) holds.

Note that \( F(x_{k+1}) = F(x_{k+1}) - F(x_k) - F^{(1)}(x_k)(x_{k+1} - x_k). \) By Taylor’s Theorem 9.48, \( \exists t_1 \in (0,1) \subset \mathbb{R} \) such that \( \| F'(x_{k+1}) \| \leq \frac{1}{2} \| F^{(2)}(t_1 x_{k+1} + (1 - t_1)x_k) \| \| x_{k+1} - x_k \| \leq K \eta_k^2 / 2. \) Then, \( \| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \| \leq \beta_k K \eta_k^2 / (2 - 2h_k) = \frac{h_k \eta_k}{2(1-h_k)} = \eta_{k+1}. \) Hence, (b) holds.

Note that \( \| x_{k+2} - x_{k+1} \| = \| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \| \leq \eta_{k+1} = \frac{h_k}{2(1-h_k)} \eta_k \leq \left( \frac{h_k}{2(1-h_k)} \right)^k \eta_1. \) Then, \( \| x_{k+2} - x_1 \| \leq \| x_{k+2} - x_{k+1} \| + \| x_{k+1} - x_1 \| \leq \left( \frac{h_k}{2(1-h_k)} \right)^k \eta_1 + \frac{2-2h_k}{2-3h_k} \left( 1 - \left( \frac{h_k}{2(1-h_k)} \right)^k \right) \eta_1 = \frac{2-2h_k}{2-3h_k} \left( 1 - \left( \frac{h_k}{2(1-h_k)} \right)^{k+1} \right) \eta_1. \) Hence, (c) holds.

Note that \( h_{k+1} = \beta_k \eta_{k+1} K = \frac{h_k^2}{2(1-h_k)}. \) Since \( 0 \leq h_k \leq 1/2, \) then \( 0 \leq \frac{h_k}{2(1-h_k)} \leq 1. \) This implies that \( 0 \leq h_{k+1} \leq h_k \leq 1/2. \) Hence, (d) holds.

This completes the induction process and the proof of the claim. □

Then, the sequence \( \{x_n\}_{n=1}^\infty \subseteq D \) is well-defined. \( \forall n \in \mathbb{N}, \| x_{n+2} - x_{n+1} \| = \| (F^{(1)}(x_{n+1}))^{-1} F(x_{n+1}) \| \leq \eta_{n+1} \leq \left( \frac{h_k}{2(1-h_k)} \right)^n \eta_1. \) Then, \( \{x_n\}_{n=1}^\infty \) is a Cauchy sequence since \( 0 \leq \frac{h_k}{2(1-h_k)} \leq 1/2 < 1. \) It then converges to \( x_{opt} \in D \) since \( D \) is closed and \( X \) is complete. Note that \( F(x_n) = F^{(1)}(x_n)(x_n - x_{n+1}), \forall n \in \mathbb{N}. \) By Proposition 9.7, we have
15.1. NEWTON’S METHOD

$F$ and $F^{(1)}$ are continuous at $x_{opt}$. Then, $F(x_{opt}) = \lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} F^{(1)}(x_n)(x_n - x_{n+1}) = F^{(1)}(x_{opt})(x_{opt} - x_{opt}) = 0$, where the first equality follows from Proposition 3.66; and the third equality follows from Propositions 3.66, 3.67, 7.23, and 7.65. This completes the proof of the proposition.

It is numerically expensive to use Newton’s Method due to the necessity of inverting $F^{(1)}(x_n)$ at each step of the iteration. But the benefit of this method is the quadratic convergence of $(x_n)_{n=1}^\infty$ to $x_{opt}$, which is summarized in the following proposition.

**Proposition 15.2** Let $X$ and $Y$ be Banach spaces over $K$, $D_1 \subseteq X$, $F : D_1 \to Y$, and $x_1 \in X$. Assume that $\exists \beta_1, \eta_1, K, M_1, M_2, M_4 \in [0, \infty) \subset K$ such that

1. $h_1 := \beta_1 \eta_1 K \leq \frac{1}{2}$;
2. $\text{dom}(F^{(3)}) \supseteq D_2 \supseteq \overline{B}_X \left( x_1, \frac{2 - 2h_1}{2 - 2h_1} \eta_1 \right)$ : $D_2 \subseteq D_1$ is an open set in $X$, $\forall x \in D_2$, $F^{(1)}(x)$ is bijective, and, $\forall x \in D_1$,

$$\|F^{(2)}(x)\| \leq K, \quad \|(F^{(1)}(x))^{-1}\| \leq M_1, \quad \|F^{(3)}(x)\| \leq M_2,$$

3. $\|F(x)\| \leq M_4$;

Then, $\exists (x_n)_{n=1}^\infty \subseteq D_1$, defined by $x_{n+1} = x_n - (F^{(1)}(x_n))^{-1}F(x_n)$, $\forall n \in \mathbb{N}$, that converges to $x_{opt} \in D$ with $F(x_{opt}) = 0$. Furthermore, $\forall n \in \mathbb{N}$, $\|x_{n+1} - x_{opt}\| \leq c \|x_n - x_{opt}\|^2$, where $c := (M_1 K + 2M_1^2 K^2 M_4 + M_1^2 M_2 M_4)/2$.

**Proof** By Proposition 15.1, $(x_n)_{n=1}^\infty \subseteq D$ is well defined and converges to $x_{opt} \in D$ with $F(x_{opt}) = 0$. Define $T : D_2 \to X$ by $T(x) = x - (F^{(1)}(x))^{-1}F(x)$, $\forall x \in D_2$. Then, by Propositions 9.34, 9.44, 9.45, 9.41, and 9.55, $T$ is twice Fréchet differentiable. This leads to $T^{(1)}(x)(h_1) = h_1 - (F^{(1)}(x))^{-1} F^{(1)}(x)(h_1) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x)) = (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x))$, $\forall x \in D_2$, $\forall h_1 \in X$. Then, $T^{(1)}(x_{opt})(h_1) = 0$, $\forall h_1 \in X$, since $F(x_{opt}) = 0$. This implies that $T^{(1)}(x_{opt}) = \theta_{B(\mathcal{X})}$. The second order derivative of $T$ is given by, $\forall x \in D_2$, $\forall h_1, h_2 \in X$,

$$T^{(2)}(x)(h_2)(h_1) = -(F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x))$$

$$+ (F^{(1)}(x))^{-1} F^{(3)}(x)(h_2)(h_1)((F^{(1)}(x))^{-1} F(x)) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)(- (F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F(x))$$

$$+ (F^{(1)}(x))^{-1} F^{(3)}(x)(h_2)(h_1)((F^{(1)}(x))^{-1} F(x))$$

$$= -(F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x))$$

$$+ (F^{(1)}(x))^{-1} F^{(3)}(x)(h_2)(h_1)((F^{(1)}(x))^{-1} F(x))$$
\[-(F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F(x))) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)(h_2)\]

Then, \(\forall x \in D, \|T^{(2)}(x)\| = \sup_{h_1, h_2 \in X, \|h_1\| \leq 1, \|h_2\| \leq 1} \|T^{(2)}(x)(h_2)(h_1)\| \leq 2 M_2^3 K^2 M_4 + M_2^2 M_2 M_4 + M_1 K = 2c\). \(\forall n \in \mathbb{N}\), by Taylor’s Theorem 9.48, \(\exists t_0 \in (0, 1) \subset \mathbb{R}\) such that \(\|x_{n+1} - x_{opt}\| = \|T(x_n) - T(x_{opt})\| = \|T(x_n) - T(x_{opt}) - T^{(1)}(x_{opt})(x_n - x_{opt})\| \leq \frac{1}{2} \|T^{(2)}(t_0 x_n + (1 - t_0) x_{opt})\| \|x_n - x_{opt}\|^2 \leq c \|x_n - x_{opt}\|^2\). This completes the proof of the proposition. \(\Box\)
Appendix A

Elements in Calculus

A.1 Some Formulas

Proposition A.1 \( \sum_{i=0}^{n} \prod_{j=1}^{m} (i+j) = \frac{1}{m+1} \frac{(n+m+1)!}{n!}, \forall n, m \in \mathbb{Z}^+. \)

Proof We will use mathematical induction on \( n \in \mathbb{Z}^+ \) to prove the result. 1\(^{\circ}\) \( n = 0. \forall m \in \mathbb{Z}^+, \) we have LHS = m! = RHS. This case is proved.

2\(^{\circ}\) Assume the result holds for \( n = k \in \mathbb{Z}^+. \)

3\(^{\circ}\) Consider the case when \( n = k + 1 \in \mathbb{N}. \forall m \in \mathbb{Z}^+, \)

\[
\text{LHS} = \sum_{i=0}^{k+1} \prod_{j=1}^{m} (i+j) = \sum_{i=0}^{k} \prod_{j=1}^{m} (i+j) + \prod_{j=1}^{m} (k+1+j) \\
= \frac{1}{m+1} \frac{(k+m+1)!}{k!} + \frac{(k+m+1)!}{(k+1)!} = \frac{1}{m+1} \frac{(k+m+2)!}{(k+1)!}
\]

where the third equality follow from the inductive assumption. This completes the induction process.

Hence, the result holds. \( \square \)

A.2 Convergence of Infinite Sequences

Proposition A.2 (Raabe’s Test) Let \( (x_n)_{n=1}^{\infty} \subseteq (0, \infty) \subset \mathbb{R}. \) Then, the following statements hold.

(i) If there exists \( a \in (1, \infty) \subset \mathbb{R} \) and \( n_0 \in \mathbb{N} \) such that \( \frac{x_{n+1}}{x_n} \leq 1 - \frac{a}{n}, \forall n \geq n_0, \) then \( \sum_{n=1}^{\infty} x_n \in (0, \infty) \subset \mathbb{R}. \)

(ii) If there exists \( a \in (\infty, 1] \subset \mathbb{R} \) and \( n_0 \in \mathbb{N} \) such that \( \frac{x_{n+1}}{x_n} \geq 1 - \frac{a}{n}, \forall n \geq n_0, \) then \( \sum_{n=1}^{\infty} x_n = \infty. \)
**Proof (Bartle)**

(i) Under the assumption, \( \forall n \geq n_0 \),

\[
x x_{n+1} \leq (n - 1)x_n - (a - 1)x_n
\]

This implies that

\[
(n - 1)x_n - nx_{n+1} \geq (a - 1)x_n > 0, \quad \forall n \geq n_0
\]

Summing the above for \( n = n_0, \ldots, m, \forall m \geq n_0 \), we have

\[
0 < (a - 1) \sum_{n=n_0}^{m} x_n \leq (n_0 - 1)x_{n_0} - mx_{m+1} < (n_0 - 1)x_{n_0} \in \mathbb{R}
\]

Hence, we have \( 0 < \sum_{n=n_0}^{m} x_n < \frac{(n_0 - 1)x_{n_0}}{a - 1} \in (0, \infty) \subset \mathbb{R}, \forall m \geq n_0 \).

Hence, \( \sum_{n=1}^{\infty} x_n \in (0, \sum_{n=1}^{n_0-1} x_n + \frac{(n_0 - 1)x_{n_0}}{a - 1}] \subset \mathbb{R} \).

(ii) Under the assumption, we have

\[
x x_{n+1} \geq (n - a)x_n \geq (n - 1)x_n, \forall n \geq n_0
\]

Then, \( nx_{n+1} \geq n_0 x_{n_0+1} =: c \in (0, \infty) \subset \mathbb{R}, \forall n \geq n_0 \). This implies that \( x_{n+1} \geq c/n, \forall n \geq n_0 \), and hence \( \sum_{n=1}^{\infty} x_n = \infty \).

This completes the proof of the proposition. \( \Box \)
List of Figures

6.1 The sum of two sets. ........................................ 143
6.2 Convex and nonconvex sets. ................................. 146
6.3 Convex hulls. .................................................. 146
6.4 Cones. ........................................................... 147

7.1 Sequence for Example 7.28. ................................. 163

8.1 Fenchel duality. ................................................ 240

13.1 Projection onto a subspace. ............................... 671
13.2 Dual projection problems. .................................. 689
13.3 Reformulation of the Projection Theorem. ............... 690
13.4 Projection to a convex set. ................................ 691

14.1 The first eight Haar functions. ............................ 783
Bibliography


URL: http://www.loc.gov/catdir/description/dover031/95051359.html


**Notations**

- $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$: the sets of natural number, integers, and rational numbers, respectively.
- $\mathbb{R}$ and $\mathbb{C}$: the sets of real numbers and complex numbers, respectively.
- $\mathbb{K}$: either $\mathbb{R}$ or $\mathbb{C}$.
- $\mathbb{Z}_+$, $\mathbb{Z}_-$: $\mathbb{N} \cup \{0\}$, $\mathbb{Z} \setminus \mathbb{N}$, respectively.
- $\mathbb{R}_+ \setminus \mathbb{Z}$, $\mathbb{R}_- \setminus \mathbb{Z}$, $\mathbb{C}_+ \setminus \mathbb{Z}$, $\mathbb{C}_-$: $(0, \infty) \subset \mathbb{R}$, $(-\infty, 0) \subset \mathbb{R}$, the open right half of the complex plane, the open left half of the complex plane, respectively.
- $\mathbb{R}_+ \cup \mathbb{Z}$, $\mathbb{R}_- \cup \mathbb{Z}$, $\mathbb{C}_+ \cup \mathbb{Z}$, $\mathbb{C}_-$: $[0, \infty) \subset \mathbb{R}$, $(-\infty, 0] \subset \mathbb{R}$, the closed right half of the complex plane, the closed left half of the complex plane, respectively.

- $\in$: belong to.
- $\notin$: not belong to.
- $\subseteq$: contained in.
- $\supseteq$: contains.
- $\subset$: strict subset of.
- $\supset$: strict super set of.
- $\forall$: for all.
- $\exists$: exists.
- $\exists!$: exists a unique.
- $\therefore$: because.
- $\therefore$: therefore.
- $\ni$: such that.
- $(x_n)_{n=1}^{\infty}$: the sequence $x_1, x_2, \ldots$.
- $(x_\alpha)_{\alpha \in \Lambda}$: the ordered collection.
- $\text{id}_A$: the identity map on a set $A$.
- $|\lambda|$: the absolute value of a real or complex number $\lambda$.
- $\overline{\lambda}$: the complex conjugate of a complex number $\lambda$.
- $\text{Re}(\lambda)$: the real part of a complex number $\lambda$.
- $\text{Im}(\lambda)$: the imaginary part of a complex number $\lambda$.
- $a \lor b$: the maximum of two real numbers $a$ and $b$.
- $a \land b$: the minimum of two real numbers $a$ and $b$.
- $\text{sgn}(x)$: $\text{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}, \forall x \in \mathbb{R}$.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>the empty set; See Page 23.</td>
</tr>
<tr>
<td>${x, y}$</td>
<td>an unordered pair; See Page 23.</td>
</tr>
<tr>
<td>$X_2$</td>
<td>the collection of all subsets of $X$; See Page 23.</td>
</tr>
<tr>
<td>$\cup$</td>
<td>the set union; See Page 23.</td>
</tr>
<tr>
<td>$(x, y)$</td>
<td>an ordered pair; See Page 24.</td>
</tr>
<tr>
<td>$X \times Y$</td>
<td>the Cartesian or direct product of sets $X$ and $Y$; See Page 24.</td>
</tr>
<tr>
<td>$A = { x \in B \mid P(x) }$</td>
<td>definition of a set $A$; See Page 24.</td>
</tr>
<tr>
<td>$x \sim y$</td>
<td>$x$ and $y$ are related in a relation; See Page 24.</td>
</tr>
<tr>
<td>$X/\equiv$</td>
<td>the quotient of the set $X$ with respect to an equivalence relation $\equiv$; See Page 24.</td>
</tr>
<tr>
<td>$f : X \to Y$</td>
<td>a function of $X$ to $Y$; ${(x, f(x)) \in X \times Y \mid \forall x \in X}$ is the graph of $f$; See Page 25.</td>
</tr>
<tr>
<td>$\text{graph}(f)$</td>
<td>the graph of a function $f$; See Page 25.</td>
</tr>
<tr>
<td>$\text{dom}(f)$</td>
<td>the domain of $f$; See Page 25.</td>
</tr>
<tr>
<td>$f(A)$</td>
<td>the image of $A \subseteq X$ under $f$; See Page 25.</td>
</tr>
<tr>
<td>$\text{range}(f)$</td>
<td>the range of $f$, equals to $f(X)$; See Page 25.</td>
</tr>
<tr>
<td>$f_{\text{inv}}(A)$</td>
<td>the preimage of $A \subseteq Y$ under $f$; See Page 25.</td>
</tr>
<tr>
<td>onto, surjective</td>
<td>$f(X) = Y$; See Page 25.</td>
</tr>
<tr>
<td>1-1, injective</td>
<td>$f(x_1) \neq f(x_2)$ if $x_1, x_2 \in X$ and $x_1 \neq x_2$; See Page 25.</td>
</tr>
<tr>
<td>bijective</td>
<td>both surjective and injective; See Page 25.</td>
</tr>
<tr>
<td>$f_{\text{inv}}$</td>
<td>the inverse function of $f$; See Page 25.</td>
</tr>
<tr>
<td>$g \circ f$</td>
<td>the composition of $g : Y \to Z$ with $f : X \to Y$; See Page 25.</td>
</tr>
<tr>
<td>$f_{\mid A}$</td>
<td>the restriction of $f$ to $A$; See Page 25.</td>
</tr>
<tr>
<td>$X^Y$</td>
<td>the set of all functions of $X$ to $Y$; See Page 25.</td>
</tr>
<tr>
<td>$\cap$</td>
<td>the set intersection; See Page 26.</td>
</tr>
<tr>
<td>$\complement A$</td>
<td>the compliment of a set $A$, where the whole set is clear from context; See Page 26.</td>
</tr>
<tr>
<td>$\setminus$</td>
<td>set minus; See Page 26.</td>
</tr>
<tr>
<td>$A \triangle B$</td>
<td>the symmetric difference of $A$ and $B$, equals to $(A \setminus B) \cup (B \setminus A)$; See Page 26.</td>
</tr>
<tr>
<td>$\text{card}(X)$</td>
<td>the number of elements in the finite set $X$; See Page 27.</td>
</tr>
<tr>
<td>$\Pi_{\alpha \in A} X_\alpha$</td>
<td>the Cartesian or direct product of $X_\alpha$'s; See Page 27.</td>
</tr>
<tr>
<td>$\pi_\alpha(x)$</td>
<td>the projection of an element in a Cartesian product space to one of the coordinates; See Page 27.</td>
</tr>
<tr>
<td>$\overline{A}$</td>
<td>the closure of a set $A$, where the whole set is clear from context; See Page 39.</td>
</tr>
<tr>
<td>$A^\circ$</td>
<td>the interior of a set $A$, where the whole set is clear from context; See Page 39.</td>
</tr>
<tr>
<td>$\partial A$</td>
<td>the boundary of a set $A$, where the whole set is clear from context; See Page 39.</td>
</tr>
<tr>
<td>Expression</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>$\prod_{\alpha \in \Lambda} (X_{\alpha}, \mathcal{O}_{\alpha})$</td>
<td>the product topological space; See Page 46.</td>
</tr>
<tr>
<td>$T_1$</td>
<td>Tychonoff space; See Page 51.</td>
</tr>
<tr>
<td>$T_2$</td>
<td>Hausdorff space; See Page 51.</td>
</tr>
<tr>
<td>$T_3$</td>
<td>regular space; See Page 51.</td>
</tr>
<tr>
<td>$T_4$</td>
<td>normal space; See Page 51.</td>
</tr>
<tr>
<td>$T_3^k$</td>
<td>completely regular space; See Page 62.</td>
</tr>
<tr>
<td>$(x_{\alpha})_{\alpha \in A}$</td>
<td>a net; See Page 63.</td>
</tr>
<tr>
<td>$\lim_{\alpha \in A} x_{\alpha}$</td>
<td>the limit of a net; See Page 63.</td>
</tr>
<tr>
<td>$\lim_{x \to x_0} f(x)$</td>
<td>the limit of $f(x)$ as $x \to x_0$; See Page 67.</td>
</tr>
<tr>
<td>$\mathbb{R}_e$</td>
<td>the set of extended real numbers, which equals to $\mathbb{R} \cup {-\infty, +\infty}$; See Page 70.</td>
</tr>
<tr>
<td>$\limsup_{\alpha \in A} x_{\alpha}$</td>
<td>the limit superior of a real-valued net; See Page 71.</td>
</tr>
<tr>
<td>$\liminf_{\alpha \in A} x_{\alpha}$</td>
<td>the limit inferior of a real-valued net; See Page 71.</td>
</tr>
<tr>
<td>$\limsup_{x \to x_0} f(x)$</td>
<td>the limit superior of $f(x)$ as $x \to x_0$; See Page 72.</td>
</tr>
<tr>
<td>$\liminf_{x \to x_0} f(x)$</td>
<td>the limit inferior of $f(x)$ as $x \to x_0$; See Page 72.</td>
</tr>
<tr>
<td>$B_X(x_0, r)$</td>
<td>the open ball centered at $x_0$ with radius $r$; See Page 77.</td>
</tr>
<tr>
<td>$\overline{B}_X(x_0, r)$</td>
<td>the closed ball centered at $x_0$ with radius $r$; See Page 77.</td>
</tr>
<tr>
<td>$\text{dist}(x_0, S)$</td>
<td>the distance from a point $x_0$ to a set $S$ in a metric space; See Page 79.</td>
</tr>
<tr>
<td>$(X, \rho_X) \times (Y, \rho_Y)$</td>
<td>the finite product metric space; See Page 84.</td>
</tr>
<tr>
<td>$(\prod_{i=1}^{\infty} X_i, \rho)$</td>
<td>the countably infinite product metric space; See Page 87.</td>
</tr>
<tr>
<td>$\text{supp}(f)$</td>
<td>the support of a function; See Page 121.</td>
</tr>
<tr>
<td>$\beta(X)$</td>
<td>the Stone-Čech compatification of a completely regular topological space $X$; See Page 132.</td>
</tr>
<tr>
<td>$(\mathcal{M}(A, Y), \mathcal{F})$</td>
<td>the vector space of $Y$-valued functions of a set $A$ over the field $\mathcal{F}$; See Page 140.</td>
</tr>
<tr>
<td>$\vartheta_X$</td>
<td>the null vector of a vector space $X$; See Page 142.</td>
</tr>
<tr>
<td>$\mathcal{N}(A)$</td>
<td>the null space of a linear operator $A$; See Page 142.</td>
</tr>
<tr>
<td>$\mathcal{R}(A)$</td>
<td>the range space of a linear operator $A$; See Page 142.</td>
</tr>
<tr>
<td>$\alpha S$</td>
<td>the scalar multiplication by $\alpha$ of a set $S$ in a vector space; See Page 143.</td>
</tr>
<tr>
<td>$S + T$</td>
<td>the sum of two sets $S$ and $T$ in a vector space; See Page 143.</td>
</tr>
<tr>
<td>$\text{span}(A)$</td>
<td>the subspace generated by the set $A$; See Page 144.</td>
</tr>
<tr>
<td>$\nu(P)$</td>
<td>the linear variety generated by a nonempty set $P$; See Page 144.</td>
</tr>
<tr>
<td>$\text{co}(S)$</td>
<td>the convex hull generated by $S$ in a vector space; See Page 145.</td>
</tr>
<tr>
<td>$|x|$</td>
<td>the norm of a vector $x$; See Page 151.</td>
</tr>
<tr>
<td>$</td>
<td>x</td>
</tr>
</tbody>
</table>
NOTATIONS

\( C_1([a, b]) \) the normed linear space of continuously differentiable real-valued functions on the interval \([a, b]\); See Page 152.
\( l_p \) the normed linear space of real-valued sequences with finite \( p \)-norm, \( 1 \leq p \leq +\infty \); See Page 153.
\( l_p(X) \) the normed linear space of \( X \)-valued sequences with finite \( p \)-norm, \( 1 \leq p \leq +\infty \); See Page 157.
\( V(P) \) the closed linear variety generated by a nonempty set \( P \); See Page 160.
\( ^oP \) the relative interior of a set \( P \); See Page 160.
\( X \times Y \) the finite Cartesian product normed linear space; See Page 160.
\( C([a, b]) \) the Banach space of continuous real-valued functions on the interval \([a, b]\); See Page 164.
\( C(K, X) \) the normed linear space of continuous \( X \)-valued functions on a compact space \( K \); See Page 164.
\( X_\mathbb{R} \) the real normed linear space induced by a complex normed linear space \( X \); See Page 168.
\([x]\) the coset of a vector \( x \) in a quotient space; See Page 169.
\( X/M \) the quotient space of a vector space \( X \) modulo a subspace \( M \); See Page 169.
\( X/M \) the quotient space of a normed linear space \( X \) modulo a closed subspace \( M \); See Page 170.
\( C_v(X, Y) \) the vector space of continuous \( Y \)-valued functions on \( X \); See Page 172.
\( B(X, Y) \) the set of bounded linear operators of \( X \) to \( Y \); See Page 179.
\( X^* \) the dual of \( X \); See Page 183.
\( x_* \) a vector in the dual; See Page 183.
\( \langle\langle x_*, x \rangle\rangle \) the evaluation of a bounded linear functional \( x_* \) at the vector \( x \), that is \( x_*(x) \); See Page 183.
\( c_0(X) \) the subspace of \( l_\infty(X) \) consisting of \( X \)-valued sequences with limit \( \vartheta_X \); See Page 187.
\( X^{**} \) the second dual of \( X \); See Page 196.
\( S^\perp \) the orthogonal complement of the set \( S \); See Page 198.
\( ^\perp S \) the pre-orthogonal complement of the set \( S \); See Page 198.
\( A' \) the adjoint of a linear operator \( A \); See Page 207.
\( A'' \) the adjoint of the adjoint of a linear operator \( A \); See Page 208.
\( \mathcal{O}_{\text{weak}}(X) \) the weak topology on a normed linear space \( X \); See Page 210.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{X}_{\text{weak}}$</td>
<td>the weak topological space associated with a normed linear space $\mathcal{X}$; See Page 211.</td>
<td>211</td>
</tr>
<tr>
<td>$\mathcal{O}_{\text{weak}^*}$</td>
<td>the weak* topology on the dual of a normed linear space $\mathcal{X}$; See Page 213.</td>
<td>213</td>
</tr>
<tr>
<td>$\mathcal{X}^<em>_{\text{weak}^</em>}$</td>
<td>the weak* topological space associated with a normed linear space $\mathcal{X}$; See Page 214.</td>
<td>214</td>
</tr>
<tr>
<td>$K_{\text{supp}}$</td>
<td>the support of a convex set $K$; See Page 226.</td>
<td>226</td>
</tr>
<tr>
<td>$[f, C]$</td>
<td>the epigraph of a convex function $f : C \to \mathbb{R}$; See Page 229.</td>
<td>229</td>
</tr>
<tr>
<td>$\text{conjugate convex set}$</td>
<td></td>
<td>233</td>
</tr>
<tr>
<td>$C_{\text{conjugate}}$</td>
<td></td>
<td>233</td>
</tr>
<tr>
<td>$f_{\text{conjugate}}$</td>
<td></td>
<td>233</td>
</tr>
<tr>
<td>$[f, C]_{\text{conjugate}}$</td>
<td>the epigraph of the conjugate convex functional; See Page 234.</td>
<td>234</td>
</tr>
<tr>
<td>$\text{conjugate pre-convex set}$</td>
<td></td>
<td>236</td>
</tr>
<tr>
<td>$\text{conjugate pre-convex functional}$</td>
<td></td>
<td>236</td>
</tr>
<tr>
<td>$[\varphi, \Gamma]_{\text{conjugate}}$</td>
<td>the epigraph of the pre-conjugate convex functional; See Page 236.</td>
<td>236</td>
</tr>
<tr>
<td>$\geq$</td>
<td>greater than or equal to (with respect to the positive cone); See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$\leq$</td>
<td>less than or equal to (with respect to the positive cone); See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$&gt;$</td>
<td>greater than (with respect to the positive cone); See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$&lt;$</td>
<td>less than (with respect to the positive cone); See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$S^\oplus$</td>
<td>the positive conjugate cone of a set $S$; See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$S^\ominus$</td>
<td>the negative conjugate cone of a set $S$; See Page 247.</td>
<td>247</td>
</tr>
<tr>
<td>$A_D(x_0)$</td>
<td>the set of admissible deviations in $D$ at $x_0$; See Page 257.</td>
<td>257</td>
</tr>
<tr>
<td>$f^{(1)}(x_0), Df(x_0)$</td>
<td>the Fréchet derivative of $f$ at $x_0$; See Page 258.</td>
<td>258</td>
</tr>
<tr>
<td>$Df(x_0; u)$</td>
<td>the directional derivative of $f$ at $x_0$ along $u$; See Page 258.</td>
<td>258</td>
</tr>
<tr>
<td>$\frac{\partial f}{\partial y}(x_0, y_0)$</td>
<td>partial derivative of $f$ with respect to $y$ at $(x_0, y_0)$;</td>
<td>259</td>
</tr>
<tr>
<td>$\text{ro}(A)(B)$</td>
<td>right operate: $\text{ro}(A)(B) = BA$; See Page 264.</td>
<td>264</td>
</tr>
<tr>
<td>$B_k(\mathcal{X}, \mathcal{Y})$</td>
<td>the set of bounded multi-linear $\mathcal{Y}$-valued functions on $\mathcal{X}^k$; See Page 271.</td>
<td>271</td>
</tr>
<tr>
<td>$B_{\mathcal{S} k}(\mathcal{X}, \mathcal{Y})$</td>
<td>the set of symmetric bounded multi-linear $\mathcal{Y}$-valued functions on $\mathcal{X}^k$; See Page 271.</td>
<td>271</td>
</tr>
<tr>
<td>$D^k f(x_0), f^{(k)}(x_0)$</td>
<td>the $k$th order Fréchet derivative of $f$ at $x_0$; See Page 271.</td>
<td>271</td>
</tr>
<tr>
<td>$\mathcal{C}<em>k, \mathcal{C}</em>\infty$</td>
<td>$k$-times and infinite-times continuously differentiable functions, respectively; See Page 272.</td>
<td>272</td>
</tr>
</tbody>
</table>
\[ \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \]  
This represents the \(k\)th-order partial derivative of \(f\); See Page 277.

\[ S_m(D) = \bigcup_{\alpha \in \mathbb{K}} \alpha D \]  
This denotes the normed linear space of \(k\)-times continuously differentiable \(\mathbb{K}\)-valued functions on a compact set \(\Omega \subseteq \mathcal{X}\); See Page 277.

\[ \mathcal{C}_k(\Omega, \mathbb{Y}) \]  
This denotes the normed linear space of \(k\)-times continuously differentiable \(\mathbb{Y}\)-valued functions on a compact set \(\Omega \subseteq \mathcal{X}\); See Page 313.

\[ \mathcal{C}_b(\mathcal{X}, \mathbb{Y}) \]  
This denotes the normed linear space of bounded continuous \(\mathbb{Y}\)-valued functions on a topological space \(\mathcal{X}\); See Page 313.

\[ \mathcal{C}^b_k(\Omega, \mathbb{Y}) \]  
This denotes the normed linear space of \(k\)-times bounded continuously differentiable \(\mathbb{Y}\)-valued functions on a set \(\Omega \subseteq \mathcal{X}\); See Page 314.

\[ A^{T_{n_1 \cdots n_m}} \]  
This denotes the transpose of an \(m\)th order tensor \(A\) with the permutation \((n_1, \ldots, n_m)\); See Page 315.

\[ A \otimes B \]  
This denotes the outer product of two tensors; See Page 316.

\[ 0_{m_1 \times \cdots \times m_n} \]  
This denotes an \(n\)th order \(\mathbb{K}\)-valued tensor in \(B(\mathbb{K}^{m_1}, \ldots, B(\mathbb{K}^{m_n}, \mathbb{K}) \cdots)\) with all elements equal to 0; See Page 317.

\[ 1_{m_1 \times \cdots \times m_n} \]  
This denotes an \(n\)th order \(\mathbb{K}\)-valued tensor in \(B(\mathbb{K}^{m_1}, \ldots, B(\mathbb{K}^{m_n}, \mathbb{K}) \cdots)\) with all elements equal to 1; See Page 317.

\[ \exp(A) \]  
This denotes the exponential function of a linear operator \(A\) on a Banach space; See Page 328.

\[ S_{+\mathcal{X}}, S_{pzd\mathcal{X}} \]  
These denote sets of positive definite and positive semi-definite operators over the real normed linear space \(\mathcal{X}\), respectively; See Page 330.

\[ S_{-\mathcal{X}}, S_{nsd\mathcal{X}} \]  
These denote sets of negative definite and negative semi-definite operators over the real normed linear space \(\mathcal{X}\), respectively; See Page 330.

\[ S_\mathcal{X} \]  
This denotes \(\mathbb{B}_{2}(\mathcal{X}, \mathbb{R})\); See Page 330.

\[ (\mathbb{R}, \mathcal{B}_L, \mu_L) \]  
This denotes the Lebesgue measure space; See Page 364.

\[ \mu_{Lo} \]  
This denotes the Lebesgue outer measure; See Page 365.

\[ \mathcal{B}_B(\mathcal{X}) \]  
This denotes the Borel sets; See Page 365.

\[ \mathcal{A}(\mathcal{X}) \]  
This denotes the algebra generated by the topology of a topological space \(\mathcal{X}\); See Page 365.

\[ \mu_B \]  
This denotes the Borel measure on \(\mathbb{R}\); See Page 366.

\[ \mathcal{P} \text{ a.e. in } \mathcal{X} \]  
This denotes \(P\) holds almost everywhere in \(\mathcal{X}\); See Page 374.

\[ \mathcal{P}(x) \text{ a.e. } x \in \mathcal{X} \]  
This denotes \(P\) holds almost everywhere in \(\mathcal{X}\); See Page 374.

\[ \chi_{A,\mathcal{X}} \]  
This denotes the indicator function of \(A \subseteq \mathcal{X}\). Function of \(X\) to \(\{0,1\}\); See Page 383.

\[ \mathcal{P} \circ f \]  
This denotes \(\mathcal{P} \circ f : X \rightarrow [0, \infty) \subseteq \mathbb{R}\) defined by \(\mathcal{P} \circ f(x) = \|f(x)\|\), \(\forall x \in \mathcal{X}\); See Page 384.

\[ \mathcal{I}(\mathcal{X}) \]  
This denotes the collection of all representation of \(\mathcal{X}\); See Page 387.

\[ \mathcal{J}(\mathcal{X}) \]  
This denotes the integration system on \(\mathcal{X}\); See Page 387.
\[ \int_X f \, d\mu \] the integral of a function \( f \) on a set \( X \) with respect to measure \( \mu \); See Page 388.

\[ \int_X f(x) \, d\mu(x) \] the integral of a function \( f \) on a set \( X \) with respect to measure \( \mu \); See Page 388.

\( \mathcal{P} \circ \mu \) the total variation of a Banach space valued pre-measure \( \mu \); See Page 418.

\( \mathcal{P} \circ \mu \) the total variation of a Banach space valued measure \( \mu \); See Page 423.

\( \mu_1 + \mu_2 \) the \( \mathcal{Y} \)-valued measure that equals to the sum of two \( \mathcal{Y} \)-valued measures on the same measurable space; See Page 453.

\( \alpha \mu \) the \( \mathcal{Y} \)-valued measure that equals to the scalar product of \( \alpha \in \mathbb{K} \) and a \( \mathcal{Y} \)-valued measure \( \mu \); See Page 454.

\( \mu y \) the \( \mathcal{Y} \)-valued measure that equals to scalar product of a \( \mathbb{K} \)-valued measure \( \mu \) and \( y \in \mathcal{Y} \); See Page 454.

\( A \mu \) the \( \mathcal{Z} \)-valued measure that equals to product of an bounded linear operator \( A \) and a \( \mathcal{Y} \)-valued measure \( \mu \); See Page 454.

\( M_\sigma(X, \mathcal{B}, \mathcal{Y}) \) the vector space of \( \sigma \)-finite \( \mathcal{Y} \)-valued measures on the measurable space \( (X, \mathcal{B}) \); See Page 460.

\( M_f(X, \mathcal{B}, \mathcal{Y}) \) the normed linear space of finite \( \mathcal{Y} \)-valued measures on the measurable space \( (X, \mathcal{B}) \); See Page 461.

\[ \lim_{n \in \mathbb{N}} \mu_n = \nu \] a sequence of \( \sigma \)-finite (\( \mathcal{Y} \)-valued) measures \( (\mu_n)_{n=1}^\infty \) converges to a \( \sigma \)-finite (\( \mathcal{Y} \)-valued) measure \( \nu \); See Page 464.

\( \mu_1 \leq \mu_2 \) the measures \( \mu_1 \) and \( \mu_2 \) on the measurable space \( (X, \mathcal{B}) \) can be compared if \( \mu_1(E) \leq \mu_2(E), \forall E \in \mathcal{B} \); See Page 464.

\[ \begin{bmatrix} \mu_{1,1} & \cdots & \mu_{1,m} \\ \vdots & \ddots & \vdots \\ \mu_{n,1} & \cdots & \mu_{n,m} \end{bmatrix} \] the vector measure; See Page 467.

\( \mu_1 \ll \mu_2 \) the measure \( \mu_1 \) is absolutely continuous with respect to the measure \( \mu_2 \); See Page 477.

\( \mu_1 \perp \mu_2 \) the measures \( \mu_1 \) and \( \mu_2 \) are mutually singular; See Page 477.

\[ \frac{d\nu}{d\mu} \] the Radon-Nikodym derivative of the \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu \) with respect to the \( \sigma \)-finite \( \mathbb{K} \)-valued measure \( \mu \); See Page 487.

\( \mathcal{P}_p \circ f \) \( \text{ess sup} \) the function \( \| f(\cdot) \|^p \); See Page 497.

\( \text{ess sup} \) the essential supremum; See Page 499.

\[ \lim_{n \in \mathbb{N}} z_n = z \text{ in } \mathbb{L}_p \] the sequence \( (z_n)_{n=1}^\infty \subseteq \mathbb{L}_p \) converges to \( z \in \mathbb{L}_p \) in \( \mathbb{L}_p \) pseudo-norm; See Page 501.

\( \mathcal{M}_f(X, \mathcal{Y}) \) the normed linear space of finite \( \mathcal{Y} \)-valued topological measures on \( X \); See Page 518.
\( M_{\sigma t}(\mathcal{X}, \mathcal{Y}) \) the vector space of \( \sigma \)-finite \( \mathcal{Y} \)-valued topological measures on \( \mathcal{X} \); See Page 518.

\( M_{\sigma}(X, \mathcal{B}) \) the set of \( \sigma \)-finite measures on the measurable space \((X, \mathcal{B})\); See Page 519.

\( M_f(X, \mathcal{B}) \) the set of finite measures on the measurable space \((X, \mathcal{B})\); See Page 519.

\( M_{\sigma t}(\mathcal{X}) \) the set of \( \sigma \)-finite topological measures on the topological space \( \mathcal{X} \); See Page 519.

\( M_{ft}(\mathcal{X}) \) the set of finite topological measures on the topological space \( \mathcal{X} \); See Page 519.

\( C_c(X, \mathcal{Y}) \) the normed linear space of \( \mathcal{Y} \)-valued continuous function of a finite-dimensional Banach space, that converges to \( \vartheta \) as \( \| x \|_X \to \infty \); See Page 537.

\( r_{x_1, x_2} \) the semi-open rectangle in \( \mathbb{R}^m \) with corners \( x_1 \) and \( x_2 \); See Page 582.

\( \mathcal{P}(\Omega) \) the principal of a region \( \Omega \); See Page 583.

\( \text{VRect}_{i, x_1, x_2} \) the set of vertexes of \( r_{x_1, x_2} \) with \( i \) coordinates equal to that of \( x_1 \); See Page 585.

\( \Delta F(r_{x_1, x_2}) \) the increment of \( F \) on \( r_{x_1, x_2} \); See Page 585.

\( T_F(r_{x_1, x_2}) \) the total variation of a function \( F \) on the semi-open rectangle \( r_{x_1, x_2} \); See Page 585.

\( (\mathbb{R}^m, \mathcal{B}_m, \mu_{Lm}) \), \( \mu_{Lm} \)-dimensional Lebesgue measure space and the \( m \)-dimensional Lebesgue outer measure; See Page 604.

\( f_U \) the integral of function \( f \) with respect to \( \mu_{B}\); See Page 605.

\( e_{m,i} \) the \( i \)th unit vector in \( \mathbb{R}^m \); See Page 626.
$S_{n-1}$: The finite compact separable metric measure space on the $n$-dimensional unit sphere, where the measure of a set is equal to the “surface area” of the set; See Page 659.

$W_{p,1,x_0}(X,Y)$: 1st order Sobolev space of $X$ to $Y$ under $p$-norm with respect to $x_0 \in X$; See Page 661.

$\langle x, y \rangle$: the inner product of vectors $x$ and $y$; See Page 667.

$x \perp y$: the vectors $x$ and $y$ are orthogonal in a pre-Hilbert space; See Page 669.

$x \perp S$: the vector $x$ is orthogonal to the set $S$ in a pre-Hilbert space; See Page 669.

$x^*$: $x^* \in X^*$ that satisfies $\langle \langle x^*, y \rangle \rangle = \langle y, x \rangle$, $\forall y \in X$; See Page 673.

$M \oplus N$: the direct sum of $M$ and $N$; See Page 675.

$A^*$: the Hermitian adjoint of $A$; See Page 676.

$\text{Gram}(y_1, \ldots, y_n)$: the Gram matrix of $y_1, \ldots, y_n$; See Page 680.

$\text{gram}(y_1, \ldots, y_n)$: the Gram determinant of $y_1, \ldots, y_n$; See Page 680.

$S_X$: set of Hermitian operators on a real or complex Hilbert space; See Page 693.

$S_{+X}$, $S_{psd}X$: sets of positive definite and positive semi-definite operators over real or complex Hilbert space $X$, respectively; See Page 693.

$S_{-X}$, $S_{nsd}X$: sets of negative definite and negative semi-definite operators over real or complex Hilbert space $X$, respectively; See Page 693.

$A^\dagger$: the pseudoinverse of $A \in B(X,Y)$; See Page 698.

$K(X,Y)$: The set of compact linear bounded operators of $X$ to $Y$; See Page 704.

$E(x)$: expectation of $x$; See Page 721.

$E(x|\hat{\mathcal{B}})$: conditional expectation of $x$ given $\hat{\mathcal{B}}$; See Page 724.

$L_x$: the law for the random variable $x$; See Page 738.

$x \sim N(\bar{x},K)$: Gaussian (normal) random variable (vector) with mean $\bar{x}$ and covariance $K$; See Page 742.

$\mathcal{N} := (\mathbb{Z}_+, \mathbb{Z}_+^2, \mu)$: the topological measure space on $\mathbb{Z}_+$ with the counting measure and the partial ordering of $\leq$; See Page 744.

$P(F|\hat{\mathcal{B}})$: the conditional probability of event $F$ happens given $\hat{\mathcal{B}}$, which is defined to be $E(\chi_{F,\Omega}|\hat{\mathcal{B}})$; See Page 746.

$\Psi_x$: the characteristic function of the random variable $x$; See Page 757.
Index

Absolute continuity of measure, 477
Absolute integrability, 409
  Banach space valued measure space, 442
Adjoint operator, 207
Admissible deviation, 257
Alaoglu Theorem, 216
Alexandroff One-Point Compactification, 124
Algebra, 172
  Boolean, 27
  Generated by, 27
Alignment, 198
Almost everywhere, 374
  Banach space valued measure, 423
  Banach space valued pre-measure, 419
Analytic radius, 317
Antisymmetric, 24
Ascoli-Arzelá Theorem, 115

Baire, 90
  Category Theorem, 91
Ball
  Closed, 77
  Open, 77
Banach measure space, 369
  Banach space valued, 518
Banach space, 162
Banach space valued pre-measure, 418
  On algebra, 548
Basis
  Topological space, 45

Vector space, 149
Bernstein Approximation Theorem, 532
Bernstein function, 531
Bessel’s Inequality, 682
Bijective, 25
Bolzano-Weierstrass property, 108
Borel measure, 604
Borel sets, 365
Borel-Cantelli Lemma, 355, 723
Borel-Lebesgue Theorem, 112
Bound
  Greatest lower, 29
  Least upper, 29
  Lower, 29
  Upper, 29
Bounded
  Function, 157
  Linear operator, 179
  Set, 157
Bounded Convergence Theorem, 395
Bourbaki, 116

Carathéodory Extension Theorem, 360, 544
Cartesian metric, 84
Cartesian product, 37
  Normed linear space, 160
  Vector space, 140
Category
  First, 52
  Second, 52
  Everywhere, 52
Cauchy
  Net, 91
Sequence, 80
In measure, 381
Cauchy-Schwarz Inequality, 154, 503, 667
Central Limit Theorem, 771
Cesàro, 717
Chain Rule, 265
Change of Variable, 553, 646
Chebyshev Inequality, 774
Choice, 30
Closed Graph Theorem, 205, 206
Commutative
Ring, 137
Compact, 101, 167
Countably, 107
Locally, 118
Sequentially, 108
σ-, 127
Complete, 80
Banach space valued measure space, 423
Banach space valued pre-measure space, 419
Measure space, 355
Metric measure space, 369
Banach space valued, 517
Sequence, 683
Completely regular, 62
Completion of a Banach space valued measure space, 425
Completion of a Banach space valued pre-measure space, 420
Completion of a measure space, 356
Component, 55
Composition, 25
Concave functional, 229
Conditional expectation, 723
Conditional probability, 746
Cone, 147
Conjugate, 247
Negative, 247
Positive, 247, 248
Conic segment, 147
Conjugate
Group, 136
Connected, 54
Arcwise, 58
Locally, 56
Simply, 58
Continuity, 42
Absolute, 605
One-sided, 581
Semi-, 44
Continuum, 106
Contraction index, 292
Contraction mapping, 292
Contraction Mapping Theorem, 293
Convergence
In distribution, 541
Random Variables, 763
In measure, 380
In measure
Banach space valued measure, 423
Convex
Functional, 229
Local, 218
Mapping, 248
Convex combination, 146
Convex hull, 145
Coset, 169
Countability
First axiom of, 45
Second axiom of, 45
Countable, 27
Covering, 46
Cube, 97
Hilbert, 97
Cumulative distribution function, 586
With origin, 597
Curve, 58
Closed, 58
Dense, 52
Nowhere, 52
Derivative
Directional, 258
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fréchet</td>
<td>258</td>
</tr>
<tr>
<td>Higher-order</td>
<td>271</td>
</tr>
<tr>
<td>Partial</td>
<td>259</td>
</tr>
<tr>
<td>Higher-order</td>
<td>277</td>
</tr>
<tr>
<td>Diameter</td>
<td>622</td>
</tr>
<tr>
<td>Dimension</td>
<td>149</td>
</tr>
<tr>
<td>Co-</td>
<td>690</td>
</tr>
<tr>
<td>Dini’s Lemma</td>
<td>110</td>
</tr>
<tr>
<td>Direct sum</td>
<td>675</td>
</tr>
<tr>
<td>Directed system</td>
<td>63</td>
</tr>
<tr>
<td>Distance</td>
<td>79</td>
</tr>
<tr>
<td>Domain</td>
<td>25</td>
</tr>
<tr>
<td>Doob’s Forward Convergence Theorem</td>
<td>752</td>
</tr>
<tr>
<td>Doob’s Optional Stopping Theorem</td>
<td>747</td>
</tr>
<tr>
<td>Doob’s Upcrossing Lemma</td>
<td>754</td>
</tr>
<tr>
<td>Dual functional</td>
<td>251</td>
</tr>
<tr>
<td>Dual space</td>
<td>183</td>
</tr>
<tr>
<td>Second</td>
<td>196</td>
</tr>
<tr>
<td>Egoroff’s Theorem</td>
<td>379</td>
</tr>
<tr>
<td>Eidelheit Separation Theorem</td>
<td>225</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>701</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>706</td>
</tr>
<tr>
<td>Eigenvector</td>
<td>701</td>
</tr>
<tr>
<td>Empty set</td>
<td>23</td>
</tr>
<tr>
<td>Epigraph</td>
<td>229</td>
</tr>
<tr>
<td>Equicontinuity</td>
<td>84</td>
</tr>
<tr>
<td>Equivalence</td>
<td>24</td>
</tr>
<tr>
<td>Equivalent</td>
<td></td>
</tr>
<tr>
<td>Metric spaces</td>
<td>79</td>
</tr>
<tr>
<td>Normed linear spaces</td>
<td>167</td>
</tr>
<tr>
<td>Essential supremum</td>
<td>499</td>
</tr>
<tr>
<td>Exhaustion</td>
<td>127</td>
</tr>
<tr>
<td>Expectation</td>
<td>721</td>
</tr>
<tr>
<td>Exponential of linear operator</td>
<td>328</td>
</tr>
<tr>
<td>Extensionality</td>
<td>23</td>
</tr>
<tr>
<td>Extreme point</td>
<td>330</td>
</tr>
<tr>
<td>Extreme subset</td>
<td>330</td>
</tr>
<tr>
<td>$F_{\sigma}$</td>
<td>522</td>
</tr>
<tr>
<td>Fatou’s Lemma</td>
<td>398, 471</td>
</tr>
<tr>
<td>For sets</td>
<td>722</td>
</tr>
<tr>
<td>Fejér</td>
<td>716</td>
</tr>
<tr>
<td>Fenchel Duality Theorem</td>
<td>240</td>
</tr>
<tr>
<td>Field</td>
<td>138</td>
</tr>
<tr>
<td>Sub-</td>
<td>138</td>
</tr>
<tr>
<td>Filtration</td>
<td>744</td>
</tr>
<tr>
<td>Natural</td>
<td>744</td>
</tr>
<tr>
<td>Finite</td>
<td>27</td>
</tr>
<tr>
<td>Banach space valued measure</td>
<td>423</td>
</tr>
<tr>
<td>Banach space valued pre-measure</td>
<td>419</td>
</tr>
<tr>
<td>Locally</td>
<td>128</td>
</tr>
<tr>
<td>Banach space valued topological measure</td>
<td>517</td>
</tr>
<tr>
<td>Topological measure space</td>
<td>369</td>
</tr>
<tr>
<td>Measure</td>
<td>355</td>
</tr>
<tr>
<td>Star</td>
<td>130</td>
</tr>
<tr>
<td>Finite intersection property</td>
<td>101</td>
</tr>
<tr>
<td>Fourier series</td>
<td>686</td>
</tr>
<tr>
<td>Fredholm alternative</td>
<td>718</td>
</tr>
<tr>
<td>Fredholm alternative</td>
<td>705</td>
</tr>
<tr>
<td>Fubini Theorem</td>
<td>570, 572, 580</td>
</tr>
<tr>
<td>Function</td>
<td>25</td>
</tr>
<tr>
<td>Analytic</td>
<td>317</td>
</tr>
<tr>
<td>Characteristic</td>
<td>757</td>
</tr>
<tr>
<td>Linear</td>
<td>142</td>
</tr>
<tr>
<td>Measurable</td>
<td>372</td>
</tr>
<tr>
<td>Of bounded variation</td>
<td>585</td>
</tr>
<tr>
<td>Functional</td>
<td>142</td>
</tr>
<tr>
<td>Sublinear</td>
<td>188</td>
</tr>
<tr>
<td>Fundamental Theorem of Calculus</td>
<td>633, 637</td>
</tr>
<tr>
<td>Gå</td>
<td>522</td>
</tr>
<tr>
<td>Generation process</td>
<td>431</td>
</tr>
<tr>
<td>Gram determinant</td>
<td>680</td>
</tr>
<tr>
<td>Gram determinant</td>
<td>680</td>
</tr>
<tr>
<td>Gram matrix</td>
<td>680</td>
</tr>
<tr>
<td>Gram-Schmidt Procedure</td>
<td>679</td>
</tr>
<tr>
<td>Graph</td>
<td>25</td>
</tr>
<tr>
<td>Greatest element</td>
<td>29</td>
</tr>
<tr>
<td>Group</td>
<td>135</td>
</tr>
<tr>
<td>abelian</td>
<td>136</td>
</tr>
<tr>
<td>Semi</td>
<td>136</td>
</tr>
<tr>
<td>abelian</td>
<td>136</td>
</tr>
<tr>
<td>Haar functions</td>
<td>782</td>
</tr>
</tbody>
</table>
Hahn Decomposition Theorem, 482
Hahn-Banach Theorem
   Extension form, 189
      Complex, 191
   Geometric form, 224
      Simple version, 192
Hausdorff
   Maximal Principle, 30
   Space, 51
Heine-Borel, 103
Helly’s Lemma, 768
Hermite polynomials, 800
Hermitian adjoint, 676
Hermitian operator, 676
Hilbert space, 669
Hölder’s Inequality, 153, 501
Homeomorphical isomeasure, 549
Homeomorphism, 44
Homomorphism, 136
   Ring, 137
   Vector space, 142
Homotopic, 58
Hyperplane, 221
   Supporting, 225

Ideal, 137
Identity element, 137
Image, 25
Implicit Function Theorem, 299, 301
   Analytic, 327
Increment on semi-open rectangle, 585
Indicator function, 383
Induced measure, 550
Infinity, 23
Injective, 25
Injective Mapping Theorem, 294
Inner product, 667
Integral
   Finite Banach space valued measure space, 433
   Finite measure space, 388
   Infinite Banach space valued measure space, 433
   Infinite measure space, 389
   Nonnegative extended real-valued functions, 398
Integration by Parts, 643, 657
Integration system, 387
Interval, 363
Inverse Function Theorem, 297
   Analytic, 324
   Global, 308, 309
Inverse image, 25
Invert, 304
Invertible, 25
Isomeasure, 549
Isometry, 79
Isomorphism, 136
   Isometrical, 161
   Ring, 137
   Vector space, 142
Iterated Limit Theorem, 98
Itô’s Formula, 801
Jensen’s Inequality, 417
Joint Limit Theorem, 97
Jordan Decomposition Theorem, 483, 484
Kernel
   For measure, 428
Kolmogorov 0-1 Law, 755
Krein-Milman, 331
Kuhn-Tucker Theorem, 343
L_p(X, Y), 498, 501
Lagrange multiplier, 253, 338, 339
Lattice, 172
Law, 738
Law of iterated logarithm, 794
Law of Large Numbers
   Strong, 743
   L_p(X, Y), 498, 501
Least element, 28
Lebesgue Decomposition Theorem, 495
Lebesgue Dominated Convergence Theorem, 406, 409, 441, 444, 471, 472
Lebesgue measure, 364, 604
Lebesgue number, 111
Lebesgue outer measure, 365, 604
Lebesgue point, 624
Lebesgue-Stieltjes integral, 605
Legendre polynomials, 683
Lévy’s Convergence Theorem, 769
Lévy’s Inversion Formula, 759
Limit, 63, 67
  Inferior, 71, 72
  Joint, 75
  Superior, 71, 72
Linear combination, 144
Linear variety, 144
  Generated by, 144
  Closed, 160
Linearly independent, 148
Vectors, 148
Lipschitz continuity, 612

m-dimensional Lebesgue measure space, 604
Maximal, 28
Mazur’s Theorem, 224
Meager, 52
  Co-, 52
  Non-, 52
Mean Value Theorem, 55, 267, 268
Measurable
  Set, 353
  Space, 353
  With respect to an outer measure, 359
Measure, 353
  Banach space valued, 423
  Induced outer, 360
  Outer, 359
  Product, 557
  Space, 353
  Vector, 467
Measure on algebra, 360
  Banach space valued, 543
Metric, 77
  Pseudo-, 93
Metric measure space, 369
  Banach space valued, 517
Metric space, 77
  Product, 84
Metrizable, 96
Minimal, 28
Minkowski functional, 222
Minkowski’s Inequality, 154, 498
Mode of convergence, 763
Monotone class, 554
Monotone Class Lemma, 554
Monotone Convergence Theorem, 399
Mutually singular measures, 477

Natural mapping, 196
Natural metric, 158
Negative definite, 330, 693
Negative semi-definite, 330, 693
Net, 63
  Joint, 75
  Sub-, 66
Newton’s Method, 803
Norm
  Pseudo, 171
Normal
  Subgroup, 136
Normal equation, 681
Normal space, 51
Normed linear measure space, 369
  Banach space valued, 518
Null space, 142
Open mapping, 203
Open Mapping Theorem, 204, 295
Operator
  Affine, 142
  Compact, 704
  Linear, 142
  Spectral theory, 705
  Pseudoinverse, 698
Order
Group, 136
Ordering
  Partial, 28
  Total, 28
Orthogonal, 198, 669
Orthogonal complement, 198
  Pre-, 198
Orthonormal, 678
Outer product of tensors, 316
Pairing, 23
Paracompact, 129
Parallelogram Law, 668
Partition of Unity, 122
\( \pi \)-system, 554
Point
  Accumulation, 39
  At infinity, 124
  Beginning, 58
  Boundary, 39
  Cluster, 63
  End, 58
  Exterior, 39
  Interior, 39
  Relative, 160
  Of closure, 39
Point of minimum, 329
Point of relative extremum, 329
Point of relative minimum, 329
Positive definite, 329, 693
Positive linear functional, 524
Positive semi-definite, 330, 693
Power set, 23
Pre-Hilbert space, 667
Primal functional, 249
Principal of a region, 583
Probability density function, 738
Probability measure space, 721
  Filtered, 744
Projection, 37, 47
Projection operator, 697
Projection Theorem, 672, 689
Proper, 125
  Countably, 125
Pythagorean Theorem, 669
Quotient, 24
Quotient space, 93, 169–171
Raabe’s Test, 807
Radon-Nikodym derivative, 486
Radon-Nikodym Theorem, 490, 492
Random variable, 721
  Gaussian, 741
Range, 25
Range space, 142
Rectangle in \( \mathbb{R}^m \), 582, 583
Refinement, 104
Reflexive, 24
  Banach space, 196
Region, 583
Regular point, 338
Regular space, 51
Regularity, 23
Relation, 24
Replacement, 23
Residual, 52
Restriction, 25
Riemann-Lebesgue Theorem, 655
Riesz Representation Theorem, 511,
  525, 535, 536, 538, 651, 652
Riesz-Fréchet Theorem, 673
Ring, 137
  Sub-, 137
Semialgebra, 370
Separable, 52
Separation, 54
Set
  Boundary, 39
  Closed, 39
  Closure, 39
  Relative, 41
  Complement, 39
  Convex, 145
  Exterior, 39
  Interior, 39
  Negative, 481
  Open, 39
  Relative, 160
INDEX

Positive, 481
\(\sigma\)-algebra, 27
  Generated by, 27
  Independence, 723
  Tail, 755
\(\sigma\)-finite
  Banach space valued measure, 423
  Measure, 355
Simple function, 383
  Banach space valued measure space, 435
Singular value decomposition, 706
Skorokhod Representation, 764
Sobolev space
  1st order, 661
Stationary point, 339
Stochastic process, 744
  Adapted, 744
  Banach space valued Martingale, 752
Martingale, 744, 791
  Stopped, 747
Sub Martingale, 745, 791
Super Martingale, 745, 791
Stone-Cech compactification, 132
Stopping time, 746
Stream of a absolutely continuous function, 638
Subordinate, 121
Subset, 24
Subspace, 141
  Banach space valued measure space, 427
  Banach space valued pre-measure space, 422
  Generated by, 144
  Measure space, 358
  Proper, 141
  Topological measure space, 370
Support
  Of a convex set, 226
  Of a function, 121, 524
Support functional, 227
Surjective, 25
  Mapping Theorem, 294
  Symmetric, 24
Taylor Series, 317
Taylor's Theorem, 289
Tensor, 315
  Transpose, 315
Tietze's Extension Theorem, 60
Tonelli Theorem, 569
Topological measure space, 369
  Banach space valued, 517
  Product, 574
Topological space, 39
Topology, 39
  Generated by, 42
  Natural, 77
  Of pointwise convergence, 61
  Product, 46
  Stronger, 42
  Subset, 41
  Weak, 61, 210
  Weak*, 213
  Weaker, 42
Total variation, 418, 423, 543, 548
  Of function, 585
Totally bounded, 83
Transitive, 24
Tychonoff
  Space, 51
  Theorem, 116
Uniform
  Boundedness Principle, 54
  Cauchy sequence, 84
  Continuity, 82
  Convergence, 84
  Equivalent
    Metric spaces, 82
    Homeomorphism, 82
Uniform growth boundedness, 764
Uniform integrability, 772
Union, 23
Urysohn Metrization Theorem, 97
Urysohn's Lemma, 59
Vector space, 138
Vitali, 623
Vitali’s Lemma, 623, 645

Well-ordering, 31
   Principle, 31
Wiener integral, 801
Wiener process, 787, 788

Zero-element, 136
Zorn’s Lemma, 31