Reading Notes of “Real Analysis” 3rd Edition
by H. L. Royden

Zigang Pan

January 10, 2018
Preface

This is a reading note of the book Royden (1988) and the MATH 441 & 442 notes by Prof. Peter Leob of University of Illinois at Urbana-Champaign. In Chapter 6, I have also included some material from the book Maunder (1996). The Chapters 6–10 include a significant amount of material from the book Luenberger (1969). Chapter 9 also referenced Bartle (1976). 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) does offer some clues as to how to invent the wheel and the book Bartle (1976) is sometimes used to validate the result. The book Spivak (1965) provides the goal of the Fundamental Theorem of Calculus that Chapter 12 is to prove. But the result in Spivak (1965) may not be correct to begin with. Now, Chapter 12 is complete.
## Contents

Preface .......................................................... 3

1  Notations ..................................................... 9

2  Set Theory .................................................... 19
    2.1 Axiomatic Foundations of Set Theory .................... 19
    2.2 Relations and Equivalence ............................... 20
    2.3 Function ............................................... 21
    2.4 Set Operations ......................................... 22
    2.5 Algebra of Sets ........................................ 23
    2.6 Partial Ordering and Total Ordering .................... 24
    2.7 Basic Principles ....................................... 26

3  Topological Spaces .......................................... 35
    3.1 Fundamental Notions .................................... 35
    3.2 Continuity .............................................. 38
    3.3 Basis and Countability .................................. 41
    3.4 Products of Topological Spaces .......................... 42
    3.5 The Separation Axioms .................................. 47
    3.6 Category Theory ........................................ 48
    3.7 Connectedness ......................................... 50
    3.8 Continuous Real-Valued Functions ....................... 55
    3.9 Nets and Convergence .................................... 59

4  Metric Spaces ................................................. 73
    4.1 Fundamental Notions .................................... 73
    4.2 Convergence and Completeness ........................... 75
    4.3 Uniform Continuity and Uniformity ....................... 78
    4.4 Product Metric Spaces ................................... 80
    4.5 Subspaces .............................................. 85
    4.6 Baire Category .......................................... 86
    4.7 Completion of Metric Spaces ............................. 87
    4.8 Metrization of Topological Spaces ....................... 92
    4.9 Interchange Limits ...................................... 93
# CONTENTS

5 Compact and Locally Compact Spaces ........................................ 97
  5.1 Compact Spaces .................................................. 97
  5.2 Countable and Sequential Compactness .......................... 103
  5.3 Real-Valued Functions and Compactness ....................... 105
  5.4 Compactness in Metric Spaces ................................. 107
  5.5 The Ascoli-Arzelá Theorem ................................... 109
  5.6 Product Spaces ................................................ 111
  5.7 Locally Compact Spaces ....................................... 114
    5.7.1 Fundamental notion ....................................... 114
    5.7.2 Partition of unity ........................................ 117
    5.7.3 The Alexandroff one-point compactification .......... 120
    5.7.4 Proper functions ......................................... 121
  5.8 σ-Compact Spaces ............................................. 123
  5.9 Paracompact Spaces ............................................ 124
  5.10 The Stone-Čech Compactification .............................. 128

6 Vector Spaces .......................................................... 131
  6.1 Group .......................................................... 131
  6.2 Ring ............................................................ 133
  6.3 Field ........................................................... 134
  6.4 Vector Spaces .................................................. 134
  6.5 Product Spaces ................................................ 136
  6.6 Subspaces ....................................................... 137
  6.7 Convex Sets ..................................................... 141
  6.8 Linear Independence and Dimensions .......................... 144

7 Banach Spaces .......................................................... 147
  7.1 Normed Linear Spaces ........................................... 147
  7.2 The Natural Metric ............................................. 153
  7.3 Product Spaces ................................................ 156
  7.4 Banach Spaces .................................................. 158
  7.5 Compactness ..................................................... 163
  7.6 Quotient Spaces ................................................ 165
  7.7 The Stone-Weierstrass Theorem ................................. 167
  7.8 Linear Operators ............................................... 175
  7.9 Dual Spaces ..................................................... 179
    7.9.1 Basic concepts ........................................... 179
    7.9.2 Duals of some common Banach spaces .................. 180
    7.9.3 Extension form of Hahn-Banach Theorem .............. 184
    7.9.4 Second dual space ....................................... 192
    7.9.5 Alignment and orthogonal complements ............... 194
  7.10 The Open Mapping Theorem ................................... 199
  7.11 The Adjoints of Linear Operators ............................. 203
  7.12 Weak Topology ................................................. 206
8 Global Theory of Optimization 217
- 8.1 Hyperplanes and Convex Sets .................................................. 217
- 8.2 Geometric Form of Hahn-Banach Theorem ................................. 220
- 8.3 Duality in Minimum Norm Problems .......................................... 222
- 8.4 Convex and Concave Functionals .............................................. 225
- 8.5 Conjugate Convex Functionals .................................................. 229
- 8.6 Fenchel Duality Theorem .......................................................... 236
- 8.7 Positive Cones and Convex Mappings ........................................ 243
- 8.8 Lagrange Multipliers ............................................................... 245

9 Differentiation in Banach Spaces 253
- 9.1 Fundamental Notion ............................................................... 253
- 9.2 The Derivatives of Some Common Functions .............................. 257
- 9.3 Chain Rule and Mean Value Theorem ........................................ 261
- 9.4 Higher Order Derivatives ........................................................ 267
  - 9.4.1 Basic concept ................................................................. 267
  - 9.4.2 Interchange order of differentiation .................................... 273
  - 9.4.3 High order derivatives of some common functions ............... 278
  - 9.4.4 Properties of high order derivatives ................................ 282
- 9.5 Mapping Theorems ............................................................... 288
- 9.6 Global Inverse Function Theorem ............................................ 300
- 9.7 Interchange Differentiation and Limit ....................................... 307
- 9.8 Tensor Algebra ................................................................. 311
- 9.9 Analytic Functions ............................................................... 313

10 Local Theory of Optimization 325
- 10.1 Basic Notion ........................................................................ 325
- 10.2 Unconstrained Optimization .................................................. 331
- 10.3 Optimization with Equality Constraints ................................... 334
- 10.4 Inequality Constraints .......................................................... 339

11 General Measure and Integration 349
- 11.1 Measure Spaces ..................................................................... 349
- 11.2 Outer Measure and the Extension Theorem ............................. 355
- 11.3 Measurable Functions ........................................................... 368
- 11.4 Integration ............................................................................ 383
- 11.5 General Convergence Theorems ............................................ 398
- 11.6 Banach Space Valued Measures ............................................ 414
- 11.7 Calculation With Measures .................................................... 446
- 11.8 The Radon-Nikodym Theorem ................................................ 477
- 11.9 $L_p$ Spaces ........................................................................ 493
- 11.10 Dual of $C(\mathcal{X}, \mathcal{Y})$ ....................................................... 509
12 Differentiation and Integration 533
12.1 Carathéodory Extension Theorem . . . . . . . . . . . . . . 533
12.2 Change of Variable . . . . . . . . . . . . . . . . . . . . . . 539
12.3 Product Measure . . . . . . . . . . . . . . . . . . . . . . . 544
12.4 Functions of Bounded Variation . . . . . . . . . . . . . . . 571
12.5 Absolute and Lipschitz Continuity . . . . . . . . . . . . . . 595
12.6 Fundamental Theorem of Calculus . . . . . . . . . . . . . . 612
12.7 Representation of $(C_k(\Omega, Y))^*$ . . . . . . . . . . . . 641

13 Hilbert Spaces 647
13.1 Fundamental Notions . . . . . . . . . . . . . . . . . . . . . . 647
13.2 Projection Theorems . . . . . . . . . . . . . . . . . . . . . . 651
13.3 Dual of Hilbert Spaces . . . . . . . . . . . . . . . . . . . . . 653
13.4 Hermitian Adjoints . . . . . . . . . . . . . . . . . . . . . . . 656
13.5 Approximation in Hilbert Spaces . . . . . . . . . . . . . . . 658
13.6 Other Minimum Norm Problems . . . . . . . . . . . . . . . . 668
13.7 Positive Definite Operators on Hilbert Spaces . . . . . . . . 672
13.8 Pseudoinverse Operator . . . . . . . . . . . . . . . . . . . . . 676
13.9 Spectral Theory of Linear Operators . . . . . . . . . . . . . 681

14 Probability Theory 689
14.1 Fundamental Notions . . . . . . . . . . . . . . . . . . . . . . 689
14.2 Gaussian Random Variables and Vectors . . . . . . . . . . . 707
14.3 Law of Large Numbers . . . . . . . . . . . . . . . . . . . . . 708
14.4 Martingale Theory . . . . . . . . . . . . . . . . . . . . . . . 708
14.5 Central Limit Theorem . . . . . . . . . . . . . . . . . . . . . 713

15 Numerical Methods 715
15.1 Newton’s Method . . . . . . . . . . . . . . . . . . . . . . . . 715

A Elements in Calculus 719
A.1 Some Formulas . . . . . . . . . . . . . . . . . . . . . . . . . . 719
A.2 Convergence of Infinite Sequences . . . . . . . . . . . . . . . 719

Bibliography 721

Index 722
Chapter 1

Notations

\( \mathbb{N}, \mathbb{Z}, \text{ and } \mathbb{Q} \) the sets of natural number, integers, and rational numbers, respectively

\( \mathbb{R} \text{ and } \mathbb{C} \) 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

\( \mathbb{R}_+, \mathbb{R}_-, \mathbb{C}_+, \mathbb{C}_- \) \((0, \infty) \subset \mathbb{R}, (-\infty, 0) \subset \mathbb{R}\), the open right half of the complex plane, the open left half of the complex plane, respectively

\( \in \) 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 \( b \)
<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>the empty set; See Page 19.</td>
</tr>
<tr>
<td>${x, y}$</td>
<td>an unordered pair; See Page 19.</td>
</tr>
<tr>
<td>$\mathcal{P}(X)$</td>
<td>the collection of all subsets of $X$; See Page 19.</td>
</tr>
<tr>
<td>$\cup$</td>
<td>the set union; See Page 19.</td>
</tr>
<tr>
<td>$(x, y)$</td>
<td>an ordered pair; See Page 20.</td>
</tr>
<tr>
<td>$X \times Y$</td>
<td>the Cartesian or direct product of sets $X$ and $Y$; See Page 20.</td>
</tr>
<tr>
<td>$A = {x \in B \mid P(x)}$</td>
<td>Definition of a set $A$; See Page 20.</td>
</tr>
<tr>
<td>$x \sim y$</td>
<td>$x$ and $y$ are related in a relation; See Page 20.</td>
</tr>
<tr>
<td>$X/\equiv$</td>
<td>the quotient of the set $X$ with respect to an equivalence relation $\equiv$; See Page 20.</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 21.</td>
</tr>
<tr>
<td>graph($f$)</td>
<td>the graph of a function $f$; See Page 21.</td>
</tr>
<tr>
<td>dom($f$)</td>
<td>the domain of $f$; See Page 21.</td>
</tr>
<tr>
<td>$f(A)$</td>
<td>the image of $A \subseteq X$ under $f$; See Page 21.</td>
</tr>
<tr>
<td>range($f$)</td>
<td>the range of $f$, equals to $f(X)$; See Page 21.</td>
</tr>
<tr>
<td>$f_{\text{inv}}(A)$</td>
<td>the preimage of $A \subseteq Y$ under $f$; See Page 21.</td>
</tr>
<tr>
<td>onto, surjective</td>
<td>$f(X) = Y$; See Page 21.</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 21.</td>
</tr>
<tr>
<td>bijective</td>
<td>both surjective and injective; See Page 21.</td>
</tr>
<tr>
<td>$f_{\text{inv}}$</td>
<td>the inverse function of $f$; See Page 21.</td>
</tr>
<tr>
<td>$g \circ f$</td>
<td>the composition of $g: Y \to Z$ with $f: X \to Y$; See Page 21.</td>
</tr>
<tr>
<td>$f</td>
<td>_A$</td>
</tr>
<tr>
<td>$Y^X$</td>
<td>the set of all functions of $X$ to $Y$; See Page 21.</td>
</tr>
<tr>
<td>$\cap$</td>
<td>the set intersection; See Page 22.</td>
</tr>
<tr>
<td>$\overline{A}$</td>
<td>the compliment of a set $A$, where the whole set is clear from context; See Page 22.</td>
</tr>
<tr>
<td>$\setminus$</td>
<td>set minus; See Page 22.</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 22.</td>
</tr>
<tr>
<td>card($X$)</td>
<td>the number of elements in the finite set $X$; See Page 23.</td>
</tr>
<tr>
<td>$\Pi_{\alpha \in \Lambda}X_\alpha$</td>
<td>the Cartesian or direct product of $X_\alpha$’s; See Page 33.</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 33.</td>
</tr>
<tr>
<td>$\overline{A}$</td>
<td>the closure of a set $A$, where the whole set is clear from context; See Page 35.</td>
</tr>
<tr>
<td>$A^\circ$</td>
<td>the interior of a set $A$, where the whole set is clear from context; See Page 35.</td>
</tr>
<tr>
<td>$\partial A$</td>
<td>the boundary of a set $A$, where the whole set is clear from context; See Page 35.</td>
</tr>
</tbody>
</table>
$\prod_{\alpha \in \Lambda} (X_\alpha, O_\alpha)$ the product topological space; See Page 42.

$T_1$ Tychonoff space; See Page 47.

$T_2$ Hausdorff space; See Page 47.

$T_3$ regular space; See Page 47.

$T_4$ normal space; See Page 47.

$T_{3\frac{1}{2}}$ completely regular space; See Page 58.

$(x_\alpha)_{\alpha \in A}$ a net; See Page 59.

$\lim_{\alpha \in A} x_\alpha$ the limit of a net; See Page 59.

$\lim_{x \to x_0} f(x)$ the limit of $f(x)$ as $x \to x_0$; See Page 63.

$\mathbb{R}_e$ the set of extended real numbers, which equals to $\mathbb{R} \cup \{-\infty, +\infty\}$; See Page 66.

$\limsup_{\alpha \in A} x_\alpha$ the limit superior of a real-valued net; See Page 67.

$\liminf_{\alpha \in A} x_\alpha$ the limit inferior of a real-valued net; See Page 67.

$\limsup_{x \to x_0} f(x)$ the limit superior of $f(x)$ as $x \to x_0$; See Page 68.

$\liminf_{x \to x_0} f(x)$ the limit inferior of $f(x)$ as $x \to x_0$; See Page 68.

$B_X(x_0, r)$ the open ball centered at $x_0$ with radius $r$; See Page 73.

$\bar{B}_X(x_0, r)$ the closed ball centered at $x_0$ with radius $r$; See Page 73.

$\text{dist}(x_0, S)$ the distance from a point $x_0$ to a set $S$ in a metric space; See Page 75.

$(X, \rho_X) \times (Y, \rho_Y)$ the finite product metric space; See Page 80.

$(\prod_{i=1}^{\infty} X_i, \rho)$ the countably infinite product metric space; See Page 83.

$\text{supp} f$ the support of a function; See Page 117.

$\beta(X)$ the Stone-Čech compatification of a completely regular topological space $X$; See Page 128.

$(\mathcal{M}(A, Y), \mathcal{F})$ the vector space of $Y$-valued functions of a set $A$ over the field $\mathcal{F}$; See Page 136.

$\vartheta_X$ the null vector of a vector space $X$; See Page 138.

$\mathcal{N}(A)$ the null space of a linear operator $A$; See Page 138.

$\mathcal{R}(A)$ the range space of a linear operator $A$; See Page 138.

$\alpha S$ the scalar multiplication by $\alpha$ of a set $S$ in a vector space; See Page 139.

$S + T$ the sum of two sets $S$ and $T$ in a vector space; See Page 139.

$\text{span}(A)$ the subspace generated by the set $A$; See Page 140.

$v(P)$ the linear variety generated by a nonempty set $P$; See Page 140.

$\text{co}(S)$ the convex hull generated by $S$ in a vector space; See Page 141.

$\|x\|$ the norm of a vector $x$; See Page 147.

$|x|$ the Euclidean norm of a vector $x$; See Page 147.
$C_1([a,b])$ the normed linear space of continuously differentiable real-valued functions on the interval $[a,b]$; See Page 148.

$l_p$ The normed linear space of real-valued sequences with finite $p$-norm, $1 \leq p \leq +\infty$; See Page 149.

$l_p(X)$ the normed linear space of $X$-valued sequences with finite $p$-norm, $1 \leq p \leq +\infty$; See Page 153.

$V(P)$ the closed linear variety generated by a nonempty set $P$; See Page 156.

$^oP$ the relative interior of a set $P$; See Page 156.

$X \times Y$ the finite Cartesian product normed linear space; See Page 156.

$C([a,b])$ the Banach space of continuous real-valued functions on the interval $[a,b]$; See Page 160.

$C(K,X)$ the normed linear space of continuous $X$-valued functions on a compact space $K$; See Page 160.

$X_{IR}$ the real normed linear space induced by a complex normed linear space $X$; See Page 164.

$[x]$ the coset of a vector $x$ in a quotient space; See Page 165.

$X/M$ the quotient space of a vector space $X$ modulo a subspace $M$; See Page 165.

$X/M$ the quotient space of a normed linear space $X$ modulo a closed subspace $M$; See Page 166.

$C_v(X,Y)$ the vector space of continuous $Y$-valued functions on $X$; See Page 168.

$B(X,Y)$ the set of bounded linear operators of $X$ to $Y$; See Page 175.

$X^*$ the dual of $X$; See Page 179.

$x_*$ a vector in the dual; See Page 179.

$\langle\langle x_*, x \rangle\rangle$ the evaluation of a bounded linear functional $x_*$ at the vector $x$, that is $x_*(x)$; See Page 179.

$c_0(X)$ the subspace of $l_\infty(X)$ consisting of $X$-valued sequences with limit $\vartheta_X$; See Page 183.

$X^{**}$ the second dual of $X$; See Page 192.

$S^\perp$ the orthogonal complement of the set $S$; See Page 194.

$\perp S$ the pre-orthogonal complement of the set $S$; See Page 194.

$A'$ the adjoint of a linear operator $A$; See Page 203.

$A''$ the adjoint of the adjoint of a linear operator $A$; See Page 204.

$\mathcal{O}_{\text{weak}}(X)$ the weak topology on a normed linear space $X$; See Page 206.
\(X_{\text{weak}}\) the weak topological space associated with a normed linear space \(X\); See Page 207.

\(O_{\text{weak}^*}(X^*)\) the weak* topology on the dual of a normed linear space \(X\); See Page 209.

\(X^*_{\text{weak}^*}\) the weak* topological space associated with a normed linear space \(X\); See Page 210.

\(K_{\text{supp}}\) the support of a convex set \(K\); See Page 222.

\([f, C]\) the epigraph of a convex function \(f : C \rightarrow \mathbb{R}\); See Page 225.

\(C_{\text{conj}}\) the conjugate convex set; See Page 229.

\(f_{\text{conj}}\) the conjugate convex functional; See Page 229.

\([f, C]_{\text{conj}}\) the epigraph of the conjugate convex functional; See Page 230.

\(\text{conj} \Gamma\) the pre-conjugate convex set; See Page 232.

\(\text{conj} \varphi\) the pre-conjugate convex functional; See Page 232.

\(\text{conj}[\varphi, \Gamma]\) the epigraph of the pre-conjugate convex functional; See Page 232.

\(\succ\) greater than or equal to (with respect to the positive cone); See Page 243.

\(\preceq\) less than or equal to (with respect to the positive cone); See Page 243.

\(\succ\) greater than (with respect to the positive cone); See Page 243.

\(\prec\) less than (with respect to the positive cone); See Page 243.

\(S^{\oplus}\) the positive conjugate cone of a set \(S\); See Page 243.

\(S^{\ominus}\) the negative conjugate cone of a set \(S\); See Page 243.

\(A_D(x_0)\) the set of admissible deviations in \(D\) at \(x_0\); See Page 253.

\(f^{(1)}(x_0), Df(x_0)\) the Fréchet derivative of \(f\) at \(x_0\); See Page 254.

\(Df(x_0; u)\) the directional derivative of \(f\) at \(x_0\) along \(u\); See Page 254.

\(\partial f\partial y(x_0, y_0)\) partial derivative of \(f\) with respect to \(y\) at \((x_0, y_0)\); See Page 255.

\(\text{ro}(A)(B)\) right operate; \(\text{ro}(A)(B) = BA\); See Page 260.

\(B_k(X, Y)\) the set of bounded multi-linear \(Y\)-valued functions on \(X^k\); See Page 267.

\(B_{S,k}(X, Y)\) the set of symmetric bounded multi-linear \(Y\)-valued functions on \(X^k\); See Page 267.

\(D^k f(x_0), f^{(k)}(x_0)\) the \(k\)th order Fréchet derivative of \(f\) at \(x_0\); See Page 267.

\(\mathcal{C}_k, \mathcal{C}_\infty\) \(k\)-times and infinite-times continuously differentiable functions, respectively; See Page 268.
\[ \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \]  

\( \partial \) \( k \)th-order partial derivative of \( f \); See Page 273.

\[ \text{Sm} (D) = \bigcup_{\alpha \in \mathbb{K}} \alpha D \]  

\( \text{Sm} \) \( (D) \) \( = \bigcup_{\alpha \in \mathbb{K}} \alpha D \); See Page 273.

\[ \mathcal{C}_k (\Omega, \mathcal{Y}) \]  

the normed linear space of \( k \)-times continuously differentiable \( \mathcal{Y} \)-valued functions on a compact set \( \Omega \subseteq \mathcal{X} \); See Page 309.

\[ \mathcal{C}_b (\mathcal{X}, \mathcal{Y}) \]  

the normed linear space of bounded continuous \( \mathcal{Y} \)-valued functions on a topological space \( \mathcal{X} \); See Page 309.

\[ \mathcal{C}_{b,k} (\Omega, \mathcal{Y}) \]  

the normed linear space of \( k \)-times bounded continuously differentiable \( \mathcal{Y} \)-valued functions on a set \( \Omega \subseteq \mathcal{X} \); See Page 310.

\[ A^{T_{n_1 \cdots n_m}} \]  

the transpose of an \( m \)th order tensor \( A \) with the permutation \( (n_1, \ldots, n_m) \); See Page 311.

\[ A \otimes B \]  

an \( n \)th order \( \mathbb{K} \)-valued tensor in \( B(\mathbb{K}^{m_n}, \ldots, B(\mathbb{K}^{m_1}, \mathbb{K}) \cdots) \) with all elements equal to 0; See Page 313.

\[ 1_{m_1 \cdots m_n} \]  

an \( n \)th order \( \mathbb{K} \)-valued tensor in \( B(\mathbb{K}^{m_n}, \ldots, B(\mathbb{K}^{m_1}, \mathbb{K}) \cdots) \) with all elements equal to 1; See Page 313.

\[ \exp(A) \]  

the exponential function of a linear operator \( A \) on a Banach space; See Page 324.

\[ S_+ \mathcal{X}, S_{\text{psd}} \mathcal{X} \]  

sets of positive definite and positive semi-definite operators over the real normed linear space \( \mathcal{X} \), respectively; See Page 326.

\[ S_- \mathcal{X}, S_{\text{nsd}} \mathcal{X} \]  

sets of negative definite and negative semi-definite operators over the real normed linear space \( \mathcal{X} \), respectively; See Page 326.

\[ S_{\mathcal{X}} \]  

\( B_{S_2}(\mathcal{X}, \mathbb{R}) \); See Page 326.

\[ (\mathbb{R}, \mathcal{B}_L, \mu_L) \]  

Lebesgue measure space; See Page 360.

\( \mu_{L_0} \)  

Lebesgue outer measure; See Page 361.

\[ \mathcal{B}_B(\mathcal{X}) \]  

Borel sets; See Page 361.

\( \mathcal{A}(\mathcal{X}) \)  

the algebra generated by the topology of a topological space \( \mathcal{X} \); See Page 361.

\[ \mu_B \]  

the Borel measure on \( \mathbb{R} \); See Page 362.

\( \mathbb{R} \)  

\( (\mathbb{R}, \mathcal{B}(\mathbb{R}) \cup \mathcal{B}_B(\mathbb{R}), \mu_B) \); See Page 366.

\( P \)  

a.e. in \( \mathcal{X} \)  

\( P \) holds almost everywhere in \( \mathcal{X} \); See Page 370.

\( P(x) \)  

a.e. \( x \in \mathcal{X} \)  

\( P \) holds almost everywhere in \( \mathcal{X} \); See Page 370.

\( \mathcal{P} \circ f \)  

\( \mathcal{P} \circ f : \mathcal{X} \to [0, \infty) \subseteq \mathbb{R} \) defined by \( \mathcal{P} \circ f(x) = \| f(x) \|, \forall x \in \mathcal{X} \); See Page 380.

\( \mathcal{R}(\mathcal{X}) \)  

the collection of all representation of \( \mathcal{X} \); See Page 383.

\( \mathcal{I}(\mathcal{X}) \)  

the integration system on \( \mathcal{X} \); See Page 383.

\[ \int_{\mathcal{X}} f \, d\mu \]  

the integral of a function \( f \) on a set \( \mathcal{X} \) with respect to measure \( \mu \); See Page 384.
\[ \int_X f(x) \, d\mu(x) \] the integral of a function \( f \) on a set \( X \) with respect to measure \( \mu \); See Page 384.

\( \mathcal{P} \circ \mu \) the total variation of a Banach space valued pre-measure \( \mu \); See Page 414.

\( \mathcal{P} \circ \mu \) the total variation of a Banach space valued measure \( \mu \); See Page 419.

\( \mu_1 + \mu_2 \) the \( Y \)-valued measure that equals to the sum of two \( Y \)-valued measures on the same measurable space; See Page 449.

\( \alpha \mu \) the \( Y \)-valued measure that equals to the scalar product of \( \alpha \in \mathbb{K} \) and \( Y \)-valued measure \( \mu \); See Page 450.

\( \mu y \) the \( Y \)-valued measure that equals to scalar product of a \( \mathbb{K} \)-valued measure \( \mu \) and \( y \in Y \); See Page 450.

\( A \mu \) the \( Z \)-valued measure that equals to product of an bounded linear operator \( A \) and a \( Y \)-valued measure \( \mu \); See Page 450.

\( M_\sigma(X, \mathcal{B}, Y) \) the vector space of \( \sigma \)-finite \( Y \)-valued measures on the measurable space \((X, \mathcal{B})\); See Page 456.

\( M_f(X, \mathcal{B}, Y) \) the normed linear space of finite \( Y \)-valued measures on the measurable space \((X, \mathcal{B})\); See Page 457.

\( \lim_{n \in \mathbb{N}} \mu_n = \nu \) a sequence of \( \sigma \)-finite \( (Y \text{-valued}) \) measures \((\mu_n)_{n=1}^{\infty} \) converges to a \( \sigma \)-finite \( (Y \text{-valued}) \) measure \( \nu \); See Page 460.

\( \mu_1 \leq \mu_2 \) 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 460.

\[
\begin{bmatrix}
\mu_{1,1} & \cdots & \mu_{1,m} \\
\vdots & & \vdots \\
\mu_{n,1} & \cdots & \mu_{n,m}
\end{bmatrix}
\]
the vector measure; See Page 463.

\( \mu_1 \ll \mu_2 \) the measure \( \mu_1 \) is absolutely continuous with respect to the measure \( \mu_2 \); See Page 473.

\( \mu_1 \perp \mu_2 \) the measures \( \mu_1 \) and \( \mu_2 \) are mutually singular; See Page 473.

\( \frac{d \nu}{d \mu} \) the Radon-Nikodym derivative of the \( \sigma \)-finite \( Y \)-valued measure \( \nu \) with respect to the \( \sigma \)-finite \( \mathbb{K} \)-valued measure \( \mu \); See Page 483.

\( \mathcal{P}_p \circ f \) the function \( \| f(\cdot) \|^p \); See Page 493.

\( \text{ess sup} \) the essential supremum; See Page 495.

\( \lim_{n \in \mathbb{N}} z_n = z \text{ in } \overline{L}_p \) the sequence \((z_n)_{n=1}^{\infty} \subseteq \overline{L}_p \) converges to \( z \in \overline{L}_p \) in \( \overline{L}_p \) pseudo-norm; See Page 497.

\( M_{f_1}(X, Y) \) the normed linear space of finite \( Y \)-valued topological measures on \( X \); See Page 514.

\( M_{\sigma f}(X, Y) \) the vector space of \( \sigma \)-finite \( Y \)-valued topological measures on \( X \); See Page 514.
CHAPTER 1. NOTATIONS

\( \mathcal{M}_\sigma (X, B) \) the set of \( \sigma \)-finite measures on the measurable space \((X, B)\); See Page 515.

\( \mathcal{M}_f (X, B) \) the set of finite measures on the measurable space \((X, B)\); See Page 515.

\( \mathcal{M}_{\sigma t} (X) \) the set of \( \sigma \)-finite topological measures on the topological space \(X\); See Page 515.

\( \mathcal{M}_{ft} (X) \) the set of finite topological measures on the topological space \(X\); See Page 515.

\( \prod_{i=1}^m \mu_i \) the product measure of \( \mu_1, \ldots, \mu_m \); See Page 547.

\( \prod_{i=1}^m X_i \) product topological measure space of \(X_1, \ldots, X_m\), where \(m \in \mathbb{N}\); See Page 564.

\( r_{x_1, x_2} \) the semi-open rectangle in \(\mathbb{R}^m\) with corners \(x_1\) and \(x_2\); See Page 572.

\( t_{x_1, x_2} \) the closed rectangle in \(\mathbb{R}^m\) with corners \(x_1\) and \(x_2\); See Page 573.

\( r^0_{x_1, x_2} \) the open rectangle in \(\mathbb{R}^m\) with corners \(x_1\) and \(x_2\); See Page 573.

\( P(\Omega) \) the principal of a region \(\Omega\); See Page 573.

\( V_{\text{Rect}} i, x_1, x_2 \) the set of vertexes of \(r_{x_1, x_2}\) with \(i\) coordinates equal to that of \(x_1\); See Page 575.

\( \Delta F (r_{x_1, x_2}) \) the increment of \(F\) on \(r_{x_1, x_2}\); See Page 575.

\( T F (r_{x_1, x_2}) \) the total variation of a function \(F\) on the semi-open rectangle \(r_{x_1, x_2}\); See Page 575.

\( (\mathbb{R}^m, \mathcal{B}_L m, \mu_L m) \) \(m\)-dimensional Lebesgue measure space and the \(m\)-dimensional Lebesgue outer measure; See Page 594.

\( \int_U g(x) \, dF(x) \) the integral of function \(g\) with respect to \(Y\)-valued measure space \((P(\Omega), \mathcal{B}_B (P(\Omega)), \mu)\) whose distribution function is \(F: \Omega \rightarrow Y\) over the set \(U \in P(\Omega)\); See Page 595.

\( \Pi_{i=0}^m \) \(\pi_{i_0} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1} \pi_{i_0} (x) = (\pi_1 (x), \ldots, \pi_{i_0-1} (x), \pi_{i_0+1} (x), \ldots, \pi_m (x)), \forall x \in \mathbb{R}^m \); See Page 595.

\( \int_a^b f(x) \, dx \) the integral of function \(f\) from \(a \in \mathbb{R}\) to \(b \in \mathbb{R}\) with respect to \(\mu_B\); See Page 606.

\( \text{dia} (S) \) the diameter of a subset of a metric space; See Page 613.

\( e_{m,i} \) the \(i\)th unit vector in \(\mathbb{R}^m\); See Page 616.

\( \langle x, y \rangle \) the inner product of vectors \(x\) and \(y\); See Page 647.

\( x \perp y \) the vectors \(x\) and \(y\) are orthogonal in a pre-Hilbert space; See Page 649.

\( x \perp S \) the vector \(x\) is orthogonal to the set \(S\) in a pre-Hilbert space; See Page 649.

\( x^* \) \(x^* \in \mathbb{X}\) that satisfies \(\langle x^*, y \rangle = \langle y, x \rangle, \forall y \in \mathbb{X}\); See Page 653.

\( M \oplus N \) the direct sum of \(M\) and \(N\); See Page 655.

\( A^* \) the Hermitian adjoint of \(A\); See Page 656.
Gram($y_1, \ldots, y_n$) the Gram matrix of $y_1, \ldots, y_n$; See Page 660.
gram($y_1, \ldots, y_n$) the Gram determinant of $y_1, \ldots, y_n$; See Page 660.
$S_X$ set of Hermitian operators on a real or complex Hilbert space; See Page 672.
$S_{+X}, S_{psdX}$ sets of positive definite and positive semi-definite operators over real or complex Hilbert space $X$, respectively; See Page 672.
$S_{-X}, S_{nsdX}$ sets of negative definite and negative semi-definite operators over real or complex Hilbert space $X$, respectively; See Page 672.
$A^\dagger$ the pseudoinverse of $A \in B(X, Y)$; See Page 678.
$E(x)$ expectation of $x$; See Page 689.
$E(x|\tilde{B})$ conditional expectation of $x$ given $\tilde{B}$; See Page 692.
$L_x$ the law for the random variable $x$; See Page 707.
$x \sim N(\bar{x}, K)$ Gaussian (normal) random variable (vector) with mean $\bar{x}$ and covariance $K$; See Page 707.
$P(F|\tilde{B})$ The conditional probability of event $F$ happens given $\tilde{B}$, which is defined to be a version of $E(\chi_{F,\Omega}|\tilde{B})$; See Page 710.
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 \( A^2 \).

**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 \( (x, y) := \{\{ \{ x \} \}, 1, \{\{ y \} \}, 2 \} \). 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 \( A_s \) of a well-defined set \( A \) is again a set, which is called a subset of \( A \), we will write \( A_s \subseteq A \). 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 \( A \times B \). \( \forall x \in A, \forall y \in B \), we say \( x \sim y \) 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

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

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 an 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 the set $\text{dom}(f) = D$. \( \forall A \subseteq X \), 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 the set $\text{range}(f) = f(X)$. \( \forall B \subseteq Y \), the inverse image under $f$ of $B$ is $f^{-1}(B) := \{x \in D \mid f(x) \in B\}$, which is a subset of $D$. $f$ is said to be surjective if $f(X) = Y$; and $f$ is said to be injective if $f(x_1) \neq f(x_2)$, \( \forall x_1, x_2 \in D \) with $x_1 \neq x_2$; $f$ is said to be bijective if it is both surjective and injective, in which case it is invertible and the inverse function is denoted by $f^{-1} : Y \to D$. 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, $f \in X \times Y^2$. The collection of all functions of $X$ to $Y$ is then a set given by

\[
Y^X := \{ f \in X \times Y^2 \mid \forall x \in X, \exists y \in Y \ni (x, y) \in f \}
\]

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 $\phi \circ \psi_i = \text{id}_Y$ 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. \( \forall y \in Y \), $\phi \circ \psi_i(y) = \text{id}_Y(y) = y$, 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, $\psi_1 = \psi_2 = \phi^{-1}$. 

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

\[
\psi_1 = \text{id}_X \circ \psi_1 = (\phi_{\text{inv}} \circ \phi) \circ \psi_1 = \phi_{\text{inv}} \circ (\phi \circ \psi_1) = \phi_{\text{inv}} \\
\psi_2 = \psi_2 \circ \text{id}_Y = \psi_2 \circ (\phi \circ \phi_{\text{inv}}) = (\psi_2 \circ \phi) \circ \phi_{\text{inv}} = \text{id}_X \circ \phi_{\text{inv}} = \phi_{\text{inv}}
\]

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

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, $\mathcal{P}(X)$ is the set consisting of all subsets of $X$. For any $A, B \subseteq X$, 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 $A, B, D, A_\lambda \in \mathcal{P}(X)$, $f : D \to Y$, $C, E, C_\lambda \in \mathcal{P}(Y)$, 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. $\emptyset = X$, $\tilde{\emptyset} = A$, $A \cup \tilde{A} = X$, and $A \cap \tilde{A} = \emptyset$, if, and only if, $\tilde{B} \subseteq A$;
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. $B \cup \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) = \left(\bigcap_{\lambda \in \Lambda} (B \cup A_\lambda)\right)$ and $B \cap \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) = \left(\bigcup_{\lambda \in \Lambda} (B \cap A_\lambda)\right)$;
7. $f \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) = \bigcup_{\lambda \in \Lambda} f(A_\lambda)$ and $f \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) \subseteq \bigcap_{\lambda \in \Lambda} f(A_\lambda)$;
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 \( \mathcal{A} \subseteq \mathcal{P}(X) \). \( \mathcal{A} \) is said to be an algebra of sets on \( X \) (or a Boolean algebra on \( X \)) if

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

(ii) \( \forall A, B \in \mathcal{A}, A \cup B \in \mathcal{A} \) and \( \sim A \in \mathcal{A} \).

\( \mathcal{A} \) is said to be a \( \sigma \)-algebra on \( X \) if it is an algebra on \( X \) and countable unions of sets in \( \mathcal{A} \) is again in \( \mathcal{A} \).

Let \( \mathcal{M} \subseteq \mathcal{P}(X) \), where \( X \) is a set, then, there exists a smallest algebra on \( X \), \( \mathcal{A}_0 \subseteq \mathcal{P}(X) \), containing \( \mathcal{M} \), which means that any algebra on \( X \), \( \mathcal{A}_1 \subseteq \mathcal{P}(X) \), that contains \( \mathcal{M} \), we have \( \mathcal{A}_0 \subseteq \mathcal{A}_1 \). This algebra \( \mathcal{A}_0 \) is said to be the algebra on \( X \) generated by \( \mathcal{M} \). Also, there exists a smallest \( \sigma \)-algebra on \( X \), \( \mathcal{A} \subseteq \mathcal{P}(X) \), containing \( \mathcal{M} \), which is said to be the \( \sigma \)-algebra on \( X \) generated by \( \mathcal{M} \).

Proposition 2.8 Let \( X \) be a set, \( \mathcal{E} \) be a nonempty collection of subsets of \( X \), \( \mathcal{A} \) be the algebra on \( X \) generated by \( \mathcal{E} \), and

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

Then, \( \mathcal{A} = \mathcal{A} \).

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

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

Note that $\tilde{A} = \bigcap_{i_1=1}^{\tilde{E}} \cap_{i_2=1}^{\tilde{E}} \cdots \cap_{i_{2n_A}=1}^{\tilde{E}} (F^{(A)}_{i_1,\ldots,i_{2n_A}})^\sim$. Let $n = n_A + 1$, $m = m_A$, $\forall i_1, \ldots, i_{2n} \in \{1, \ldots, m\}$, $G_{i_1,\ldots,i_{2n}} = (F^{(A)}_{i_1,\ldots,i_{2n-1}})^\sim$. Then, $\tilde{A} = \bigcup_{i_1=1}^{m} \cap_{i_2=1}^{m} \cdots \cap_{i_{2n}=1}^{m} G_{i_1,\ldots,i_{2n}} \in \tilde{A}$.

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

$$G_{i_1,\ldots,i_{2n}} = \begin{cases} F^{(A)}_{i_1,\ldots,i_{2n}} & \text{if } i_1 \leq m_A \\ F^{(B)}_{\tilde{i}_1,\ldots,\tilde{i}_{2n}} & \text{if } i_1 > m_A \end{cases}$$

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

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

### 2.6 Partial Ordering and Total Ordering

**Definition 2.9** Let $A$ be a set and $\leq$ be a relation on $A$. $\leq$ 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 $x \leq y$ or $y \leq x$ (not both).

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

**Definition 2.10** Let $A$ be a set with a partial ordering “$\leq$”.

1. $a \in A$ is said to be minimal if, $\forall x \in A$, $x \leq a$ implies $a \leq x$;
2. $a \in A$ is said to be the least element if, $\forall x \in A$, $a \leq x$, and $x \leq a$ implies that $x = a$.
3. $a \in A$ is said to be maximal if, $\forall x \in A$, $a \leq x$ implies $x \leq a$;
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 \( \leq \), and \( E \subseteq A \).

1. \( a \in A \) is said to be an upper bound of \( E \) if \( x \leq a, \ \forall x \in E \). 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 \( a \leq x, \ \forall x \in E \). 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 \( \leq \). 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 \( \geq \) by \( \forall x, y \in A, \ x \geq y \) if \( y \leq x \). Then, \( \geq \) is a partial ordering on \( A \). Furthermore, \( \geq \) is antisymmetric if \( \leq \) is antisymmetric.

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

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

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

**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 \( a_1 \leq a_2 \). 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 \( x \leq x \) implies \( y \geq x \), then \( \geq \) is reflexive. If \( x \geq y \) and \( y \geq z \), we have \( y \leq x \) and \( z \leq y \), which implies \( z \leq x \), and hence
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. $\square$

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 exists 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 $z \prec x$, 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 arbitraryness 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.

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 \( z \prec e_0 \), implies that \( z \) must not be an upper bound of \( F \). Therefore, \( \exists x_0 \in F \) such that \( x_0 \not\preceq z \). By (v), we have \( z \prec x_0 \). Hence, by \( x_0 \in F \subseteq P \), \( z \in A \), and \( z \prec x_0 \), 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 \( x_1 \preceq x_2 \), then \( x_1 \) and \( x_2 \) are related through \( \preceq \). If \( f(x_2) \preceq x_1 \), then \( x_2 \leq f(x_2) \preceq x_1 \), 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 \( \preceq \) 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 \( \preceq \) is antisymmetric.

This completes the proof of the lemma.

\[ \blacksquare \]

**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

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

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

\( \forall A \in \mathcal{E} \), define a collection

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

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

Clearly, \( C \subseteq \mathcal{E} \). We will show \( \preceq \) is a total ordering on \( C \). Since \( \preceq \) is a partial ordering on \( \mathcal{E} \), 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 \( B_2 \subseteq B \subseteq \mathcal{E} \), 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, \( C \in \mathcal{E} \). This shows that \( B \) admits least upper bound \( C \) in \( \mathcal{E} \) with respect to \( \subseteq \). By the definition of \( T \), it is clear that \( A \subseteq T(A) \), \( \forall A \in \mathcal{E} \). By Lemma 2.17, \( T \) has a fixed point on \( \mathcal{E} \), i.e., \( \exists A_0 \in \mathcal{E} \) such that \( T(A_0) = A_0 \).

By the definitions of \( T \) and \( A_{A_0} \), there does not exist \( B \in \mathcal{E} \) such that \( A_0 \subseteq B \). Hence, by Proposition 2.12 (vii), \( A_0 \) is maximal in \( \mathcal{E} \) 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 \( e_0 \in F \).

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 \( x_2 \preceq x_1 = e_0 \). In Case 4, we have \( x_1 = e_0 \preceq e_0 = x_2 \). 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 \( e_0 \preceq e_1 \). \( \forall x \in E \), 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

\[
\mathcal{E} := \{ (A_\alpha, \preceq_\alpha) \mid A_\alpha \subseteq E, \ A_\alpha \text{ is well-ordered by } \preceq_\alpha \}
\]
2.7. BASIC PRINCIPLES

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) \[ \preceq_2 = \preceq_1 \text{ on } A_1 \]; (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 $E$. $\forall (A_1, \preceq_1), (A_2, \preceq_2), (A_3, \preceq_3) \in E$. Clearly, $(A_1, \preceq_1) \preceq (A_1, \preceq_1)$. Hence, $\preceq$ is reflexive. If $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$ and $(A_2, \preceq_2) \preceq (A_3, \preceq_3)$, we have $A_1 \subseteq A_2 \subseteq A_3$, and (i) holds; (ii) $\preceq_3 = \preceq_2$ on $A_2$ and $\preceq_2 = \preceq_1$ on $A_1$ implies that $\preceq_3 = \preceq_1$ 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 $x_1 \in A_2$ and $x_1 \preceq x_2$, thus, we have $x_1 \preceq x_2$ in both cases. Therefore, $(A_1, \preceq_1) \preceq (A_3, \preceq_3)$ and hence $\preceq$ is transitive. If $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$ and $(A_2, \preceq_2) \preceq (A_1, \preceq_1)$, then $A_1 \subseteq A_2 \subseteq A_1 \Rightarrow A_1 = A_2$ and $\preceq_2 = \preceq_1$ on $A_1$. Hence, $(A_1, \preceq_1) = (A_2, \preceq_2)$, 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 $A = \{ (A_0, \preceq_0) \mid \alpha \in \Lambda \}$ where $\Lambda \neq \emptyset$ is an index set. Define $A := \bigcup_{\alpha \in \Lambda} A_\alpha$. Define an ordering $\preceq$ on $A$ by: $\forall x_1, x_2 \in A$, $\exists (A_1, \preceq_1), (A_2, \preceq_2) \in A$ such that $x_1 \in A_1$ and $x_2 \in A_2$, without loss of generality, assume that $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$ 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 $(A_2, \preceq_2) \in A$. Let $(A_3, \preceq_3) \in A$ 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: $(A_3, \preceq_3) \preceq (A_2, \preceq_2)$; Case 2: $(A_2, \preceq_2) \preceq (A_3, \preceq_3)$. In Case 1, we have $A_3 \subseteq A_2$ and $\preceq_3 = \preceq_2$ on $A_3$, which implies that $x_1 \preceq x_2 \iff x_1 \preceq x_2 \iff x_1 \preceq x_2$. In Case 2, we have $A_2 \subseteq A_3$ and $\preceq_3 = \preceq_2$ on $A_2$, which implies that $x_1 \preceq x_2 \iff x_1 \preceq x_2 \iff x_1 \preceq x_2$. Hence, the ordering $\preceq$ is well-defined on $A$.

Next, we will show that $\preceq$ is a total ordering on $A$. $\forall x_1, x_2, x_3 \in A$, $\exists (A_i, \preceq_i) \in A$ 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 $(A_1, \preceq_1) \preceq (A_2, \preceq_2) \preceq (A_3, \preceq_3)$. Then, $x_1, x_2, x_3 \in A_3$. Clearly, $x_1 \preceq x_1$ since $x_1 \preceq x_1$, 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 $\preceq_3$ 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 $\preceq_3$ is antisymmetric on $A_3$. This shows that $\preceq$ is antisymmetric. Since $\preceq_3$ is a well-ordering on $A_3$, then we must have $x_1 \preceq x_2 \iff x_1 \preceq x_2$ or $x_2 \preceq x_1 \iff x_2 \preceq x_1$. Hence, $\preceq$ defines a total ordering on $A$.

Next, we will show that $\preceq$ is a well-ordering on $A$. $\forall B \subseteq A$ with $B \neq \emptyset$. Fix $x_0 \in B$. Then, $\exists (A_1, \preceq_1) \in A$ 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 $\preceq_1$, then $\exists e \in B \cap A_1$, which is the least element of $B \cap A_1$. $\forall y \in B \subseteq A$, $\exists (A_2, \preceq_2) \in A$ 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, $e \preceq y$ since $e$ is the least element.
of $B \cap A_1$, which implies that $e \leq y$. In Case 2, since $\mathcal{A}$ 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 $\leq$ is a total ordering on $A$, then $e$ is the least element of $B$. Therefore, $\preceq$ is a well ordering on $A$, which implies $(A, \preceq) \in \mathcal{E}$.

For all $(A_1, \preceq_1) \in \mathcal{A}$, (i) $A_1 \subseteq A$. (ii) $\forall x_1, x_2 \in A_1$, $x_1 \preceq_1 x_2 \iff x_1 \preceq x_2$; hence $\preceq = \preceq_1$ on $A_1$. (iii) $\forall x_1 \in A_1$, $\forall x_2 \in A \setminus A_1$, $\exists (A_2, \preceq_2) \in \mathcal{A}$ such that $x_2 \in A_2 \setminus A_1$; since $\mathcal{A}$ is totally ordered by $\preceq$, then, we must have $(A_1, \preceq_1) \preceq (A_2, \preceq_2)$; hence $x_1 \preceq_2 x_2$ and $x_1 \preceq x_2$. Therefore, we have shown $(A_1, \preceq_1) \preceq (A, \preceq)$. Hence, $(A, \preceq) \in \mathcal{E}$ is an upper bound of $A$.

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 $\exists x_0 \in E \setminus F$. Define an ordering $\leq_H$ on $H$ by: $\forall x_1, x_2 \in H$, if $x_1, x_2 \in F$, we say $x_1 \leq_H x_2$ if $x_1 \leq_F x_2$; if $x_1 \notin F$ and $x_2 \in H$, then we let $x_1 \leq_H x_2$; if $x_1 = x_2 = x_0$, we let $x_1 \leq_H x_2$. Now, we will show that $\leq_H$ is a well ordering on $H$. $\forall x_1, x_2, x_3 \in H$. If $x_1 \in F$, then, $x_1 \leq_F x_1$ and $x_1 \leq_H x_1$; if $x_1 = x_0$, then $x_1 \leq_H x_1$. Hence, $\leq_H$ is reflexive. If $x_1 \leq_H x_2$ and $x_2 \leq_H x_3$, we have 4 exhaustive and mutually exclusive cases: Case 1: $x_1, x_3 \in F$; Case 2: $x_1 \in F$ and $x_3 = x_0$; Case 3: $x_3 \in F$ and $x_1 = x_0$; Case 4: $x_1 = x_3 = x_0$. In Case 1, we must have $x_2 \in F$ and then $x_1 \leq_F x_2$ and $x_2 \leq_F x_3$, which implies that $x_1 \leq_F x_3$, and hence $x_1 \leq_H x_3$. In Case 2, we have $x_1 \leq_H x_3$. In Case 3, we must have $x_2 = x_0$, which leads to a contradiction $x_0 \leq_H x_3$, hence, this case is impossible. In Case 4, we have $x_1 \leq_H x_3$. In all cases except that is impossible, we have $x_1 \leq_H x_3$. Hence, $\leq_H$ is transitive. If $x_1 \leq_H x_2$ and $x_2 \leq_H x_1$. 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 $x_1 \leq_F x_2$ and $x_2 \leq_F x_1$, which implies that $x_1 = x_2$ since $\leq_F$ is antisymmetric on $F$. In Case 2, we have $x_0 \leq_H x_1$, which is a contradiction, and hence this case is impossible. In Case 3, we have $x_0 \leq_H x_2$, 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, $\leq_H$ is antisymmetric. When $x_1, x_2 \in F$, then, we must have $x_1 \leq_F x_2$ or $x_2 \leq_F x_1$ since $\leq_F$ is a well ordering on $F$, and hence $x_1 \leq_H x_2$ or $x_2 \leq_H x_1$. When $x_1 \in F$ and $x_2 = x_0$, then $x_1 \leq_H x_2$. When $x_2 \in F$ and $x_1 = x_0$, then $x_2 \leq_H x_1$. When $x_1 = x_2 = x_0$, then $x_1 \leq_H x_2$. This shows that $\leq_H$ 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 $\leq_F$. $\forall x \in B$, if $x \in B \setminus \{x_0\}$, then $e_0 \leq_F x$ and hence $e_0 \leq_H x$; if $x = x_0$, then $e_0 \leq_H x$. Hence, $e_0$ is the least element of $B$ since $\leq_H$ is a total ordering on $H$. Therefore, $\leq_H$ is a well ordering on $H$ and $(H, \leq_H) \in \mathcal{E}$.

Clearly, $F \subset H$, $\leq_F = \leq_H$ on $F$, and $\forall x_1 \in F$ and $\forall x_2 \in H \setminus F$, we have $x_2 = x_0$ and $x_1 \leq_H x_2$. This implies that $(F, \leq_F) \preceq (H, \leq_H)$. Since $(F, \leq_F)$
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 $\preceq$. $\forall \lambda \in \Lambda$, $A_\lambda \subseteq A$ 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.

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.
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

1. \(\emptyset, X \in \mathcal{O}\);
2. \(\forall O_1, O_2 \in \mathcal{O}, \) we have \(O_1 \cap O_2 \in \mathcal{O}\);
3. \(\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O}, \) where \(\Lambda\) 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 \(\overline{F} := \bigcap_{F \subseteq B} B, \) which is clearly a closed set. The interior of \(F\) is given by \(F^\circ := \bigcup_{B \in \mathcal{O} \cap \overline{F}} B, \) 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 \(\forall O \in \mathcal{O} \) with \(x \in O, \) 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 \(F^\circ\), where \(F^\circ\) is called the exterior of \(F\). An accumulation point of \(F\) is a point \(x \in X\) such that \(\forall O \in \mathcal{O} \) with \(x \in O, \) 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) $E \subseteq \overline{E}$, $\overline{E} = E^\circ$, $E^\circ \subseteq E$, $(E^\circ)^\circ = E^\circ$, 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) $\overline{A \cup B} = \overline{A} \cup \overline{B}$, $(A \cap B)^\circ = A^\circ \cap B^\circ$;

(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, $\overline{E} = \bigcap_{C \supseteq E, C \in \mathcal{O}} C = \overline{E}$. Hence, we have $\overline{E} = \overline{E}$.

Clearly, $E^\circ \subseteq E$. Note that

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

Furthermore,

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

(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” $\forall x \in \overline{E} = \overline{E} \in \mathcal{O}$. 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 $E^\circ \subseteq O$.

“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, $O_1 \cap O_2 \in \mathcal{B}_B$. 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,

$$(A \cap B)^\circ = \bigcup_{O \subseteq A \cap B} O = \bigcup_{O_1 \subseteq A \cap B} (O_1 \cap O_2)$$

$$= \left( \bigcup_{O_1 \subseteq A} O_1 \right) \cap \left( \bigcup_{O_2 \subseteq B} O_2 \right) = A^\circ \cap B^\circ$$
3.1. FUNDAMENTAL NOTIONS

We also have
\[
\overline{A \cup B} = \overline{(A \cup B)^{\circ}} = \overline{(A \cap \overline{B})^{\circ}} = \overline{(\overline{A}) \cap (\overline{B})^{\circ}}
\]

(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 \(X = \overline{E \cup F} = E^{\circ} \cup \overline{D E \cup E^{\circ}}\). By (iii) and Definition 3.2, \(E^{\circ}\) and \(\partial E\) are disjoint. It is obvious that \(\overline{E}\) is disjoint with \(E^{\circ} \cup \partial E = \overline{F}\). Hence, the result holds. \(\square\)

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 \(\mathcal{X} := (X, \mathcal{O})\). 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 \(\mathcal{O}_X\). When it is clear from the context, we will neglect the subscript \(\mathcal{X}\).

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

**Proof** Clearly, \(\mathcal{O}_A\) is a collection of subsets of \(A\). \(\emptyset = \emptyset \cap A \in \mathcal{O}_A\) and \(A = X \cap A \in \mathcal{O}_A\). \(\forall A_1, A_2 \in \mathcal{O}_A, \exists \Omega_1, \Omega_2 \in \mathcal{O}\) such that \(A_1 = \Omega_1 \cap A\) and \(A_2 = \Omega_2 \cap A\). Then, \(\Omega_1 \cap \Omega_2 \in \mathcal{O}\) since \(\mathcal{O}\) is a topology. Then, \(O_{A_1} \cap O_{A_2} = (\Omega_1 \cap \Omega_2) \cap A \in \mathcal{O}_A\). \(\forall (O_{A_a})_{a \in \Lambda} \subseteq \mathcal{O}_A\), where \(\Lambda\) is an index set, we have, \(\forall \alpha \in \Lambda, \exists \Omega_{\alpha} \in \mathcal{O}\) such that \(O_{A_{a}} = \Omega_{\alpha} \cap A\). Then, \(\bigcup_{\alpha \in \Lambda} O_{A_{a}} \in \mathcal{O}\) since \(\mathcal{O}\) is a topology. Therefore, \(\bigcup_{\alpha \in \Lambda} O_{A_{a}} = (\bigcup_{\alpha \in \Lambda} O_{\alpha}) \cap A \in \mathcal{O}_A\).

Hence, \(\mathcal{O}_A\) is a topology on \(A\). \(\square\)

Let \((X, \mathcal{O})\) 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 \((X, \mathcal{O})\), that is, this property may change if we consider the subset topology \((A, \mathcal{O}_A)\).

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

1. \(E\) is closed in \(\mathcal{O}_A\) if, and only if, \(E = A \cap F\), where \(F \subseteq \mathcal{X}\) is closed in \(\mathcal{O}_X\);

2. the closure of \(E\) relative to \((A, \mathcal{O}_A)\) (the closure of \(E\) in \(\mathcal{O}_A\)) 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” \(A \setminus E = A \setminus (A \cap F) = A \cap \overline{A \cap F} = A \cap \overline{F}\). Since \(F\) is closed in \(\mathcal{O}_X\), then \(\overline{F} \in \mathcal{O}_X\). Then, \(A \setminus E \in \mathcal{O}_A\). Hence, \(E\) is closed in \(\mathcal{O}_A\).
“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, $E = A \setminus (A \setminus E) = A \cap A \cap O = A \cap O$. 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 $x \in A$ is a point of closure of $E$ relative to $A$, 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 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, $\mathcal{O}_2$ is said to be weaker (coarser) than $\mathcal{O}_1$.

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

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

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

(ii) $\forall A_1, A_2 \in \mathcal{O}$, we have $A_1 \cap A_2 \in \mathcal{X}$, $\forall \mathcal{X} \in \mathcal{M}$. 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 $\forall _\alpha \in \Lambda, \forall \mathcal{X} \in \mathcal{M}$, $A_\alpha \in \mathcal{X}$. Then, $\bigcup_{\alpha \in \Lambda} A_\alpha \in \mathcal{X}$, $\forall \mathcal{X} \in \mathcal{M}$. 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 $\mathcal{A}$. \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 $\forall O_Y \in \mathcal{O}_Y$, 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, \( \exists \tilde{U} \in \mathcal{O}_X \) such that \( U = \tilde{U} \cap D \). Then, \( x_0 \in \tilde{U} \). By Proposition 2.5, \( f(\tilde{U}) = f(U) \subseteq O_Y \). 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 \( f : \mathcal{X} \rightarrow \mathcal{Y} \). \( f \) is continuous if, and only if, \( \forall B \subseteq \mathcal{Y} \) with \( \overline{B} \in \mathcal{O}_Y \), 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, \( f_{\text{inv}}(O) = f_{\text{inv}}(\overline{O}) \in \mathcal{O}_X \). Hence, \( f \) is continuous.

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

This completes the proof of the proposition.

**Theorem 3.11** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be topological spaces, \( f : \mathcal{X} \rightarrow \mathcal{Y} \), 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 \( f|_{X_1} : X_1 \rightarrow \mathcal{Y} \) and \( f|_{X_2} : X_2 \rightarrow \mathcal{Y} \) are continuous. Then, \( f \) is continuous.

**Proof** Consider the case that \( X_1 \) and \( X_2 \) are both open. \( \forall x_0 \in \mathcal{X} \), \( \forall O \in \mathcal{O}_Y \) 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 \( X_1 \in \mathcal{O}_X \), 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 arbitrariness 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 \( f_{\text{inv}}(B) \subseteq \mathcal{X} \). 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, \( f : \mathcal{X} \rightarrow \mathcal{Y} \), \( g : \mathcal{Y} \rightarrow \mathcal{Z} \), 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, \( g \circ f : \mathcal{X} \rightarrow \mathcal{Z} \) 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 said to be

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 $\forall \epsilon > 0, \exists U \in \mathcal{O}_X$ 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. \□

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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U_f \in \mathcal{O}_X$ 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 $g$, $\exists U_g \in \mathcal{O}_X$ with $x_0 \in U_g$ such that $g(x) < g(x_0) + \epsilon$, $\forall x \in U_g$. Then, $\exists U := U_f \cap U_g \in \mathcal{O}_X$ 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 $x_0$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists U_f \in \mathcal{O}_X$ 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 $f$, $\exists U_1 \in \mathcal{O}_X$ with $x_0 \in U_1$ such that $f(x) < f(x_0) + \epsilon$, $\forall x \in U_1$. By the lower semicontinuity of $f$, $\exists U_2 \in \mathcal{O}_X$ with $x_0 \in U_2$ such that $f(x) > f(x_0) - \epsilon$, $\forall x \in U_2$. Then, $\exists U := U_1 \cap U_2 \in \mathcal{O}_X$ 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. \□
3.3 Basis and Countability

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

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

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

\[
\begin{align*}
(i) \quad & \forall x \in X, \exists B \in B \text{ such that } x \in B; \text{ (that is, } \bigcup_{B \in B} B = X;) \\
(ii) \quad & \forall B_1, B_2 \in B, \forall x \in B_1 \cap B_2, \exists B_3 \in B \text{ such that } x \in B_3 \subseteq B_1 \cap B_2.
\end{align*}
\]

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

“If” Define \(\bar{O} := \{O \subseteq X \mid \forall x \in O, \exists B \in B\) such that \(x \in B \subseteq O\}\}. Clearly, \(B \subseteq \bar{O}\) and \(\emptyset \in \bar{O}\). \(X \in \bar{O}\) since (i). \(\forall O_1, O_2 \in \bar{O}, \forall x \in O_1 \cap O_2\), we have \(x \in O_1\) and \(x \in O_2\). Then, by the definition of \(\bar{O}\), \(\exists B_1, B_2 \in B\) 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), \(\exists B_3 \in B\) such that \(x \in B_3 \subseteq B_1 \cap B_2 \subseteq O_1 \cap O_2\). Then, \(O_1 \cap O_2 \in \bar{O}\) where \(\Lambda\) is an index set, let \(O = \bigcup_{\alpha \in \Lambda} O_\alpha\). \(\forall \alpha \in O, \exists \alpha \in \Lambda\) such that \(x \in O_\alpha\). By the definition of \(\bar{O}\), \(\exists B \in B\) such that \(x \in B \subseteq O_\alpha \subseteq O\). Hence, \(O \in \bar{O}\). Therefore, \(\bar{O}\) is a topology on \(X\). By the definition of \(\bar{O}\), \(B\) is a basis for \(\bar{O}\). Note that \(\bar{O} \supseteq O\) since \(O\) is the weakest topology containing \(B\). Then, \(B\) is a basis for \(\bar{O}\).

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

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

**Definition 3.20** A topological space \((X, O)\) is said to satisfy the first axiom of countability if there exists a countable basis at each \(x \in X\), i.e., \(\forall x \in X, \exists B_x \subseteq O\), which is countable, such that \(\forall O \in O\) with \(x \in O\), we have \(\exists B \in B_x\) with \(x \in B \subseteq O\); and \(\forall B \in B_x\), 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 \(B\) for \(O\), 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 \((B_i)_{i \in N}\) 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. For all \( \alpha \in \Lambda \), \( O_\alpha = \bigcup_{i \in N_\alpha} B_i \), 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 \( \bar{N} \subset N \) 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 \( \mathcal{A}_i := \{ O_\alpha \mid \alpha \in \Lambda, B_i \subseteq O_\alpha \} \) is nonempty. Then, by Axiom of Choice, there exists an assignment \((O_{\alpha_i})_{i \in \bar{N}}\) such that \( O_{\alpha_i} \in \mathcal{A}_i \), \( \forall i \in \bar{N} \). 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 \((X_\alpha, O_\alpha) = (X, \mathcal{O}) =: (X, \mathcal{O})\), \( \forall \alpha \in \Lambda \), we will denote \( \prod_{\alpha \in \Lambda}(X_\alpha, O_\alpha) \) by \( X^\Lambda \).
3.4. PRODUCTS OF TOPOLOGICAL SPACES

\textbf{Proof} \hspace{1em} 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_{1\alpha}, O_{1\alpha} \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda$, and $O_{1\alpha} = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_1$, where $\Lambda_1 \subseteq \Lambda$ is a finite set; $B_2 = \prod_{\alpha \in \Lambda} O_{2\alpha}, O_{2\alpha} \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda$, and $O_{2\alpha} = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_2$, where $\Lambda_2 \subseteq \Lambda$ is a finite set. Let $B_3 = B_1 \cap B_2 = \prod_{\alpha \in \Lambda} (O_{1\alpha} \cap O_{2\alpha})$. Clearly, $O_{1\alpha} \cap O_{2\alpha} \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda$, and $O_{1\alpha} \cap O_{2\alpha} = X_\alpha, \forall \alpha \in \Lambda \setminus (\Lambda_1 \cup \Lambda_2)$, 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. \hfill \square

\textbf{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 $\pi_\alpha : \mathcal{Y} \rightarrow X_\alpha, \forall \alpha \in \Lambda$, by $\pi_\alpha(y) = y_\alpha$ for $y = (y_\alpha)_{\alpha \in \Lambda} \in \mathcal{Y}$.

\textbf{Proposition 3.27} Let $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is an index set, and $(\mathcal{Y}, \mathcal{O}) = \prod_{\alpha \in \Lambda} (X_\alpha, \mathcal{O}_\alpha)$ 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.

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

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

$$\mathcal{A} := \left\{ O \subseteq \mathcal{Y} \mid O = \prod_{\alpha \in \Lambda} O_\alpha, \exists \alpha_0 \in \Lambda \text{ such that } O_{\alpha_0} \in \mathcal{O}_{\alpha_0} \text{ and }\right.$$ \vspace{1em}

$$O_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_0\} \right\}$$

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

\textbf{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 $\mathcal{X} := (X, \mathcal{O}) := \prod_{i \in N} X_i$ is second countable.

\textbf{Proof} \hspace{1em} \forall $i \in N$, let $B_i \subseteq \mathcal{O}_i$ be a countable basis for $X_i$, without loss of generality, assume $X_i \in B_i$. Let $\mathcal{B}$ be the basis for the product topological space $\mathcal{X}$ as defined in Proposition 3.25. Define

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

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

**Proof** Note that $\pi_\alpha : \mathcal{X} \to \mathcal{X}_\alpha$ is the projection function, $\forall \alpha \in \Lambda$. $\forall \alpha \in \Lambda$, by Definition 3.2, $\pi_\alpha$ is continuous. Then, $\pi_{\alpha \text{inv}}(\overline{A_\alpha}) \subseteq \overline{A_\alpha}$. By Proposition 3.27, $\pi_\alpha$ is continuous. Then, $\prod_{\alpha \in \Lambda} \overline{A_\alpha}$ is a closed set in $\mathcal{O}$. Clearly, $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \overline{A_\alpha}$. Then, we have $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \overline{A_\alpha}$.

On the other hand, $\forall x \in \prod_{\alpha \in \Lambda} A_\alpha$, $\forall$ basis open set $O \in \mathcal{O}$ with $x \in O$, by Proposition 3.25, we have $O = \prod_{\alpha \in \Lambda} O_\alpha$, where $O_\alpha \in \mathcal{O}_\alpha$, $\forall \alpha \in \Lambda$, and $O_\alpha = X_\alpha$ for all but finitely many $\alpha$'s. $\forall \alpha \in \Lambda$, $\pi_\alpha(x) \in O_\alpha$ and $\pi_\alpha(x) \in \overline{A_\alpha}$. 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, $x \in \prod_{\alpha \in \Lambda} \overline{A_\alpha}$. Hence, we have $\prod_{\alpha \in \Lambda} A_\alpha \subseteq \prod_{\alpha \in \Lambda} \overline{A_\alpha}$.

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

An immediate consequence of the above proposition is that if $A_\alpha \subseteq \mathcal{X}_\alpha$ is closed, $\forall \alpha \in \Lambda$, where $\mathcal{X}_\alpha$'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 $(X_\alpha, \mathcal{O}_\alpha)$ be a topological space, $\forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$. $\forall \beta \in \Gamma$, $(\prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}(\beta))$ is the product space with the product topology. $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}(\Gamma))$ is the product space of product spaces with the product topology. Let $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O})$ be the product space with the product topology. Then, $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \mathcal{O}(\Gamma))$ and $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha, \mathcal{O})$ are homeomorphic.

**Proof** Define a mapping $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\alpha} X_\alpha \to \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta} X_\alpha$ by, $\forall x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha, \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_\beta$, $\exists! \beta_\alpha \in \Gamma : \exists! \alpha \in \Lambda_\beta, \pi_\alpha(E(x)) = \pi_{\alpha \text{inv}}(\pi_{\beta_\alpha}(x))$. 

CHAPTER 3. TOPOLOGICAL SPACES
3.4. PRODUCTS OF TOPOLOGICAL SPACES

\[ \forall y \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\alpha}} X_{\alpha} \implies \forall \beta \in \Gamma, \text{ define } x_\beta \in \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \text{ be such that } \pi_{\alpha}^{(\beta)}(x_\beta) = \pi_{\alpha}(y), \forall \alpha \in \Lambda_{\beta}. \text{ Define } x \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \text{ by } \pi_{\beta}^{(\Gamma)}(x) = x_\beta, \forall \beta \in \Gamma. \text{ Then, } \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}, \exists! \beta_{\alpha} \ni \alpha \in \Lambda_{\beta_{\alpha}}, \text{ and } \pi_{\alpha}(E(x)) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x)) = \pi_{\alpha}(y) \text{ and hence } E(x) = y. \text{ Hence, } E \text{ is surjective.} \]

\[ \forall x, z \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \text{ with } E(x) = E(z). \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}, \exists! \beta_{\alpha} \ni \alpha \in \Lambda_{\beta_{\alpha}}. \text{ Then, } \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x)) = \pi_{\alpha}(E(x)) = \pi_{\alpha}(E(z)) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(z)). \]

Hence, \( \forall \beta \in \Gamma, \forall \alpha \in \Lambda_{\beta}, \pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(x)) = \pi_{\alpha}^{(\beta)}(\pi_{\beta}^{(\Gamma)}(z)) \) which implies that \( \pi_{\beta}^{(\Gamma)}(x) = \pi_{\beta}^{(\Gamma)}(z) \). Hence, \( x = z \). This shows that \( E \) is injective.

Hence, \( E \) is bijective and admits an inverse \( E_{\text{inv}} \).

Next, we show that \( E \) is continuous by showing that \( E \) is continuous at \( x_0 \), \( \forall x_0 \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \). \( \forall x_0 \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \). Let \( y_0 = E(x_0) \).

\[ \forall B \in \mathcal{O} \text{ which is a basis open set with } y_0 \in B. \text{ Then, } B = \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}} B_{\alpha}, \text{ where } B_{\alpha} \in \mathcal{O}_{\alpha}, \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}, \text{ and } B_{\alpha} = X_{\alpha} \text{ for all } \alpha \text{'s except finitely many } \alpha \text{'s, say } \alpha \in \Lambda_{\alpha}. \\forall \alpha \in \Lambda_{\alpha}, \forall \alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}, \exists! \beta_{\alpha} \in \Gamma \ni \alpha \in \Lambda_{\beta_{\alpha}}. \text{ Let } \Gamma_{\alpha} = \{ \beta_{\alpha} \mid \alpha \in \Lambda_{\alpha} \}. \text{ Then, } \Gamma_{\alpha} \text{ is a finite set. } \forall \beta \in \Gamma_{\alpha}, \forall \alpha \in \Lambda_{\alpha}, \exists! \beta_{\alpha} \in \Gamma \ni \alpha \in \Lambda_{\beta_{\alpha}}. \text{ Then, } \Lambda_{\alpha} = \bigcup_{\beta \in \Gamma_{\alpha}} \Lambda_{\alpha}. \text{ Define } B^{(\beta)} := \prod_{\alpha \in \Lambda_{\alpha}} B_{\alpha}, \forall \beta \in \Gamma. \text{ Then, } \forall \beta \in \Gamma, B^{(\beta)} \text{ is a basis open set in } \bigcup_{\beta \in \Gamma_{\alpha}} \Lambda_{\alpha}. \text{ Define } B^{(\Gamma)} := \prod_{\beta \in \Gamma} B^{(\beta)}, \forall \beta \in \Gamma \setminus \Gamma_{\alpha}, B^{(\beta)} = \prod_{\alpha \in \Lambda_{\alpha}} X_{\alpha}. \text{ Clearly, } B^{(\Gamma)} \text{ is a basis open set in } \bigcup_{\beta \in \Gamma} \Lambda_{\beta}. \forall \beta \in \Gamma, \forall \alpha \in \Lambda_{\alpha}, \forall \beta \in \Gamma_{\alpha}, \forall \alpha \in \Lambda_{\alpha}, \forall \alpha \in \Lambda_{\beta}. \text{ We have } \pi_{\alpha}(E(x)) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x)) \in \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(B^{(\beta_{\alpha})})) = \pi_{\alpha}^{(\beta_{\alpha})}(B^{(\beta_{\alpha})}) = B_{\alpha}. \text{ Hence, } E(x) \in B, \text{ which implies that } E(B^{(\beta)}) \subseteq B. \text{ Hence, } E \text{ is continuous at } x_0. \text{ Then, by the arbitrariness of } x_0 \text{ and Proposition 3.9, } E \text{ is continuous.}

Finally, we will show \( E_{\text{inv}} \) is continuous by showing that \( \forall y_0 \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}} X_{\alpha}, \text{ } E_{\text{inv}} \text{ is continuous at } y_0 \). \( \forall y_0 \in \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}} X_{\alpha} \). Let \( x_0 = E_{\text{inv}}(y_0) \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \). For any basis open set \( B^{(\Gamma)} \in \mathcal{O}^{(\Gamma)} \) with \( x_0 \in B^{(\Gamma)} \). Then, \( B^{(\Gamma)} = \prod_{\beta \in \Gamma} \tilde{B}^{(\beta)}, \tilde{B}^{(\beta)} \in \mathcal{O}^{(\beta)}, \forall \beta \in \Gamma, \text{ and } \tilde{B}^{(\beta)} = \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \text{ for all } \beta \in \Gamma \text{ except finitely many } \beta \text{'s, say } \beta \in \Gamma_{\beta}. \forall \beta \in \Gamma_{\beta}, \pi_{\beta}^{(\Gamma)}(x_0) \in \tilde{B}^{(\beta)}. \text{ Then, there exists a basis open set } B^{(\beta)} \in \mathcal{O}^{(\beta)} \text{ such that } \pi_{\beta}^{(\Gamma)}(x_0) \in B^{(\beta)} \subseteq \tilde{B}^{(\beta)}. \text{ Then, } B^{(\beta)} = \prod_{\alpha \in \Lambda_{\beta}} B_{\alpha}, \text{ where } B_{\alpha} \in \mathcal{O}_{\alpha}, \forall \alpha \in \Lambda_{\beta}, \text{ and } B_{\alpha} = X_{\alpha} \text{ for all } \alpha \text{'s except finitely many } \alpha \text{'s, say } \alpha \in \Lambda_{\alpha}. \forall \alpha \in \Lambda_{\alpha}, \forall \alpha \in \Lambda_{\beta}, \pi_{\alpha}(y_0) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x_0)) \in B_{\alpha}. \text{ Let } \Lambda_{\alpha} = \bigcup_{\beta \in \Gamma_{\alpha}} \Lambda_{\alpha}, \text{ which is a finite set and the union is pairwise disjoint.} \text{ Let } B_{\alpha} = X_{\alpha}, \forall \alpha \in \Lambda_{\beta}, \forall \beta \in \Gamma \setminus \Gamma_{\alpha}. \text{ Define } B := \prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}} B_{\alpha}. \text{ Clearly, } B \text{ is a basis open set in } \mathcal{O} \text{ and } y_0 \in B. \forall y \in B. \text{ Let } x = E_{\text{inv}}(y), \text{ then, } y = E(x). \forall \beta \in \Gamma, \forall \alpha \in \Lambda_{\beta}, \pi_{\alpha}(y) = \pi_{\alpha}(E(x)) = \pi_{\alpha}^{(\beta_{\alpha})}(\pi_{\beta_{\alpha}}^{(\Gamma)}(x)) \in B_{\alpha}. \text{ Then, } \forall \beta \in \Gamma_{\beta}, \pi_{\beta}^{(\Gamma)}(x) \in \prod_{\alpha \in \Lambda_{\beta}} B_{\alpha} = B^{(\beta)} \subseteq \tilde{B}^{(\beta)}. \text{ Then, } x \in B^{(\Gamma)}.
Hence, $E_{\text{inv}}(B) \subseteq B^{(\alpha)}$. 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 $(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}, \mathcal{O}^{(\Gamma)})$ to $(\prod_{\alpha \in \bigcup_{\beta \in \Gamma} \Lambda_{\beta}} X_{\alpha}, \mathcal{O})$, 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** $\forall \alpha \in \Lambda$, 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 $H_{\alpha}^{-1}$ are continuous. Define mapping $H : X \to Y$ by, $\forall x \in \prod_{\alpha \in \Lambda} X_{\alpha}$, $\pi_{\alpha}^{(X)}(H(x)) = H_{\alpha}(\pi_{\alpha}^{(X)}(x))$, $\forall \alpha \in \Lambda$.

$\forall y \in Y$, define $x \in X$ by $\pi_{\alpha}^{(X)}(x) = H_{\alpha}(\pi_{\alpha}^{(Y)}(y))$, $\forall \alpha \in \Lambda$. 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 $\pi_{\alpha}^{(X)}(x_1) \neq \pi_{\alpha}^{(X)}(x_2)$. Then, $\pi_{\alpha}^{(Y)}(H(x_1)) = H_{\alpha}(\pi_{\alpha}^{(X)}(x_1)) \neq H_{\alpha}(\pi_{\alpha}^{(X)}(x_2)) = \pi_{\alpha}^{(Y)}(H(x_2))$, since $H_{\alpha}$ is injective. Hence, $H(x_1) \neq H(x_2)$. Therefore, $H$ is injective. Therefore, $H$ is invertible 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 $Y_{\alpha} = Y_{\alpha}$ for all $\alpha$’s except finitely many $\alpha$’s, say $\alpha \in \Lambda_N$. Then, $H_{\text{inv}}(O_Y) = \prod_{\alpha \in \Lambda} H_{\alpha}(O_{Y_{\alpha}}) \in \mathcal{O}_X$. 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 $X_{\alpha} = X_{\alpha}$ 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, $Y_{\alpha} := (Y_{\alpha}, \mathcal{O}_{\alpha})$ 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 $f(x_0) \in B \subseteq O$, where $B_{\alpha} \in \mathcal{O}_{\alpha}, \forall \alpha \in \Lambda$, and $B_{\alpha} = Y_{\alpha}$ for all $\alpha$’s except finitely many $\alpha$’s, say $\alpha \in \Lambda_N$. $\forall \alpha \in \Lambda_N$, $f_{\alpha}(x_0) = \pi_{\alpha}(f(x_0)) \in \pi_{\alpha}(B) = B_{\alpha}$. By the continuity of $f_{\alpha}$ at $x_0$, $\exists U_{\alpha} \in \mathcal{O}_{X_{\alpha}}$ 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 $x_0 \in U$. $\forall x \in U$, $f(x) \in O$. Hence, $f$ is continuous at $x_0$.
3.5. THE SEPARATION AXIOMS

∀α ∈ Λ, πα(f(x)) = fα(x) ∈ fα(U) ⊆ fα(Uα) ⊆ Bα. Hence, f(x) ∈ B and f(U) ⊆ B ⊆ O. Then, f is continuous at x₀.

“Necessity” ∀α ∈ Λ, the projection function πα is continuous, by Proposition 3.27. Note that fα = πα ◦ f. Then, fα is continuous at x₀ 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

T₁ (Tychonoff): ∀x, y ∈ X with x ≠ y, ∃O ∈ O such that x ∈ O and y ∈ O̅.

T₂ (Hausdorff): ∀x, y ∈ X with x ≠ y, ∃O₁, O₂ ∈ O such that x ∈ O₁, y ∈ O₂, and O₁ ∩ O₂ = ∅.

T₃ (regular): it is Tychonoff and, ∀x ∈ X, ∀F ⊆ X with F being closed and x /∈ F, ∃O₁, O₂ ∈ O such that x ∈ O₁, F ⊆ O₂, and O₁ ∩ O₂ = ∅.

T₄ (normal): it is Tychonoff and, ∀F₁, F₂ ⊆ X with F₁ and F₂ being closed and F₁ ∩ F₂ = ∅, ∃O₁, O₂ ∈ O such that F₁ ⊆ O₁, F₂ ⊆ O₂, and O₁ ∩ O₂ = ∅.

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 ∈ O̅. Then, it is clear that T₄ ⇒ T₃ ⇒ T₂ ⇒ T₁.

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. 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 ⊆ O, ∃U ∈ O such that F ⊆ U ⊆ O̅ ⊆ O.

Proof “Necessity” Since (X, O) is normal, F, O̅ are closed, and F ∩ O̅ = ∅, then ∃O₁, O₂ ∈ O such that F ⊆ O₁, O̅ ⊆ O₂, and O₁ ∩ O₂ = ∅.

Then, we have F ⊆ O₁ ⊆ O̅ ⊆ O, and O₁ ∩ O₂ = ∅. Since O₂ is closed, then, O₁ ⊆ O₂.

Therefore, we have F ⊆ O₁ ⊆ O̅ ⊆ O₂ ⊆ O. So, U = O₁.
“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 $F_1 \subseteq U \subseteq U \subseteq F_2$. 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. $\Box$

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, $\forall O \in \mathcal{O}$, $O \cap Y = \overline{O}$.

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

**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 $\overline{F_n} = F_n = \overline{O_n} = \mathcal{X}$. Note that $\bigcap_{n=1}^{\infty} O_n = \left(\bigcup_{n=1}^{\infty} F_n\right)^\circ$. Now, suppose that $\bigcap_{n=1}^{\infty} O_n$ is not dense. Then, $\exists O \in \mathcal{O}_X$ which is nonempty such that $(\bigcap_{n=1}^{\infty} O_n) \cap O = \emptyset$. Then, $\left((\bigcup_{n=1}^{\infty} F_n)^\circ\right) \cap O = \emptyset$. Hence, $O = \bigcup_{n=1}^{\infty} (F_n \cap O)$ and 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” $\forall O \in \mathcal{O}$ with $O \neq \emptyset$. For any countable collection $(E_\alpha)_{\alpha \in \Lambda}$ of nowhere dense subsets of $\mathcal{X}$. $\forall \alpha \in \Lambda$, $\overline{E_\alpha}$ is open and dense. Then, $\bigcap_{\alpha \in \Lambda} \overline{E_\alpha}$
is dense. Then, \( O \cap \left( \bigcap_{\alpha \in \Lambda} \overline{E_\alpha} \right) \neq \emptyset \), which further implies that \( O \not\subseteq \bigcup_{\alpha \in \Lambda} E_\alpha \). Hence, \( O \not\subseteq \bigcup_{\alpha \in \Lambda} E_\alpha \). Hence, \( O \) is of second category by the arbitrariness of \( (E_\alpha)_{\alpha \in \Lambda} \). Hence, \( X \) 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 \( F \setminus \overline{F^\circ} \) 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

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

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

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

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

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, \( \overline{F} \supseteq \bigcup_{n=1}^m \overline{F_n} = \bigcup_{n=1}^m \overline{F_n} \), by Proposition 3.3. Therefore,

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

This implies \( \overline{F} = \bigcup_{n=1}^\infty \overline{F_n} \). Hence, we have \( \overline{F} \cap \bigcap_{n=1}^\infty \overline{F_n} \), 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 \( U \subseteq O \setminus E \).
Proof  (i) “Only if” Suppose \( \exists U \in \mathcal{O} \) with \( U \neq \emptyset \) and \( U \subseteq F \). Then, \( \bar{F} \cap U = \emptyset \). This contradicts with the fact that \( \bar{F} = F = X \), by Proposition 3.3. Hence, the result holds.

“If” \( \forall x \in X, \forall U \in \mathcal{O} \) with \( x \in U \). Then, \( U \nsubseteq F \) implies \( U \cap \bar{F} \neq \emptyset \). This implies that \( x \) is a point of closure of \( \bar{F} \), by Proposition 3.3. Hence, we have \( \bar{F} = F = X \). Hence, \( F \) is nowhere dense.

(ii) “Only if” \( \forall O \in \mathcal{O} \) with \( O \neq \emptyset \). Note that

\[
O \setminus E = O \cap \bar{E} \supseteq O \cap \bar{E}
\]

Since \( E \) is nowhere dense, then \( \bar{E} \) is open and dense, and hence \( O \cap \bar{E} \) is open and nonempty by the nonemptiness of \( O \) and Proposition 3.3. Then, \( \exists U = O \cap \bar{E} \).

“If” \( \forall x \in X, \forall O \in \mathcal{O} \) with \( x \in O \). Then, \( \exists U \in \mathcal{O} \) with \( U \neq \emptyset \) such that \( U \subseteq O \setminus E \). Note that \( U = U^\circ \subseteq \left( O \cap \bar{E} \right)^\circ = O^\circ \cap \bar{E}^\circ = O \cap \bar{E} \), where we have made use of Proposition 3.3. Hence, \( O \cap \bar{E} \) is nonempty. By the arbitrariness of \( O \) and Proposition 3.3, \( x \) is a point of closure for \( \bar{F} \). By the arbitrariness of \( x, E \) is nowhere dense.

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

Theorem 3.41 (Uniform Boundedness Principle) Let \((X, \mathcal{O})\) be a topological space that is second category everywhere. Let \( F \) be a family of continuous real-valued functions of \( X \). \( \forall x \in X, \exists M_x \in [0, \infty) \subset \mathbb{R} \) such that \( |f(x)| \leq M_x, \forall f \in \mathcal{F} \). Then, \( \exists U \in \mathcal{O} \) with \( U \neq \emptyset \) and \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( |f(x)| \leq M, \forall x \in U \) and \( \forall f \in \mathcal{F} \).

Proof  Let \( E_{m,f} := f_{\text{inv}}([-m, m]) \subseteq X, \forall f \in \mathcal{F} \) and \( \forall m \in \mathbb{Z}_+ \). Since \( f \) is continuous and \([-m, m]\) is a closed interval, then, by Proposition 3.10, \( E_{m,f}'s \) are closed. \( \forall m \in \mathbb{Z}_+ \), let \( E_m := \bigcap_{f \in \mathcal{F}} E_{m,f} \), which is closed. \( \forall x \in X, \exists m \in \mathbb{Z}_+ \) with \( m \geq M_x \) such that \( |f(x)| \leq m, \forall f \in \mathcal{F} \). Then, \( x \in E_{m,f}, \forall f \in \mathcal{F} \), and hence \( x \in E_m \). Therefore, \( X = \bigcup_{m=0}^{\infty} E_m \). Since \( X \) is second category everywhere, then \( E_m \)'s are not all nowhere dense. Then, \( \exists n \in \mathbb{Z}_+ \) such that \( E_n \) is not nowhere dense, that is \( \bar{E}_n = E_n \neq X \). Then, by Proposition 3.40, \( E_n \) contains a nonempty open subset \( U \in \mathcal{O} \). Then, \( U \subseteq E_n \) and \( U \neq \emptyset \). \( \forall x \in U, \forall f \in \mathcal{F} \), we have \( |f(x)| \leq n \). This completes the proof of the theorem. \( \square \)

3.7 Connectedness

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, \( f^{-1}(O_1) \cup f^{-1}(O_2) = f^{-1}(O_1 \cup O_2) = \mathcal{X} \) 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. \( \square \)

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. \( \square \)

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 \( \mathcal{O}_V \). 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, \( x_i \in \overline{U} \cap O_i, \ i = 1, 2 \). By Proposition 3.4, \( O_{U_1}, O_{U_2} \in \mathcal{O}_U \). Note that \( O_{U_1} \cap O_{U_2} = U \cap (O_1 \cap O_2) = U \cap O_1 \cap O_2 = U \cap V \cap O_1 \cap O_2 = U \cap (O_1 \cap V) \cap (O_2 \cap V) = U \cap O_1 \cap O_2 \neq \emptyset \) 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. \( \square \)

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.
**Proof** 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. 

**Proposition 3.48** 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).

**Proof** Suppose $A$ is not connected. Then, $\exists O_1, O_2 \in \mathcal{O}$ such that $O_1 \cap A \neq \emptyset \neq O_2 \cap A$, $(O_1 \cap A) \cap (O_2 \cap A) = \emptyset$, and $(O_1 \cap A) \cup (O_2 \cap A) = (O_1 \cup O_2) \cap A = A$. 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.

**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.

**Proof** Let $\mathcal{X}$ be a locally connected topological space and $\mathcal{B}$ be a basis made up of connected sets. $\forall x_0 \in \mathcal{X}$, let $A$ be the component containing $x_0$. By Proposition 3.47, $A$ is closed and connected. $\forall x \in A$, $\exists B \in \mathcal{B}$ such that $x \in B$. Since $B$ is connected, then, by Definition 3.46, $B \subseteq A$. Hence, $A$ is open. This completes the proof of the proposition.

**Proposition 3.51** Let $(X_\alpha, \mathcal{O}_\alpha)$ be a connected topological space, $\forall \alpha \in \Gamma$, where $\Gamma$ is an index set. Let $(X, \mathcal{O})$ be the product topological space $\prod_{\alpha \in \Gamma} (X_\alpha, \mathcal{O}_\alpha)$. Then, $(X, \mathcal{O})$ is connected.

**Proof** 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 $\Lambda_1$ and $\Lambda_2$ such that $O_1 = \bigcup_{\lambda \in \Lambda_1} B_{\lambda}$ and $O_2 = \bigcup_{\lambda \in \Lambda_2} C_{\lambda}$, where $B_{\lambda} \in B$, $\forall \lambda \in \Lambda_1$, and $C_{\lambda} \in B$, $\forall \lambda \in \Lambda_2$. $\forall \lambda \in \Lambda_1, B_{\lambda} = \prod_{\alpha \in \Gamma} B_{\lambda \alpha}$, where $B_{\lambda \alpha} \in O_{\alpha}$, $\forall \alpha \in \Gamma$, and $B_{\lambda \alpha} = X_{\alpha}$ for all except finitely many $\alpha$’s, say $\alpha \in \Gamma_{\lambda}$. Note that $\Gamma_{\lambda} \neq \emptyset$, since $B_{\lambda} \subseteq O_1 \subset X$. $\forall \lambda \in \Lambda_2, C_{\lambda} = \prod_{\alpha \in \Gamma} C_{\lambda \alpha}$, where $C_{\lambda \alpha} \in O_{\alpha}$, $\forall \alpha \in \Gamma$, and $C_{\lambda \alpha} = X_{\alpha}$ for all except finitely many $\alpha$’s, say $\alpha \in \Gamma_{\lambda}$. Note that $\Gamma_{\lambda} \neq \emptyset$, since $C_{\lambda} \subseteq O_2 \subset X$. Note that

$$\emptyset = O_1 \cap O_2 = \bigcup_{\lambda \in \Lambda_1} \bigcup_{\gamma \in \Lambda_2} (B_{\lambda} \cap C_{\gamma})$$

Therefore, $B_{\lambda} \cap C_{\gamma} = \emptyset$, $\forall \lambda \in \Lambda_1$, $\forall \gamma \in \Lambda_2$, which implies that $\exists \alpha_{\lambda \gamma} \in \Gamma$ $\exists^* B_{\lambda \alpha_{\lambda \gamma}} \cap C_{\gamma \alpha_{\lambda \gamma}} = \emptyset$.

Fix $x_1 \in O_1$ and $x_2 \in O_2$. Then, $x_1 \neq x_2$, $\exists \lambda_1 \in \Lambda_1$ such that $x_1 \in B_{\lambda_1}$.

Let $\Gamma_1 := \Gamma_{\lambda_1}$, $\exists \lambda_2 \in \Lambda_2$ such that $x_2 \in C_{\lambda_2}$. Let $\Gamma_2 := \Gamma_{\lambda_2}$. Note that $\forall \lambda \in \Lambda_1 \cup \Lambda_2$ with $\pi_{\alpha}(x) = \pi_{\alpha}(x_1)$, $\forall \alpha \in \Gamma_1$, we have $x \in B_{\lambda_1} \subseteq O_1$. Similarly, $\forall \lambda \in \Lambda_1 \cup \Lambda_2$ with $\pi_{\alpha}(x) = \pi_{\alpha}(x_2)$, $\forall \alpha \in \Gamma_2$, we have $x \in C_{\lambda_2} \subseteq O_2$. Therefore, starting with $x_1 \in O_1$ and switch, 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 $x_3 \in O_2$. Therefore, there must exist a step in this process such that switching one coordinate $\pi_{\alpha_0}(x_1)$ to $\pi_{\alpha_0}(x_2)$, 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

$$\Lambda_{10} := \{ \lambda \in \Lambda_1 \mid B_{\lambda} = \prod_{\alpha \in \Gamma} B_{\lambda \alpha}, \pi_{\alpha}(x_1) \in B_{\lambda \alpha}, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

and

$$\Lambda_{20} := \{ \lambda \in \Lambda_2 \mid C_{\lambda} = \prod_{\alpha \in \Gamma} C_{\lambda \alpha}, \pi_{\alpha}(x_2) \in C_{\lambda \alpha}, \forall \alpha \in \Gamma \setminus \{\alpha_0\} \}$$

Let $M := \prod_{\alpha \in \Gamma} M_{\alpha} \subseteq X$ where $M_{\alpha} = \{ \pi_{\alpha}(x_1) \}$, $\forall \alpha \in \Gamma \setminus \{\alpha_0\}$, and $M_{\alpha_0} = X_{\alpha_0}$. Note that $M \subseteq X = O_1 \cup O_2 = \left( \bigcup_{\lambda \in \Lambda_1} B_{\lambda} \right) \cup \left( \bigcup_{\gamma \in \Lambda_2} C_{\gamma} \right)$ implies that

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

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

$$X_{\alpha_0} = \bigcup_{\lambda \in \Lambda_{10}} \bigcup_{\gamma \in \Lambda_{20}} (B_{\lambda \alpha_0} \cup C_{\gamma \alpha_0}) = \bigcup_{\lambda \in \Lambda_{10}} B_{\lambda \alpha_0} \cup \bigcup_{\gamma \in \Lambda_{20}} C_{\gamma \alpha_0} =: D_1 \cup D_2$$
Claim 3.54.1

**Proof of claim:** Let $O$ be a closed curve that is connected. Then, there exists a curve $O$ such that $O \cap \bar{O} = \emptyset$. Hence, $O$ is said to be arcwise connected.

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

This completes the proof of the proposition. \qed

**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 $\gamma : I \to X$. $\gamma(0)$ 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 $\phi : I \times I \to X$ 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 $t := \sup \{ s \in [0,1] \subset \mathbb{R} | \gamma([0,s]) \subseteq O_1 \}$. 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, $s \notin \{ s \in [0,1] \subset \mathbb{R} | \gamma([0,s]) \subseteq O_1 \}$. 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. \qed

Since $\gamma(1) = x_2 \in O_2$, then $t < 1$. By the continuity of $\gamma$, $\exists t_2 \in (t, 1] \subset \mathbb{R}$ such that $\gamma([t,t_2]) \subseteq O_1$. $\forall s \in [0,t)$, $\exists s_1 \in (s, t]$ such that $\gamma(s_1) = x_2$. Therefore, we have $D_1 \cap D_2 = \bigcup_{\lambda \in \Lambda_{10}} \{ \gamma \in \Lambda_{12} \cap C_{\gamma_{\alpha_0}} = \emptyset \}$. Note that, $\forall \alpha \in \Gamma \setminus \{ \alpha_0 \}$, $\pi_{\alpha} (\bar{x}_1) = \pi_{\alpha} (\bar{x}_2) \in B_{\alpha} \cap C_{\gamma_{\alpha_0}} \neq \emptyset$. Then, we must have $B_{\lambda_{\alpha_0}} \cap C_{\gamma_{\alpha_0}} = \emptyset$. Therefore, we have $D_1 \cap D_2 = \bigcup_{\lambda \in \Lambda_{10}} \{ \gamma \in \Lambda_{12} \cap C_{\gamma_{\alpha_0}} = \emptyset \}$.
3.8. CONTINUOUS REAL-VALUED FUNCTIONS

\( s_1 \in \{ s \in [0, 1] \subset \mathbb{R} \mid \gamma([0, s]) \subseteq O_1 \} \). Then, \( \gamma([0, s_1]) \subseteq O_1 \) and hence \( \gamma(s) \in O_1 \). Therefore, \( \gamma([0, t]) \subseteq O_1 \). This coupled with \( \gamma([t, t_2]) \subseteq O_1 \), we have \( \gamma([0, t_2]) \subseteq O_1 \). Then, \( t \geq t_2 \). This contradicts with \( t < t_2 \).

Therefore, \((X, \mathcal{O})\) is connected. This completes the proof of the proposition.

\[ \square \]

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 \( f : X \to [0, 1] \subseteq \mathbb{R} \) such that \( f(x) = 0 \), \( \forall x \in A \), and \( f(x) = 1 \), \( \forall x \in B \).

Proof Since the set \( Q := \mathbb{Q} \cap [0, 1] \) 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. \( \forall x_0 \in X \), 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 \( a_0 \in U \), \( \exists a_1, a_2, a_3, a_4 \in \mathbb{Q} \) such that \( a_1 < a_2 < a_0 < a_3 < a_4 \) and \( (a_1, a_4) \subseteq U \). Let \( \bar{a}_2 = \max\{a_2, 0\} \) and \( \bar{a}_3 = \min\{a_3, 1\} \). Then, we must have \( a_1 < \bar{a}_2 < a_0 < a_3 < a_4 \) 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 \( f(x) \leq \bar{a}_3 \). Also, \( x \in \overline{O_{\bar{a}_2}} \) implies that \( f(x) \geq \bar{a}_2 \). Hence, \( f(V) \subseteq [\bar{a}_2, \bar{a}_3] \subseteq (a_1, a_4) \subseteq U \).

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 \( x_0 \in V \). \( \forall x \in V \), \( 0 \leq f(x) \leq \bar{a}_3 \). Hence, \( f(V) \subseteq [0, \bar{a}_3] \subseteq (a_1, a_4) \subseteq U \). 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 \( a_1 < \bar{a}_2 < a_0 = 1 < a_4 \). Take \( V = \overline{O_{\bar{a}_2}} \in \mathcal{O} \). Since \( f(x_0) = a_0 = 1 \), then \( x_0 \in O_{\bar{a}_2} \subseteq \overline{O_{\bar{a}_2}} = V \). \( \forall x \in V \),
\[ f(x) \geq \bar{a}_2. \] Hence, \( f(V) \subseteq [\bar{a}_2, 1] \subset (a_1, a_4) \subseteq U. \) 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. \( \square \)

**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 \( U_1 := f_{\text{inv}}(O_1) \in \mathcal{O}_X \) and \( U_2 := f_{\text{inv}}(O_2) \in \mathcal{O}_X. \) Note that \( x \in U_1 \cap U_2 \in \mathcal{O}_X \) and \( x \in \mathcal{D}, \) 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. \( \square \)

**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 \( |l(x)| \leq c_1 \in \mathbb{R}, \forall x \in A, \) 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. \( \square \)

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, \( |f_i(x)| \leq \frac{2^{i-1}}{3^r}, \forall x \in X, \) and \( \sum_{k=1}^{i} f_k(x) \leq \frac{2^i}{3^r}, \forall x \in A. \)

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 \( |g(x)| \leq \sum_{i=1}^{\infty} \frac{2^{i-1}}{3^r} = 1, \forall x \in X. \forall x_0 \in X. \)
Proposition 3.59

Let $k$ be a collection of continuous functions of $g \forall \forall$ is the weak topology generated by $O$ in Proposition 3.10. Note that $f \in F$ and $\sum_{i=1}^{N} f_i(x) = g(x)$.

The topology that $F$ inherits as a subspace of $I^X$ is called the topology of pointwise convergence. Now, $X$ can be identified with a subset of $I^F$ by, $x \in X, \pi_f(x) = f(x)$, $f \in F$. Then, the topology of $X$ as a subset of $I^F$ is the weak topology generated by $F$.

Proposition 3.59

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

Proof

Fix a basis open set $O$ in $I^F$ with $E(x_0) \in O$. By Proposition 3.25, $O = \bigcap f \in F O_f$, where $O_f \in \mathcal{O}_I$ with $O_f$ being the subset topology on $I$, $\forall f \in F$, and $O_f = I$ for all $f$'s except finitely many $f$'s, say...
$f \in \mathcal{F}_N$. 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 $x_0 \in U$.

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

Under the additional assumption on $\mathcal{F}$, we need to show that $E$ is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X})$. $\forall x, y \in \mathcal{X}$ 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, $E : \mathcal{X} \rightarrow E(\mathcal{X})$ is injective. Clearly, $E : \mathcal{X} \rightarrow E(\mathcal{X})$ is surjective. Then, $E : \mathcal{X} \rightarrow E(\mathcal{X})$ is bijective and admits inverse $E_{\text{inv}} : E(\mathcal{X}) \rightarrow \mathcal{X}$. $\forall x_0 \in \mathcal{X}$, we will show that $E_{\text{inv}}$ is continuous at $E(x_0)$. $\forall O \in \mathcal{O}_X$ with $x_0 \in O$. $\bar{O}$ is closed and $x_0 \notin \bar{O}$. Then, $\exists f_0 \in \mathcal{F}$ such that $f_0(x_0) = 1$ and $f_0|_{\bar{O}} = 0$. Define $U = \bigcap_{f \in \mathcal{F}} U_f \subseteq I^p$ by $U_f = I$, $\forall f \in \mathcal{F} \setminus \{f_0\}$ and $U_{f_0} = (1/2, 1] \in O_I$. Clearly, $U$ is open in $I^p$. Clearly, $E(x_0) \in U$. $\forall x \in \mathcal{X}$ with $E(x) \in U$, we have $\pi_{f_0}(E(x)) = f_0(x) > 1/2$. Then, $x \notin \bar{O}$ and $x \in O$. This shows that $E_{\text{inv}}(E(\mathcal{X}) \cap U) \subseteq O$. Hence, $E_{\text{inv}}$ is continuous at $E(x_0)$. By the arbitrariness of $x_0$ and Proposition 3.9, $E_{\text{inv}} : E(\mathcal{X}) \rightarrow \mathcal{X}$ is continuous. This implies that $E : \mathcal{X} \rightarrow E(\mathcal{X})$ is a homeomorphism.

This completes the proof of the proposition.

**Definition 3.60** A topological space $\mathcal{X}$ is said to be completely regular (or $T_{3\frac{1}{2}}$) if it is Tychonoff and $\forall x_0 \in \mathcal{X}$ and $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, there exists a continuous real-valued function $f : \mathcal{X} \rightarrow [0, 1]$ 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 $\mathcal{X}$ be a normal topological space. Then, $\mathcal{X}$ is Tychonoff. $\forall x_0 \in \mathcal{X}$ and $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, we have $\{x_0\}$ is closed, by Proposition 3.34. By Urysohn’s Lemma, there exists a continuous real-valued function $f : \mathcal{X} \rightarrow [0, 1]$ such that $f(x_0) = 1$ and $f|_{F} = 0$. Hence, $\mathcal{X}$ is completely regular.

Let $\mathcal{X}$ be a completely regular topological space. Then, $\mathcal{X}$ is Tychonoff. $\forall x_0 \in \mathcal{X}$ and $\forall$ closed set $F \subseteq \mathcal{X}$ with $x_0 \notin F$, there exists a continuous real-valued function $f : \mathcal{X} \rightarrow [0, 1]$ such that $f(x_0) = 1$ and $f|_{F} = 0$. Hence, $\mathcal{X}$ is completely regular.

**Corollary 3.62** Let $\mathcal{X}$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, and $\mathcal{F} := \{f : \mathcal{X} \rightarrow I \mid f \text{ is continuous}\}$. Then, the equivalence map: $E : \mathcal{X} \rightarrow I^\mathcal{F}$ defined by $\pi_f(E(x)) = f(x)$, $\forall x \in \mathcal{X}$, $\forall f \in \mathcal{F}$, is a homeomorphism between $\mathcal{X}$ and $E(\mathcal{X}) \subseteq I^\mathcal{F}$.
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 $\prec_A$, 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 $x \in O$, $\exists \alpha_0 \in A \ni \forall \alpha \in A$ 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 $\alpha \prec \beta \ni x_\beta \in O$.

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.

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
∀α₂ ∈ A, ∀α ∈ A with α₂ < α, we have xₐ ∈ O₂. Since A is a directed system, ∃O₃ ∈ A such that α₁ < α₃ and α₂ < α₃. Then, we have xₐ₃ ∈ O₁ and xₐ₃ ∈ O₂, which implies that O₁ ∩ O₂ ≠ ∅, which is a contradiction. Therefore, every net in X has at most one limit point.

"If" Suppose X is not Hausdorff. Then, ∃xₐ, xₜ ∈ X with xₐ ≠ xₜ such that ∀Oₐ, Oₜ ∈ O with xₐ ∈ Oₐ and xₜ ∈ Oₜ, we have Oₐ ∩ Oₜ ≠ ∅. Let Λ := \{ (Oₐ, Oₜ) | xₐ ∈ Oₐ ∈ O, xₜ ∈ Oₜ ∈ O \}. Clearly, (X, X) ∈ Λ, then Λ ≠ ∅. Define a relation ≺ on Λ by, ∀(Oₐ₁, Oₜ₁), (Oₐ₂, Oₜ₂) ∈ Λ, we say \( (Oₐ₁, Oₜ₁) ≺ (Oₐ₂, Oₜ₂) \) if Oₐ₁ ⊇ Oₐ₂ and Oₜ₁ ⊇ Oₜ₂. Clearly, ≺ is transitive on Λ. ∀(Oₐ₁, Oₜ₁), (Oₐ₂, Oₜ₂) ∈ Λ, we have xₐ ∈ Oₐ₁ ∩ Oₐ₂ ∈ O and xₜ ∈ Oₜ₂ ∩ Oₜ₁ ∈ O. Then, we have (Oₐ₃, Oₜ₃) ∈ Λ, \( (Oₐ₁, Oₜ₁) ≺ (Oₐ₃, Oₜ₃) \), and \( (Oₐ₂, Oₜ₂) ≺ (Oₐ₃, Oₜ₃) \). Hence, \( A := (Λ, ≺) \) is a directed system. ∀(Oₐ, Oₜ) ∈ Λ, Oₐ ∩ Oₜ ≠ ∅. By Axiom of Choice, we may have a mapping \( x_{(Oₐ, Oₜ)} \) ∈ Oₐ ∩ Oₜ, ∀(Oₐ, Oₜ) ∈ Λ.

Then, the net \( \{ x_{(Oₐ, Oₜ)} \} \) with \( xₐ ∈ Oₐ \). Fix Oₐ₁ := X ∈ O with xₐ ∈ Oₐ₁. Fix Oₐ₁ := X ∈ O with xₜ ∈ Oₐ₁. Then, (Oₐ₁, Oₜ₁) ∈ Λ. ∀(Oₐ₂, Oₜ₂) ∈ Λ with \( (Oₐ₁, Oₜ₁) ≺ (Oₐ₂, Oₜ₂) \), we have \( x_{(Oₐ₂, Oₜ₂)} ∈ Oₐ₂ \cap Oₜ₂ \subseteq Oₐ₁ \cap Oₜ₁ = Oₐ₁ \). Hence, xₜ is a limit point of \( \{ x_{(Oₐ, Oₜ)} \} \) with \( xₜ ∈ Oₐ₁ \). Then, \( (Oₐ₂, Oₜ₂) ∈ Λ \), \( (Oₐ₁, Oₜ₁) ≺ (Oₐ₂, Oₜ₂) \), \( (Oₐ₃, Oₜ₃) ≺ (Oₐ₂, Oₜ₂) \), we have \( x_{(Oₐ₃, Oₜ₃)} ∈ Oₐ₃ \cap Oₜ₃ \subseteq Oₐ₁ \cap Oₜ₁ = Oₐ₁ \). Hence, xₜ is a limit point of \( \{ x_{(Oₐ, Oₜ)} \} \) with \( xₜ ∈ Oₐ₃ \).

This completes the proof of the proposition.

**Proposition 3.66** Let X and Y be topological spaces, D ⊆ X with subset topology Oₐ, and \( f : D \to Y \). Then, the following are equivalent.

(i) \( f \) is continuous at \( x₀ ∈ D \);

(ii) \( \forall \) net \( \{ xₐ \}_{α ∈ A} \subseteq D \) with \( x₀ \) as a limit point, we have that the net \( \{ f(xₐ) \}_{α ∈ A} \) has a limit point \( f(x₀) \).

(iii) \( \forall \) net \( \{ xₐ \}_{α ∈ A} \subseteq D \) with \( x₀ \) as a cluster point, we have that the net \( \{ f(xₐ) \}_{α ∈ A} \subseteq Y \) has a cluster point \( f(x₀) \).

**Proof** (i) ⇒ (ii). Fix a net \( \{ xₐ \}_{α ∈ A} \subseteq D \) with \( x₀ ∈ D \) as a limit point. ∀Oₐ ∈ Oₐ with \( f(x₀) ∈ Oₐ \). By the continuity of \( f \) at \( x₀ \), ∃Oₓ ∈ Oₓ with \( x₀ ∈ Oₓ \) such that \( f(Oₓ) ⊆ Oₐ \). Since \( x₀ \) is a limit point of \( \{ xₐ \}_{α ∈ A} \),
then \( \exists \alpha_0 \in \mathcal{A} \) such that, \( \forall \alpha \in \mathcal{A} \) with \( \alpha_0 \prec \alpha \), we have \( x_\alpha \in O_\mathcal{X} \). 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, \( \forall O_\mathcal{X} \in \mathcal{O}_X \) with \( x_0 \in O_\mathcal{X} \), we have \( f(O_\mathcal{X}) \nsubseteq O_Y \). 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 \( x_\alpha \in O \cap D \) with \( f(x_\alpha) \notin O_Y \). Clearly, \( x_0 \) is a limit point of \( (x_\alpha)_{\alpha \in \mathcal{A}} \). Yet, \( f(x_0) \in O_Y \) and \( f(x_\alpha) \notin O_Y, \forall O \in \mathcal{M} \). Then, \( f(x_0) \) 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 \( x_0 \), \( \exists U \in \mathcal{O}_X \) 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, \( \mathcal{A} := (\mathcal{M}, \subseteq) \) is a directed system. \( \exists O_Y \in \mathcal{O}_Y \) with \( f(x_0) \in O_Y \) such that \( \forall U \in \mathcal{M}, f(U) \nsubseteq O_Y \). 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, \( f(x_0) \) 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, \( x_\alpha \in X \) 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 \( x_\alpha \in X \), 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 \( x_\alpha \in X \) 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 \( x_\alpha \in B \) such that, \( \forall \beta \in \mathcal{A}, \exists \beta \in \mathcal{A} \) with \( \beta \in B \), we have \( x_\beta \notin B \). Then, \( B = \bigcap_{\alpha \in \Lambda} O_\alpha \), \( O_\alpha \in \mathcal{O}_\alpha \), \( \forall \alpha \in \Lambda \), 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 \( \alpha \beta \in \Lambda_N \). Then, by an argument of contradiction, we may show that \( \exists \alpha_0 \in \Lambda_N \) such that, \( \forall \beta_0 \in \mathcal{A}, \exists \beta \in \mathcal{A} \) with \( \beta_0 \prec \beta \), we have \( \pi_\alpha(x_\beta) \notin O_\alpha \). Hence, \( \pi_\alpha(x_\alpha) \in O_\alpha \) is not the limit of the net.
(π₀(π₆)) ∈ A. This contradicts with the assumption. Hence, we have x₀ is a limit point of (π₆) ∈ A.

This completes the proof of the proposition. □

**Proposition 3.68** Let (X, O) be a topological space, E ⊆ X, and x ∈ X, x ∈ E if, and only if, ∃ a net (xₐ)ₐ∈A ⊆ E such that x is a limit point of the net.

**Proof** “Only if” Let M := {O ∈ O | x ∈ O}. Clearly, X ∈ M, then M ≠ ∅. Clearly, A := (M, ⊇) is a directed system. Since x ∈ E, then, by Proposition 3.3, ∀O ∈ A, O ⊇ E ≠ ∅. By Axiom of Choice, ∃ a net (xₐ)O∈A ⊆ E such that x₀ ∈ O ∩ E, ∀O ∈ A. ∀O ∈ O with x ∈ O, then O ∈ A. ∀O₁ ∈ A with O ⊇ O₁, we have x₀ ∈ O₁ ∩ E ⊆ O. Hence, x is a limit point of (xₐ)O∈A.

“If” Let (xₐ)ₐ∈A ⊆ E be the net such that x is a limit point of the net. ∀O ∈ O with x ∈ O, ∃α₀ ∈ A, ∀α ∈ A with α₀ ≺ α, we have x₀ ∈ E ∩ O. Since (xₐ)ₐ∈A is a net, then ∃α₁ ∈ A with α₀ ≺ α₁. Then, x₀₁ ∈ O ∩ E ≠ ∅. By Proposition 3.3, x ∈ E. This completes the proof of the proposition. □

**Definition 3.69** Let (X, O) be a topological space, A := (A, ≺) be a directed system, and (xₐ)ₐ∈A ⊆ X be a net. Let Aₛ ⊆ A be a subset with the same relation ≺ as A such that ∀α ∈ A, ∃αₛ ∈ Aₛ such that α ≺ αₛ. Then, Aₛ := (Aₛ, ≺) is a directed system and (xₐ)ₐ∈Aₛ is a net, which is called a subnet of (xₐ)ₐ∈A.

**Proposition 3.70** Let (X, O) be a topological space and (xₐ)ₐ∈A ⊆ X be a net. Then, x₀ ∈ X is a limit point of (xₐ)ₐ∈A if, and only if, any subnet (xₐ)ₐ∈Aₛ has a limit point x₀.

**Proof** “Only if” Since x₀ ∈ X is a limit point of (xₐ)ₐ∈A, then ∀O ∈ O with x₀ ∈ O, ∃α₁ ∈ A such that ∀α ∈ A with α₁ ≺ α, we have x₀ ∈ O. Let (xₐ)ₐ∈Aₛ be a subnet. Then, ∃αₛ ∈ Aₛ such that αₐ < αₛ. ∀αₛ ∈ Aₛ with αₐ < αₛ, we have αₐ < αₛ and xₐ ∈ O. Hence, x₀ is a limit point of the subnet.

“If” Since (xₐ)ₐ∈A is a subnet of itself, then it has limit x₀.

This completes the proof of the proposition. □

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

**Proposition 3.71** Let (X, O) be a topological space and (xₐ)ₐ∈A ⊆ X be a net. Then, x₀ ∈ X is a limit point of (xₐ)ₐ∈A if, and only if, for every subnet (xₐ)ₐ∈Aₛ of (xₐ)ₐ∈A, there exists a subnet (xₐ)ₐ∈Aₛₛ that has a limit point x₀.

**Proof** “Sufficiency” We assume that every subnet (xₐ)ₐ∈Aₛ of (xₐ)ₐ∈A, there exists a subnet (xₐ)ₐ∈Aₛₛ that has a limit point x₀. We
will prove the result using an argument of contradiction. Suppose \( x_0 \) is not a limit point of \( (x_\alpha)_{\alpha \in A} \). Then, \( \exists O_0 \in \mathcal{O} \) with \( x_0 \in O_0, \forall \alpha_0 \in A, \exists \alpha \in A \) with \( \alpha_0 < \alpha \) such that \( x_\alpha \in \overline{O_0} \). Define \( A_\alpha := \{ \alpha \in A \mid x_\alpha \in \overline{O_0} \} \), \(<\). Clearly, \((x_\alpha)_{\alpha \in A_\alpha}\) is a subnet of \((x_\alpha)_{\alpha \in A}\). Any subnet \((x_\alpha)_{\alpha \in A_{ss}}\) of \((x_\alpha)_{\alpha \in A_\alpha}, \forall \alpha_{ss} \in A_{ss}\), we have \( x_{\alpha_{ss}} \in O_0 \). Then, \( x_0 \) is not a limit of \((x_\alpha)_{\alpha \in A_{ss}}\). This contradicts the assumption. Therefore, \( x_0 \) is a limit point of \((x_\alpha)_{\alpha \in A}\).

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

This completes the proof of the proposition.

\[ \square \]

\textbf{Definition 3.72} Let \( \mathcal{X} := (X, \mathcal{O}_X) \) and \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq \mathcal{X} \), \( f : D \to \mathcal{Y} \), 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 \( x \to x_0 \) 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 \( x \to x_0 \).

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

\textbf{Proposition 3.73} Let \( \mathcal{X} := (X, \mathcal{O}_X) \) and \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) be topological spaces, \( D \subseteq \mathcal{X} \), \( f : D \to \mathcal{Y} \), 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 \( x \to x_0 \). In this case, we will write \( \lim_{x \to x_0} f(x) = y_0 \in \mathcal{Y} \) when the limit exists.

\textbf{Proof} Suppose \( f(x) \) admits limit points \( y_A, y_B \in Y \) as \( x \to x_0 \) 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 \( x \to x_0 \), 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 \( x \to x_0 \), 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 \( x \in (D \cap V) \setminus \{x_0\} \) 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.

\[ \square \]

\textbf{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 \( \mathcal{O}_D \), \( f : D \to \mathcal{Y} \), 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 \( x \to x_0 \).
Proof  
(i) ⇒ (ii). This is straightforward.

(ii) ⇒ (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 $D$.  
$\exists V \in \mathcal{O}_X$ with $x_0 \in V$ such that $V \cap D = \{x_0\}$.  
$\forall U \in \mathcal{O}_Y$ 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 $D$.  
$\forall U \in \mathcal{O}_Y$ with $f(x_0) \in U$,  
$\exists V \in \mathcal{O}_X$ 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. □

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

Proof  
$\forall O_Z \in \mathcal{O}_Z$ with $g(y_0) \in O_Z$, by the continuity of $g$ at $y_0$,  
$\exists O_Y \in \mathcal{O}_Y$ with $y_0 \in O_Y$ such that $g(O_Y) \subseteq O_Z$. Since $y_0$ is the limit of $f(x)$ as $x \rightarrow x_0$, then $\exists O_X \in \mathcal{O}_X$ 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 $x \rightarrow x_0$. This completes the proof of the proposition. □

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, $D \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of $D$, $f : D \rightarrow \bar{D}$, and $g : \bar{D} \rightarrow \mathcal{Z}$. Assume that

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

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

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

Proof  
$\forall O_Z \in \mathcal{O}_Z$ with $z_0 \in O_Z$, by $\lim_{y \rightarrow y_0} g(y) = z_0$,  
$\exists O_Y \in \mathcal{O}_Y$ 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 $\lim_{x \rightarrow x_0} f(x) = y_0$,  
$\exists O_X \in \mathcal{O}_X$ with $x_0 \in O_X$ such that $f(O_X \setminus \{x_0\}) = f((O_X \cap D) \setminus \{x_0\}) \subseteq O_Y$. Let $O_1 := O_0 \cap O_X \in \mathcal{O}_X$. 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, $g((f(O_1 \cap D) \setminus \{y_0\})) \subseteq O_Z$. Hence, $\lim_{x \rightarrow x_0} g(f(x)) = z_0$. This completes the proof of the proposition. □

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 $D$, $D \subseteq \mathcal{Y}$, $y_0 \in \mathcal{Y}$ be an accumulation point of $D$, $f : \bar{D} \rightarrow \bar{D}$ be bijective, and $g : \bar{D} \rightarrow \mathcal{Z}$. Assume that $\lim_{x \rightarrow x_0} f(x) = y_0$ and $\lim_{y \rightarrow y_0} f^{-1}(y) = x_0$. Then, $\lim_{x \rightarrow x_0} g(f(x)) = \lim_{y \rightarrow y_0} g(y)$ whenever one of the limits exists in $\mathcal{Z}$. 
3.9. NETS AND CONVERGENCE

Proof We will prove the result by distinguishing two exhaustive cases:
Case 1: \( \lim_{y \to y_0} g(y) = z_0 \in \mathbb{Z} \); Case 2: \( \lim_{x \to x_0} g(f(x)) = z_0 \in \mathbb{Z} \).

Case 1: \( \lim_{y \to y_0} g(y) = z_0 \in \mathbb{Z} \). We will further distinguish three exhaustive and mutually exclusive subcases: Case 1a: \( y_0 \notin D \); Case 1b: \( y_0 \in D \) and \( x_0 = f_{\text{inv}}(y_0) \); Case 1c: \( y_0 \in D \) and \( x_0 \neq f_{\text{inv}}(y_0) \). Case 1a: \( y_0 \notin D \). Then, (i) of Proposition 3.76 is satisfied with \( O_0 := \mathcal{X} \). By Proposition 3.76, we have \( \lim_{x \to x_0} g(f(x)) = z_0 \in \mathbb{Z} \). Case 1b: \( y_0 \in D \) and \( x_0 = f_{\text{inv}}(y_0) \). Let \( O_0 := \mathcal{X} \). \( \forall x \in (O_0 \cap \bar{D}) \setminus \{x_0\} \), 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, \( \lim_{x \to x_0} g(f(x)) = z_0 \in \mathbb{Z} \) by Proposition 3.76. Case 1c: \( y_0 \in D \) and \( x_0 \neq f_{\text{inv}}(y_0) \). By \( \mathcal{X} \) being Hausdorff, \( \exists O_0 \in \mathcal{O}_\mathcal{X} \) such that \( x_0 \in O_0 \) and \( x_0 \notin O_0 \). This leads to \( \forall x \in (O_0 \cap D) \setminus \{x_0\} \), we have \( D \ni f(x) \neq f(x_0) = y_0 \) by \( f \) being bijective. This implies that \( f(O_0 \setminus \{x_0\}) \subseteq D \setminus \{y_0\} \). Then, \( \lim_{x \to x_0} g(f(x)) = z_0 \in \mathbb{Z} \) by Proposition 3.76. Hence, \( \lim_{x \to x_0} g(f(x)) = lim_{y \to y_0} g(y) = z_0 \in \mathbb{Z} \) in all three subcases. Hence, the result holds in this case.

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

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

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

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

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

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

Proof (i) \( \Rightarrow \) (ii). Fix any net \( (x_\alpha)_{\alpha \in \mathcal{A}} \subseteq D \setminus \{x_0\} \) 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 \( f(V \setminus \{x_0\}) \subseteq U \). \( \exists A_0 \in \mathcal{A} \) such that \( \forall \alpha \in \mathcal{A} \) with \( \alpha_0 < \alpha \), we have \( x_\alpha \in V \). Then, \( x_\alpha \in (D \cap V) \setminus \{x_0\} \) and \( f(x_\alpha) \in U \). Hence, \( (f(x_\alpha))_{\alpha \in \mathcal{A}} \) 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, $\exists U_0 \in O_Y$ with $y_0 \in U_0$, $\forall V \in O_X$ 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 $f(x_V) \notin U_0$. Let $M := \{V \in O_X \mid x_0 \in V\}$ and $A := (M, \supseteq)$. Clearly, $A$ is a directed system. By Axiom of Choice, we may construct a net $(x_V)_{V \in A} \subseteq D \setminus \{x_0\}$. Clearly, $x_0$ is a limit point of this net. But, $\forall V \in A$, $f(x_V) \not\subseteq U_0$. Hence, $y_0$ is not a limit point of the net $(f(x_V))_{V \in A}$. 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: $\forall x \in \mathbb{R}$,

$$
\begin{align*}
  x + \infty &= \infty & x - \infty &= -\infty \\
  x \cdot \infty &= \infty; & \text{if } x > 0 \\
  x \cdot (-\infty) &= -\infty; & \text{if } x < 0 \\
  \infty + \infty &= \infty & -\infty - \infty &= -\infty \\
  \infty \cdot \infty &= \infty & \infty \cdot (-\infty) &= -\infty
\end{align*}
$$

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. $0, \mathbb{R}, \mathbb{R}_e \in \mathcal{B}_{\mathbb{R}_e}$. $\forall r_1, r_2 \in \mathbb{Q}$ 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 $O_{\mathbb{R}_e}$, which is the usual topology on $\mathbb{R}_e$. It is easy to show that $(\mathbb{R}_e, O_{\mathbb{R}_e})$ 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 $O$ that equals to the usual topology $O_{\mathbb{R}}$ 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) $\sup_{x \in X} (f_1 + f_2)(x) \leq \sup_{x \in X} f_1(x) + \sup_{x \in X} f_2(x)$;
\section{Nets and Convergence}

\begin{itemize}
  \item \(\forall \alpha \in (0, \infty) \subseteq \mathbb{R}, \sup_{x \in X}(\alpha g_1(x)) = \alpha \sup_{x \in X} g_1(x); \forall \alpha \in [0, \infty) \subseteq \mathbb{R}, \sup_{x \in X}(\alpha f_1(x)) = \alpha \sup_{x \in X} f_1(x) \) when \(\sup_{x \in X} f_1(x) \in \mathbb{R};\)
  \item \(\sup_{x \in X}(g_1 + g_2)(x) \leq \sup_{x \in X} g_1(x) + \sup_{x \in X} g_2(x), \) when the right-hand-side is well defined.
\end{itemize}

\textbf{Proof} \quad This is straightforward, and is therefore omitted. \hfill \Box

\begin{definition}
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

\[
\limsup_{\alpha \in A} x_\alpha = \inf_{\alpha \in A} \sup_{\beta \in A \text{ with } \alpha < \beta} x_\beta \in \mathbb{R}_e
\]
\[
\liminf_{\alpha \in A} x_\alpha = \sup_{\alpha \in A} \inf_{\beta \in A \text{ with } \alpha < \beta} x_\beta \in \mathbb{R}_e
\]
\end{definition}

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

\begin{enumerate}[(i)]
  \item \(\lim_{\alpha \in A} \inf_{\alpha \in A} x_\alpha \leq \limsup_{\alpha \in A} x_\alpha;\)
  \item \(- \liminf_{\alpha \in A} x_\alpha = \limsup_{\alpha \in A} (-x_\alpha);\)
  \item \(\liminf_{\alpha \in A} x_\alpha = \limsup_{\alpha \in A} x_\alpha = L \in \mathbb{R}_e \) if, and only if, \(\lim_{\alpha \in A} x_\alpha = L;\)
  \item \(\limsup_{\alpha \in A} (y_\alpha + z_\alpha) \leq \limsup_{\alpha \in A} y_\alpha + \limsup_{\alpha \in A} z_\alpha, \) when the right-hand-side makes sense;
  \item if \(\lim_{\alpha \in A} y_\alpha = y \in \mathbb{R}, \) then \(\limsup_{\alpha \in A} (y_\alpha + z_\alpha) = y + \limsup_{\alpha \in A} z_\alpha.\)
\end{enumerate}
\end{proposition}

\textbf{Proof} \quad 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 \(V_\alpha \supseteq V_\beta, \forall \alpha, \beta \in A \text{ with } \alpha < \beta.\)

1. Let \(l := \liminf_{\alpha \in A} x_\alpha \in \mathbb{R}_e \) and \(L := \limsup_{\alpha \in A} x_\alpha \in \mathbb{R}_e. \forall m \in \mathbb{R}\) with \(m < l, \sup_{\alpha \in A} \inf_{\beta \in A \text{ with } \alpha < \beta} x_\beta > m\) 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 \(L \geq l.\)

2. Note that, by Proposition 3.81,

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

3. “If” \(\forall m \in \mathbb{R}\) with \(m > L, \exists \alpha_0 \in A \) such that \(x_\alpha \in (-m, m), \forall \alpha \in A \text{ with } \alpha_0 < \alpha.\) 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 $\alpha \in A$ be an accumulation point of $x_\alpha$ and $\lambda$ the right-hand-side makes sense. Hence, we have $\limsup_{\alpha \in A} x_\alpha = \limsup_{\alpha \in A} x_\alpha \\ < L$.

"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}, \limsup_{\alpha \in A} x_\alpha < m$ implies that $\exists \alpha_0 \in A$ such that $\sup_{\beta \in V_{\alpha_0}} x_\beta < m$. Then, $x_\beta \in (-\infty, m)$, \forall $\beta \in V_{\alpha_0}$. Hence, we have $\lim_{\alpha \in A} x_\alpha = -\infty = L$.

Case 2: $L \in \mathbb{R}$. \forall $\epsilon \in (0, \infty) \subset \mathbb{R}, L = \liminf_{\alpha \in A} x_\alpha$ implies that $\exists \alpha_1 \in A$ such that $\inf_{\beta \in V_{\alpha_1}} x_\beta > L - \epsilon$. Then, $x_\beta \in (L - \epsilon, +\infty) \subset \mathbb{R}, \forall \beta \in V_{\alpha_1}$. $L = \limsup_{\alpha \in A} x_\alpha$ implies that $\exists \alpha_2 \in A$ such that $\sup_{\beta \in V_{\alpha_2}} x_\beta < L + \epsilon$. Then, $x_\beta \in (-\infty, L + \epsilon) \subset \mathbb{R}, \forall \beta \in V_{\alpha_2}$. Let $\alpha_0 \in A$ with $\alpha_1 < \alpha_0$ and $\alpha_2 < \alpha_0$. Then, $x_\beta \in (L - \epsilon, L + \epsilon), \forall \beta \in A$ with $\alpha_0 < \beta$. Therefore, $\lim_{\alpha \in A} x_\alpha = L$. Case 3: $L = +\infty$. \forall $\alpha \in A$, \liminf_{\alpha \in A} x_\alpha > M$ implies that $\exists \alpha_0 \in A$ such that $\inf_{\beta \in V_{\alpha_0}} x_\beta > M$. Then, $x_\beta \in (M, \infty), \forall \beta \in V_{\alpha_0}$. Hence, we have $\lim_{\alpha \in A} x_\alpha = +\infty = L$.

(iv) Note that, \forall $\alpha \in A$, by Proposition 3.81,

$$\sup_{\beta \in V_{\alpha}} (y_\beta + z_\beta) \leq \sup_{\beta \in V_{\alpha}} y_\beta + \sup_{\beta \in V_{\alpha}} z_\beta$$

Then, by Proposition 3.81, we have, \forall $\alpha \in A$,

$$\lim_{\alpha \in A} \sup_{\beta \in V_{\alpha}} (y_\alpha + z_\alpha) \leq \sup_{\beta \in V_{\alpha}} y_\beta + \sup_{\beta \in V_{\alpha}} z_\beta$$

\forall $\gamma \in A$, $\exists \alpha_0 \in A$ with $\alpha < \alpha_0$ and $\gamma < \alpha_0$. Then,

$$\lim_{\alpha \in A} \sup_{\beta \in V_{\alpha_0}} (y_\alpha + z_\alpha) = \sup_{\beta \in V_{\alpha_0}} y_\beta + \sup_{\beta \in V_{\alpha_0}} z_\beta$$

Hence, we have $\limsup_{\alpha \in A} (y_\alpha + z_\alpha) \leq \limsup_{\alpha \in A} y_\alpha + \limsup_{\alpha \in A} z_\alpha$, when the right-hand-side makes sense.

(v) By (iv) and (iii), we have $\limsup_{\alpha \in A} (y_\alpha + z_\alpha) \leq y + \limsup_{\alpha \in A} z_\alpha$. Note that $\limsup_{\alpha \in A} z_\alpha = \limsup_{\alpha \in A} (y_\alpha + z_\alpha - y_\alpha) \leq \limsup_{\alpha \in A} (y_\alpha + z_\alpha - y_\alpha) + \limsup_{\alpha \in A} (-y_\alpha)$. Then, we have $\limsup_{\alpha \in A} (y_\alpha + z_\alpha) \geq y + \limsup_{\alpha \in A} z_\alpha$. Hence, we have $\limsup_{\alpha \in A} (y_\alpha + z_\alpha) = y + \limsup_{\alpha \in A} z_\alpha$.

This completes the proof of the proposition.

\begin{proof}
\end{proof}

\begin{definition}
Let $\mathcal{X} := (X, \mathcal{O}_X)$ be topological spaces, $D \subseteq \mathcal{X}$, $f : D \to \mathbb{R}$, and $x_0 \in \mathcal{X}$ be an accumulation point of $D$. Then the limit superior and limit inferior of $f(x)$ as $x \to x_0$ are defined by

$$\limsup_{x \to x_0} f(x) = \inf_{O \in \mathcal{O} \text{ with } x_0 \in O} \sup_{x \in (D \cap \overline{O}) \setminus \{x_0\}} f(x) \in \mathbb{R}$$

$$\liminf_{x \to x_0} f(x) = \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \cap \overline{O}) \setminus \{x_0\}} f(x) \in \mathbb{R}$$

\end{definition}

\begin{proposition}
Let $\mathcal{X}$ be a topological space, $D \subseteq \mathcal{X}$, $x_0 \in \mathcal{X}$ be an accumulation point of $D$, $f : D \to \mathbb{R}$, and $g : D \to \mathbb{R}$. Then, we have

\begin{enumerate}
  \item ...
  \item ...
\end{enumerate}

\end{proposition}
(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 \( m < l \), \( \sup_{x \in O} \inf_{x \in (D \cap O) \setminus \{x_0\}} f(x) > m \) implies that \( \exists U \in \mathcal{O} \) with \( x_0 \in U \) such that \( \inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) > m \). \( \forall O \in \mathcal{O} \) 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, \( m < \inf_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq \inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap U) \setminus \{x_0\}} f(x) \leq \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \). Hence, \( L \geq m \). By the arbitrariness of \( m \), we have \( L \geq l \).

(ii) Note that, by Proposition 3.81,
\[
\limsup_{x \to x_0} (-f)(x) = \inf_{x \in (D \cap V) \setminus \{x_0\}} \left( -\inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) \right),
\]
\[
= \inf_{x \in (D \cap V) \setminus \{x_0\}} \left( \inf_{x \in (D \cap U) \setminus \{x_0\}} f(x) \right),
\]
\[
= \limsup_{x \to x_0} f(x).
\]

(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, \( \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) \leq m \) 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: \( L = -\infty \). \( \forall m \in \mathbb{R} \), \( \limsup_{x \to x_0} f(x) < m \) implies that \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( \sup_{x \in (D \cap V) \setminus \{x_0\}} f(x) < m \). Then, \( f(V \setminus \{x_0\}) \subseteq (-\infty, m) \). Hence, we have \( \lim_{x \to x_0} f(x) = -\infty = L \). Case 2: \( L \in \mathbb{R} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( L = \liminf_{x \to x_0} f(x) \) implies that \( \exists V_1 \in \mathcal{O} \) with \( x_0 \in V_1 \) such that \( \inf_{x \in (D \cap V_1) \setminus \{x_0\}} f(x) > L - \epsilon \). Then, \( f(V_1 \setminus \{x_0\}) \subseteq (L - \epsilon, +\infty) \subset \mathbb{R} \). \( L = \limsup_{x \to x_0} f(x) \) implies that \( \exists V_2 \in \mathcal{O} \) with \( x_0 \in V_2 \) such that \( \sup_{x \in (D \cap V_2) \setminus \{x_0\}} f(x) < L + \epsilon \). 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, \( \forall \mathcal{O} \in \mathcal{O} \) with \( x_0 \in \mathcal{O} \), by Proposition 3.81,

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

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

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

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

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

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

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. \( \Box \)

**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 \( x_0 \), \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists \mathcal{O} \in \mathcal{O}_X \) with \( x_0 \in \mathcal{O} \) such that \( f(x) < f(x_0) + \epsilon \), \( \forall x \in \mathcal{O} \cap D \). Then, \( \sup_{x \in (\mathcal{O} \cap D) \setminus \{x_0\}} f(x) \leq f(x_0) + \epsilon \), 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 \( D \).

\( \exists V \in \mathcal{O}_X \) with \( x_0 \in V \) such that \( V \cap D = \{x_0\} \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), we have
Proposition 3.87 Let $X$ be a topological space, $D \subseteq X$, $x_0 \in X$ be an accumulation point of $D$, and $f : D \rightarrow \mathbb{R}$. 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 $\liminf_{\alpha \in A} f(x_\alpha) = c \in \mathbb{R}_e$. Then, $\liminf_{x \rightarrow x_0^+} f(x) \leq c$.

Proof By Definitions 3.84 and 3.82, we have

\[
\liminf_{x \rightarrow x_0^+} f(x) = \sup_{O \in \mathcal{O} \text{ with } x_0 \in O} \inf_{x \in (D \setminus O) \setminus \{x_0\}} f(x)
\]

\[
\liminf_{\alpha \in A} f(x_\alpha) = \sup_{\alpha \in A} \inf_{\beta \in A \text{ with } \alpha \prec \beta} f(x_\beta) = c
\]

\[\forall O \in \mathcal{O} \text{ with } x_0 \in O, \text{ since } \lim_{\alpha \in A} x_\alpha = x_0, \text{ then } \exists \alpha_0 \in A \text{ such that } \forall \alpha \in A \text{ with } \alpha_0 \prec \alpha, \text{ we have } x_\alpha \in O. \text{ Then, } x_\alpha \in (O \cap D) \setminus \{x_0\}. \text{ This leads to }
\]

\[
\inf_{x \in (D \setminus O) \setminus \{x_0\}} f(x) \leq \inf_{\beta \in A \text{ with } \alpha_0 \prec \beta} f(x_\beta) \leq c
\]

Hence, we have $\liminf_{x \rightarrow x_0^+} f(x) \leq c$. 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 $X := (X, \mathcal{O})$ 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 $A := (A_1 \times A_2, \prec) := A_1 \times A_2$ is a directed system. A joint net is a mapping of the directed system $A_1 \times A_2$ to $X$, denoted by $(x_{\alpha, \beta})_{(\alpha, \beta) \in A_1 \times A_2}$. The joint net is said to admit a (joint) limit point $\hat{x} \in X$ if it admits a limit point $\hat{x}$ when viewed as a net over the directed system $A$. When $X$ is Hausdorff, by Proposition 3.65, the joint net admits at most one joint limit point, which will be denoted by $\lim_{(\alpha, \beta) \in A_1 \times A_2} x_{\alpha, \beta} \in X$ if it exists.

Proposition 3.89 Let $X := [0, \infty) \subseteq \mathbb{R}_e$ with subset topology $\mathcal{O}$, $X := (X, \mathcal{O})$, and $c \in (0, \infty) \subseteq \mathbb{R}$. Then, $+ : X \times X \rightarrow X$ and $c : X \rightarrow X$ 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 \( \forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2 \), 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 \( \forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2 \), 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 \( \forall (\bar{x}_1, \bar{x}_2) \in V_1 \times V_2 \), 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. \( \square \)
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 \(\mathcal{X} := (X, \rho)\) 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 \(\mathcal{O}\) 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 \( \rho(x_3, x_i) \leq \rho(x_3, x) + \rho(x, x_i) < r_1 \) and \( \rho(x_3, x_2) \leq \rho(x_3, x) + \rho(x, x_2) < r_2 \). 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 \( x \in \mathcal{B}_X(x, r_1) \subseteq \overline{\mathcal{B}}_X(x_0, r) \). 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 \( \rho \) referring the the components of \( \{B_{x,r} \} \). 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 \( \mathcal{M} \subseteq \mathcal{O} \). \( \forall r \in \mathcal{O}, \forall x \in O, \exists r \in (0, \infty) \cap \mathbb{Q} \) such that \( \mathcal{B}(x, r) \subseteq O \). 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 \( x \in \mathcal{B}(x_1, r_1) \subseteq \mathcal{B}(x, r/2) \subseteq O \) 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 \( x_B \in B, \forall B \in \mathcal{B} \) with \( B \neq \emptyset \). Let \( D = \{ x_B \in X \mid B \in \mathcal{B}, B \neq \emptyset \} \). Then, \( D \) is countable. \( \forall x \in X \), \( \forall O \in \mathcal{O} \) with \( x \in O \), \( \exists B \in \mathcal{B} \) such that \( x \in B \subseteq O \). Then, \( x_B \in D \cap B \subseteq D \cap O \neq \emptyset \). 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 \( f : \mathcal{X} \rightarrow \mathcal{Y}, g : \mathcal{Y} \rightarrow \mathcal{Z}, h : \mathcal{Y} \rightarrow \mathcal{X}, x_0 \in \mathcal{X}, \text{ and } y_0 \in \mathcal{Y} \). 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 $h(y) \in U$, $\forall y \in \mathcal{B}_Y(y_0, \delta)$.

**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 \( O_1 = \emptyset \in \mathcal{O} \) and \( O_2 = \mathcal{X} \in \mathcal{O} \), then \( F_1 \subseteq O_1, F_2 \subseteq O_2, O_1 \cap O_2 = \emptyset \). Case 2: \( F_1 \neq \emptyset \) and \( F_2 \neq \emptyset \).

\[ \forall x \in F_1, \, \text{dist}(x, F_2) \in (0, \infty) \subset \mathbb{R}, \text{ by Proposition 4.10.} \]

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

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 \( x_0 \in B(x_1, \text{dist}(x_1, F_2) / 3) \), \( \exists x_2 \in F_2 \) such that \( x_0 \in B(x_2, \text{dist}(x_2, F_1) / 3) \), 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, \( \exists O_1, O_2 \in \mathcal{O} \) such that \( F_1 \subseteq O_1, F_2 \subseteq O_2, O_1 \cap O_2 = \emptyset \). Hence, \( \mathcal{X} \) 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 \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \forall n, m \geq N, \rho(x_n, x_m) < \epsilon. \)

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

"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, \( \mathcal{Y} := (Y, \mathcal{O}) \) be a topological space, \( f : \mathcal{X} \rightarrow \mathcal{Y}, \) 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 \( x \rightarrow x_0; \)

(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) \(\Leftrightarrow\) (ii). This follows from Proposition 3.74.

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

(iii) \(\Rightarrow\) (iv). \(\forall (x_n)_{n=1}^{\infty} \subseteq \mathcal{X}\) with \(x_0\) as a cluster point. Since \(\mathcal{X}\) 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 \((f(x_{n_i}))_{i=1}^{\infty}\) and therefore a cluster point of \((f(x_n))_{n=1}^{\infty}\).

(iv) \(\Rightarrow\) (i). Suppose \(f\) is not continuous at \(x_0\). \(\exists O_{Y_0} \in \mathcal{O}_Y\) with \(f(x_0) \in O_{Y_0}\) 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 \(f(x_n) \in O_{Y_0}\). 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 \(O_{Y_0} \ni f(x_0)\), \(\forall n \in \mathbb{N}\), \(f(x_n) \in O_{Y_0}\). 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.

\[\Box\]

**Proposition 4.16** Let \(\mathcal{X} := (X, \rho)\) be a metric space, \(\mathcal{Y} := (Y, \mathcal{O})\) be a topological space, \(D \subseteq \mathcal{X}\), \(f : D \to \mathcal{Y}\), \(x_0 \in \mathcal{X}\) 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 \(\forall \varepsilon \in (D \cap \mathcal{B}_X(x_0, \delta)) \setminus \{x_0\}, f(x) \in O\). \(\forall (x_n)_{n=1}^{\infty} \subseteq D \setminus \{x_0\}\) 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, \(x_n \in (D \cap \mathcal{B}_X(x_0, \delta)) \setminus \{x_0\}\) 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 \(y_0 \in O_0\), \(\forall n \in \mathbb{N}\), \(\exists x_n \in (D \cap \mathcal{B}_X(x_0, 1/n)) \setminus \{x_0\}\) such that \(f(x_n) \in \mathcal{X} \setminus O_0\). 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.

\[\Box\]

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

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

\[
\liminf_{x \to x_0} f(x) = \sup_{\varepsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap \mathcal{B}(x_0, \varepsilon)) \setminus \{x_0\}} f(x)
\]
4.3 Uniform Continuity and Uniformity

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

Clearly, a function $f$ is uniformly continuous implies that it is continuous.

**Definition 4.19** Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces and $f : \mathcal{X} \to \mathcal{Y}$. $f$ 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** Let $L = \limsup_{x \to x_0} f(x) \in \mathbb{R}$ and $\bar{L} := \inf_{x \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap \mathcal{B}(x_0, \rho)) \setminus \{x_0\}} f(x) \in \mathbb{R}$. $\forall m \in \mathbb{R}$ with $m < L$, we have $m < \sup_{x \in (D \cap \mathcal{B}(x_0, \rho)) \setminus \{x_0\}} f(x)$, $\forall V \in \mathcal{O}_X$ with $x_0 \in V$. Then, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\sup_{x \in (D \cap \mathcal{B}(x_0, \rho)) \setminus \{x_0\}} f(x) > m$. Hence, $m \leq \bar{L}$. By the arbitrariness of $m$, we have $L \leq \bar{L}$. Hence, $L = \bar{L}$.

Note that, by Propositions 3.85 and 3.81,

$$\liminf_{x \to x_0} f(x) = -\limsup_{x \to x_0} (-f(x)) = -\inf_{\varepsilon \in (0, \infty) \subset \mathbb{R}} \sup_{x \in (D \cap \mathcal{B}(x_0, \rho)) \setminus \{x_0\}} (-f(x)) = \sup_{\varepsilon \in (0, \infty) \subset \mathbb{R}} \inf_{x \in (D \cap \mathcal{B}(x_0, \rho)) \setminus \{x_0\}} f(x).$$

This completes the proof of the proposition.  

Properties preserved under uniform homeomorphisms are called uniform properties. These includes Cauchy sequences, completeness, uniform continuity, and total boundedness.
we have
\[
\sigma(x_1, x_2) \leq \frac{\rho(x_1, x_3) + \rho(x_3, x_2)}{1 + \rho(x_1, x_3) + \rho(x_3, x_2)} \leq \frac{\rho(x_1, x_3) + \rho(x_3, x_2)}{1 + \rho(x_1, x_3)} + \frac{\rho(x_3, x_2)}{1 + \rho(x_1, x_3)}
\]
\[
= \sigma(x_1, x_3) + \sigma(x_3, x_2)
\]

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

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

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

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

**Proposition 4.23** Let \( X, Y, W, \) and \( Z \) be metric spaces, \( (x_n)_{n=1}^{\infty} \subseteq X \) be a Cauchy sequence, \( f : X \rightarrow Y \) and \( g : Y \rightarrow Z \) be uniformly continuous functions, and \( h : X \rightarrow W \) 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 \( X \) is complete, then \( W \) is complete.

(iv) If \( X \) is totally bounded and \( f \) is surjective, then \( Y \) 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 \( \rho_Y(f(x_a), f(x_b)) < \epsilon, \forall x_a, x_b \in X \) with \( \rho_X(x_a, x_b) < \delta \). Since \( (x_n)_{n=1}^{\infty} \) is Cauchy, then \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_X(x_n, x_m) < \delta, \forall n, m \geq N \). Then, \( \rho_Y(f(x_n), f(x_m)) < \epsilon \). 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 \( \rho_Z(g(y_1), g(y_2)) < \epsilon, \forall y_1, y_2 \in Y \) with \( \rho_Y(y_1, y_2) < \delta_1 \). By the uniform continuity of \( f \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( \rho_X(f(x_a), f(x_b)) < \delta_1, \forall x_a, x_b \in X \) with \( \rho_X(x_a, x_b) < \delta \). Then, we have \( \rho_Z(g(f(x_a)), g(f(x_b))) < \epsilon \). Hence, \( g \circ f \) is uniformly continuous.

(iii) \( \forall \) Cauchy sequence \( (w_i)_{i=1}^{\infty} \subseteq W \). By (i), \( (h_{inv}(w_i))_{i=1}^{\infty} \subseteq X \) is a Cauchy sequence. Since \( X \) is complete, then \( \lim_{i \in \mathbb{N}} h_{inv}(w_i) = x_0 \in X \). By Proposition 3.66, we have \( \lim_{i \in \mathbb{N}} w_i = \lim_{i \in \mathbb{N}} h(h_{inv}(w_i)) = h(x_0) \in Y \). Hence, \( W \) 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 \( \rho_Y(f(x_a), f(x_b)) < \epsilon, \forall x_a, x_b \in X \) with \( \rho_X(x_a, x_b) < \delta \). By the total boundedness of \( X \), there exists a finite set \( X_N \subseteq X \), such that
$\bigcup_{x \in X} B_x(x, \delta) = X$. Then, by the surjectiveness of $f$ and Proposition 2.5, we have $\bigcup_{x \in X} f(B_x(x, \delta)) = Y$. Note that $f(B_x(x, \delta)) \subseteq B_{\rho_x}(f(x), \epsilon), \forall x \in X$. 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. \hfill \square

**Definition 4.24** Let $X$ be a set and $Y := (Y, \rho)$ be a metric space. Let $(f_{\alpha})_{\alpha \in \mathcal{A}}$ 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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \alpha_0 \in \mathcal{A}, \forall \alpha \in \mathcal{A}$ with $\alpha_0 < \alpha$, we have $\rho(f_{\alpha_0}(x), f(x)) < \epsilon, \forall x \in X$.

**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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \forall n, m \in \mathbb{N}$ 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** $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, \forall x \in \mathcal{X}$, 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. \hfill \square

**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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists U \in \mathcal{O}$ 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.
4.4. PRODUCT METRIC SPACES

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 $\mathcal{M}_m := \{B_{X \times Y}((x, y), r) \mid x \in X, y \in Y, r \in (0, \infty) \subset \mathbb{R}\}$.

Thus, $\forall B_{X \times Y}((x_0, y_0), r) \in \mathcal{M}_m$. $\forall (x, y) \in B_{X \times Y}((x_0, y_0), r)$. 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 $\rho((\bar{x}, \bar{y}), (x_0, y_0)) \leq \rho((\bar{x}, \bar{y}), (x, y)) + \rho((x, y), (x_0, y_0)) = ((\rho_X(\bar{x}, x))^2 + (\rho_Y(\bar{y}, y))^2)^{1/2} = \sqrt{d^2 + r^2}$. 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 $B_Y(x, d_1) \subseteq O_X$ 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 $\delta > ((\rho_X(x_1, y_1))^2 + (\rho_X(x_2, y_2))^2)^{1/2} \geq (\rho_X(x_1, y_1) + \rho_X(x_2, y_2))/\sqrt{2}$. Note that

$$
\rho_X(x_1, x_2) \leq \rho_X(x_1, y_1) + \rho_X(y_1, x_2) \\
\rho_X(y_1, x_2) \leq \rho_X(x_1, x_2) + \rho_X(x_1, y_1) \\
\rho_X(y_1, y_2) \leq \rho_X(x_1, y_2) + \rho_X(x_2, y_2) \\
\rho_X(y_1, y_2) \leq \rho_X(x_2, y_1) + \rho_X(x_2, y_2)
$$

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. \hfill \Box

**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 $\{ (x_n, y_n) \}_{n=1}^{\infty} \subseteq Z$. \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N} \) 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 $Y$, $\exists x_0 \in X$ and $\exists y_0 \in Y$ such that $\lim_{n \to \infty} x_n = x_0$ and $\lim_{n \to \infty} y_n = y_0$. By Proposition 3.67, we have $\lim_{n \to \infty} (x_n, y_n) = (x_0, y_0) \in Z$. Hence, $Z$ is complete. This completes the proof of the proposition. \hfill \Box

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 $\prod_{\emptyset} X_\alpha = (\{\emptyset\}, \rho)$, 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 $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \rightarrow \prod_{\alpha \in \Lambda} X_\alpha$ by, $\forall x \in X^{(\Gamma)}$, $\forall \alpha \in \Lambda$, $\exists \beta_\alpha \in \Gamma$ s.t. $\alpha \in \Lambda_{\beta_\alpha}$, $\pi_\alpha(E(x)) = \pi_{\beta_\alpha}^{(\Gamma)}(\pi_{\beta_\alpha}^{(\Gamma)}(x))$. By Proposition 3.30, $E$ is a homeomorphism. $\forall x, y \in X^{(\Gamma)}$, we have

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

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

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

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

$$
= \left( \sum_{\beta \in \Gamma} \left( \rho_{(\beta)}(\pi_{\beta}^{(\Gamma)}(x)), \pi_{\beta}^{(\Gamma)}(\pi_{\beta}^{(\Gamma)}(y))) \right)^2 \right)^{1/2} = \rho^{(\Gamma)}(x, y)
$$

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

Proof Let \( F_\alpha : X_\alpha \to Y_\alpha \) be a uniform homeomorphism, \( \forall \alpha \in \Lambda \). Define \( F : X \to Y \) by, \( \forall x \in X \), \( \pi^{(Y)}(x) = F_\alpha(\pi^{(X)}(x)) \), \( \forall \alpha \in \Lambda \). By Propositions 4.29 and 3.31, \( F \) is a homeomorphism between \( X \) and \( Y \). 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 \( \Lambda \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall \alpha \in \Lambda, \exists \delta_\alpha \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall x_{\alpha 1}, x_{\alpha 2} \in X_\alpha \) with \( \rho_\alpha(x_{\alpha 1}, x_{\alpha 2}) < \delta_\alpha \), we have \( \rho_\alpha(F_\alpha(x_{\alpha 1}), F_\alpha(x_{\alpha 2})) < \frac{\epsilon}{\sqrt{1 + m}} \), by the uniform continuity of \( F_\alpha \).

Let \( \delta = \min\{\inf_{\alpha \in \Lambda}\delta_\alpha, 1\} \in (0, \infty) \subset \mathbb{R} \). \( \forall x_1, x_2 \in X \) with \( \rho_X(x_1, x_2) < \delta \), we have, \( \forall \alpha \in \Lambda \),

\[
\rho_\alpha(\pi^{(X)}(x_1), \pi^{(X)}(x_2)) < \delta \leq \delta_\alpha
\]

This implies that \( \rho_Y(F_\alpha(\pi^{(X)}(x_1)), F_\alpha(\pi^{(X)}(x_2))) < \frac{\epsilon}{\sqrt{1 + m}} \). Hence, we have

\[
\rho_Y(F(x_1), F(x_2)) = \left( \sum_{\alpha \in \Lambda} \rho_\alpha(\pi^{(X)}(F(x_1)), \pi^{(Y)}(F(x_2))) \right)^{1/2} = \left( \sum_{\alpha \in \Lambda} \rho_\alpha(F_\alpha(\pi^{(X)}(x_1)), F_\alpha(\pi^{(X)}(x_2))) \right)^{1/2} < \epsilon
\]

Hence, \( F \) is uniformly continuous.

It is easy to see that \( F_{\text{inv}} : Y \to X \) is given by, \( \forall y \in Y \), \( \pi^{(X)}(F_{\text{inv}}(y)) = F_{\text{inv}}(\pi^{(Y)}(y)) \), \( \forall \alpha \in \Lambda \). 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. \( \square \)

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 \( X := (X, \rho) \) 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 \( \rho_i(\pi_i(x), \pi_i(y)) = 0 \), \( \forall i \in \mathbb{N} \),

\[
\rho(x, y) = \sum_{i=1}^{\infty} 2^{-i} \frac{\rho_i(\pi_i(x), \pi_i(y))}{1 + \rho_i(\pi_i(x), \pi_i(y))} = 0
\]
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

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

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 \( r \in (0, \infty) \subset \mathbb{R} \). \( \forall y \in B \). 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 \( C_{y_i} = B_{X_i}(\pi_i(y), \delta/2), i = 1, \ldots, N \), and \( C_{y_i} = X_i, i = N+1, N+2, \ldots \). Clearly, \( y \in C_y \in \mathcal{O} \). \( \forall z \in C_y \), we have \( \rho(y, z) < \sum_{i=1}^{N} 2^{-i} \delta/2 + \sum_{i=N+1}^{\infty} 2^{-i} < \delta/2 + 2^{-N} < \delta \). Then, \( \rho(x, z) \leq \rho(x, y) + \rho(y, z) < r \). 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 \( C_i \in \mathcal{C}_i, i = 1, \ldots, N \), 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, \( \forall i = 1, \ldots, N, \exists \delta_i \in (0, \infty) \subset \mathbb{R} \) such that \( B_{X_i}(\pi_i(x), \delta_i) \subseteq C_i \). Let \( \delta_x := \min_{1 \leq i \leq N} 2^{-i} \delta_i > 0 \). Let \( B = B_{X}(x, \delta_x) \in \mathcal{O}_m \). \( \forall y \in B \), \( \rho(x, y) < \delta_x \) implies that \( 2^{-i} \rho_i(\pi_i(x), \pi_i(y)) < 2^{-i} \frac{\delta_i}{1+\delta_i} \) 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, \( \mathcal{O}_m \subseteq \mathcal{O} \).

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 \( \mathcal{Y} := (Y, \rho_Y) := (\prod_{i=2}^{\infty} Y_i, \rho_Y) \), where \( \rho_X \) and \( \rho_Y \) are defined as in Proposition 4.34. Then, \( \mathcal{X} \) and \( \mathcal{Y} \) are uniformly homeomorphic.

**Proof** Let \( F_i : \mathcal{X}_i \to \mathcal{Y}_i \) be a uniform homeomorphism, \( \forall i \in \mathbb{N} \). Define \( F : \mathcal{X} \to \mathcal{Y} \) by, \( \forall x \in \mathcal{X}, \pi_{Y_i}(F(x)) = F_i(\pi_{X_i}(x)), \forall i \in \mathbb{N} \). By
Propositions 4.34 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.

\[ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N} \text{ such that } 2^{-N} < \epsilon/2. \forall i = 1, \ldots, N, \exists \delta_i \in (0, \infty) \subset \mathbb{R} \text{ such that, } \forall x_{i1}, x_{i2} \in X_i \text{ with } \rho_{X_i}(x_{i1}, x_{i2}) < \delta_i, \text{ we have } \rho_{Y_i}(F_i(x_{i1}), F_i(x_{i2})) < \epsilon/2, \text{ by the uniform continuity of } F_i. \]

Let \( \delta = \min_{1 \leq i \leq N} 2^{-i} \frac{\delta_i}{1 + \delta_i} > 0. \forall x_1, x_2 \in \mathcal{X} \text{ with } \rho_{\mathcal{X}}(x_1, x_2) < \delta, \text{ we have, } \forall i = 1, \ldots, N, \]

\[ \frac{\rho_{X_i}(\pi_i^{(X)}(x_1), \pi_i^{(X)}(x_2))}{1 + \rho_{X_i}(\pi_i^{(X)}(x_1), \pi_i^{(X)}(x_2))} < 2^i \delta \leq \frac{\delta_i}{1 + \delta_i}. \]

This implies that \( \rho_{X_i}(\pi_i^{(X)}(x_1), \pi_i^{(X)}(x_2)) < \delta_i, \) which further implies that \( \rho_{Y_i}(F_i(\pi_i^{(X)}(x_1))), F_i(\pi_i^{(X)}(x_2))) < \epsilon/2. \) Hence, we have

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

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

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

Hence, $F$ is uniformly continuous.

It is easy to see that $F_{\text{inv}} : \mathcal{Y} \to \mathcal{X}$ is given by, $\forall y \in \mathcal{Y}, \pi_i^{(X)}(F_{\text{inv}}(y)) = F_{\text{inv}}(\pi_i^{(Y)}(y)), \forall i \in \mathbb{N}$. 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. \( \square \)

### 4.5 Subspaces

**Proposition 4.37** Let $\mathcal{X} := (X, \rho)$ be a metric space with natural topology $\mathcal{O}$ and $A \subseteq \mathcal{X}$ be a subset. $A := (A, \rho)$ is a metric space with natural topology $\mathcal{O}_A$. A may also be endowed with the subset topology $\mathcal{O}_s$ with respect to $(X, \mathcal{O})$. Then, $\mathcal{O}_A = \mathcal{O}_s$.

**Proof** A basis for the natural topology $\mathcal{O}_A$ is

\[ \mathcal{B} := \{ \mathcal{B}_A (x, r) \mid x \in A, r \in (0, \infty) \subseteq \mathbb{R} \} \]

\[ = \{ \mathcal{B}_X (x, r) \cap A \mid x \in A, r \in (0, \infty) \subseteq \mathbb{R} \} \]

$\mathcal{O}_s = \{ O \cap A \mid O \in \mathcal{O} \}$. Clearly, we have $\mathcal{B} \subseteq \mathcal{O}_s$. Then, $\mathcal{O}_A \subseteq \mathcal{O}_s$. On the other hand, $\forall \mathcal{O}_s \subseteq \mathcal{O}_s$, $\mathcal{O}_s = O \cap A$ with $O \in \mathcal{O}$. $\forall x \in \mathcal{O}_s$, $x \in A$ and $\exists r \in (0, \infty) \subset \mathbb{R}$ such that $\mathcal{B}_X (x, r) \subseteq O$. Then, $\mathcal{B}_A (x, r) \subseteq \mathcal{O}_s$. Then, $\mathcal{O}_s \subseteq \mathcal{O}_A$. Hence, $\mathcal{O}_A = \mathcal{O}_s$.

This completes the proof of the proposition. \( \square \)
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 3.5, \( S \subseteq \overline{D} \). By Proposition 3.3, \( \overline{D} \subseteq \overline{S} \subseteq \overline{D} \). Hence, \( \overline{D} = \overline{S} \) and \( D \) is dense in \( \overline{S} \). Hence, \( \overline{S} \) is separable. This completes the proof of the proposition.

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, \( (S, \rho) \) 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.

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

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, \( \forall n \in \mathbb{N} \),

\[
\rho(x_n, x_0) = \lim_{m \in \mathbb{N}} \rho(x_n, x_m) < \frac{1}{2}(r_n + 2^{-1} r_n + 2^{-2} r_n + \cdots) = r_n
\]

Then, \( x_0 \in \mathcal{B}(x_n, r_n) \subseteq O_n \). 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. \( \square \)
4.7. COMPLETION OF METRIC SPACES

Theorem 4.41 (Baire Category Theorem) Let \( X \) be a complete metric space, then it is second category everywhere, that is no nonempty open subset of \( X \) 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 \( X \). When \( \Lambda \) is finite, we may add the open dense set \( X \) 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 \( X \). Then, by Proposition 3.38, \( X \) is second category everywhere. This completes the proof of the theorem. \qed

4.7 Completion of Metric Spaces

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

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

This defines a function \( f : X \to Y \).

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

This completes the proof of the proposition. \qed

Definition 4.43 Let \( X \) be a metric space. A net \( (x_\alpha)_{\alpha \in \Lambda} \subseteq X \) is said to be Cauchy if \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \alpha_0 \in \Lambda, \forall \alpha_1, \alpha_2 \in \Lambda \) 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 \( X \) be a metric space. Then, \( X \) is complete if, and only if, every Cauchy net has a limit in \( X \).

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

“Only if” Let \( (x_\alpha)_{\alpha \in \Lambda} \subseteq X \) be a Cauchy net. \( \exists \alpha_1 \in \Lambda \ni \forall \alpha_1, \alpha_2 \in \Lambda \) with \( \alpha_1 < \alpha_2 \), we have \( \rho(x_{\alpha_1}, x_{\alpha_2}) < 1 \). \( \forall n \in \mathbb{N} \) with \( n > 1 \), \( \exists \alpha_n \in \Lambda \) with \( \alpha_{n-1} < \alpha_n \), \( \forall \alpha_1, \alpha_2 \in \Lambda \) 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 \( \alpha_n < \alpha_{n+1} < \cdots < \alpha_{m_1} \) and \( \alpha_n < \alpha_{n+1} < \cdots < \alpha_{m_2} \). Then, \( \rho(x_{\alpha_{m_1}}, x_{\alpha_{m_2}}) < 1/n \). Hence, \( (x_\alpha)_{\alpha=1}^\infty \subseteq X \) is a Cauchy sequence. By the completeness of \( X \), \( \exists x_0 \in X \) such that \( \lim_{n \in \mathbb{N}} x_{\alpha_n} = x_0 \). \( \forall n \in \mathbb{N}, \forall \alpha \in \Lambda \)
with $\alpha_n < \alpha$, $\forall m \in \mathbb{N}$ with $m > n$, we have $\alpha_n < \alpha_{n+1} < \cdots < \alpha_m$ and $\rho(x_n, x_{n_m}) < 1/n$. Then, by Propositions 4.30, 3.66, and 3.67, we have $\rho(x_n, x_0) = \lim_{m \in \mathbb{N}} \rho(x_n, x_{n_m}) \leq 1/n$. Hence, we have $\lim_{n \in \mathbb{N}} x_n = x_0$.

This completes the proof of the proposition. \qed

**Proposition 4.45** Let $X := (X, \mathcal{O})$ be a topological space, $\mathcal{Y} := (Y, \rho)$ be a complete metric space, $D \subseteq X$, $x_0 \in X$ be an accumulation point of $D$, and $f : D \to \mathcal{Y}$. Then, $\lim_{x \to x_0} f(x) \in \mathcal{Y}$ if, and only if, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \mathcal{O} \subseteq \mathcal{O}$ with $x_0 \in \mathcal{O}$, $\forall \bar{x}, \hat{x} \in (D \cap \mathcal{O}) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$.

**Proof** “Sufficiency” Assume that $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \mathcal{O} \subseteq \mathcal{O}$ with $x_0 \in \mathcal{O}$, $\forall \bar{x}, \hat{x} \in (D \cap \mathcal{O}) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$. Define $\mathcal{M} := \{\mathcal{O} \subseteq \mathcal{O} \mid x_0 \in \mathcal{O}\}$. Clearly, $X \in \mathcal{M}$ and $\mathcal{M} \neq \emptyset$. It is easy to see that $\mathcal{A} := (\mathcal{M}, \supseteq)$ is a directed system. Since $x_0$ is an accumulation point of $D$, then $\forall \mathcal{O} \in \mathcal{A} \subseteq (D \cap \mathcal{O}) \setminus \{x_0\} \neq \emptyset$. By Axiom of Choice, $\exists$ a net $(x_0)_{\mathcal{O} \in \mathcal{A}} \subseteq X$ such that $x_0 \in (D \cap \mathcal{O}) \setminus \{x_0\}$, $\forall \mathcal{O} \in \mathcal{A}$. Also, we define a net $(f(x_0))_{\mathcal{O} \in \mathcal{A}} \subseteq Y$ by Axiom of Replacement. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, by the assumption, $\exists \mathcal{O} \in \mathcal{A}$, $\forall \bar{x}, \hat{x} \in (D \cap \mathcal{O}) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon$. Define $\mathcal{O}_1, \mathcal{O}_2 \subseteq \mathcal{A}$ with $\mathcal{O} \supseteq \mathcal{O}_1$ and $\mathcal{O} \supseteq \mathcal{O}_2$, $x_{i_0} \in (D \cap \mathcal{O}_1) \setminus \{x_0\}$, $\forall \mathcal{O} \subseteq \mathcal{O}$, we have $\rho(x_{i_0} f(\bar{x}), f(\hat{x})) < \epsilon$. This shows that the net $(f(x_0))_{\mathcal{O} \in \mathcal{A}} \subseteq Y$ is Cauchy. By Proposition 4.44, $\lim_{\mathcal{O} \in \mathcal{A}} f(x_0) = y_0 \in Y$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \mathcal{O}_1 \in \mathcal{A}$, $\forall \mathcal{O} \in \mathcal{A}$ with $\mathcal{O} \supseteq \mathcal{O}_1$, we have $\rho(y_0, f(x_0)) < \epsilon/2$. By the assumption, $\exists \mathcal{O}_2 \in \mathcal{A}$, $\forall \bar{x}, \hat{x} \in (D \cap \mathcal{O}_2) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), f(\hat{x})) < \epsilon/2$. Let $\mathcal{O}_3 := \mathcal{O}_1 \cap \mathcal{O}_2 \subseteq \mathcal{A}$. Then, $\forall \mathcal{O} \supseteq \mathcal{O}_3$ and $x_{i_0} \in (D \cap \mathcal{O}_3) \setminus \{x_0\}$, we have $\rho(y_0, f(x_0)) \leq \rho(y_0, f(x_{i_0}))) + \rho(f(x_{i_0}), f(x)) < \epsilon$. Hence, we have $\lim_{x \to x_0} f(x) = y_0 \in \mathcal{Y}$.

“Necessity” Let $\lim_{x \to x_0} f(x) = y_0 \in \mathcal{Y}$. Then, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \mathcal{O} \subseteq \mathcal{O}$ with $x_0 \in \mathcal{O}$, $\forall \bar{x} \in (D \cap \mathcal{O}) \setminus \{x_0\}$, we have $\rho(f(\bar{x}), y_0) < \epsilon/2$. Hence, $\forall \bar{x}, \hat{x} \in (D \cap \mathcal{O}) \setminus \{x_0\}$, $\rho(f(\bar{x}), f(\hat{x})) \leq \rho(f(\bar{x}), y_0) + \rho(y_0, f(\hat{x})) < \epsilon$.

This completes the proof of the proposition. \qed

**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 $\lim_{i \in \mathbb{N}} x_i = x_0 \in X$, 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 $\lim_{i \in \mathbb{N}} f(x_i) = y_0$. Let $(\bar{x}_i)_{i=1}^{\infty} \subseteq E$ be any other sequence with $\lim_{i \in \mathbb{N}} \bar{x}_i = x_0$. Then, the sequence $(x_1, x_2, \ldots)$ converges to $x_0$. By Proposition 4.23 and the uniform continuity of $f$, $(f(x_1), f(\bar{x}_1), f(x_2), f(\bar{x}_2), \ldots)$ 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 $\lim_{i \in \mathbb{N}} f(\bar{x}_i) = y_0$. Hence, $y_0$ is dependent only on $x_0$ but not on the sequence $(x_i)_{i=1}^{\infty}$. 


4.7. COMPLETION OF METRIC SPACES

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, \( \exists (x_i^{(1)}), (x_i^{(2)}) \subseteq \overline{E} \) such that \( \lim_{i \to \infty} x_i^{(j)} = x_j, j = 1, 2. \) By the definition of \( g \), we have \( \lim_{i \to \infty} f(x_i^{(j)}) = g(x_j), j = 1, 2. \) By Proposition 3.67, we have \( \lim_{i \to \infty} (f(x_i^{(1)}), f(x_i^{(2)})) = (g(x_1), g(x_2)). \) By Propositions 3.66 and 4.30, we have \( \lim_{i \to \infty} \rho_Y(f(x_i^{(1)}), f(x_i^{(2)})) = \rho_Y(g(x_1), g(x_2)). \) 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 \( h|_E = f. \) \( \forall x \in \overline{E}, \) by Proposition 4.13, \( \exists (x_i) \subseteq \overline{E} \) such that \( \lim_{i \to \infty} x_i = x. \) By continuity of \( h \) and \( g \) and Proposition 3.66, we have \( h(\bar{x}) = \lim_{i \to \infty} h(x_i) = \lim_{i \to \infty} f(x_i) = \lim_{i \to \infty} g(x_i) = g(\bar{x}). \) 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. \( \Box \)

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

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) \( \rho_Y(Y_1, Y_2) = 0 \Leftrightarrow \forall x_1 \in Y_1, \forall x_2 \in Y_2, \rho(x_1, x_2) = 0 \) and \( x_1 \equiv x_2 \leftrightarrow Y_1 = Y_2; \) (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. \( \Box \)
Theorem 4.49 Let \((X, \rho)\) be an metric space. Then,

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

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

Proof  
(i) Let \(\mathcal{X} := \{ (x_n)_{n=1}^{\infty} \subseteq X \mid (x_n)_{n=1}^{\infty} \text{ is a Cauchy sequence} \}\). Define \(\sigma : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}\) by \(\sigma\left( (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \right) = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)})\), \(\forall (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \in \mathcal{X}\). The mapping \(\sigma\) is well defined by Propositions 4.30 and 4.23 and the fact that \((x_n^{(1)}, x_n^{(2)})_{n=1}^{\infty}\) is a Cauchy sequence.

\[
\forall (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty}, (x_n^{(3)})_{n=1}^{\infty} \in \mathcal{X}, \quad \sigma\left( (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \right) = \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)}) = \lim_{n \in \mathbb{N}} \rho(x_n^{(2)}, x_n^{(1)}) = \sigma\left( (x_n^{(2)})_{n=1}^{\infty}, (x_n^{(1)})_{n=1}^{\infty} \right);
\]

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

By Lemma 4.48, we can define the equivalence relation \(\equiv\) on \(\mathcal{X}\) by \((x_n^{(1)})_{n=1}^{\infty} \equiv (x_n^{(2)})_{n=1}^{\infty} \iff \lim_{n \in \mathbb{N}} \rho(x_n^{(1)}, x_n^{(2)}) = 0\), \(\forall (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \in \mathcal{X}\). Then, define the set \(\bar{X} := \mathcal{X}/\equiv = \{ F \subseteq \mathcal{X} \mid F \neq \emptyset, \forall (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \in F, (x_n^{(1)})_{n=1}^{\infty} \equiv (x_n^{(2)})_{n=1}^{\infty} \forall (x_n^{(3)})_{n=1}^{\infty} \in \mathcal{X} \setminus F, (x_n^{(1)})_{n=1}^{\infty} \not\equiv (x_n^{(3)})_{n=1}^{\infty} \}\). A metric on \(\bar{X}\) is \(\bar{\rho}\), defined by \(\bar{\rho}\left(F_1, F_2\right) = \sigma\left( (x_n^{(1)})_{n=1}^{\infty}, (x_n^{(2)})_{n=1}^{\infty} \right), \forall F_1, F_2 \in \bar{X}\), for some \((x_n^{(1)})_{n=1}^{\infty} \in F_1\) and \((x_n^{(2)})_{n=1}^{\infty} \in F_2\).

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

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

Next, we show that \((\bar{X}, \bar{\rho})\) is complete. Fix a Cauchy sequence \( \left( F_n \right)_{n=1}^{\infty} \subseteq \bar{X} \). \( \forall n \in \mathbb{N} \), let \( \left( x^{(n)}_{i} \right)_{i=1}^{\infty} \in F_n \). Note that \( \left( x^{(n)}_{i} \right)_{i=1}^{\infty} \subseteq X \) is a Cauchy sequence. \( \forall n, m \in \mathbb{N} \), \( \bar{\rho}(F_n, F_m) = \sigma \left( \left( x^{(n)}_{i} \right)_{i=1}^{\infty}, \left( x^{(m)}_{i} \right)_{i=1}^{\infty} \right) \).

Define a sequence \( \left( x^{(i)}_{i} \right)_{i=1}^{\infty} \subseteq X \) as following. Set \( n_0 = 0 \). \( \forall i \in \mathbb{N} \), choose \( n_i \in \mathbb{N} \) 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 \( x^{(i)}_{0} = x^{(i)}_{n_i} \).

**Claim 4.49.1** \( \left( x^{(i)}_{i} \right)_{i=1}^{\infty} \subseteq X \).

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

\[
\rho(x^{(i)}_{i}, x^{(j)}_{j}) \leq \rho(x^{(i)}_{n_i}, x^{(i)}_{i_0}) + \rho(x^{(i)}_{i_0}, x^{(j)}_{i_0}) + \rho(x^{(j)}_{i_0}, x^{(j)}_{n_j}) < 1/i + 2\epsilon + 1/j < 4\epsilon
\]

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

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

**Claim 4.49.2** \( \lim_{n \in \mathbb{N}} F_n = F_0 \).

**Proof of claim:** \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists N \in \mathbb{N} \) 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 \( i_0 := n_m \geq m \geq N \). \( \forall i \geq i_0 \), \( \exists 1 > n_i \geq i \geq n_m \) such that \( 0 \leq \rho(x^{(m)}_{i}, x^{(i)}_{i}) < \bar{\rho}(F_{n_m}, F_{i}) + \epsilon < 2\epsilon \). Then,

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

Hence, \( 0 \leq \bar{\rho}(F_{n_m}, F_{0}) = \lim_{i \in \mathbb{N}} \rho(x^{(m)}_{i}, x^{(i)}_{i}) \leq 4\epsilon \). Therefore, we have \( \bar{\rho}(F_{n_m}, F_{0}) = 0 \). This completes the proof of the claim. \( \square \)

Hence, we have shown that any Cauchy sequence \( \left( F_n \right)_{n=1}^{\infty} \subseteq \bar{X} \) 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 \( \left( x_n \right)_{n=1}^{\infty} \subseteq F \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists N \in \mathbb{N} \) 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 \( T : X \rightarrow \bar{X} \) is uniformly continuous. By Proposition 4.46, \( \exists G : X_c \rightarrow \bar{X} \) such that \( G|_X = T \) and \( G \) is uniformly continuous. Note that \( T_{inv} : X_I \rightarrow X \subseteq Y \) is uniformly continuous, since \( T_{inv} \) is metric preserving. Then, by Propositions 4.39 and 4.46, \( \exists H : \bar{X} \rightarrow X_c \) such that \( H|_{X_I} = T_{inv} \) and \( H \) is uniformly continuous. \( G \circ H : \bar{X} \rightarrow \bar{X} \) is uniformly continuous, by Proposition 4.23. Note that \( (G \circ H)|_X = \text{id}_X \). By Proposition 3.56, we have \( G \circ H = \text{id}_X \). \( H \circ G : X_c \rightarrow X_c \) is uniformly continuous, by Proposition 4.23. Note that \( (H \circ G)|_X = \text{id}_X \). By Proposition 3.56, we have \( H \circ G = \text{id}_X \). Therefore, \( G : X_c \rightarrow \bar{X} \) is bijective with inverse \( H : \bar{X} \rightarrow X_c \), by Proposition 2.4. Hence, \( X_c \) and \( \bar{X} \) are uniformly homeomorphic.

\[ \forall y_1, y_2 \in X_c, \exists \left( \left( x_i^{(1)} \right)_{i=1}^\infty, \left( x_i^{(2)} \right)_{i=1}^\infty \right) \in \mathcal{X} \text{ such that } \lim_{i \in \mathbb{N}} x_i^{(1)} = y_j, j = 1, 2, \text{ by Proposition 4.13. Then, by Proposition 3.66, we have } \]

\[ G(y_1) = \lim_{i \in \mathbb{N}} G(x_i^{(1)}) = \lim_{i \in \mathbb{N}} T(x_i^{(1)}) = \lim_{i \in \mathbb{N}} \rho(x_i^{(1)}, x_i^{(2)}) = \lim_{i \in \mathbb{N}} \rho(x_i^{(1)}, x_i^{(2)}) = \rho(y_1, y_2) \]  

Hence, \( G : X_c \rightarrow \bar{X} \) 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 \((X, O)\) is said to be metrizable if there is a metric \( \rho \) on \( X \) whose natural topology is exactly the same as \( O \).

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

**Proof** \( \forall \alpha \in \Lambda \), let \( \rho_\alpha \) be a metric on \( X_\alpha \) whose natural topology is exactly \( O_\alpha \). 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, \( (X, O) = (\{\emptyset\}, \{\emptyset, \{\emptyset\}\}) \). 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 \( O \). Hence, \( \mathcal{X} \) is metrizable.
4.9 Interchange Limits

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 \( f_{\alpha_1, \alpha_2} : X \to [0, 1] := I \subseteq \mathbb{R} \) such that \( f_{\alpha_1, \alpha_2}|_{\overline{B_{\alpha_1}}} = 0 \) and \( f_{\alpha_1, \alpha_2}|_{\overline{B_{\alpha_2}}} = 1 \).

Define \( \mathcal{F} := \{ f_{\alpha_1, \alpha_2} : X \to I \mid \alpha_1, \alpha_2 \in \Lambda, B_{\alpha_1} \cap \overline{B_{\alpha_2}} = \emptyset \} \). Clearly, \( \mathcal{F} \) is a countable set. \( \forall x_0 \in X \), \forall closed set \( F \subseteq \mathcal{X} \) with \( x_0 \notin F \), we have \( \exists \alpha_1 \in \Lambda \) such that \( x_0 \in B_{\alpha_1} \subseteq \overline{F} \). 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 \( \{ x_0 \} \subseteq B_{\alpha_2} \subseteq \overline{B_{\alpha_2}} \subseteq B_{\alpha_1} \), and hence \( B_{\alpha_1} \cap \overline{B_{\alpha_2}} = \emptyset \). 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 \( \forall x_0 \in X \), \forall closed set \( F \subseteq \mathcal{X} \) with \( x_0 \notin F \), \exists \alpha \in \Lambda \) such that \( f_{\alpha}(x_0) = 1 \) 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 \( E : \mathcal{X} \to I^\mathcal{X} \) 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 \( \rho_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \) 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. \( \square \)

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 \( \lim_{(\alpha, \beta) \in A_1 \times A_2} y_{\alpha, \beta} = y \) and, \( \forall \alpha \in A_1 \), \( \lim_{\beta \in A_2} y_{\alpha, \beta} = y_{\alpha} \in \mathcal{Y} \). Then, \( \lim_{\alpha \in A_1} y_{\alpha} = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha, \beta} = \lim_{(\alpha, \beta) \in A_1 \times A_2} y_{\alpha, \beta} \).

Proof Let \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), 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 \( (\alpha_0, \beta_0) \prec (\alpha, \beta) \), we have
\[ \rho(y_{\alpha,\beta}, \hat{y}) < \epsilon/2. \]
Fix any \( \alpha \in A_1 \) with \( \alpha_0 \prec_1 \alpha \). By the fact that 
\[ \lim_{\beta \in A_2} y_{\alpha,\beta} = \hat{y}_{\alpha} \in \mathcal{Y}, \exists \beta_{\alpha} \in A_2, \forall \beta \in A_2 \text{ with } \beta_{\alpha} \prec_2 \beta, \text{ we have } \rho(y_{\alpha,\beta}, \hat{y}_{\alpha}) < \epsilon/2. \]
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 
\[ \rho(y_{\alpha,\beta}, \hat{y}) \leq \rho(y_{\alpha,\beta}, \hat{y}_{\alpha}) + \rho(y_{\alpha,\beta}, \hat{y}) < \epsilon. \]
This shows that \( \lim_{\alpha \in A_1} \hat{y}_{\alpha} = \hat{y} \). Hence, we have 
\[ \lim_{\alpha \in A_1} \hat{y}_{\alpha} = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta}. \]
This completes the proof of the theorem. \( \square \)

**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 
\[ \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \hat{y} \in \mathcal{Y}, \forall \alpha \in A_1, \lim_{\beta \in A_2} y_{\alpha,\beta} = \hat{y}_{\alpha} \in \mathcal{Y}, \text{ and } \forall \beta \in A_2, \lim_{\alpha \in A_1} y_{\alpha,\beta} = \hat{y}_{\beta} \in \mathcal{Y}. \]
Then, 
\[ \lim_{\alpha \in A_1} \hat{y}_{\alpha} = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \lim_{\beta \in A_2} \lim_{\alpha \in A_1} y_{\alpha,\beta}. \]

**Proof** This is straightforward by Joint Limit Theorem 4.54. \( \square \)

**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 
(i) \( \forall \alpha \in A_1, \lim_{\beta \in A_2} y_{\alpha,\beta} = \hat{y}_{\alpha} \in \mathcal{Y} \);
(ii) the nets \( (y_{\alpha,\beta})_{\alpha \in A_1} \) converge uniformly to \( \hat{y}_{\beta} \in \mathcal{Y}, \forall \beta \in A_2 \).
Then, 
\[ \lim_{\alpha \in A_1} \hat{y}_{\alpha} = \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \lim_{\beta \in A_2} \lim_{\alpha \in A_1} y_{\alpha,\beta}. \]

**Proof** By (ii), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \alpha_0 \in A_1, \forall \alpha \in A_1 \) with 
\( \alpha_0 \prec_1 \alpha, \forall \beta \in A_2 \), we have 
\( \rho(y_{\alpha,\beta}, \hat{y}_{\beta}) < \epsilon/4. \) Fix an \( \alpha_1 \in A_1 \) with 
\( \alpha_0 \prec_1 \alpha_1 \). Since \( \lim_{\beta \in A_2} y_{\alpha_1,\beta} = \hat{y}_{\alpha_1}, \exists \beta_0 \in A_2, \forall \beta \in A_2 \) with 
\( \beta_0 \prec_2 \beta \), we have 
\( \rho(y_{\alpha_1,\beta}, \hat{y}_{\alpha_1}) < \epsilon/4. \) Then, \( \forall \beta_1, \beta_2 \in A_2 \) with 
\( \beta_1 \prec_2 \beta_2 \) and \( \beta_0 \prec_2 \beta_1 \), we have 
\[ \rho(y_{\beta_1,\beta_2}, \hat{y}_{\beta_1} + \rho(y_{\alpha_1,\beta_1}, \hat{y}_{\alpha_1}) + \rho(y_{\beta_1,\beta_2}, \hat{y}_{\beta_2}) < \epsilon. \]
Hence, the net \( (\beta_{\beta})_{\beta \in A_2} \) is a Cauchy net.
By Proposition 4.44, \( \lim_{\beta \in A_2} \hat{y}_{\beta} = \hat{y} \in \mathcal{Y}. \)
Then, \( \exists \beta \in A_2, \forall \beta \in A_2 \) with \( \beta \prec_2 \beta \), we have 
\( \rho(\hat{y}_{\beta}, \hat{y}) < \epsilon/2. \) Note that 
\( \forall (\alpha, \beta) \in A_1 \times A_2 \), we have 
\( \rho(y_{\alpha,\beta}, \hat{y}) \leq \rho(y_{\alpha,\beta}, \hat{y}_{\beta}) + \rho(\hat{y}_{\beta}, \hat{y}) < \epsilon. \)
Hence, \( \lim_{(\alpha,\beta) \in A_1 \times A_2} y_{\alpha,\beta} = \hat{y}. \) Finally, by Joint Limit Theorem 4.54, we have 
\( \lim_{\alpha \in A_1} \lim_{\beta \in A_2} y_{\alpha,\beta} = \lim_{\alpha \in A_1} \hat{y}_{\alpha} = \hat{y}. \)
This completes the proof of the theorem. \( \square \)

**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 
\( \lim_{x \to x_0} f_{\alpha}(x) = y_0 \in \mathcal{Y}, \forall \alpha \in A. \) Then, 
\[ \lim_{x \to x_0} y_0 = \lim_{\alpha \in A} \lim_{x \to x_0} f_{\alpha}(x) = \lim_{x \to x_0} \lim_{\alpha \in A} f_{\alpha}(x) = \lim_{x \to x_0} f(x) \in \mathcal{Y}. \]
Fix any net \((x_\beta)_{\beta \in \mathcal{A}} \subseteq D\) with \(x_0\) as a limit. Such net exists by Proposition 3.68. By the assumption that \((f_\alpha)_{\alpha \in \mathcal{A}}\) converges uniformly to \(f\), then, for the joint net \((f_\alpha(x_\beta))_{(\alpha, \beta) \in \mathcal{A} \times \mathcal{A}}\), we have the nets \((f_\alpha(x_\beta))_{\alpha \in \mathcal{A}}\) converge uniformly to \(f(x_\beta) \in \mathcal{Y}, \forall \beta \in \mathcal{A}\). By Proposition 3.79, we have \(\lim_{\beta \in \mathcal{A}} f_\alpha(x_\beta) = y_\alpha, \forall \alpha \in \mathcal{A}\). By Iterated Limit Theorem 4.56, we have \(\lim_{\alpha \in \mathcal{A}} y_\alpha = \lim_{\alpha \in \mathcal{A}} \lim_{\beta \in \mathcal{A}} f_\alpha(x_\beta) = \lim_{(\alpha, \beta) \in \mathcal{A} \times \mathcal{A}} f_\alpha(x_\beta) = \lim_{\alpha \in \mathcal{A}} \lim_{\beta \in \mathcal{A}} f_\alpha(x_\beta) = \lim_{\beta \in \mathcal{A}} f(x_\beta) = y \in \mathcal{Y}\). By the arbitrariness of \((x_\beta)_{\beta \in \mathcal{A}}\) and Proposition 3.79, we have \(y = \lim_{x \to x_0} f(x)\). Hence, we have \(\lim_{\alpha \in \mathcal{A}} y_\alpha = \lim_{\alpha \in \mathcal{A}} \lim_{x \to x_0} f_\alpha(x) = y = \lim_{x \to x_0} f(x) = \lim_{x \to x_0} \lim_{\alpha \in \mathcal{A}} f_\alpha(x) \in \mathcal{Y}\). 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 \(\forall (O_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{O} \text{ with } X \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha\), 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, \((\bar{O}_\alpha)_{\alpha \in \Lambda}\) is a collection of closed sets with the finite intersection property. Yet, \(\bigcap_{\alpha \in \Lambda} \bar{O}_\alpha = \left(\bigcup_{\alpha \in \Lambda} O_\alpha\right)^\sim = \emptyset\). 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 \(\bigcup_{F \in \mathcal{F}} F = \left(\bigcap_{F \in \mathcal{F}} F\right)^\sim = \mathcal{X}\). By the compactness of \(\mathcal{X}\), there exists a finite collection \(\mathcal{F}_N \subseteq \mathcal{F}\) such that
arbitrariness of assumption, covered by any finite subcollection of sets in \( \bigcap_{B \in \Lambda} F_\alpha \neq \emptyset \). Hence, \( F \) 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 \( (x_\alpha)_{\alpha \in A} \subseteq \mathcal{X} \) has a cluster point in \( \mathcal{X} \).

**Proof**  
"Only if" Fix any net \( (x_\alpha)_{\alpha \in A} \subseteq \mathcal{X} \). \( \forall \alpha \in A \), let \( F_\alpha := \{ \beta \mid 0 \leq \alpha \wedge B \in A \} \). Let \( F_\alpha \) be the closure of \( F_\alpha \). For any finite set \( A \subseteq \mathcal{X} \), \( \exists \alpha_0 \in A \) such that \( \alpha \prec \alpha_0, \forall \alpha \in A \), since \( A \) is a directed system. Then, \( x_\alpha \in F_\alpha, \forall \alpha \in A \). Hence, we have \( x_\alpha \in \bigcap_{\alpha \in A} F_\alpha \neq \emptyset \). Hence, \( \bigcap_{\alpha \in A} F_\alpha \) 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 \( \bigcap_{\alpha \in A} F_\alpha \neq \emptyset \). Fix an \( x \in \bigcap_{\alpha \in A} F_\alpha, \forall O \in \mathcal{O} \) with \( x \in O, \forall \alpha \in A \), we have \( x \in F_\alpha \) and \( O \cap F_\alpha \neq \emptyset \). Hence, \( \exists \alpha_0 \in A \) with \( \alpha_0 \prec \alpha \) such that \( x_\alpha \in O \cap F_\alpha \neq \emptyset \). 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 \( X = (\bigcup_{\alpha \in \Lambda} O_\alpha) \cup K \), which forms an open covering of \( X \). By the compactness of \( X \), there exists a finite set \( \Lambda_N \subseteq \Lambda \) such that \( X = (\bigcup_{\alpha \in \Lambda_N} O_\alpha) \cup K \). 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, \( \exists O_1^{(x)}, O_2^{(x)} \in \mathcal{O} \) such that \( x_0 \in O_1^{(x)} \), \( x \in O_2^{(x)} \), and \( O_1^{(x)} \cap O_2^{(x)} = \emptyset \). Then, \( (O_2^{(x)})_{x \in K} \) forms an open covering of \( K \). By the compactness of \( K \), \( \exists \) a finite set \( K_N \subseteq K \) such that \( (O_2^{(x)})_{x \in K_N} \) forms a finite open covering of \( K \).
forms an open covering of $K$. Then, $x_0 \in O := \bigcap_{x \in K_N} O_1^{(x)} \in \mathcal{O}$, and $O \cap K \subseteq O \cap \left( \bigcup_{x \in K_N} O_2^{(x)} \right) = \emptyset$. This contradicts with $x_0 \in \overline{R}$, by Proposition 3.3. Hence, $K$ is closed.

This completes the proof of the proposition.

**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, $c \in A$. $\exists \alpha_0 \in \Lambda$ such that $c \in O_{\alpha_0}$. $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $[c-\delta, c+\delta] \subseteq O_{\alpha_0}$. By the definition of $c$, $\exists \delta \in B$ 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, $\exists \alpha_0 \in \Lambda$ such that $[a, c+\delta]$ is covered by finitely many sets in $(O_\alpha)_{\alpha \in \Lambda}$. Now, adding $O_{\alpha_0}$ to this finitely many sets, then $[a, c+\delta]$ is covered by finitely many sets in $(O_\alpha)_{\alpha \in \Lambda}$. 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 $(\bigcup_{\alpha \in \Lambda} O_\alpha) \cup A \supseteq [a, b]$. By the compactness of $[a, b]$, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $[a, b] \subseteq (\bigcup_{\alpha \in \Lambda_N} O_\alpha) \cup A$. 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 $I_n := (-n, n) \subseteq \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, $\bigcup_{n \in \mathbb{N}} I_n \supseteq A$. By the compactness of $A$, $\exists N \in \mathbb{N}$ such that $A \subseteq \bigcup_{1 \leq \ell \leq N} I_\ell$. Hence, $A$ is bounded.

This completes the proof of the theorem. \qed

**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 $(O_\gamma)_{\gamma \in \Lambda} \subseteq \mathcal{O}_\mathcal{Y}$ be an arbitrary open covering of $f(\mathcal{X})$, where $\Lambda$ is an index set. Then, $(f_{\text{inv}}(O_\gamma))_{\gamma \in \Lambda} \subseteq \mathcal{O}_\mathcal{X}$ is an open covering of $\mathcal{X}$, by the continuity of $f$. By the compactness of $\mathcal{X}$, there exists a finite set $\Lambda_N \subseteq \Lambda$ such that $(f_{\text{inv}}(O_\gamma))_{\gamma \in \Lambda_N}$ is a subcovering of $\mathcal{X}$. Then, by Proposition 2.5, $(O_\Gamma)_{\Gamma \in \Lambda_N}$ is a finite subcovering of $f(\mathcal{X})$. Hence, $f(\mathcal{X})$ is compact. This completes the proof of the proposition. \qed

**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. \( \Box \)

**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 \( \forall \mathcal{V} \in \mathcal{V}, \exists \mathcal{U} \in \mathcal{U} \) such that \( \mathcal{V} \subseteq \mathcal{U} \).

**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 \( \mathcal{U} \). \( \forall \mathcal{V} \in \mathcal{V}, \exists \mathcal{U}_V \in \mathcal{U} \) such that \( \mathcal{V} \subseteq \mathcal{U}_V \). By Axiom of Choice and Axiom of Replacement, we may define a set \( \mathcal{U} := \{ \mathcal{U}_V \in \mathcal{U} \mid V \in \mathcal{V} \} \). Clearly, \( \mathcal{U} \subseteq \mathcal{U} \) is a finite subcovering of \( X \). Hence, \( \mathcal{X} \) is compact.

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

**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 \( K_{n_0} \subseteq O \).

**Proof** By Proposition 5.5, \( K_n \) is closed, \( \forall n \in \mathbb{N} \). Let \( O_n := O \cup \bar{K}_n \), \( \forall n \in \mathbb{N} \). Then, we have

\[
\bigcup_{n=1}^{\infty} O_n = O \cup \left( \bigcup_{n=1}^{\infty} \bar{K}_n \right) = O \cup \left( \bigcap_{n=1}^{\infty} K_n \right)^\sim = X \supseteq K_1
\]

By the compactness of \( K_1 \), \( \exists N \in \mathbb{N} \) such that \( K_1 \subseteq \bigcup_{n=1}^{N} O_n = O \cup \left( \bigcap_{n=1}^{N} K_n \right)^\sim = O \cup \bar{K}_N \). Then, \( K_N \subseteq O \). This completes the proof of the proposition. \( \Box \)

**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 \( \bigcup_{\alpha \in \Lambda} \bar{K}_\alpha = \left( \bigcap_{\alpha \in \Lambda} K_\alpha \right)^\sim = \mathcal{X} \supseteq K_{\alpha_0} \), for some \( \alpha_0 \in \Lambda \). By the compactness of \( K_{\alpha_0} \), there exists a finite set \( \Lambda_N \subseteq \Lambda \) such that \( K_{\alpha_0} \subseteq \bigcup_{\alpha \in \Lambda_N} \bar{K}_\alpha \). Let \( \Lambda_N := \Lambda_N \cup \{ \alpha_0 \} \), which is a finite set. Then, \( \bigcap_{\alpha \in \Lambda_N} K_\alpha = K_{\alpha_0} \cap \left( \bigcap_{\alpha \in \Lambda_N} K_\alpha \right) = \emptyset \). This completes the proof of the proposition. \( \Box \)
5.1. COMPACT SPACES

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 \( \exists O_{x_1}^{(1)}, O_{x_2}^{(2)} \in \mathcal{O} \) such that \( x_j \in O_{x_j}^{(j)}, j = 1, 2, \) and \( O_{x_1}^{(1)} \cap O_{x_2}^{(2)} = \emptyset \). Then, \( \forall x_1 \in K_1, K_2 \subseteq \bigcup_{x_2 \in K_2} O_{x_2}^{(2)} \). By the compactness of \( K_2 \), there exists a finite set \( K_{2,x_1} \subseteq K_2 \) such that \( K_2 \subseteq \bigcup_{x_2 \in K_{2,x_1}} O_{x_2}^{(2)} =: O_{x_1}^{(1)} \in \mathcal{O} \). It is clear that \( x_1 \in \bigcap_{x_2 \in K_{2,x_1}} O_{x_2}^{(2)} =: O_{x_1}^{(1)} \in \mathcal{O} \) and \( O_{x_1}^{(1)} \cap O_{x_2}^{(2)} = \emptyset \). Then, \( K_1 \subseteq \bigcup_{x_2 \in K_{1,n}} O_{x_2}^{(2)} \). By the compactness of \( K_1 \), there exists a finite set \( K_{1,n} \subseteq K_1 \) such that \( K_1 \subseteq \bigcup_{x_2 \in K_{1,n}} O_{x_2}^{(2)} =: O_1 \in \mathcal{O} \). Note that \( K_2 \subseteq \bigcap_{x_2 \in K_{1,n}} O_{x_2}^{(2)} =: O_2 \in \mathcal{O} \). Clearly, \( O_1 \cap O_2 = \emptyset \). This completes the proof of the proposition. \( \square \)

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. \( \square \)

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 \( \mathcal{U}_1 \subseteq \mathcal{O}_1 \) 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 \( \mathcal{U}_N \subseteq \mathcal{U}_1 \) of \((X, \mathcal{O})\). Clearly \( \mathcal{U}_N \subseteq \mathcal{O}_1 \) is a finite subcovering of \((X, \mathcal{O}_1)\). Therefore, \((X, \mathcal{O}_1)\) is compact. This completes the proof of the proposition. \( \square \)

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. \( \square \)

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 \( \mathcal{O}_2 \supseteq \mathcal{O} \) is not compact.
Proof Since $O_1 \subset O$, then $\exists O_0 \in O \setminus O_1$. Then, $\tilde{O_0}$ is closed in $(X,O)$. By Proposition 5.5, $\tilde{O_0}$ is compact in $(X,O)$. By Proposition 5.15, $\tilde{O_0}$ is compact in $(X,O_1)$. Suppose $(X,O_1)$ is Hausdorff. Then, by Proposition 5.5, $\tilde{O_0}$ is closed in $(X,O_1)$. This implies that $O_0 \in O_1$, which contradicts with $O_0 \in O \setminus O_1$. Hence, $(X,O_1)$ is not Hausdorff.

Since $O_2 \supset O$, then $\exists O_0 \in O_2 \setminus O$. Suppose $(X,O_2)$ is compact. Then, by Proposition 5.5, $\tilde{O_0}$ is compact in $(X,O_2)$. By Proposition 5.15, $\tilde{O_0}$ is compact in $(X,O)$. By Proposition 5.5, $\tilde{O_0}$ is closed in $(X,O)$. Then, $O_0 \in O$, which contradicts with $O_0 \in O_2 \setminus O$. Hence, $(X,O_2)$ 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 $\forall O_2 \subseteq Z, g_{inv}(O_2) = f(T_{inv}(O_2)).$

Proof of claim: $\forall y_0 \in g_{inv}(O_2) \subseteq \mathcal{Y}$, $\exists x_0 \in \mathcal{X}$ such that $y_0 = f(x_0)$ since $f$ is surjective. Then, $g(y_0) = g(f(x_0)) = T(x_0) \in O_2$. This implies that $x_0 \in T_{inv}(O_2)$. Then, $y_0 \in f(T_{inv}(O_2))$. Hence, $g_{inv}(O_2) \subseteq f(T_{inv}(O_2))$.

On the other hand, $\forall y_0 \in f(T_{inv}(O_2))$, $\exists x_0 \in T_{inv}(O_2)$ such that $y_0 = f(x_0)$. Then, $O_2 \ni T(x_0) = g(f(x_0)) = g(y_0)$. This implies that $y_0 \in g_{inv}(O_2)$. Hence, $f(T_{inv}(O_2)) \subseteq g_{inv}(O_2)$.

Therefore, we have $g_{inv}(O_2) = f(T_{inv}(O_2))$. This completes the proof of the claim. □

$\forall$ closed set $F \subseteq Z$. $T_{inv}(F)$ is closed in $\mathcal{X}$ by Proposition 3.10 and the continuity of $T$. Then, $T_{inv}(F)$ is compact by Proposition 5.5. This implies that, by Proposition 5.7, $f(T_{inv}(F))$ is compact. By Proposition 5.5, $g_{inv}(F) = f(T_{inv}(F))$ 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 $\mathcal{X} := (X,O)$ 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 $(X, O)$ 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. □

**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. □

**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 $(O_\alpha)_{\alpha \in \Lambda} \subseteq O_Y$ 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. □
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 (i) 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, \( x_0 \in (x_n)_{n=1}^{\infty} \subseteq X \). Since \( x_0 \in (x_i)_{i=n}^{\infty} \), \( (x_i)_{i=n}^{\infty} \subseteq \bigcap_{i=n}^{\infty} \overline{O}_i \), and \( \bigcap_{i=1}^{n} \overline{O}_i \) is closed, then \( x_0 \in \bigcap_{i=n}^{\infty} \overline{O}_i \). 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 \( B_n := (x_i)_{i=n}^{\infty} \subseteq X \). Suppose \( \bigcap_{n=1}^{\infty} B_n = \emptyset \). Then, \( \left( B_n \right)_{n \in \mathbb{N}} \) 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 \( x_0 \in \bigcap_{n=1}^{\infty} B_n \subseteq \mathcal{X} \). \( \forall O \in \mathcal{O} \) with \( x_0 \in O \), \( \forall n \in \mathbb{N} \), \( x_0 \in B_n \) 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.

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
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 \( x_{n_i} \in \bigcap_{j=1}^{n_i} B_j \in \mathcal{O} \). 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 \( f : \mathcal{X} \rightarrow \mathbb{R} \) be continuous. Then, \( f \) is bounded, that is, \( \exists a, b \in \mathbb{R} \) such that \( a \leq f(x) \leq b, \quad \forall x \in \mathcal{X} \). Furthermore, if \( \mathcal{X} \neq \emptyset \), \( \exists x_M, x_M \in \mathcal{X} \) such that \( f(x_M) = \min_{x \in \mathcal{X}} f(\bar{x}) \leq f(x) \leq \max_{x \in \mathcal{X}} f(\bar{x}) = f(x_M), \quad \forall x \in \mathcal{X} \), 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 3.6, 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 \( f : \mathcal{X} \rightarrow \mathbb{R} \) be upper semicontinuous. Then, \( f \) is bounded from above, that is, \( \exists b \in \mathbb{R} \) such that \( f(x) \leq b, \quad \forall x \in \mathcal{X} \). Furthermore, if \( \mathcal{X} \neq \emptyset \), then \( \exists x_M \in \mathcal{X} \) such that \( f(x) \leq \max_{x \in \mathcal{X}} f(\bar{x}) = f(x_M), \quad \forall x \in \mathcal{X} \), 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, $-\infty < M \leq N$. \forall n \in \mathbb{N}, \exists x_n \in X$ such that $M - 1/n < f(x_n) \leq M$. We thus obtain a sequence $(x_n)_{n=1}^{\infty} \subseteq X$. 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 $(f(x_n))_{n=1}^{\infty}$ 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 $f(x_M) = M$. 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 \in (0, \infty) \subset \mathbb{R}$, let $O_{\epsilon} := \{ x \in X \mid f_n(x) < \epsilon \}, \forall n \in \mathbb{N}$. By the upper semicontinuity of $f_n$, $O_{\epsilon} \subseteq \mathcal{O}$, $\forall n \in \mathbb{N}$. Since $\lim_{n \in \mathbb{N}} f_n(x) = 0$, $\forall x \in X$, then $X = \bigcup_{n \in \mathbb{N}} O_{\epsilon}$. By the countable compactness of $X$, $\exists N \in \mathbb{N}$ such that $X = \bigcup_{n=1}^{N} O_{\epsilon}$. $\forall n \in \mathbb{N}$ with $n \geq N$, $\forall x \in X$, $\exists i \in \{1, \ldots, N\}$ such that $x \in O_{\epsilon}$, 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 $f_{inv}((-\infty, a)) \subseteq f_{j_{inv}}(\{\infty, a\}), \forall i \leq j$, then $f_{j_{inv}}(\{\infty, a\}) = \bigcup_{n \in \mathbb{N}} f_{n_{inv}}(\{\infty, a\})$. By the upper semicontinuity of $f_n$, $f_{n_{inv}}(\{\infty, a\}) \in \mathcal{O}$, $\forall n \in \mathbb{N}$. Hence, $f_{j_{inv}}(\{\infty, a\}) \in \mathcal{O}$. 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

**Proof** \( \forall x_0 \in X \). \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}, \) we have \( |f_N(x) - f(x)| < \varepsilon/3 \), \( \forall x \in X \). Since \( f_N \) is upper semicontinuous, then, by Proposition 3.15, \( \exists U \in O \) with \( x_0 \in U \), \( \forall x \in U \), we have \( f_N(x) < f_N(x_0) + \varepsilon/3 \). Then, \( \forall x \in U \), we have

\[
f(x) - f(x_0) = f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0) < \varepsilon
\]

Therefore, \( f \) is upper semicontinuous at \( x_0 \). By Proposition 3.15, \( f \) is upper semicontinuous.

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

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, \( \exists x_0 \in (0, \infty) \subset \mathbb{R}, X \) can not be covered by finitely many open balls with radius \( \varepsilon_0 \). Clearly, \( X \neq \emptyset \). Fix an \( x_1 \in X \). \( \forall n \in \mathbb{N} \) with \( n \geq 2, \exists x_n \in X \setminus \left( \bigcup_{i=1}^{n-1} B(x_i, \varepsilon_0) \right) \). The sequence \( (x_n)_{n=1}^{\infty} \) is such that \( \rho(x_n, x_m) \geq \varepsilon_0, \forall n, m \in \mathbb{N} \) 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. \( \square \)

**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

\[
\varepsilon := \sup \{ \delta \in (0, \infty) \subset \mathbb{R} \mid \forall x \in X, \exists \bar{x} \in X \ni \rho(x, \bar{x}) \geq \delta
\text{ and } B(x, \delta) \subseteq O_{\alpha} \text{ for some } \alpha \in \Lambda \}
\]

**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 \( \varepsilon \) of \((O_{\alpha})_{\alpha \in \Lambda}\) is positive and belongs to \( \mathbb{R} \).

**Proof** Define a function \( \phi : X \to \mathbb{R} \) 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 \( \exists \bar{x} \in X \) 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 $\tilde{x} \neq x$ such that $\rho(\tilde{x}, x) =: c_2 > 0$. Then, $c := \min\{c_1, c_2\} > 0$ and $\phi(x) \geq c$. Hence, $0 < \phi(x) < +\infty$, $\forall x \in X$.

Next, we will show that $\phi$ is continuous. $\forall x, y \in X$, 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 $\phi(y) \geq 0 \geq \phi(x) - \rho(x, y)$.

Case 2: $\phi(x) - \rho(x, y) > 0$. Then, $\forall \delta \in (0, l)$, $\exists \alpha \in \Lambda$ and $\exists x \in X$ such that $B(x, \phi(x) - \delta) \subseteq O_\alpha$ and $\rho(x, \bar{x}) \geq \phi(x) - \delta$.

Then, $\rho(\bar{x}, y) \geq \rho(x, \bar{x}) - \rho(x, y) \geq \phi(x) - \delta - \rho(x, y) = l - \delta$. $\forall \bar{y} \in B(y, l - \delta)$, we have $\rho(x, \bar{y}) \leq \rho(x, y) + \rho(y, \bar{y}) < l - \delta + \rho(x, y) = \phi(x) - \delta$. Then, $\bar{y} \in B(x, \phi(x) - \delta) \subseteq O_\alpha$. This implies that $B(y, l - \delta) \subseteq O_\alpha$. 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 $X$, $X \neq \emptyset$, and Propositions 5.28 and 5.29, $\exists x_m \in X$ such that $\phi(x_m) = \min_{x \in X} \phi(x) \in (0, \infty) \subseteq \mathbb{R}$. Then, the Lebesgue number $\epsilon = \phi(x_m)$ is positive and belongs to $\mathbb{R}$. This completes the proof of the lemma. $\square$

**Theorem 5.37 (Borel-Lebesgue Theorem)** Let $X$ be a metric space. Then, the following are equivalent.

(i) $X$ is compact.

(ii) $X$ is countably compact.

(iii) $X$ is sequentially compact.

(iv) $X$ has the Bolzano-Weierstrass property.

**Proof** (i) $\Rightarrow$ (ii). This follows directly from Definitions 5.1 and 5.21.

(ii) $\Rightarrow$ (iii). Clearly, $X$ is first countable. By Proposition 5.28, $X$ is sequentially compact.

(iii) $\Rightarrow$ (i). We will distinguish three exhaustive and mutually exclusive cases: Case 1: $X = \emptyset$; Case 2: $X$ is a singleton set; Case 3: $X$ contains at least two distinct points. Case 1: $X = \emptyset$. Clearly, $X$ is compact. Case 2: $X$ is a singleton set. Clearly, $X$ is compact. Case 3: $X$ contains at least two distinct points. Fix any open covering $(O_\alpha)_{\alpha \in \Lambda} \subseteq O$ of $X$. Let $\epsilon$ be the Lebesgue number of the covering, which is a positive real number by Lemma 5.36. By Proposition 5.34, $X$ is totally bounded. Then, there exists a finite set $X_N \subseteq X$ such that $X = \bigcup_{x \in X_N} B(x, \epsilon/2)$. $\forall x \in X_N$, $B(x, \epsilon/2) \subseteq O_{\alpha_x}$ for some $\alpha_x \in \Lambda$. Hence, $X = \bigcup_{x \in X_N} O_{\alpha_x}$, which is a finite subcovering. Therefore, $X$ is compact.

(ii) $\Leftrightarrow$ (iv). This is proved in Proposition 5.26.

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

**Proposition 5.38** A metric space $X$ 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 \((x_n)_{n=1}^{\infty}\) that converges to \(x_0 \in X\). Then, it is straightforward to show that \(\lim_{n \in \mathbb{N}} x_n = x_0\). (Here, the notation \(\lim_{n \in \mathbb{N}} x_n\) makes sense since metric spaces are Hausdorff.) Hence, \(X\) is complete. By Borel-Lebesgue Theorem and Proposition 5.34, \(X\) is totally bounded.

“\(f\)” Fix any sequence \((x_n)_{n=1}^{\infty} \subseteq X\). We will construct subsequences as following. Take \((x_n^{(1)})_{n=1}^{\infty} = (x_n)_{n=1}^{\infty}\). \(\forall m \in \mathbb{N}\) with \(m \geq 2\), since \(X\) is covered by finite many balls of radius \(1/m\), we may find a subsequence \((x_n^{(m)})_{n=1}^{\infty}\) of \((x_n^{(m-1)})_{n=1}^{\infty}\) such that \((x_n^{(m)})_{n=1}^{\infty} \subseteq B(x^{(m)}, 1/m)\) for some \(x^{(m)} \in X\). Now, consider the sequence of diagonal elements \((x_n^{(n)})_{n=1}^{\infty} \subseteq X\), which is a subsequence of the original sequence \((x_n)_{n=1}^{\infty}\). Clearly, this sequence is Cauchy by construction. By the completeness of \(X\), \(\lim_{n \in \mathbb{N}} x_n = x_0 \in X\). 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 \(\forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \forall x \in X\), by the continuity of \(f\), \(\exists \delta_x \in (0, \infty) \subseteq \mathbb{R}\) such that \(\sigma(f(x), f(\bar{x})) < \varepsilon/2, \forall \bar{x} \in B(x, \delta_x)\). 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 \(\delta = \min\{\inf_{x \in X_N} \delta_x/2, 1\} \in (0, \infty) \subseteq \mathbb{R}\). \(\forall x_1, x_2 \in X\) with \(\rho(x_1, x_2) < \delta\), \(\exists x_0 \in X_N\) such that \(x_1 \in B(x_0, \delta x_0/2)\). Then, \(x_2 \in B(x_0, \delta x_0)\). This implies that \(x_1, x_2 \in B(x_0, \delta x_0)\). 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 \(\left( f_n^{(0)} \right)_{n=1}^{\infty} = (f_n)_{n=1}^{\infty}\). \(\forall k = 1, \ldots, m, \left( f_n^{(k-1)}(x_k) \right)_{n \in \mathbb{N}} \subseteq (f_n(x_k))_{n \in \mathbb{N}}\) is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence \(\left( f_n^{(k)} \right)_{n=1}^{\infty}\) of \(\left( f_n^{(k-1)} \right)_{n=1}^{\infty}\) such that \(\left( f_n^{(k)}(x_k) \right)_{n=1}^{\infty}\) converges. Hence, the sequence \(\left( f_n^{(m)} \right)_{n=1}^{\infty}\) is the subsequence we seek. Case 3: \(D\) is countably infinite. Take \(D = (x_m)_{m=1}^{\infty}\). Let \(\left( f_n^{(0)} \right)_{n=1}^{\infty} = (f_n)_{n=1}^{\infty}\). \(\forall k \in \mathbb{N}, \left( f_n^{(k-1)}(x_k) \right)_{n \in \mathbb{N}} \subseteq (f_n(x_k))_{n \in \mathbb{N}}\) is a closed set and therefore a compact set by Proposition 5.5. By Borel-Lebesgue Theorem, there exists a subsequence \(\left( f_n^{(k)} \right)_{n=1}^{\infty}\) of \(\left( f_n^{(k-1)} \right)_{n=1}^{\infty}\) such that \(\left( f_n^{(k)}(x_k) \right)_{n=1}^{\infty}\) converges. Now, the diagonal sequence \(\left( f_k^{(k)} \right)_{k=1}^{\infty}\) 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** \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x_0 \in \mathcal{X},\) 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 \(\rho(f_n(x_0), f_0(x_0)) < \epsilon\).

\(\forall x_0 \in \mathcal{X}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall x \in U_{x_0, \epsilon},\) 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.

\(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \mathcal{X} \subseteq \bigcup_{x \in X} U_{x, \epsilon}\). By the compactness of \(\mathcal{X}\), there exists a finite set \(X_N \subseteq X\) 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.

\(\square\)
5.6. PRODUCT SPACES

Lemma 5.43 Let $X := (X, \mathcal{O})$ be a topological space, $Y := (Y, \rho)$ be a metric space, $(f_n)_{n=1}^{\infty}$ be an equicontinuous sequence of functions of $X$ to $Y$, and $(f_n(x))_{n=1}^{\infty}$ converge at every $x \in D$, where $D \subseteq X$ is dense. Assume that $\forall x \in X$, $(f_n(x))_{n=1}^{\infty} \subseteq Y$ is complete. Then, $\exists$ continuous function $f_0 : X \rightarrow Y$ such that $\lim_{n \in \mathbb{N}} f_n(x) = f_0(x)$, $\forall x \in X$.

Proof $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\forall x \in X$, since $(f_n)_{n=1}^{\infty}$ is equicontinuous, then $\exists O \in \mathcal{O}$ with $x \in O$ such that, $\forall \bar{x} \in O$, $\rho(f_n(x), f_n(\bar{x})) < \epsilon$, $\forall n \in \mathbb{N}$.

Since $D$ is dense, then, $\exists x_0 \in D \cap O$. $(f_n(x_0))_{n=1}^{\infty} \subseteq Y$ 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, $(f_n(x))_{n=1}^{\infty} \subseteq (f_n(x))_{n=1}^{\infty} \subseteq Y$ is a Cauchy sequence, which converges since $(f_n(x))_{n=1}^{\infty}$ is complete. Then, we may define $f_0 : X \rightarrow Y$ by $f_0(x) = \lim_{n \in \mathbb{N}} f_n(x)$.

By Proposition 4.30, 3.66, and 3.67, we have, $\forall \bar{x} \in O$,

$$\rho(f_0(x), f_0(\bar{x})) = \lim_{n \in \mathbb{N}} \rho(f_n(x), f_n(\bar{x})) \leq \epsilon$$

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. □

Theorem 5.44 (Ascoli-Arzela Theorem) Let $X := (X, \mathcal{O})$ be a separable topological space, $Y := (Y, \rho)$ be a metric space, $F$ be an equicontinuous family of functions of $X$ to $Y$, and $(f_n)_{n=1}^{\infty} \subseteq F$ be such that $(f_n(x))_{n \in \mathbb{N}} \subseteq Y$ is compact, $\forall x \in X$. Then, there exists a subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ that converges pointwise to a continuous function $f_0 : X \rightarrow Y$ and the convergence is uniform on any compact subset of $X$.

Proof Let $D \subseteq X$ 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 $\bar{f} : D \rightarrow Y$. Note that $\forall x \in X$, $(f_n(x))_{n \in \mathbb{N}} \subseteq Y$ 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 $f_0 : X \rightarrow Y$ 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 $X$. This completes the proof of the theorem. □

5.6 Product Spaces

Lemma 5.45 Let $A$ 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 \( A \subseteq M \), \( M \) has the finite intersection property, and \( M \) is maximal with respect to this property, that is \( \forall C \subseteq X^2 \) with the finite intersection property and \( M \subseteq C \), we have \( C = M \).

**Proof** Clearly, \( X \neq \emptyset \) by the fact that \( A \) has the finite intersection property. Define \( M := \{ C \subseteq X^2 \mid A \subseteq C \text{ and } C \text{ has the finite intersection property} \} \) Set containment \( \subseteq \) forms an antisymmetric partial ordering on \( M \) and \( A \in M \neq \emptyset \). Let \( N \subseteq M \) be any nonempty totally ordered subset. Let \( B := \bigcup_{C \in N} C \subseteq X^2 \). Fix any finite subcollection \( B_N \subseteq B \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( B_N = \emptyset \); Case 2: \( B_N \neq \emptyset \). Case 1: \( B_N = \emptyset \). Then, \( \overline{B} := \bigcap_{B \in B_N} B = X \neq \emptyset \). 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 \( \bigcap_{i=1}^n B_i \neq \emptyset \). 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 \( A \subseteq B \). Hence, \( B \in M \) and is an upper bound of \( N \). By Zorn’s Lemma, there exists a maximal element \( M \) of \( M \). This completes the proof of the lemma.

**Lemma 5.46** Let \( B \) 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 \( B \) is again in \( B \), and each set that meets every set in \( B \) is itself in \( B \).

**Proof** Fix any finite subcollection \( B_N \subseteq B \). Let \( \overline{B} := \bigcap_{B \in B_N} B \). Then, \( \overline{B} \neq \emptyset \). Let \( \overline{C} := B \cup \{ \overline{B} \} \supseteq B \). It is easy to see that \( \overline{C} \) has the finite intersection property. Then, by maximality of \( B \), we have \( \overline{B} = B \). Hence, \( \overline{B} \in B \).

Let \( C \subseteq X \) be such that \( C \cap B \neq \emptyset, \forall B \in B \). Let \( \overline{B} := B \cup \{ C \} \supseteq B \). Fix any finite subcollection \( M \subseteq B \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( C \notin M \); Case 2: \( C \in M \). Case 1: \( C \notin M \). Then, \( M \subseteq B \). Then, \( \bigcap_{B \in M} B \neq \emptyset \) since \( B \) has the finite intersection property. Case 2: \( C \in M \). Then, \( M := M \setminus \{ C \} \subseteq B \) and is a finite set. Then, \( \bigcap_{B \in M} B \in B \). Then, \( \bigcap_{B \in M} B = C \cap ( \bigcap_{B \in M} B ) \neq \emptyset \). Hence, in both cases, we have \( \bigcap_{B \in M} B \neq \emptyset \). This shows that \( \overline{B} \) has the finite intersection property. By the maximality of \( B \), we have \( \overline{B} = B \). Hence, \( C \in B \).

This completes the proof of the lemma.

**Theorem 5.47 (Tychonoff Theorem)** Let \( X_\alpha := (X_\alpha, O_\alpha) \) be a compact topological space, \( \forall \alpha \in \Lambda \), where \( \Lambda \) is an index set. Then, the product topological space \( X := (X, O) := \prod_{\alpha \in \Lambda} X_\alpha \) 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

Case 2: $\Lambda \neq \emptyset$. We will prove this case by Proposition 5.3. Note that a basis for $X$ is given by $B := \{ \prod_{\alpha \in \Lambda} O_\alpha \subseteq X \mid O_\alpha \in \mathcal{O}_\alpha, \forall \alpha \in \Lambda, \text{ and } O_\alpha = X_\alpha \text{ for all } \alpha \text{'s except finitely many } \alpha \text{'s} \}$. Fix a collection $\mathcal{A}$ of closed sets in $X$ with the finite intersection property. By Lemma 5.45, let $\mathcal{C}$ be a collection of subsets of $X$ such that $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{C}$ is maximal with respect to the finite intersection property.

$\forall \alpha \in \Lambda$, let $C_\alpha := \{ C_\alpha \subseteq X_\alpha \mid \exists \mathcal{C} \ni \pi_\alpha(C) = C_\alpha \}$. $\forall$ finite subcollection $C_{\alpha,N} \subseteq C_\alpha$, $\forall C_{\alpha,N} \in C_{\alpha,N}$, $\exists \mathcal{C}_{\alpha} \in \mathcal{C}$ such that $C_{\alpha} = \pi_\alpha(C_{\alpha})$. Since $\mathcal{C}$ has the finite intersection property, then $\bigcap_{\alpha \in C_{\alpha,N}} C_{\alpha} \neq \emptyset$. Then, by Proposition 2.5, $\bigcap_{\alpha \in C_{\alpha,N}} C_{\alpha} = \bigcap_{\alpha \in C_{\alpha,N}} \pi_\alpha(C_{\alpha}) \supseteq \pi_\alpha \left( \bigcap_{\alpha \in C_{\alpha,N}} C_{\alpha} \right) \neq \emptyset$. Hence, $\mathcal{C}_{\alpha}$ has finite intersection property. Let $\mathcal{C}_\alpha := \{ C_{\alpha} \mid C_{\alpha} \in \mathcal{C}_{\alpha} \}$. Then, $\mathcal{C}_\alpha$ has finite intersection property. By the compactness of $X_\alpha$ and Proposition 5.3, $\exists x_\alpha \in \bigcap_{\alpha \in C_{\alpha,N}} \mathcal{C}_\alpha$.

Let $x \in X$ be given by $\pi_\alpha(x) = x_\alpha$, $\forall \alpha \in \Lambda$. Consider a set $S$ of the form $S = \pi_\alpha(O_\alpha)$ for some $\alpha \in \Lambda$ and for some $O_\alpha \in \mathcal{O}_\alpha$ with $x_\alpha \in O_\alpha$. 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 $\exists O_{\alpha_i} \in \mathcal{O}_{\alpha_i}, i = 1, \ldots, n$, such that $x_{\alpha_i} \in O_{\alpha_i}$ and $B = \bigcap_{i=1}^n \pi_{\alpha_i}(O_{\alpha_i})$. 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, $X$ is compact.

This completes the proof of the theorem. \Box

**Proposition 5.48** Let $X_\alpha := (X_\alpha, \mathcal{O}_\alpha)$ be a sequentially compact topological space, $\forall \alpha \in \Lambda$, where $\Lambda$ is a countable index set. Then, the product topological space $X := (X, \mathcal{O}) := \prod_{\alpha \in \Lambda} X_\alpha$ 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, $X$ 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 $\left( x_n^{(0)} \right)_{n=1}^\infty \subseteq X$. $\forall k \in \{1, \ldots, m\}$, $\left( \pi_k(x_n^{(k-1)}) \right)_{n=1}^\infty \subseteq X_k$. By the sequential compactness of $X_k$, there exists a subsequence $\left( x_n^{(k)} \right)_{n=1}^\infty$ of $\left( x_n^{(k-1)} \right)_{n=1}^\infty$ such that $\left( \pi_k(x_n^{(k)}) \right)_{n=1}^\infty$ converges to $x_k \in X_k$. Let $x \in X$ be given by $\pi_k(x) = x_k$, $\forall k \in \{1, \ldots, m\}$. The sequence $\left( x_n^{(m)} \right)_{n=1}^\infty \subseteq X$ is a subsequence of $\left( x_n^{(0)} \right)_{n=1}^\infty$. 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, $\left( x_n^{(m)} \right)_{n=1}^\infty$ converges to $x$. Hence, $X$ is sequentially compact.
Case 3: \( \Lambda \) is countably infinite. Without loss of generality, assume that \( \Lambda = \mathbb{N} \). Fix a sequence \( \left( x_n^{(0)} \right)_{n=1}^\infty \subseteq \mathcal{X} \). \( \forall k \in \mathbb{N}, \left( \pi_k(x_n^{(k-1)}) \right)_{n=1}^\infty \subseteq \mathcal{X}_k \).

By the sequential compactness of \( \mathcal{X}_k \), there exists a subsequence \( \left( x_n^{(k)} \right)_{n=1}^\infty \) of \( \left( x_n^{(k-1)} \right)_{n=1}^\infty \) 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 \( \pi_k(x) = x_k, \forall k \in \mathbb{N} \). Now consider the diagonal sequence \( \left( x_n^{(n)} \right)_{n=1}^\infty \subseteq \mathcal{X} \), which is a subsequence of \( \left( x_n^{(0)} \right)_{n=1}^\infty \). Clearly, \( \forall k \in \mathbb{N}, \left( \pi_k(x_n^{(n)}) \right)_{n=1}^\infty \) converges to \( x_k \). By Proposition 3.67, \( \left( x_n^{(n)} \right)_{n=1}^\infty \) converges to \( x \). Hence, \( \mathcal{X} \) is sequentially compact.

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

### 5.7 Locally Compact Spaces

#### 5.7.1 Fundamental notion

**Definition 5.49** A topological space \( \mathcal{X} \) is locally compact if \( \forall x \in \mathcal{X}, \exists O \in \mathcal{O} \) with \( x \in O \) 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 \( \mathcal{B}_1 := \{ O \in \mathcal{O} \mid \overline{O} \text{ is compact} \} \) 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 \( x \in B \). \( \forall O \in \mathcal{O}, \forall x \in O, \exists B \in \mathcal{B}_1 \) such that \( x \in B \) and \( \overline{B} \) is compact. Let \( B_1 := O \cap B \). Then, \( x \in B_1 \in \mathcal{O} \) and \( B_1 \subseteq O \). Note that \( \overline{B_1} \subseteq \overline{B} \) and is closed. By Proposition 5.5, \( \overline{B_1} \) is compact. Then, \( B_1 \in \mathcal{B}_1 \). Therefore, \( \mathcal{B}_1 \) is a basis for \( \mathcal{O} \).

“If” This is straightforward.

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

**Proposition 5.52** Let \( \mathcal{X} \) be a locally compact topological space and \( K \subseteq \mathcal{X} \) be compact. Then, \( \exists O \in \mathcal{O}, \forall x \in O, \exists K \subseteq \overline{O} \subseteq \mathcal{O} \) such that \( K \subseteq \overline{O} \) and \( \overline{O} \) is compact.

**Proof** \( \forall x \in K, \) by the local compactness of \( \mathcal{X} \), \( \exists O_x \in \mathcal{O} \) such that \( x \in O_x \subseteq \overline{O_x} \) and \( \overline{O_x} \) is compact. Then, \( K \subseteq \bigcup_{x \in K} O_x \). 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. \( \square \)

**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 \( x \in B \subseteq \overline{B} \subseteq \mathcal{O} \) 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: \( O = X \); Case 2: \( O \subset X \). Case 1: \( O = X \). The result holds by Definition 5.49. Case 2: \( O \subset X \). \( \forall y \in \tilde{O} \), since \( X \) 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, \( \exists B_y \in \mathcal{B}_1 \) 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. This implies that \( y \notin \overline{B_y} \). Note that \( \tilde{O} \) is closed in \( X \). Clearly, \( \tilde{O} \cap \overline{B_y} \) is compact. Note that \( \bigcap_{y \in \tilde{O}} (\tilde{O} \cap \overline{B_y}) = \emptyset \). By Proposition 5.12, there exists a finite set \( D_N \subseteq \tilde{O} \) such that \( \bigcap_{y \in D_N} (\tilde{O} \cap \overline{B_y}) = \emptyset \). Clearly, \( D_N \neq \emptyset \) since \( O \subset X \neq \emptyset \). Then, we have \( x \in \bigcap_{y \in D_N} B_y \subseteq \bigcap_{y \in D_N} \overline{B_y} \subseteq O \). Since \( \bigcap_{y \in D_N} B_y \in \mathcal{O} \), then \( \exists B \in \mathcal{B}_1 \) such that \( x \in B \subseteq \bigcap_{y \in D_N} B_y \) and \( \overline{B} \) is compact. Note that \( \overline{B} \subseteq \bigcap_{y \in D_N} \overline{B_y} \subseteq O \). This completes the proof of the proposition. \( \square \)

Proposition 5.54 Let \( X \) be a locally compact Hausdorff topological space, \( U \in \mathcal{O} \), and \( K \subseteq U \) be compact. 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, \( \overline{V} = \bigcup_{x \in K_N} \overline{V_x} \subseteq U \) and is compact. This completes the proof of the proposition. \( \square \)

Proposition 5.55 Let \( X \) be a locally compact space and \( F \subseteq X \). Then, \( F \) is closed if, and only if, for any closed and compact set \( K \subseteq X \), we have \( F \cap K \) is closed.

Proof “Only if” Let \( F \) be closed. For any closed and compact \( K \subseteq X \), \( F \cap K \) is closed.

“\( \forall x \in F \)”. By local compactness of \( X \), \( \exists O \in \mathcal{O} \) 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 \( x \in O \cap U \in \mathcal{O} \). 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, \( \overline{O} \subseteq F \) and \( F \) is closed.

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

Proposition 5.56 Let \( X \) be a locally compact Hausdorff topological space and \( (O_n)_{n=1}^{\infty} \subseteq \mathcal{O} \) be a sequence of open dense subsets in \( X \). Then, \( \bigcap_{n=1}^{\infty} O_n \) is dense. Therefore, \( X \) 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, \( (\overline{V_n})^{\infty}_{n=1} \) is a sequence of nonempty closed sets which is nonincreasing (that is \( V_{n+1} \subseteq \overline{V_n} \), \( \forall n \in \mathbb{N} \)). Clearly, this sequence have the finite intersection property. By the compactness of \( \overline{V_n} \) and Proposition 5.3, \( \bigcap^{\infty}_{n=1} \overline{V_n} \neq \emptyset \). Clearly, \( \bigcap^{\infty}_{n=1} \overline{V_n} \subseteq U \cap (\bigcap^{\infty}_{n=1} O_n) \neq \emptyset \). Therefore, \( \bigcap^{\infty}_{n=1} O_n \) is dense in \( \mathcal{X} \), by the arbitrariness of \( U \).

By Proposition 3.38, \( \mathcal{X} \) is second category everywhere. This completes the proof of the proposition.

\[ \square \]

**Proposition 5.57** Let \( \mathcal{X} := (X, \mathcal{O}) \) 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 \( \mathcal{X} \).

**Proof** \( \forall y \in Y \), by the local compactness of \( \mathcal{Y} \), \( \exists U_y \in \mathcal{O}_Y \) such that \( y \in U_y \subseteq U_{yc} \subseteq Y \), where \( U_{yc} \) is the closure of \( U_y \) in \( \mathcal{Y} \), and \( U_{yc} \) is compact. Then, \( U_{yc} \) is compact in \( \mathcal{X} \) 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 \( U_y \subseteq O \cap Y \). By Proposition 3.37, we have \( U_{yc} = \overline{O} \cap Y = \overline{O} \). Hence, we have \( y \in U_y \subseteq O \subseteq \overline{O} = U_{yc} \subseteq Y \). Hence, \( Y \) is open in \( \mathcal{X} \). This completes the proof of the proposition.

\[ \square \]

**Lemma 5.58** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \subseteq \mathcal{O} \), and \( \mathcal{O}_Y \) be the subset topology on \( Y \). Then, \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) 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 \( \mathcal{Y} \) is \( \overline{U} \), by Proposition 3.5. Then, \( \overline{U} \) is compact in \( \mathcal{Y} \). Hence, \( \mathcal{Y} \) is locally compact.

This completes the proof of the lemma.

\[ \square \]

**Lemma 5.59** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \subseteq \mathcal{X} \) be closed, and \( \mathcal{O}_Y \) be the subset topology on \( Y \). Then, \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) is locally compact.

**Proof** \( \forall x \in Y \), by the local compactness of \( \mathcal{X} \), \( \exists U \in \mathcal{O} \) such that \( x \in U \subseteq \overline{U} \) and \( \overline{U} \) is compact. Then, \( U := U \cap Y \in \mathcal{O}_Y \). Let \( U_c \) be the closure of \( U \) in \( \mathcal{Y} \). Then, \( U_c \subseteq U \cap Y \subseteq \overline{U} \). Note that \( U_c \) is a closed set in \( \mathcal{Y} \) and therefore a closed set in \( \mathcal{X} \) since \( Y \) is closed in \( \mathcal{X} \) and Proposition 3.5. Then, \( U_c \) is compact by Proposition 5.5. Thus, we have \( x \in U \subseteq U_c \subseteq Y \) and \( U_c \) is compact. Hence, \( \mathcal{Y} \) is locally compact. This completes the proof of the lemma.

\[ \square \]

**Proposition 5.60** Let \( \mathcal{X} := (X, \mathcal{O}) \) be a locally compact Hausdorff topological space, \( Y \subseteq X \), and \( \mathcal{O}_Y \) be the subset topology on \( Y \). \( \mathcal{Y} := (Y, \mathcal{O}_Y) \) is locally compact if, and only if, \( Y \) is relatively open in \( \overline{Y} \).
Proof. "Only if" Let $\mathcal{O}_Y$ be the subset topology on $\overline{Y}$. Then, $(\overline{Y}, \mathcal{O}_Y)$ is Hausdorff. By Proposition 5.54, $Y$ is relative open in $\mathcal{O}_Y$.

"If" Since $\overline{Y}$ is closed in $\mathcal{X}$, 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. \qed

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 $f : \mathcal{X} \to [0,1] \subset \mathbb{R}$ such that $f|_K = 1$ and $f|_{\overline{Y}} = 0$. 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 $\bar{f} : X \to \mathbb{R}$ by

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

Clearly, $\bar{f} : X \to [0,1]$, $\bar{f}(x) = 0$, $\forall x \in O_0$, and $\bar{f}(x) = 1$, $\forall x \in \overline{O_1}$. By 1, we have $K \subseteq O_0$ and $O_1 \subseteq U$. Next, we will show that $\bar{f}$ is continuous.

Let $x_0 \in X$, we will show that $\bar{f}$ is continuous at $x_0$. Let $a_0 = \bar{f}(x_0) \in [0,1]$. $\forall B \subseteq \mathbb{R}$ 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 $a_2 = \max\{a_2, 0\}$ and $a_3 = \min\{a_3, 1\}$. Then, we must have $a_1 < a_2 < a_0 < a_3 < a_4$ and $a_2, a_3 \in Q$. 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 $a_1 < a_2 < a_0 < a_3 < a_4$. Let $V = \overline{O_{a_2}} \cap O_{a_3} \in \mathcal{O}$. $\forall x \in V$, we have $x \in O_{a_3}$ and $\bar{f}(x) \leq a_3$. Also, $x \in \overline{O_{a_2}}$ implies that $\bar{f}(x) \geq a_2$. Hence, $\bar{f}(V) \subseteq [a_2, a_3] \subseteq (a_1, a_4) \subseteq B$. $\bar{f}(x_0) = a_0 < \bar{a_3}$ implies that $x_0 \in O_{a_3}$. $\bar{f}(x_0) = a_0 > \bar{a_2} = a_2$ implies that $\exists x_0 \in \overline{O_{a_2}} \cap Q$ such that $x_0 \in \overline{O_{a_2}} \subseteq \overline{O_{a_2}}$. Therefore, $x_0 \in V$. This shows that $\exists V \in \mathcal{O}$ with $x_0 \in V$ such that $\bar{f}(V) \subseteq B$.

Case 2: $a_0 = 0$. Then, we must have $a_1 < 0 = a_0 < a_3 < a_4$. Take $V = O_{a_3} \in \mathcal{O}$. We must have $x_0 \in V$. $\forall x \in V$, $0 \leq \bar{f}(x) \leq a_3$. Hence,
CHAPTER 5. COMPACT AND LOCALLY COMPACT SPACES

\[ f(V) \subseteq [0, a_3] \subset (a_1, a_4) \subseteq B. \] Hence, \( \exists V \in \mathcal{O} \) with \( x_0 \in V \) such that \( f(V) \subseteq B. \)

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

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 arbitraryness of \( x_0 \) and Proposition 3.9, \( \bar{f} \) is continuous. Define \( f : \mathcal{X} \rightarrow [0, 1] \subset \mathbb{R} \) by \( f(x) = 1 - \bar{f}(x), \forall x \in \mathcal{X}. \) Clearly, \( f \) is continuous by Proposition 3.12, \( f|_K = 1, \) and \( f|_{\overline{U}} = 0. \)

Note that \( \text{supp}\ f = \{ x \in \mathcal{X} \mid f(x) < 1 \} \subseteq \overline{O_1} \subseteq U \) and is compact by Proposition 5.5. This completes the proof of the proposition. \( \square \)

**Theorem 5.63 (Partition of Unity)** Let \( \mathcal{X} \) be a locally compact Hausdorff topological space, \( K \subseteq \mathcal{X} \) 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 \( \mathcal{X} \) 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 \( \text{supp}\ f \subseteq M := \{ x \in \mathcal{X} \mid g(x) \neq 0 \}, \) define \( f/g : \mathcal{X} \rightarrow \mathbb{R} \) 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 \( f(x_0) = 1, \ f|_{\overline{O}} = 0, \) and \( f : \mathcal{X} \rightarrow [0, 1]. \)

Then, there exists a finite collection \( \Phi \subseteq \mathcal{F} \) of continuous nonnegative real-valued functions on \( \mathcal{X}, \) which is subordinate to \( (O_\lambda)_{\lambda \in \Lambda}, \) such that \( \left( \sum_{\phi \in \Phi} \phi \right)|_K = 1, \sum_{\phi \in \Phi} \phi(x) \in [0, 1] \subset \mathbb{R}, \forall x \in \mathcal{X}, \) and \( \text{supp}\ \phi \) 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, \( \text{supp}\ f \cup M = \mathcal{X}, \) and Theorem 3.11 implies the result. \( \diamond \)

**Proof** Let \( O := \bigcup_{\lambda \in \Lambda} O_\lambda \supseteq K. \) By Proposition 5.54, there exists \( U \in \mathcal{O} \) such that \( K \subseteq U \subseteq \overline{U} \subseteq O \) and \( \overline{U} \) is compact. \( \forall \lambda \in \Lambda, \) let \( U_\lambda := U \cap O_\lambda \in \mathcal{O}. \) Then, \( K \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda = U. \)
∀x₀ ∈ K, ∃λ₀ ∈ Λ such that x₀ ∈ U₀λ₀. By Proposition 5.53, ∃Ūλ₀ ∈ O such that x₀ ∈ Ūλ₀ ⊆ Ū₀λ₀ ⊆ U₀λ₀ and Ū₀λ₀ is compact. By (iv), there exists a continuous function fₓ₀ : X → [0, 1] in F such that fₓ₀(x₀) = 1 and fₓ₀|⁻₀λ₀ = 0. Let Vₓ₀ := {x ∈ X | fₓ₀(x) > 0}. Then, x₀ ∈ Vₓ₀ ∈ O since fₓ₀ is continuous. Note that Vₓ₀ ⊆ Ū₀λ₀, which implies that supp fₓ₀ = Vₓ₀ ⊆ Ū₀λ₀ ⊆ U₀λ₀ ⊆ O₀λ₀, which is compact by Proposition 5.5. Hence, fₓ₀ has compact support.

∀x₀ ∈ O \ K, K ∈ O, by Proposition 5.5. By (iv), there exists a continuous function gₓ₀ : X → [0, 1] in F such that gₓ₀(x₀) = 1 and gₓ₀|₀K = 0. Let Wₓ₀ := {x ∈ X | gₓ₀(x) > 0}. Then, x₀ ∈ Wₓ₀ ∈ O since gₓ₀ is continuous. Clearly, Wₓ₀ ⊆ K. Let U := (⋃x∈K Vₓ) ∪ (⋃x∈K\K Wₓ). By the compactness of U, there exist finite sets K ⊆ K and U′ ⊆ U \ K such that U ⊆ (⋃x∈K Vₓ) ∪ (⋃x∈U′ Wₓ). By the construction of Wₓ’s, K ⊆ U′ ∈ O. Let f := ∑x∈K fₓ ∈ F and g := ∑x∈U′ 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 ∈ K}, which is a finite set. Note that ∀x ∈ K, supp fₓ ⊆ U₀λ ⊆ U ⊆ U ⊆ X such that (f + g)(x) ≠ 0, for some λ ∈ Λ. Then, fₓ/(f + g) ∈ F. Hence, Φ ⊆ F. ∀φ ∈ Φ, ∃x ∈ K 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 ⊆ O₀λ, for some λ ∈ Λ. Hence, Φ is subordinate to (O₀λ)₀∈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

\[
0 ≤ \sum_{φ ∈ Φ} φ(x) = \frac{f(x)}{(f + g)}(x) ≤ 1
\]

This completes the proof of the theorem.

**Corollary 5.65** Let X be a locally compact Hausdorff topological space, K ⊆ X be compact, and (O₀λ)₀∈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 (O₀λ)₀∈Λ, such that \( \sum_{φ ∈ Φ} φ \) |₀K = 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 $X = (X, \mathcal{O})$ 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 $\emptyset \in \mathcal{O}_\omega$; $X_\omega \setminus X_\omega = \emptyset$ is compact in $X_\omega$, then $X_\omega \in \mathcal{O}_\omega$. (ii) $\forall O_1, O_2 \in \mathcal{O}_\omega$, we will distinguish four exhaustive and mutually exclusive cases: Case 1: $O_1 \cap O_2 \in \mathcal{O}$; Case 2: $X \setminus O_1$ and $X \setminus O_2$ are compact in $X$; Case 3: $X_\omega \setminus O_1$ is compact in $X$ and $O_2 \in \mathcal{O}$; Case 4: $O_1 \in \mathcal{O}$ and $X_\omega \setminus O_2$ 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 $O_1 \cap O_2 \in \mathcal{O}_\omega$. Case 2: $X \setminus O_1$ and $X \setminus O_2$ are compact in $X$. Then, $X_\omega \setminus (O_1 \cap O_2) = (X \setminus O_1) \cup (X \setminus O_2)$, which is compact in $X$. This implies that $O_1 \cap O_2 \in \mathcal{O}_\omega$. Case 3: $X_\omega \setminus O_1$ is compact in $X$ and $O_2 \in \mathcal{O}$. Let $O_1 := O_1 \setminus \{\omega\}$, then $O_1 \cap O_2 = X \setminus (X \setminus O_1)$. Since $X_\omega \setminus O_1$ is compact in $X$ and $X$ is Hausdorff, then, by Proposition 5.5, $X_\omega \setminus O_1$ is closed in $X$. Then, $O_1 \cap O_2 \in \mathcal{O}_\omega$. Note that $O_1 \cap O_2 = O_1 \cap O_2 \in \mathcal{O}_\omega$. Then, $O_1 \cap O_2 \in \mathcal{O}_\omega$. Case 4: $O_1 \in \mathcal{O}$ and $X_\omega \setminus O_2$ 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 $\cup_{\lambda \in \Lambda} O_\lambda \in \mathcal{O}_\omega$. Case B: $\exists \lambda_0 \in \Lambda$ such that $\omega \in O_{\lambda_0}$. Then, $X_\omega \setminus O_{\lambda_0}$ is compact in $X$. We may partition $\Lambda$ into two disjoint set $A_1$ and $A_2$ such that $\Lambda = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$, $\forall \lambda \in A_1$, $O_{\lambda} \in \mathcal{O}$, $\forall \lambda \in A_2$, $X_\omega \setminus O_{\lambda}$ is compact in $X$. Note that

$$X_\omega \setminus \left( \bigcup_{\lambda \in \Lambda} O_\lambda \right) = \left( \bigcap_{\lambda \in A_1} (X \setminus O_\lambda) \right) \cap \left( \bigcap_{\lambda \in A_2} (X \setminus O_\lambda) \right)$$

$$= \left( \bigcap_{\lambda \in A_1} (X \setminus O_\lambda) \right) \cap \left( \bigcap_{\lambda \in A_2} (X \setminus O_\lambda) \right)$$

$\forall \lambda \in A_1$, $X \setminus O_\lambda$ is a closed set in $X$. $\forall \lambda \in A_2$, $X_\omega \setminus O_\lambda$ is compact in $X$ and therefore closed in $X$ by Proposition 5.5. Hence, $X_\omega \setminus \left( \bigcup_{\lambda \in \Lambda} O_\lambda \right)$ is a closed subset of $X_\omega \setminus O_{\lambda_0}$ and hence compact in $X$ by Proposition 5.5. Then, $\bigcup_{\lambda \in \Lambda} O_\lambda \in \mathcal{O}_\omega$. Hence, in both cases, we have $\bigcup_{\lambda \in \Lambda} O_\lambda \in \mathcal{O}_\omega$. 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 $A_1$ and $A_2$ such that $\Lambda = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$, $\forall \lambda \in A_1$, $O_{\lambda} \in \mathcal{O}$, $\forall \lambda \in A_2$, $X_\omega \setminus O_{\lambda}$ is compact.
5.7. **LOCALLY COMPACT SPACES**

In \( X \). Since \( \omega \in X \), 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 \( O_{c\lambda} := O_{c\lambda} \setminus \{ \omega \} \), then \( O_{c\lambda} = X \setminus (X_c \setminus O_{c\lambda}) \).

By Proposition 5.5 and the compactness of \( X_c \setminus O_{c\lambda}, O_{c\lambda} \in \mathcal{O} \). 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, \( X_c \setminus O_{c\lambda_0} \subseteq \left( \bigcup_{\lambda \in \Lambda_1} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_2} O_{c\lambda} \right) \). 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 \( X_c \setminus O_{c\lambda_0} \subseteq \left( \bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_{N2}} O_{c\lambda} \right) \). Then, \( X_c = \left( \bigcup_{\lambda \in \Lambda_{N1}} O_{c\lambda} \right) \cup \left( \bigcup_{\lambda \in \Lambda_{N2}} O_{c\lambda} \right) \). 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: \( x_1, x_2 \in X \); Case 2: \( x_1 = \omega \) or \( x_2 = \omega \). Case 1: \( x_1, x_2 \in X \). 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 \( X \), \( \exists O_1 \in \mathcal{O} \) such that \( x_1 \in O_1 \subseteq \overline{O_1} \subseteq X \) and \( \overline{O_1} \) is compact. Then, \( O_1 \in \mathcal{O}_c \) and \( O_2 := X_c \setminus O_1 \in \mathcal{O}_c \). 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 \( \text{id} : X \to X_c \setminus \{ \omega \} \) 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 \( O_c \cap (X_c \setminus \{ \omega \}) = O_c \cap \mathcal{O} \in \mathcal{O} \); or we have \( X_c \setminus O_c \) is compact, which implies that \( \overline{O_c} := O_c \setminus (X_c \setminus O_c) \in \mathcal{O} \), by Proposition 5.5, and hence \( O_c \cap (X_c \setminus \{ \omega \}) = \overline{O_c} \in \mathcal{O} \). 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, \( \text{id} : X \to X_c \setminus \{ \omega \} \) 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 \( f : X \to Y \) 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 \( f : X \to Y \) 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 \( f_c : X_c \to Y_c \) 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. \( \forall O_{Y_c} \in \mathcal{O}_{Y_c} \), we will distinguish two exhaustive and mutually exclusive cases: Case 1: \( O_{Y_c} \in \mathcal{O}_Y \); Case
2: $Y_c \setminus O_{Y_c}$ is compact in $\mathcal{Y}$. Case 1: $O_{Y_c} \in \mathcal{O}_Y$. Then, $f_{\text{inv}}(O_{Y_c}) = f_{\text{inv}}(O_Y) \in \mathcal{O}_X$. This implies that $f_{\text{inv}}(O_{Y_c}) \in \mathcal{O}_{X_c}$. Case 2: $Y_c \setminus O_{Y_c}$ is compact in $\mathcal{Y}$. Then, $f_{\text{inv}}(Y_c \setminus O_{Y_c}) = f_{\text{inv}}(Y_c \setminus O_Y)$ is compact in $\mathcal{X}$ by the properness of $f$. Then, by Proposition 2.5, $f_{\text{inv}}(O_{Y_c}) = X_c \setminus f_{\text{inv}}(Y_c \setminus O_{Y_c}) \in \mathcal{O}_{X_c}$. Hence, in both cases, we have $f_{\text{inv}}(O_{Y_c}) \in \mathcal{O}_{X_c}$. Hence, $f_c$ is continuous.

"If" Let $f_c$ be continuous. Fix a compact set $K \subseteq \mathcal{Y}$. Then, $\omega_y \in Y_c \setminus K \in \mathcal{O}_{Y_c}$. By the continuity of $f_c$, we have $\omega_y \in f_{\text{inv}}(Y_c \setminus K) \in \mathcal{O}_{X_c}$. This implies that $X_c \setminus f_{\text{inv}}(Y_c \setminus K)$ is compact in $\mathcal{X}$. By Proposition 2.5, $f_{\text{inv}}(K) = f_{\text{inv}}(K) = X_c \setminus f_{\text{inv}}(Y_c \setminus K)$. Hence, $f$ is proper.

This completes the proof of the proposition. \qed

**Proposition 5.69** Let $\mathcal{X}$ be a topological space, $\mathcal{Y}$ be a metric space, $F \subseteq \mathcal{X}$ be closed, and $f : \mathcal{X} \to \mathcal{Y}$ be continuous and proper. Then, $f(F) \subseteq \mathcal{Y}$ is closed.

**Proof** \hspace{1cm} $\forall y \in f(F)$, by Proposition 4.13, $\exists (y_n)_{n=1}^{\infty} \subseteq f(F)$ such that $\lim_{n \in \mathbb{N}} y_n = y$. 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 $\mathcal{Y}$ is Hausdorff and $\lim_{n \in \mathbb{N}} y_n = y$, 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. \qed

In the above proposition, the assumption on $\mathcal{Y}$ may be relaxed to first countable Hausdorff space.

**Proposition 5.70** Let $\mathcal{X}$ be a topological space, $\mathcal{Y}$ be a locally compact Hausdorff topological space, $F \subseteq \mathcal{X}$ be closed, and $f : \mathcal{X} \to \mathcal{Y}$ be continuous and proper. Then, $f(F) \subseteq \mathcal{Y}$ is closed.

**Proof** \hspace{1cm} $\forall y \in f(F)$, by the local compactness of $\mathcal{Y}$, $\exists O \in \mathcal{O}_Y$ such that $y \in O$ and $\overline{O}$ is compact in $\mathcal{Y}$. Then, $y \in \overline{f(F)} \cap O$. \forall U \in \mathcal{O}_Y$ with $y \in U$, by Proposition 3.3, we have $\emptyset \neq (U \cap O) \cap f(F) = U \cap (O \cap f(F))$. Hence, $y \in \overline{f(F)} \cap O$. By Proposition 3.3. By Proposition 3.68, there exists a net $(y_\alpha)_{\alpha \in A} \subseteq f(F) \cap O$ such that $\lim_{\alpha \in A} y_\alpha = y$, $\forall \alpha \in A$, $\exists x_\alpha \in F$ such that $y_\alpha = f(x_\alpha)$. Then, $x_\alpha \in F \cap f_{\text{inv}}(O)$. Then, the net $(x_\alpha)_{\alpha \in A} \subseteq F \cap f_{\text{inv}}(O)$. By the properness of $f$, we have $f_{\text{inv}}(\overline{O}) \subseteq \mathcal{X}$ is compact. By Proposition 3.5, $F \cap f_{\text{inv}}(\overline{O})$ is closed relative to $f_{\text{inv}}(\overline{O})$, which further implies that $F \cap f_{\text{inv}}(\overline{O})$ is compact by Proposition 5.5. By Proposition 5.4, the net $(x_\alpha)_{\alpha \in A}$ admits a cluster point $x \in F \cap f_{\text{inv}}(\overline{O})$. 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 $\mathcal{Y}$ is Hausdorff and $\lim_{\alpha \in A} y_\alpha = y$, then $y = f(x) \in f(F)$. Hence, $\overline{f(F)} \subseteq f(F)$ and $f(F)$ is closed. This completes the proof of the proposition. \qed
5.8 \( \sigma \)-Compact Spaces

**Definition 5.71** A topological space is said to be \( \sigma \)-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 \( \sigma \)-compact.

(iii) \( \exists (O_n)_{n=1}^{\infty} \subseteq \mathcal{O} \) 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) \( \exists \phi : \mathcal{X} \rightarrow [0, \infty) \subseteq \mathbb{R} \) that is proper and continuous.

**Proof**

(i) \( \Rightarrow \) (ii). \( \forall x \in \mathcal{X} \), by local compactness of \( \mathcal{X} \), \( \exists O_x \subseteq \mathcal{O} \) 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 \( \sigma \)-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 \( \sigma \)-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 \( \sigma \)-compact. In all cases, \( \mathcal{X} \) is \( \sigma \)-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 \( O_0 := \emptyset \subseteq \mathcal{O} \). Then, \( \overline{O_0} = \emptyset \) is compact. \( \forall n \in \mathbb{N} \), \( K_n \cup \overline{O_{n-1}} \) is compact, by Proposition 5.52, \( \exists O_n \subseteq \mathcal{O} \) such that \( K_n \cup \overline{O_{n-1}} \subseteq O_n \subseteq \overline{O_n} \) and \( \overline{O_n} \) is compact. Then, \( (O_n)_{n=1}^{\infty} \) is an exhaustion of \( \mathcal{X} \) that we seek.

(iii) \( \Rightarrow \) (i). Let \( (U_n)_{n=1}^{\infty} \) be an exhaustion of \( \mathcal{X} \). Fix any open covering \( (O_n)_{\alpha \in \Lambda} \subseteq \mathcal{O} \) 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, \( \mathcal{X} = \bigcup_{n=1}^{\infty} \overline{U_n} \subseteq \bigcup_{n=1}^{\infty} \bigcup_{\alpha \in \Lambda_n} O_\alpha \), 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 5.62, there exists a continuous function \( \phi_n : \mathcal{X} \rightarrow [0, 1] \) such that \( \phi_n|_{\overline{O_n}} = 1 \) and \( \phi_n|_{\overline{O_{n+1}}} = 0 \). Let \( \phi := \sum_{n=1}^{\infty} (1 - \phi_n) \).
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 $x \notin \bigcup_{\lambda \in \Lambda} E_{\lambda}$. 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 $O \cap E_{\lambda} = \emptyset$, $\forall \lambda \in \Lambda$. Then, $O \cap E = \emptyset$. This contradicts with $x \in \overline{E}$, by Proposition 3.3. Hence, $x \in \bigcup_{\lambda \in \Lambda} E_{\lambda}$. 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, $O \cap E \neq \emptyset$, and hence $x \in E$. 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_\alpha)_{\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 \(\mathcal{X}\) is said to be paracompact if every open covering of \(\mathcal{X}\) has a locally finite open refinement.

Proposition 5.76 A closed subset of a paracompact space is paracompact in the subset topology.

Proof Let \(\mathcal{X}\) be a paracompact space, \(F \subseteq \mathcal{X}\) be closed, and \(\mathcal{O}_F\) be the subset topology on \(F\). Let \((O_{F\alpha})_{\alpha \in \Lambda} \subseteq \mathcal{O}_F\) be any open covering of \(F\), where \(\Lambda\) is an index set (open in the subset topology \(\mathcal{O}_F\)). By Proposition 3.4, \(\forall \alpha \in \Lambda\), \(\exists \alpha \in \mathcal{O}\) such that \(O_{F\alpha} = O_\alpha \cap F\). Then, \(F \subseteq \bigcup_{\alpha \in \Lambda} O_\alpha\) and \(\mathcal{X} \subseteq (\bigcup_{\alpha \in \Lambda} O_\alpha) \cup \bar{F}\). By the paracompactness of \(\mathcal{X}\), there exists a locally finite open refinement \(\mathcal{V}\) of \((O_{F\alpha})_{\alpha \in \Lambda} \cup \bar{F}\). 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 \((O_{F\alpha})_{\alpha \in \Lambda}\) (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 \(\mathcal{X}\) be a metric space. Let \((O_\alpha)_{\alpha \in \Lambda}\) be any open covering of \(\mathcal{X}\), 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 \(X_0 = \mathcal{X}\).

Step n (\(n \in \mathbb{N}\)); \(\forall \alpha \in \Lambda\), let \(P_{\alpha n} := \{x \in O_\alpha \cap X_{n-1} \mid \alpha\) is the least element of \(\{\beta \in \Lambda \mid x \in O_\beta\}\}\). Let \(Q_{\alpha n} := \{x \in P_{\alpha n} \mid B(x, 1/2^n) \subseteq O_\alpha\}\) and \(D_{\alpha n} := \bigcup_{x \in Q_{\alpha n}} B(x, 1/2^n) \in \mathcal{O}\). 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 \((O_\alpha)_{\alpha \in \Lambda}\). \(\forall x_0 \in \mathcal{X}\). 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 \(\mathcal{X}\). 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 \(x_0 \in \mathcal{X}\). 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 \(B(x_0, 2^{-j_0 - n_0}) \ni x_0\).
\[ \forall n \geq n_0 + j_0, \forall \alpha \in \Lambda, \text{ since } B(x_0, 2^{-j_0}) \subseteq D_{\alpha,n_0} \subseteq O_{\alpha}, \text{ then } \forall x \in B(x_0, 2^{-j_0-n_0}), \text{ we have } B(x, 2^{-j_0} - 2^{-j_0-n_0}) \subseteq B(x_0, 2^{-j_0}) \subseteq D_{\alpha,n_0}. \]

\[ \forall y \in D_{\alpha,n}, \text{ since } n > n_0, \text{ then } \exists \bar{x} \in Q_{\alpha} \text{ with } \bar{x} \notin D_{\alpha,n_0} \text{ such that } y \in B(\bar{x}, 2^{-n}). \text{ Then, } \exists \bar{x} \in B(y, 2^{-n}) \text{ such that } \bar{x} \in D_{\alpha} \setminus D_{\alpha,n_0}. \text{ Note that } 2^{-n} \leq 2^{-j_0-n_0} \leq 2^{-j_0} - 2^{-j_0-n_0}. \text{ Therefore, } x \notin D_{\alpha,n}. \text{ Hence, } B(x_0, 2^{-j_0-n_0}) \cap D_{\alpha,n} = \emptyset. \text{ Thus, } B(x_0, 2^{-j_0-n_0}) \text{ does not intersect any } D_{\alpha,n}, \forall \alpha \in \Lambda, \forall n \geq n_0 + j_0. \]

**Claim 5.77.1** \[ \forall n < n_0 + j_0, B(x_0, 2^{-j_0-n_0}) \text{ intersects at most one of the set } D_{\alpha,n} \text{’s, } \alpha \in \Lambda. \]

**Proof of claim:** Suppose the result is not true. Then, \( \exists \alpha, \beta \in \Lambda \text{ such that } B(x_0, 2^{-j_0-n_0}) \cap D_{\alpha,n} \neq \emptyset, B(x_0, 2^{-j_0-n_0}) \cap D_{\beta,n} \neq \emptyset, \text{ and } \alpha \neq \beta. \)

Without loss of generality, assume that \( \alpha \leq \beta. \) Let \( p \in B(x_0, 2^{-j_0-n_0}) \cap D_{\alpha,n} \) and \( q \in B(x_0, 2^{-j_0-n_0}) \cap D_{\beta,n}. \) By the definition of \( D_{\alpha,n}, \exists \tilde{p} \in Q_{\alpha} \subseteq D_{\alpha,n} \text{ such that } p \in B(\tilde{p}, 2^{-n}) \subseteq D_{\alpha,n}. \) Similarly, \( \exists \tilde{q} \in Q_{\beta} \subseteq D_{\beta,n} \text{ such that } q \in B(\tilde{q}, 2^{-n}) \subseteq D_{\beta,n}. \) Then, by the definition of \( Q_{\alpha} \), we have \( B(\tilde{p}, 3/2^n) \subseteq O_{\alpha}. \) Similarly, \( B(\tilde{q}, 3/2^n) \subseteq O_{\beta}. \) Note that

\[
\rho(\tilde{p}, \tilde{q}) \leq \rho(p, p) + \rho(p, x_0) + \rho(x_0, q) + \rho(q, \tilde{q}) < 2^{-n} + 2^{-j_0-n_0} + 2^{-j_0-n_0} + 2^{-n} \leq 3/2^n
\]

Then, \( \tilde{q} \in O_{\alpha}. \) But, \( \tilde{q} \in Q_{\beta} \subseteq P_{\beta,n}, \alpha \leq \beta, \) and \( \alpha \neq \beta \) implies that \( \tilde{q} \notin O_{\alpha}. \)

This is a contradiction. Therefore, the result must hold. This completes the proof of the claim. \( \square \)

Hence, \( B(x_0, 2^{-j_0-n_0}) \) can meet at most \( n_0 + j_0 - 1 \) sets in \( \mathcal{V}. \) Hence, \( \mathcal{V} \) is locally finite.

Therefore, we have obtained a locally finite open refinement of \( (O_{\alpha})_{\alpha \in \Lambda}. \)

Hence, \( \mathcal{X} \) is paracompact. This completes the proof of the theorem. \( \square \)

**Definition 5.78** Let \( \mathcal{X} \) be a topological space and \( (E_{\alpha})_{\alpha \in \Lambda} \) be a collection of subsets of \( \mathcal{X}, \) where \( \Lambda \) is an index set. \( (E_{\alpha})_{\alpha \in \Lambda} \) is said to be star-finite if \( \forall \alpha \in \Lambda, E_{\alpha,n_0} \text{ meets only finitely many members in the collection.} \)

It is easy to see that a star-finite open covering of \( \mathcal{X} \) is locally finite. But the converse need not hold.

**Proposition 5.79** A \( \sigma \)-compact locally compact space is paracompact.

**Proof** Let \( \mathcal{X} \) be a \( \sigma \)-compact locally compact space. Fix any open covering \( (O_{\alpha})_{\alpha \in \Lambda} \) of \( \mathcal{X}, \) where \( \Lambda \) is an index set. By Proposition 5.72, there exists an exhaustion \( (U_n)_{n=1}^{\infty} \) of \( \mathcal{X}. \) Let \( U_{-1} = \emptyset \) and \( U_0 = \emptyset. \) Let \( V_{\alpha,n} := O_{\alpha} \cap (U_{n+1} \setminus U_{n-2}) \in \mathcal{O}, \forall \alpha \in \Lambda, \forall n \in \mathbb{N}. \) It is easy to see that \( \mathcal{V} := \{ V_{\alpha,n} \mid \alpha \in \Lambda, n \in \mathbb{N} \} \) is an open refinement of \( (O_{\alpha})_{\alpha \in \Lambda} \) that covers \( \mathcal{X}. \) \( \forall n \in \mathbb{N}, \text{ by the compactness of } U_n \) and Proposition 5.5, \( U_n \setminus U_{n-1} \subseteq \bigcup_{\alpha \in \Lambda} V_{\alpha,n} \) is compact. Then, there exists a finite set \( \mathcal{V}_n \subseteq \)
\{V_\alpha \mid \alpha \in \Lambda\} such that \(\overline{U_n \setminus U_{n-1}} \subseteq \bigcup_{V \in V_n} V\). Then, \(V_L := \bigcup_{n=1}^{\infty} V_n\) is an open refinement of \((O_\alpha)_{\alpha \in \Lambda}\) that covers \(X\) since \(\bigcup_{n=1}^{\infty} (U_n \setminus U_{n-1}) = X\). \(\forall V \in V_L\), \(\exists n_0 \in \mathbb{N}\) such that \(V \subseteq V_{n_0}\). Then, \(V\) does not intersect any member of \(V_n\) with \(n \in \mathbb{N}\) and \(|n - n_0| > 2\). Hence, \(V_L\) is star-finite. Then, \(V_L\) is a locally finite open refinement of \((O_\alpha)_{\alpha \in \Lambda}\). Then, \(X\) is paracompact. This completes the proof of the proposition. \(\Box\)

**Proposition 5.80** Let \(X\) be a locally compact Hausdorff topological space. \(X\) is paracompact if, and only if, any open covering of \(X\) has a star-finite open refinement (that covers \(X\)).

**Proof** "Only if" Let \(U\) be any open covering of \(X\). \(\forall x \in X\), \(\exists U_x \in U\) such that \(x \in U_x \in \mathcal{O}\). By Proposition 5.53, \(\exists O_x \in \mathcal{O}\) such that \(x \in O_x \subseteq \overline{O_x} \subseteq U_x\) and \(O_x\) is compact. Hence, \((O_x)_{x \in X}\) is an open refinement of \(U\). By the paracompactness of \(X\), there exists a locally finite open refinement \(V\) of \((O_x)_{x \in X}\) that covers \(X\). \(\forall V \in \mathcal{V}\), \(\exists x \in X\) such that \(V \subseteq O_x\). Then, \(\overline{V} \subseteq \overline{O_x}\). By Propositions 5.5 and 3.5, we have \(\overline{V}\) is compact. By Proposition 5.74, \(\overline{V}\) meets only finitely many members of \(\mathcal{V}\). Hence, \(\mathcal{V}\) is star-finite. Therefore, \(\mathcal{V}\) is a star-finite open refinement of \(U\).

"If" Let \(U\) be any open covering of \(X\). Then, there is a star-finite open refinement \(V\) of \(U\). Then, \(V\) is a locally finite open refinement of \(U\). Hence, \(X\) is paracompact.

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

**Proposition 5.81** A paracompact Hausdorff topological space is normal.

**Proof** Let \(X\) be a paracompact Hausdorff topological space. Clearly, \(X\) is Tychonoff. \(\forall x_0 \in X\) and \(\forall\) closed set \(F \subseteq X\) with \(x_0 \not\in F\), we will show that \(\exists O_1, O_2 \in \mathcal{O}\) such that \(x_0 \in O_1\), \(F \subseteq O_2\), and \(O_1 \cap O_2 = \emptyset\). \(\forall x \in F\), then \(x \neq x_0\). Since \(X\) is Hausdorff, \(\exists O_1^{(1)}, O_2^{(1)} \in \mathcal{O}\) such that \(x_0 \in O_1^{(1)}\), \(x \in O_2^{(1)}\), and \(O_1^{(1)} \cap O_2^{(1)} = \emptyset\). Then, \(X \subseteq \overline{F} \cup \bigcup_{x \in F} O_2^{(1)}\). By the paracompactness of \(X\), there exists a locally finite open refinement \(V \subseteq \mathcal{O}\) of \(\{\overline{F}\} \cup \{O_2^{(1)} \mid x \in F\}\). Let \(\mathcal{V} := \{V \in \mathcal{V} \mid V \subseteq O_2^{(1)}\text{ for some }x \in F\}\). Then, \(\mathcal{V}\) is an open covering of \(F\), since \(\forall V \in \mathcal{V} \setminus \mathcal{V}\) we have \(V \subseteq \overline{F}\). Furthermore, \(\mathcal{V}\) is locally finite. \(\forall V \in \mathcal{V}, \exists x \in F\) such that \(V \subseteq O_2^{(1)}\), \(x_0 \in O_2^{(1)}\), and \(O_2^{(1)} \cap O_2^{(1)} = \emptyset\). Then, \(V \subseteq O_2^{(1)} \subseteq \overline{O_2^{(1)}}\). Hence, we have \(\overline{V} \subseteq O_2^{(1)}\) and \(x_0 \not\in \overline{V}\). By Proposition 5.74, \(\bigcup_{V \in \mathcal{V}} V = \bigcup_{V \in \mathcal{V}} \overline{V} =: \overline{O_1}\).

Then, \(O_1, O_2 \in \mathcal{O}\), \(F \subseteq \bigcup_{V \in \mathcal{V}} V =: O_2 \subseteq O_2 \subseteq O_1\), and \(x_0 \in O_1\). Hence, \(\exists O_1, O_2 \in \mathcal{O}\) such that \(x_0 \in O_1\), \(F \subseteq O_2\), and \(O_1 \cap O_2 = \emptyset\). Then, \(X\) is regular.

Next, we show that \(X\) is normal. Fix any closed sets \(F_1, F_2 \subseteq X\) with \(F_1 \cap F_2 = \emptyset\). \(\forall x \in F_2\), then \(x \not\in F_1\). Since \(X\) is regular, \(\exists O_1^{(1)}, O_2^{(1)} \in \mathcal{O}\) such that \(F_1 \subseteq O_1^{(1)}\), \(x \in O_2^{(1)}\), and \(O_1^{(1)} \cap O_2^{(1)} = \emptyset\). Then, \(X \subseteq \overline{F_2} \cup \overline{F_1}\).
(\bigcup_{x \in F_2}O_x^{(2)})$. By the paracompactness of $\mathcal{X}$, there exists a locally finite open refinement $\mathcal{V} \subseteq \mathcal{O}$ of $\{\widehat{F}_2\} \cup \{O_x^{(2)} \mid x \in F_2\}$. Let $\tilde{\mathcal{V}} := \{V \in \mathcal{V} \mid V \subseteq O_x^{(2)}\}$, for some $x \in F_2$. Then, $\tilde{\mathcal{V}}$ is an open covering of $F_2$, since $\forall V \in \mathcal{V} \setminus \tilde{\mathcal{V}}$ we have $V \subseteq \widehat{F}_2$. Furthermore, $\tilde{\mathcal{V}}$ is locally finite. $\forall V \in \tilde{\mathcal{V}}$, $\exists x \in F_2$ 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 $\bigcup_{V \in \tilde{\mathcal{V}}}^\bigcup_{V \in \mathcal{V}} V =: \tilde{O}_1$. Then, $O_1 \subseteq \mathcal{O}$, $F_2 \subseteq \bigcup_{V \in \tilde{\mathcal{V}}} V =: O_2 \subseteq \mathcal{O}$, $O_2 \subseteq \tilde{O}_1$, and $F_1 \cap \tilde{O}_1 = \emptyset$. Hence, $\exists O_1, O_2 \subseteq \mathcal{O}$ such that $F_1 \subseteq O_1$, $F_2 \subseteq O_2$, and $O_1 \cap O_2 = \emptyset$. Then, $\mathcal{X}$ is normal.

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

### 5.10 The Stone-Čech Compactification

**Definition 5.82** Let $\mathcal{X} := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, and $\mathcal{F}$ be the family of continuous functions of $\mathcal{X}$ to $I$. By Corollary 3.62, $\mathcal{X}$ is homeomorphic to $E(\mathcal{X}) \subseteq I^\mathcal{F}$, where $E$ is the equivalence map. Let $F = E(\mathcal{X})$ and $\mathcal{O}_F$ be the subset topology of $F$. Then, $\beta(\mathcal{X}) := (F, \mathcal{O}_F)$ is a compact Hausdorff topological space, by Tychonoff Theorem and Proposition 5.5. $\beta(\mathcal{X})$ is said to be the Stone-Čech compactification of $\mathcal{X}$.

**Proposition 5.83** Let $\mathcal{X} := (X, \mathcal{O})$ be a completely regular topological space, $I = [0, 1] \subseteq \mathbb{R}$, $\mathcal{F}$ be the family of continuous functions of $\mathcal{X}$ to $I$, and $E : \mathcal{X} \to E(\mathcal{X}) \subseteq I^\mathcal{F}$ be the equivalence map. Then, there exists a unique compact Hausdorff topological space $\beta(\mathcal{X})$ with the following properties:

(i) the space $E(\mathcal{X})$ is dense in $\beta(\mathcal{X})$;

(ii) each bounded continuous real-valued function on $E(\mathcal{X})$ extends to a bounded continuous real-valued function on $\beta(\mathcal{X})$;

(iii) if $\mathcal{X}$ is a dense subset of a compact Hausdorff topological space $\mathcal{Y}$, then, there exists a unique continuous mapping $\phi : \beta(\mathcal{X}) \to \mathcal{Y}$ such that $\phi$ is surjective and $\phi|_{E(\mathcal{X})} = E_{\text{inv}}$.

Furthermore, if $\mathcal{X}$ is locally compact, then $E(\mathcal{X})$ is an open subset of $\beta(\mathcal{X})$.

**Proof** We will first show that $\beta(\mathcal{X})$ defined in Definition 5.82 satisfies properties (i), (ii), and (iii). (i). By Proposition 3.5, $E(\mathcal{X})$ is dense in $F = E(\mathcal{X}) = \beta(\mathcal{X})$. (ii). Fix any bounded continuous real-valued function $g : E(\mathcal{X}) \to \mathbb{R}$. Then, $\exists N \in \mathbb{N}$ such that $g : E(\mathcal{X}) \to [-N, N] \subseteq \mathbb{R}$. Then, $\tilde{g} := (g + N)/(2N) \circ E \in \mathcal{F}$ by Proposition 3.12. Now, consider the function $\tilde{h} := \pi_{\beta(\mathcal{X})} : \beta(\mathcal{X}) \to [0, 1] \subseteq \mathbb{R}$, which is continuous by
Proposition 3.27. Note that \( \tilde{h} \big|_{E(X)} = \tilde{g} \circ E_{\text{inv}} = (g + N)/(2N) \). Then, by Proposition 3.12, \( h := 2N\tilde{h} - N \) is a continuous function of \( \beta(X) \) to \([-N, N]\). Furthermore, \( h \big|_{E(X)} = g \). 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 \( \mathcal{G} \) be the family of continuous real-valued functions of \( \mathcal{Y} \) to \( I \). Let \( E^{(3)} : \mathcal{Y} \to I^\mathcal{G} \) be the equivalence map and \( \pi_g^{(\mathcal{G})} : I^\mathcal{G} \to I \) be the projection function, \( \forall g \in \mathcal{G} \). By Propositions 5.14 and 3.61, \( \mathcal{Y} \) is completely regular. By Corollary 3.62, \( E^{(3)} : \mathcal{Y} \to E^{(3)}(\mathcal{Y}) \subseteq I^\mathcal{G} \) is a homeomorphism. Define a mapping \( \psi : \mathcal{G} \to \mathcal{F} \) by \( \psi(g) = g|_X \), \( \forall g \in \mathcal{G} \). By Proposition 3.56, \( \psi \) is injective. By Tychonoff Theorem, \( I^\mathcal{G} \) is a compact Hausdorff space. Define a mapping \( \Psi : I^\mathcal{F} \to I^\mathcal{G} \) by \( \pi_{\psi(g)}^{(\mathcal{G})}((\psi) f) = \pi_{\psi(g)}(f), \forall f \in I^\mathcal{F}, \forall g \in \mathcal{G}, \forall x \in X, E^{(3)}(x) \in I^\mathcal{G} \) satisfies \( \pi_{\psi(g)}^{(\mathcal{F})}(E^{(3)}(x)) = g(x) = \psi(g)(x) = \pi_{\psi(g)}(E(x)), \forall g \in \mathcal{G} \). 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 \( U \subseteq I^\mathcal{G} \). Then, \( U = \prod_{g \in \mathcal{G}} U_g \), where \( U_g \subseteq I \) is open, \( \forall g \in \mathcal{G} \), and \( U_g = I \) for all \( g \)’s except finitely many \( g \)’s, say \( g \in \mathcal{G}_N \subseteq \mathcal{G} \). Let \( V := \prod_{f \in \mathcal{F}} V_f \subseteq I^\mathcal{F} \) 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. \( \forall f \in V \), \( \forall g \in \mathcal{G}, \pi_{\psi(g)}^{(\mathcal{G})}((\psi) f) = \pi_{\psi(g)}(f) \in V_g \). Hence, we have \( (\Psi)(f) \in U \). Then, \( V \subseteq \Psi_{\text{inv}}(U) \). \( \forall f \in I^\mathcal{F} \setminus V, \exists f_0 \in \text{range}(\psi) \) such that \( \pi_{f_0}^{(\mathcal{G})}(f) \notin V_{f_0} \). Then, \( f_0 = \psi(g_0) \), where \( g_0 \in \mathcal{G}_N \). Note that \( \pi_{\psi(g_0)}^{(\mathcal{G})}(f) = \pi_{\psi(g_0)}(f) \notin V_{\psi(g_0)} = U_{g_0} \). Then, \( (\Psi)(f) \notin U \). This shows that \( I^\mathcal{F} \setminus V \subseteq \Psi_{\text{inv}}(I^\mathcal{G} \setminus U) \). Hence, we have \( V = \Psi_{\text{inv}}(U) \), by Proposition 2.5. Clearly, \( V \) is a basis open set in \( I^\mathcal{F} \). Hence, \( \Psi \) is continuous. This completes the proof of the claim. □

Since \( X \) is dense in \( \mathcal{Y} \) and \( E^{(3)} \) 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 \( I^\mathcal{G} \), by Proposition 5.7. Furthermore, by Proposition 5.5, \( E^{(3)}(\mathcal{Y}) \) is closed in \( I^\mathcal{G} \). Then, \( E^{(3)}(X) = E^{(3)}(\mathcal{Y}) \), where \( E^{(3)}(X) \) is the closure of \( E^{(3)}(X) \) in \( I^\mathcal{G} \). By the compactness of \( \beta(X) \), the continuity of \( \Psi \), and Proposition 5.7, \( \Psi(\beta(X)) \subseteq I^\mathcal{G} \) is compact. Furthermore, by Proposition 5.5, \( \Psi(\beta(X)) \) is closed in \( I^\mathcal{G} \). 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 \( E^{(3)}(\mathcal{Y}) \subseteq \Psi(\beta(X)) \), then \( \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \subseteq \beta(X) \) and \( \beta(X) \cap (I^\mathcal{F} \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y}))) \neq \emptyset \). Note that \( \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \) is closed in \( I^\mathcal{F} \) by the closedness of \( E^{(3)}(\mathcal{Y}) \), the continuity of \( \Psi \), and Proposition 3.10, which further implies that \( I^\mathcal{F} \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \in \mathcal{O}_X \). By the denseness of \( E(X) \) in \( \beta(X) \), we have \( E(X) \cap (I^\mathcal{F} \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y}))) \neq \emptyset \) and
hence \( \exists x \in \mathcal{X} \) such that \( E(x) \in I^\mathcal{X} \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) \). By Proposition 2.5, \( I^\mathcal{X} \setminus \Psi_{\text{inv}}(E^{(3)}(\mathcal{Y})) = \Psi_{\text{inv}}(I^\mathcal{Y} \setminus E^{(3)}(\mathcal{Y})) \). Then, \( \Psi(E(x)) \in I^\mathcal{Y} \setminus E^{(3)}(\mathcal{Y}) \). This contradicts with the fact that \( \Psi(E(x)) = E^{(3)}(x) \in E^{(3)}(\mathcal{Y}) \). Therefore, we must have \( \Psi(\beta(\mathcal{X})) = E^{(3)}(\mathcal{Y}) \). This completes the proof of the claim.

Define \( \phi : \beta(\mathcal{X}) \to \mathcal{Y} \) by \( \phi = E^{(3)} \circ \Psi_{\beta(\mathcal{X})} \). Clearly, \( \phi \) is continuous and surjective by Proposition 3.12. Clearly, \( \phi \circ E = E^{(3)} \circ E^{(3)}|\mathcal{X} = \text{id}|\mathcal{X} \). 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 \( \tilde{\phi}|E(\mathcal{X}) = E_{\text{inv}} = \phi|E(\mathcal{X}) \). 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.

\( \Box \)

**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.  \( \Box \)
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) \((g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)\); (associativeness)

(ii) \(e + g_1 = g_1 + e = g_1\); (unit element)

(iii) \(\exists (-g_1) \in G\) such that \(g_1 + (-g_1) = e = (-g_1) + g_1\). (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.

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. \( \square \)

**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 \((G, +, e_G)\) and \((H, +, e_H)\) be two groups, and \( T : G \rightarrow H \). \( T \) is said to be a homomorphism if, \( \forall g_1, g_2 \in G \), we have \( T(g_1 + H) = T(g_2) \). \( T \) 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 \( T : G \rightarrow H \) 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 \( f : X \rightarrow X \) 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 \( \circ : G \times G \rightarrow G \), and a unit element \( e \in G \), satisfying, \( \forall g_1, g_2, g_3 \in G \),

(i) \((g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)\); \( \text{ (associativity)} \)

(ii) \( e \circ g_1 = g_1 = g_1 \circ e \). \( \text{ (unit element)} \)

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) \((r_1 \times r_2) \times r_3 = r_1 \times (r_2 \times r_3)\); \:(\text{associativeness})

(ii) \(r_1 \times (r_2 + r_3) = r_1 \times r_2 + r_1 \times r_3\); \:(\text{right distributiveness})

(iii) \((r_1 + r_2) \times r_3 = r_1 \times r_3 + r_2 \times r_3\). \:(\text{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.

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.

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 \((R, +, \times, 0_R)\) and \((S, +, \times, 0_S)\) be two rings and \(T : R \to S\) is said to be a ring-homomorphism if it is a homomorphism and \(T(r_1) \times_S T(r_2) = T(r_1 \times_R r_2)\), \(\forall r_1, r_2 \in R\). \(T\) 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 \((F, +, \times, 0)\) 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.

6.4 Vector Spaces

Associated with every vector space is a set of scalars. This set of scalars can be any algebraic field \(\mathcal{F} := (F, +, 0, 1)\). 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 \(x \in \mathcal{F}\) when \(x \in F\).

Definition 6.16 A vector space \(X\) over a field \(\mathcal{F} := (F, +, 0, 1)\) 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 \(\forall \alpha, \beta \in F\)

\(\begin{align*}
(i) & \quad x \oplus y = y \oplus x; \quad \text{(commutative law)} \\
(ii) & \quad (x \oplus y) \oplus z = x \oplus (y \oplus z); \quad \text{(associative law)}
\end{align*}\)
We thus denote the quadruple \((X, \oplus, \otimes, \vartheta)\) by \(\mathcal{X}\). The vector space is denoted by \((\mathcal{X}, \mathcal{F})\).

For convenience, \((-1) \otimes x := \ominus x\) and called the negative of the vector \(x\). Note that

\[(\ominus x) \oplus x = x \oplus (\ominus x) = 1 \otimes x \oplus (-1) \otimes x = (1 + (-1)) \otimes x = 0 \otimes x = \vartheta\]

We will also denote \(x \ominus y := x \ominus (\ominus y)\). We will abuse the notation to say \(x \in \mathcal{X}\) when \(x \in X\). Note that \((\mathcal{X}, \oplus, \ominus)\) forms an abelian group.

**Proposition 6.17** Let \(\mathcal{X} := (X, \oplus, \otimes, \vartheta)\) be a vector space over the field \(\mathcal{F} := (F, +, \cdot, 0, 1)\). \(\forall x, y, z \in \mathcal{X}\), \(\forall \alpha, \beta \in \mathcal{F}\), we have

1. \(x \oplus y = x \oplus z \Rightarrow y = z\); (cancellation law)
2. \(\alpha \otimes x = \alpha \otimes y \text{ and } \alpha \neq 0 \Rightarrow x = y\); (cancellation law)
3. \(\alpha \otimes x = \beta \otimes x \text{ and } x \neq \vartheta \Rightarrow \alpha = \beta\); (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 \(\alpha \otimes \vartheta = \vartheta\), \(\forall \alpha \in \mathcal{F}\). Then, \((X, \oplus, \otimes, \vartheta)\) 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 \(\mathcal{F}\).

**Example 6.20** Let \(\mathcal{F} := (F, +, \cdot, 0, 1)\) be a field, \(\mathcal{Y} := (Y, \oplus_Y, \otimes_Y, \vartheta_Y)\) 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 \(z_a(u) = x(u) + y \otimes_Y u\), \(\forall u \in A\), \(z_s := \alpha \otimes x \in X\) is given by \(z_s(u) = x(u) + \vartheta_Y x(u)\), \(\forall u \in A\). Let \(\vartheta \in X\) be given by \(\vartheta(u) = \vartheta_Y\), \(\forall u \in A\).

Now, we will show that \(\mathcal{X} := (X, \oplus, \otimes, \vartheta)\) 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
Then, it is straightforward to check that
\[(x \oplus y)(u) = x(u) \oplus y(u) = y(u) \oplus x(u) = (y \oplus x)(u) \Rightarrow x \oplus y = y \oplus x; \]
\[(x \oplus y \oplus z)(u) = (x \oplus y)(u) \oplus y z(u) \Rightarrow x(u) \oplus y \oplus z = (x \oplus y)(u) \oplus y z(u); \]
\[(x \oplus \vartriangledown)(u) = x(u) \oplus \vartriangledown \vartriangledown y = x(u) \Rightarrow x \oplus \vartriangledown = x; \]
\[(\alpha \oplus (x \oplus y))(u) = \alpha \oplus (x \oplus y)(u) = \alpha \oplus y x(u) \oplus y y(u) = \alpha \oplus y x(u) \oplus y \alpha x y(u) = (\alpha \oplus x)(u) \oplus y (\alpha \oplus y)(u) = (\alpha \oplus x) \oplus y (\alpha \oplus y)(u) \Rightarrow \alpha \oplus x = \alpha \oplus y = \alpha \oplus y; \]
\[(\alpha \oplus (\beta \oplus x)) (u) = (\alpha \oplus \beta \oplus x)(u) = \alpha \oplus (\beta \oplus x)(u) = (\alpha \oplus x \oplus \beta \oplus x)(u) \Rightarrow \alpha \oplus (\beta \oplus x) = (\alpha \oplus \beta \oplus x) = \alpha \oplus (\beta \oplus x); \]
\[(0 \oplus x)(u) = 0 \oplus x(u) = \vartriangledown x \Rightarrow \vartriangledown x = (0 \oplus x)(u) = 0 \oplus x = 0; \]  
Therefore, \(X\) is a vector space over \(F\). This vector space will be denoted by \((M(A, Y), F)\).

\[\Box\]

**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 \(x \oplus y := (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n) \in X, \alpha \oplus x := (\alpha \xi_1, \ldots, \alpha \xi_n) \in X, \forall x, \alpha \in F, \forall y, \alpha \in F, \forall x, \forall y, \forall \alpha \in F\).
Then, it is straightforward to check that \(F^n := (X, \oplus, \alpha, \oplus)\) is a vector space over \(F\).

\[\Box\]

**Example 6.22** Let \(F := (F, +, \cdot, 0, 1)\) be a field. \(X = F^{n \times m} := \{m \times n\}-\text{dimensional } F \text{-valued matrices}\) with \(m, n \in \mathbb{N}\). Define vector addition and scalar multiplication by \(x \oplus y := (\xi_{ij} + \eta_{ij})_{m \times n} \in X, \alpha \oplus x := (\alpha \xi_{ij})_{m \times n} \in X, \forall x, \alpha \in F, \forall y, \alpha \in F, \forall x, \forall y, \forall \alpha \in F\).
Then, it is straightforward to check that \(F^{m \times n} := (X, \oplus, \alpha, \oplus)\) is a vector space over \(F\).

\[\Box\]

**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 \(x \oplus y := (\xi_k + \eta_k)_{k=1}^\infty \in X, \alpha \oplus x := (\alpha \xi_k)_{k=1}^\infty \in X, \forall x, \alpha \in F, \forall y, \alpha \in F, \forall x, \forall y, \forall \alpha \in F\).
Then, it is straightforward to check that \((X, \oplus, \alpha, \oplus)\) is a vector space over \(F\).

\[\Box\]

### 6.5 Product Spaces

**Proposition 6.24** Let \(X := (X, \oplus_X, \circ_X, \vartriangledown_X)\) and \(Y := (Y, \oplus_Y, \circ_Y, \vartriangledown_Y)\) be vector spaces over the field \(F := (F, +, \cdot, 0, 1)\). The Cartesian product of \(X\) and \(Y\), denoted by \(X \times Y\), is the quadruple \((X \times Y, \oplus, \circ, (\vartriangledown_X, \vartriangledown_Y))\),
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, $\forall (x_1, y_1), (x_2, y_2) \in X \times Y$, $\forall \alpha \in \mathcal{F}$, $(x_1, y_1) \oplus (x_2, y_2) := (x_1 \oplus x_2, y_1 \oplus y_2)$ and $\alpha \otimes (x_1, y_1) := (\alpha \otimes x_1, \alpha \otimes y_1)$. Then, $(\mathcal{X}, \mathcal{Y}, \mathcal{F})$ 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) $(x_1, y_1) \oplus (x_2, y_2) = (x_1 \oplus x_2, y_1 \oplus y_2) = (x_2, y_2) \oplus (x_1, y_1)$;

(ii) $((x_1, y_1) \oplus (x_2, y_2)) \oplus (x_3, y_3) = (x_1 \oplus x_2, y_1 \oplus y_2) \oplus (x_3, y_3) = ((x_1 \oplus x_2) \oplus x_3, (y_1 \oplus y_2) \oplus y_3) = (x_1 \oplus (x_2 \oplus x_3), (y_1 \oplus (y_2 \oplus y_3))) = (x_1, y_1) \oplus ((x_2, y_2) \oplus (x_3, y_3))$;

(iii) $(x_1, y_1) \odot \vartheta = (x_1 \odot x \vartheta_X, y_1 \odot y \vartheta_Y) = (x_1, y_1)$;

(iv) $\alpha \odot ((x_1, y_1) \oplus (x_2, y_2)) = \alpha \odot (x_1 \oplus x_2, y_1 \oplus y_2) = (\alpha \odot x_1 \oplus \alpha \odot x_2, \alpha \odot y_1 \oplus \alpha \odot y_2) = (\alpha \odot x_1, \alpha \odot y_1) \oplus (\alpha \odot x_2, \alpha \odot y_2) = \alpha \odot (x_1, y_1) \oplus \alpha \odot (x_2, y_2)$;

(v) $(\alpha + \beta) \odot (x_1, y_1) = ((\alpha + \beta) \odot x_1, (\alpha + \beta) \odot y_1) = (\alpha \odot x_1 \oplus \beta \odot x_1, \alpha \odot y_1 \oplus \beta \odot y_1) = (\alpha \odot x_1, \alpha \odot y_1) \oplus (\beta \odot x_1, \beta \odot y_1) = \alpha \odot (x_1, y_1) \oplus \beta \odot (x_1, y_1)$;

(vi) $(\alpha \beta) \odot (x_1, y_1) = ((\alpha \beta) \odot x_1, (\alpha \beta) \odot y_1) = (\alpha \odot (\beta \odot x_1), \alpha \odot (\beta \odot y_1)) = \alpha \odot (\beta \odot (x_1, y_1))$;

(vii) $0 \odot (x_1, y_1) = (0 \odot x \vartheta_X, 0 \odot y \vartheta_Y) = (\vartheta_X, \vartheta_Y) = \vartheta; 1 \odot (x_1, y_1) = (1 \odot x \vartheta_X, 1 \odot y \vartheta_Y) = (x_1, y_1)$.

Hence, $\mathcal{X} \times \mathcal{Y}$ is a vector space over $\mathcal{F}$. \hfill \square

With the above definition, it is easy to generalize to $\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$, where $n \in \mathbb{N}$. We will also write $\mathcal{X}^n = \underbrace{\mathcal{X} \times \cdots \times \mathcal{X}}_{n}$, where $n \in \mathbb{N}$. When $n = 0$, $\prod_{i=1}^{n} \mathcal{X}_i$ is given by the vector space defined in Example 6.18.

### 6.6 Subspaces

**Proposition 6.25** Let $\mathcal{X} := (X, \odot, \otimes, \vartheta)$ be a vector space over the field $\mathcal{F} := (\mathcal{F}, +, \cdot, 0, 1)$, and $M \subseteq X$ with $M \neq \emptyset$. Then, $\mathcal{M} := (M, \odot, \otimes, \vartheta)$ is a vector space over $\mathcal{F}$ (which will be called a subspace of $(\mathcal{X}, \mathcal{F})$) if, and only if, $\forall x, y \in M, \forall \alpha, \beta \in \mathcal{F}$, we have $\alpha \odot x \oplus \beta \odot y \in M$. We will also abuse the notation to say $M$ is a subspace of $(\mathcal{X}, \mathcal{F})$. $M$ is said to be a proper subspace of $(\mathcal{X}, \mathcal{F})$ if $M \subset X$. 


The range space $\beta A$ is said to be linear if an affine operator $y$ is defined. Let $\forall x, y \in M$, $\forall \alpha, \beta \in F$. $x \oplus y = 1 \otimes x \oplus 1 \otimes y \in M$. Hence, $M$ is closed under vector addition. $\alpha \otimes x = (\alpha + 0) \otimes x = \alpha \otimes x \otimes 0 \otimes x \in M$. Hence, $M$ is closed under scalar multiplication. Then, it is straightforward to show that $M$ 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 $M = \{\emptyset\}$ 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 $(\mathbb{R}^n, \mathbb{R})$, $n \in \mathbb{N}$. 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 $\emptyset = (0, 0, \ldots)$. By Example 6.23, $X := (X, \oplus, \otimes, \emptyset)$ is a vector space over $\mathbb{R}$. Let $M := \{(\xi_k)_{k=1}^\infty \in X \mid \lim_{k \in \mathbb{N}} \xi_k \in \mathbb{R}\}$ is a subspace.

5. Let $X := \{f : (0,1] \to \mathbb{R}^n\}$, $\oplus$ and $\otimes$ be the usual addition and scalar multiplication, and $\emptyset : (0,1] \to \mathbb{R}^n$ be given by $\emptyset(t) = \emptyset_{1 \times n}$, $\forall t \in (0, 1]$. By Example 6.20, $X := (X, \oplus, \otimes, \emptyset)$ is a vector space over $\mathbb{R}$. Let $M := \{f \in X \mid f$ 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 $\emptyset_X$. 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 $N(A) := \{x \in \mathcal{X} \mid A(x) = 0\}$. The range space of $A$ is $R(A) := \text{range}(A)$. $B : \mathcal{X} \to \mathcal{Y}$ 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 \( B^{-1} \) are linear operators, then \( BA \) is also bijective and \( (BA)^{-1} = A^{-1}B^{-1} \).

**Definition 6.30** Let \( (\mathcal{X}, F) \) be a vector space, \( \alpha \in F \), and \( S, T \subseteq \mathcal{X} \). The sets \( \alpha S \) and \( S + T \) are defined by

\[
\alpha S := \{ \alpha s \mid s \in S \}; \quad S + T := \{ s + t \mid s \in S, t \in T \}
\]

This concept is illustrated in Figure 6.1.

![Figure 6.1: The sum of two sets.](image)

We should note that \( S + T = T + S \), \( \emptyset + S = \emptyset \), \( \{ \emptyset \} + S = S \), and \( \alpha \emptyset = \emptyset \). Thus, \( S - T := S + (-T) \).

**Proposition 6.31** Let \( M \) and \( N \) be subspaces of a vector space \( (\mathcal{X}, F) \) and \( \alpha \in F \). Then, \( M \cap N, M + N, \) and \( \alpha M \) are subspaces of \( (\mathcal{X}, F) \).

**Proof** Since \( M \) and \( N \) are subspaces, then \( \emptyset \in M \) and \( \emptyset \in N \). Hence, \( \emptyset \in M \cap N \neq \emptyset \), \( \emptyset \in M + N \neq \emptyset \), and \( \emptyset = \alpha \emptyset \in \alpha M \neq \emptyset \).

For every \( x, y \in M \cap N \), \( \forall \alpha, \beta \in F \), \( \alpha x + \beta y \in M \) and \( \alpha x + \beta y \in N \). Then, \( \alpha x + \beta y \in M \cap N \). Hence, \( M \cap N \) is a subspace.

For every \( x, y \in M + N \), \( \forall \alpha, \beta \in F \), 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.
∀x, y ∈ aM, ∀α, β ∈ F, we have x = αx + y, where x, y ∈ M.
Then, αx + βy = α(αx + βy) = αx + βy ∈ aM. Hence, aM is a subspace.
This completes the proof of the proposition.
We should note that M ∪ 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 \((X, F)\) is a sum of the form \(\sum_{i=1}^{n} \alpha_i x_i := \alpha_1 x_1 + \cdots + \alpha_n x_n\), where \(\alpha_1, \ldots, \alpha_n \in F\).

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 \(\emptyset\).

Definition 6.33 Let \((X, F)\) be a vector space and \(S \subseteq X\).

\[
\text{span}(S) := \{ \sum_{i=1}^{n} \alpha_i x_i \mid x_i \in S, \alpha_i \in F, i = 1, \ldots, n, n \in \mathbb{Z}_+ \} \\
= \{ \text{linear combination of vectors in } S \}
\]

is called the subspace generated by \(S\).

Proposition 6.34 Let \((X, F)\) be a vector space and \(S \subseteq X\). Then, \(\text{span}(S)\) is the smallest subspace containing \(S\).

Proof Clearly, \(\emptyset \in \text{span}(S)\) ≠ \(\emptyset\). ∀\(x, y \in \text{span}(S)\), ∀\(α, β \in F\). It is easy to show that \(αx + βy \in \text{span}(S)\). Hence, \(\text{span}(S)\) is a subspace. Clearly \(\text{span}(S) \supseteq S\).

∀\(M \subseteq X\) such that \(S \subseteq M\) and \(M\) is a subspace. Clearly, \(\emptyset \in M\) by the proof of Proposition 6.25. \(\emptyset\) equals to a linear combination of vectors in \(S\). Then, \(y = \sum_{i=1}^{n} \alpha_i x_i\), where \(n \in \mathbb{Z}_+, x_1, \ldots, x_n \in S\), and \(\alpha_1, \ldots, \alpha_n \in F\). 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.

Definition 6.35 Let \((X, F)\) 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 \(\{x_0\} + M\). \(\forall x \in V\), \(V - x := V - \{\bar{x}\}\) is a subspace.

Definition 6.36 Let \((X, F)\) be a vector space, \(S \subseteq X\), and \(S \neq \emptyset\). The linear variety generated by \(S\), denoted by \(v(S)\), is defined as the intersection of all linear varieties in \(X\) that contain \(S\).

Proposition 6.37 Let \((X, F)\) be a vector space and \(\emptyset \neq S \subseteq X\). Then, \(v(S)\) is a linear variety given by \(v(S) = x_0 + \text{span}(S - x_0)\), where \(x_0\) is any vector in \(S\).
Proposition 6.39 Let \( X \) be a vector space and \( C \subseteq X \). Then, \( v(S) \subseteq x_0 + \text{span}(S - x_0) \) is a linear variety. Hence, \( v(S) \subseteq x_0 + \text{span}(S - x_0) \).

\[ \forall V \subseteq X \text{ such that } S \subseteq V \text{ and } V \text{ is a linear variety. Then, } x_0 \in V \text{ and } V - x_0 \text{ is a subspace. Clearly, } S - x_0 \subseteq V - x_0, \text{ which implies that, by } \text{Proposition 6.34, } \text{span}(S - x_0) \subseteq V - x_0. \text{ Therefore, } x_0 + \text{span}(S - x_0) \subseteq V. \]

Hence, \( x_0 + \text{span}(S - x_0) \subseteq v(S) \). Therefore, \( v(S) = x_0 + \text{span}(S - x_0) \).

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

6.7 Convex Sets

Denote \( K \) to be either \( \mathbb{R} \) or \( \mathbb{C} \).

Definition 6.38 Let \((X, K)\) be a vector space and \( C \subseteq X \). \( C \) is said to be convex if, \( \forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subseteq \mathbb{R}, \) we have \( \alpha x_1 + (1 - \alpha)x_2 \in C \).

Subspaces and linear varieties are convex, so is \( \emptyset \).

Proposition 6.39 Let \((X, K)\) be a vector space and \( K, G \subseteq X \) be convex sets. Then,

1. \( \lambda K \) is convex, \( \forall \lambda \in K \);

2. \( K + G \) is convex.

Proof \( \forall \lambda \in K. \forall x_1, x_2 \in \lambda K, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \exists k_1, k_2 \in K \) such that \( x_1 = \lambda k_1 \) and \( x_2 = \lambda k_2 \). Then, \( \alpha x_1 + (1 - \alpha)x_2 = \alpha \lambda k_1 + (1 - \alpha)\lambda k_2 = \lambda(\alpha k_1 + (1 - \alpha)k_2) \). Since \( K \) is convex, then \( \alpha k_1 + (1 - \alpha)k_2 \in K \). This implies that \( \alpha x_1 + (1 - \alpha)x_2 \in \lambda K \). Hence, \( \lambda K \) is convex.

\( \forall x_1, x_2 \in K + G, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \exists k_1, k_2 \in K \) and \( \exists g_1, g_2 \in G \) such that \( x_i = k_i + g_i, i = 1, 2 \). 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 \( \alpha g_1 + (1 - \alpha)g_2 \in G \). This implies that \( \alpha x_1 + (1 - \alpha)x_2 \in K + G \). Hence, \( K + G \) is convex.

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

Proposition 6.40 Let \((X, K)\) be a vector space and \( \{C_\lambda\}_{\lambda \in \Lambda} \) be a collection of convex subsets of \( X \). Then, \( C := \bigcap_{\lambda \in \Lambda} C_\lambda \) is convex.

Proof \( \forall x_1, x_2 \in C, \forall \alpha \in [0, 1] \subseteq \mathbb{R}. \forall \lambda \in \Lambda, x_1, x_2 \in C_\lambda. \) Since \( C_\lambda \) is convex, then \( \alpha x_1 + (1 - \alpha)x_2 \in C_\lambda \). This implies that \( \alpha x_1 + (1 - \alpha)x_2 \in C \). Hence, \( C \) is convex. This completes the proof of the proposition. \( \square \)

Definition 6.41 Let \((X, K)\) be a vector space and \( S \subseteq X \). The convex hull generated by \( S \), denoted by \( \text{co}(S) \), is the smallest convex set containing \( S \).
CHAPTER 6. VECTOR SPACES

Convex

Nonconvex

Figure 6.2: Convex and nonconvex sets.

Convex hulls.

Figure 6.3: Convex hulls.

Justification of the existence of convex hull rests with Proposition 6.40.

Definition 6.42 Let \((\mathcal{X}, \mathbb{K})\) be a vector space. A convex combination of vectors \(x_1, \ldots, x_n \in \mathcal{X}\), 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 \((\mathcal{X}, \mathbb{K})\) be a vector space and \(S \subseteq \mathcal{X}\). Then,

\[\text{co}(S) = \{\text{convex combinations of vectors in } S\}\]

Proof We need the follow result.

Claim 6.43.1 Let \(G \subseteq \mathcal{X}\) 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 \(\sum_{i=1}^{n} \alpha_i x_i = x_1 \in G\). 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,

\[\sum_{i=1}^{k+1} \alpha_i x_i = (\alpha_1 + \cdots + \alpha_k) \sum_{i=1}^{k} \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_k} x_i + \alpha_{k+1} x_{k+1}\]
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 \( S \subseteq G \). \( \forall p \in K \), by Claim 6.43.1, \( p \in G \) since \( S \subseteq G \). Then, \( K \subseteq G \).

The above implies that \( K \) is the smallest convex set containing \( S \). Hence, \( K = \text{co} \left( S \right) \). This completes the proof of the proposition. \( \square \)

**Definition 6.44**

Let \((X, \mathbf{K})\) be a vector space and \( C \subseteq X \). \( C \) 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 \( p \in X \) 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 \(\mathcal{F} := (\mathbb{F}, +, \cdot, 0, 1)\) and \(S \subseteq \mathcal{X}\). Then, \(S\) is a linearly independent set if, and only if, \(\forall n \in \mathbb{N}, \forall x_1, \ldots, x_n \in \mathcal{X}\) which are distinct, we have \(\sum_{i=1}^{n} \alpha_i x_i = 0\) implies that \(\alpha_i = 0\), \(i = 1, \ldots, n\).

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 \(\sum_{i=1}^{n} \alpha_i x_i = 0\) 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 \(\sum_{i=1}^{n} \alpha_i x_i = 0\) 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, \(x_1 = \sum_{i=2}^{n} (-\alpha_i^{-1} \alpha_i) x_i\). 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 \(\mathcal{F} := (\mathbb{F}, +, \cdot, 0, 1)\), \(\{x_1, \ldots, x_n\} \subseteq \mathcal{X}\) 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 \(\sum_{i=1}^{n} (\alpha_i - \beta_i) x_i = 0\). 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 \(\mathcal{F} := (\mathbb{F}, +, \cdot, 0, 1)\), \(n \in \mathbb{Z}_+, x_1, \ldots, x_n \in \mathcal{X}\). 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 $\mathcal{F} := (F, +, -, 0, 1)$. 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 form Theorem 6.46.

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

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 $\mathcal{F} := (F, +, -, 0, 1)$. 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.
∀y_1,\ldots,y_m \in X with m > n. There are two exhaustive and mutually exclusive cases: Case 1: n = 0; Case 2: n > 0. Then, S_1 = \emptyset and X contains a single vector \vartheta. Hence, y_1 = \cdots = y_m = \vartheta. Clearly, y_1,\ldots,y_m are linearly dependent. This case is proven. Case 2: n > 0. Take S_1 = \{x_1,\ldots,x_n\}. Since S_1 is a basis for the vector space, then, \forall i \in \{1,\ldots,m\}, y_i \in \text{span}(S_1), that is, \exists \alpha_{ij} \in F, j = 1,\ldots,n, such that y_i = \sum_{j=1}^{n} \alpha_{ij} x_j. Consider the m \times n-dimensional matrix A := (\alpha_{ij})_{m \times n}. Since m > n, then rank(A) < m, which implies that the row vectors of A are linearly dependent. \exists \beta_1,\ldots,\beta_m \in F, which are not all 0's, such that \sum_{i=1}^{m} \beta_i \alpha_{ij} = 0, j = 1,\ldots,n. This implies \sum_{i=1}^{m} \beta_i y_i = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_i \alpha_{ij} x_j = \sum_{j=1}^{n} (\sum_{i=1}^{m} \beta_i \alpha_{ij}) x_j = \vartheta. Hence, y_1,\ldots,y_m are linearly dependent. This case is also proven.

Hence, the dimension of the vector space is unique.

∀S_2 \subseteq X with n vectors that is a linearly independent set, we will show that S_2 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, S_2 = \emptyset and X is a singleton set consisting of the null vector, which equals to \text{span}(S_2). Hence, S_2 is a basis for the vector space. Case 2: n > 0. Take S_2 = \{z_1,\ldots,z_n\}. \forall x \in X, x, z_1,\ldots,z_n (for a total of n + 1 vectors), by the preceding proof, are linearly dependent. Then, \exists \alpha, \beta_1,\ldots,\beta_n \in F, which are not all 0's, such that \alpha x + \sum_{i=1}^{n} \beta_i z_i = \vartheta. Suppose \alpha = 0. Then, \sum_{i=1}^{n} \beta_i z_i = \vartheta. Since S_2 is a linearly independent set, by Theorem 6.46, then \beta_1 = \cdots = \beta_n = 0. This contradicts with the fact that \alpha, \beta_1,\ldots,\beta_n are not all 0's. Hence, \alpha \neq 0. Then, x = \sum_{i=1}^{n} (-\alpha^{-1} \beta_i) z_i and x \in \text{span}(S_2). Therefore, we have X = \text{span}(S_2). Hence, S_2 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 $\mathbb{K}$ to be either $\mathbb{R}$ or $\mathbb{C}$.

Definition 7.1 Let $(\mathcal{X}, \mathbb{K})$ be a vector space. A norm on the vector space is a real-valued function $\| \cdot \| : \mathcal{X} \to [0, \infty) \subset \mathbb{R}$ that satisfies the following properties: $\forall x, y \in \mathcal{X}$, $\forall \alpha \in \mathbb{K}$,

(i) $0 \leq \| x \| < +\infty$ and $\| x \| = 0 \iff x = \emptyset$;

(ii) $\| x + y \| \leq \| x \| + \| y \|$; \hspace{5mm} (triangle inequality)

(iii) $\| \alpha x \| = |\alpha| \| x \|$.

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 $(\mathcal{X}, \mathbb{K}, \| \cdot \|)$.

To simplify notation in the theory, we may later simply discuss a normed linear space $\mathcal{X} := (\mathcal{X}, \mathbb{K}, \| \cdot \|)$ without further reference to components of $\mathcal{X}$, where the operations are understood to be $\oplus_{\mathcal{X}}$ and $\otimes_{\mathcal{X}}$, the null vector is understood to be $\emptyset_{\mathcal{X}}$, and the norm is understood to be $\| \cdot \|_{\mathcal{X}}$. When it is clear from the context, we will neglect the subscript $\mathcal{X}$.

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 $| (\xi_1, \ldots, \xi_n) | = \sqrt{\sum_{i=1}^n |\xi_i|^2}$, $\forall (\xi_1, \ldots, \xi_n) \in$
$C^n$. 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, $X := (X, \oplus, \otimes, \vartheta)$ is a vector space over $\mathbb{R}$. Let $M := \{ f \in X \mid f$ is continuous $\}$ is a subspace by Proposition 6.25. Let $M := (M, \oplus, \otimes, \vartheta)$. Then, $(M, \mathbb{R})$ is a vector space. Introduce a norm on this space by $\|x\| = \max_{a \leq t \leq b} |x(t)|$, $\forall x \in M$. Now, we verify the properties of the norm. $\forall x, y \in M$, $\forall \alpha \in \mathbb{R}$.

(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, $\|x\| = 0 \iff x = \vartheta$.

(ii) $\|x + y\| = \max_{a \leq t \leq b} |x(t) + y(t)| \leq \max_{a \leq t \leq b} (|x(t)| + |y(t)|) \leq \max_{a \leq t \leq b} |x(t)| + \max_{a \leq t \leq b} |y(t)| = \|x\| + \|y\|$. 

(iii) $\|\alpha x\| = \max_{a \leq t \leq b} |\alpha x(t)| = \max_{a \leq t \leq b} |\alpha| |x(t)| = |\alpha| \|x\|$, where have made use of Proposition 3.81 and the fact that $\|x\| \in \mathbb{R}$ in the third equality.

Hence, $C([a, b]) := (M, \mathbb{R}, \|\cdot\|)$ 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, $X := (X, \oplus, \otimes, \vartheta)$ 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 $M := (M, \oplus, \otimes, \vartheta)$. Then, $(M, \mathbb{R})$ is a vector space. Introduce a norm on this space by $\|x\| = \max_{a \leq t \leq b} |x(t)| + \max_{a \leq t \leq b} |x^{(1)}(t)|$, $\forall x \in M$. Now, we verify the properties of the norm. $\forall x, y \in M$, $\forall \alpha \in \mathbb{R}$.

(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, $\|x\| = 0 \iff x = \vartheta$.

(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, $C_1([a, b]) := (M, \mathbb{R}, \|\cdot\|)$ is a normed linear space.
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 $\|x\|_p = (\sum_{k=1}^\infty |\xi_k|^p)^{1/p}, \forall x = (\xi_k)_{k=1}^\infty \in M_p$. 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: $x \in M_p, \forall p \in [1, +\infty] \subset \mathbb{R}$, we will show that $M_p := (M_p, \oplus, \ominus, \vartheta)$ is a subspace in $(X, \mathbb{R})$ and $l_p := (M_p, \mathbb{R}, \| \cdot \|_p)$ is a normed linear space.

Let $f \in \mathcal{X}_p, \forall x,y \in M_p, \forall \alpha, \beta \in \mathbb{R}$. It is easy to check that $\|\alpha x + \beta y\|_p \leq \|\alpha x\|_p + \|\beta y\|_p = |\alpha| \|x\|_p + |\beta| \|y\|_p < +\infty$. This implies that $\alpha x + \beta y \in M_p$. Hence, $M_p$ is a subspace of $(X, \mathbb{R})$. It is easy to check that $\|x\|_p \in [0, \infty) \subset \mathbb{R}, \forall x \in l_p$, and $\|x\|_p = 0 \iff x = \vartheta$. Therefore, $l_p$ 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

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b$$

where equality holds if, and only if, $a = b$.

**Proof** Define $f : [0, \infty) \rightarrow \mathbb{R}$ 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 $(0, \infty)$. $f'(t) = \lambda t^{\lambda-1} - \lambda, \forall t \in (0, \infty)$. Then, $f'(1) > 0, \forall t \in (0, 1)$ and $f'(1) < 0, \forall t \in (1, \infty)$. 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 $a^\lambda b^{1-\lambda} = 0 \leq \lambda a = \lambda a + (1-\lambda)b$, where equality holds if, and only if, $a = b = 0$. This case is proved. Case 2: $b > 0$. Since $f(a/b) \leq 0$, we immediately obtain the desired inequality, where equality holds if, and only if, $a/b = 1 \iff a = b$. This case is also proved.

This completes the proof of the lemma.

**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 $|\alpha| |\xi_k|^p = |\beta| |\eta_k|^q$, $k = 1, 2, \ldots$. 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, $y = \vartheta$ 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, $\alpha_1 |\xi_k|^p = \| y \|_q^q$, $k = 1, 2, \ldots$. 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. \qed

When $p = 2 = q$, the Hölder’s inequality becomes the well-known Cauchy-Schwarz Inequality:

$$
\sum_{k=1}^{\infty} |\xi_k \eta_k| \leq \left( \sum_{k=1}^{\infty} |\xi_k|^2 \right)^{1/2} \left( \sum_{k=1}^{\infty} |\eta_k|^2 \right)^{1/2}
$$
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 \( \forall x = (\xi_k)_{k=1}^\infty, y = (\eta_k)_{k=1}^\infty \in l_p \). We will distinguish three exhaustive and mutually exclusive cases: Case 1: \( p = 1 \); Case 2: \( p = \infty \); Case 3: \( 1 < p < \infty \).

Case 1: \( p = 1 \). \( \|x + y\|_1 = \sum_{k=1}^{\infty} |\xi_k + \eta_k| \leq \sum_{k=1}^{\infty} (|\xi_k| + |\eta_k|) = \|x\|_1 + \|y\|_1 < +\infty \). 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
\[
\frac{|\xi_k + \eta_k|^p}{(\|x\|_p + \|y\|_p)^p} \leq \lambda \frac{|\xi_k|^p}{\|x\|_p^p} + (1 - \lambda) \frac{|\eta_k|^p}{\|y\|_p^p}
\]
where equality holds \( \iff \xi_k \eta_k \geq 0 \) and
\[
\frac{|\xi_k|}{\|x\|_p} = \frac{|\eta_k|}{\|y\|_p} \iff \xi_k = \frac{\eta_k}{\|y\|_p} \|x\|_p
\]

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 \( \iff \frac{\xi_k}{\| x \|_p} = \frac{\eta_k}{\| y \|_p}, \ k = 1, 2, \ldots \iff (1/\| x \|_p)x = (1/\| y \|_p)y \). 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, \( x + y \in \ell_p \). 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, \( Y, \otimes, \otimes, \vartheta \) is a vector space over \( \mathbb{K} \), where \( \otimes, \otimes, \) and \( \vartheta \) are defined as in the example. For \( p \in [1, \infty) \subset \mathbb{R} \), let \( M_p := \{ (\xi_k)_{k=1}^{\infty} \in Y \mid \sum_{k=1}^{\infty} \| \xi_k \|_X^p < +\infty \} \). Define the norm \( \| x \|_p = (\sum_{k=1}^{\infty} \| \xi_k \|_X^p )^{1/p}, \forall x = (\xi_k)_{k=1}^{\infty} \in M_p \). Let \( M_\infty := \{ (\xi_k)_{k=1}^{\infty} \in Y \mid \sup_{k \geq 1} \| \xi_k \|_X < +\infty \} \). Define the norm \( \| x \|_\infty = \sup_{k \geq 1} \| \xi_k \|_X, \forall x = (\xi_k)_{k=1}^{\infty} \in M_\infty \).

\[ \| x + y \|_p = \begin{cases} \left( \sum_{k=1}^{\infty} \| \xi_k + \eta_k \|_X^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \| \xi_k + \eta_k \|_X & p = \infty \end{cases} \]
\[ \leq \begin{cases} \left( \sum_{k=1}^{\infty} \| \xi_k \|_X^p \right)^{1/p} + \left( \sum_{k=1}^{\infty} \| \eta_k \|_X^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \| \xi_k \|_X + \sup_{k \geq 1} \| \eta_k \|_X & p = \infty \end{cases} \]
\[ = \| x \|_p + \| y \|_p < +\infty \]
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

\[ \| \alpha x \|_p = \begin{cases} \left( \sum_{k=1}^{\infty} \| \alpha \xi_k \|_X^p \right)^{1/p} = \left( \sum_{k=1}^{\infty} |\alpha|^p \| \xi_k \|_X^p \right)^{1/p} & p \in [1, \infty) \\ \sup_{k \geq 1} \| \alpha \xi_k \|_X = \sup_{k \geq 1} |\alpha| \| \xi_k \|_X & p = \infty \end{cases} \]
Example 7.11 Let \( X \) be a vector space. Introduce a norm on this space by \( \| \cdot \|_p = \left( \sum_{k=1}^{\infty} \| \xi_k \|_X^p \right)^{1/p} \) \( p \in [1, \infty) \) and \( \| \cdot \|_\infty = \sup \{ \| \xi_k \|_X \} \) \( \forall p \in [1, \infty) \). This is straightforward and is therefore omitted.

\[ \| \xi_k \|_X \]

where we have made use of Proposition 3.81. Then, \( \forall \alpha, \beta \in \mathbb{I} \), we have \( \| \alpha \xi + \beta \eta \|_p \leq \| \alpha \|_p \| \xi \|_p + \| \beta \|_p \| \eta \|_p \). Hence, \( M_p := (m_p, \oplus, \otimes, \vartheta) \) is a subspace in \( (Y, \oplus, \otimes, \vartheta) \).

7.2 The Natural Metric

\[ \| \alpha x \|_p = \| \beta y \|_p = | \alpha | \| x \|_p + | \beta | \| y \|_p \]

Hence, \( \mathcal{M}_p := (m_p, \oplus, \otimes, \vartheta) \) is a real (complex) normed linear space, \( \forall \varphi \in [1, \infty] \in \mathbb{R} \). ⊤

Example 7.11 Let \( X := \{ f : [a, b] \rightarrow \mathbb{R} \} \), where \( a, b \in \mathbb{R} \) with \( a < b \), \( \oplus \) and \( \otimes \) be the usual vector addition and scalar multiplication, and \( \vartheta : [a, b] \rightarrow \mathbb{R} \) be given by \( \vartheta(t) = 0 \), \( \forall t \in [a, b] \). By Example 6.20, \( X := (X, \oplus, \otimes, \vartheta) \) is a vector space over \( \mathbb{R} \). Let \( M := \{ f \in X \mid f \text{ is continuous} \} \) be a subspace. Then, \( (X, \oplus, \otimes, \vartheta) \) is a vector space. Introduce a norm on this space by \( \| x \| = \int_a^b | x(t) | \, dt \), \( \forall x \in M \). Now, we verify the properties of the norm. \( \forall x, y \in M, \forall \alpha, \beta \in \mathbb{R} \).

(i) Since \( x \) is continuous, then, \( | x(t) | \) is continuous and therefore integrable on \([a, b]\). Hence, \( 0 \leq \| x \| < +\infty \). \( x = \vartheta \Rightarrow \| x \| = 0 \).

(ii) \( \| x + y \| = \int_a^b | x(t) + y(t) | \, dt \leq \int_a^b (| x(t) | + | y(t) |) \, dt = \| x \| + \| y \|. \)

(iii) \( \| \alpha x \| = \int_a^b | \alpha x(t) | \, dt = | \alpha | \| x \|. \)

Hence, \( (M, \mathbb{R}, \| \cdot \|) \) is a normed linear space. ⊤

Definition 7.12 Let \( X \) be a set, \( \mathcal{Y} \) be a normed linear space, \( f : X \rightarrow \mathcal{Y} \). \( f \) is said to be bounded if \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| f(x) \| \leq M \), \( \forall x \in X \). \( S \subseteq \mathcal{Y} \) 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 \( X := (X, \mathbb{I}, \| \cdot \|) \) be a normed linear space and \( M \subseteq X \) be a subspace. Then, \( (M, \mathbb{I}, \| \cdot \|) \) 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) \( \rho(x, y) = \| x - y \| = 0 \iff x - y = \theta \iff x = y \); (iii) \( \rho(x, y) = \| x - y \| = \| (-1)(x - y) \| = \| y - x \| = \rho(y, x) \); and (iv) Note that

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

Hence, \( \rho \) is a metric. This completes the proof of the proposition. \( \square \)

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 \( C^0 \) are convex.

Proof \( \forall x_1, x_2 \in \overline{C}, \forall x \in [0, 1] \subset \mathbb{R} \), we need to show that \( \alpha x_1 + (1 - \alpha)x_2 \in \overline{C} \). By Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R}, B(x_1, r) \cap C \neq \emptyset, i = 1, 2 \). Let \( p_i \in B(x_i, r) \cap C, i = 1, 2 \). Then, by convexity of \( C \), we have \( \alpha p_1 + (1 - \alpha)p_2 \in C \). Note that

\[
\| (\alpha x_1 + (1 - \alpha)x_2) - (\alpha p_1 + (1 - \alpha)p_2) \|
= \| \alpha(x_1 - p_1) + (1 - \alpha)(x_2 - p_2) \|
\leq \alpha \| x_1 - p_1 \| + (1 - \alpha) \| x_2 - p_2 \| < r
\]

Then, \( \alpha p_1 + (1 - \alpha)p_2 \in B(\alpha x_1 + (1 - \alpha)x_2, r) \cap C \neq \emptyset \). 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.

\( \forall x_1, x_2 \in C^0, \forall x \in [0, 1] \subset \mathbb{R} \), we need to show that \( \alpha x_1 + (1 - \alpha)x_2 \in C^0 \). \( \exists r_i \in (0, \infty) \subset \mathbb{R} \) such that \( B(x_i, r_i) \subseteq C, i = 1, 2 \). Let \( r := \min\{r_1, r_2\} > 0 \). \( \forall r \in B(\alpha x_1 + (1 - \alpha)x_2, r) \), let \( w := p - \alpha x_1 - (1 - \alpha)x_2 \). Then, \( \| w \| < r \) and \( x_i + w \in B(x_i, r_i) \subseteq C, i = 1, 2 \). By the convexity of \( C \), we have \( C \ni \alpha(x_1 + w) + (1 - \alpha)(x_2 + w) = p \). Hence, we have \( B(\alpha x_1 + (1 - \alpha)x_2, r) \subseteq C \) and \( \alpha x_1 + (1 - \alpha)x_2 \in C^0 \). Hence, \( C^0 \) is convex.

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

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, \( \overline{P} = x_0 + S \) and \( P^0 = x_0 + S^0 \).

Proof \( \forall x \in \overline{P} \), by Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R}, \exists p_0 \in P \cap B(x, r) \). Then, \( x_0 - x_0 \in S \cap B(x - x_0, r) \neq \emptyset \). Hence, by Proposition 3.3, \( x - x_0 \in S \) and \( x \in x_0 + S \). Hence, \( \overline{P} \subseteq x_0 + S \).

On the other hand, \( \forall x \in x_0 + S \), we have \( x - x_0 \in S \) and, by Proposition 3.3, \( \forall r \in (0, \infty) \subset \mathbb{R}, \exists s_0 \in S \cap B(x - x_0, r) \neq \emptyset \). Then, \( x_0 + s_0 \in P \cap B(x, r) \neq \emptyset \). By Proposition 3.3, \( x \in \overline{P} \). Hence, \( x_0 + S \subseteq \overline{P} \).
Therefore, we have $\overline{P} = x_0 + \overline{S}$.

\[ \forall x \in P^o, \exists r \in (0, \infty) \subset \mathbb{R} \text{ such that } B(x, r) \subseteq P. \] Then, $B(x-x_0, r) = B(x, r) - x_0 \subseteq P - x_0 = S$. Hence, $x - x_0 \in S^o$, $x \in x_0 + S^o$ 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, $B(x, r) = B(x-x_0, r) + x_0 \subseteq x_0 + S = P$. Therefore, $x \in P^o$ and $x_0 + S^o \subseteq P^o$.

Hence, we have $P^o = x_0 + S^o$. This completes the proof of the proposition. \hfill $\Box$

**Proposition 7.17** Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$. Then, the following statements hold.

(i) If $M \subseteq \mathcal{X}$ is a subspace, then $\overline{M}$ is a subspace.

(ii) If $V \subseteq \mathcal{X}$ is a linear variety, then $\overline{V}$ is a linear variety.

(iii) If $C \subseteq \mathcal{X}$ is a cone with vertex $p \in \mathcal{X}$, then $\overline{C}$ is a cone with vertex $p$.

(iv) Let $x_0 \in \mathcal{X}$ and $r \in (0, \infty) \subset \mathbb{R}$. Then, $\overline{B}(x_0, r) = B(x_0, r)$.

**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 $x + y \in \overline{M}$.

(ii) Note that $V = x_0 + M$, where $x_0 \in \mathcal{X}$ and $M \subseteq \mathcal{X}$ 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, $ax \in \overline{C}$. Hence, $\overline{C}$ is a cone with vertex at the origin.

For general $p \in \mathcal{X}$, 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, $B(\emptyset, r)$ is a closed set and $\overline{B}(\emptyset, r) \supseteq B(\emptyset, r)$. Then, we have $B(\emptyset, r) \subseteq \overline{B}(\emptyset, r)$. On the other hand, $\forall x \in \overline{B}(\emptyset, r)$, define $(x_k)_{k=1}^\infty \subseteq \mathcal{X}$ by $x_k := \frac{1}{k} x, \forall k \in \mathbb{N}$. Then, $\|x_k\| = \frac{1}{k} \|x\| \leq \frac{1}{k} r < r, \forall k \in \mathbb{N}$. Therefore, $(x_k)_{k=1}^\infty \subseteq B(\emptyset, r)$. It is obvious that $\lim_{k \in \mathbb{N}} x_k = x$, 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 \( V(P) \), is the intersection of all closed linear varieties containing \( P \). Then, \( V(P) = \overline{v(P)} \).

**Proof** By Proposition 7.17, \( \overline{v(P)} \) is a closed linear variety containing \( P \). Then, we have \( V(P) \subseteq \overline{v(P)} \). On the other hand, \( V(P) \) is a closed linear variety containing \( P \), and then \( \overline{v(P)} \subseteq V(P) \). Hence, \( \overline{v(P)} \subseteq V(P) \). Therefore, the result holds. This completes the proof of the proposition. \( \square \)

The justification of the definition of \( V(P) \) 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 \( V(P) \). The set of all relative interior points of \( P \) is called the relative interior of \( P \), and denoted by \( ^\circ P \). \( P \) is said to be relatively open if \( P \) is open in the subset topology of \( V(P) \).

**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, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \alpha_0 \in A, \forall \alpha \in A \) 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 \( \| \cdot \| : X \times Y \rightarrow [0, \infty) \subset \mathbb{R} \) 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**

\[ \forall (x_1, y_1), (x_2, y_2) \in X \times Y, \forall \alpha \in K, \text{ we have } \| (x_1, y_1) \| = (\| x_1 \|_X^2 + \| y_1 \|_Y^2)^{1/2} \in [0, \infty) \subset \mathbb{R} \text{ and } \| (x_1, y_1) \| = 0 \iff \| x_1 \|_X = 0 \text{ and } \| y_1 \|_Y = 0 \iff x_1 = 0 \text{ and } y_1 = 0 \iff (x_1, y_1) = (\emptyset, \emptyset); \| (x_1, y_1) + (x_2, y_2) \| = \| (x_1 + x_2, y_1 + y_2) \| = (\| x_1 + x_2 \|_X^2 + \| y_1 + y_2 \|_Y^2)^{1/2} \leq (\| x_1 \|_X + \| x_2 \|_X)^2 + (\| y_1 \|_Y + \| y_2 \|_Y)^2 \].
\[(\|y_1\|_y + \|y_2\|_y)^2 \leq (\|x_1\|_X^2 + \|y_1\|_y^2)^{1/2} + (\|x_2\|_X^2 + \|y_2\|_y^2)^{1/2} = \|(x_1, y_1)\| + \|(x_2, y_2)\|,\] 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; \(\|\alpha(x_1, y_1)\| = (\|\alpha x_1\|_X^2 + \|\alpha y_1\|_Y^2)^{1/2} = (\|\alpha^2\|_X^2 + \|\alpha y_1\|_Y^2)^{1/2} = (\|\alpha\|_Y)^2 (\|x_1\|_X^2 + \|y_1\|_Y^2)^{1/2} = (\|\alpha\|_Y ) \|(x_1, y_1)\|\). Hence, \(\|\cdot\|\) is a norm on \(X \times Y\). This completes the proof of the proposition. \(\Box\)

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 \(\prod_{\alpha \in \emptyset} X_\alpha = \{\emptyset =: \emptyset\}, K, \|\cdot\|\), where \(\|\cdot\| = 0\).

**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
\[
\|(x_1, x_2) - (y_1, y_2)\|_{X \times X} < \epsilon/\sqrt{2}
\]
we have
\[
\|(x_1 + x_2) - (y_1 + y_2)\| \leq \|x_1 - y_1\| + \|x_2 - y_2\|
\leq \sqrt{2} (\|x_1 - y_1\|^2 + \|x_2 - y_2\|^2)^{1/2} = \sqrt{2} \|(x_1, x_2) - (y_1, y_2)\|_{X \times X} < \epsilon
\]
Hence, the vector addition \(\oplus_X\) is uniformly continuous.

\(\forall a_0 \in K, \forall x_0 \in X, \forall \epsilon \in (0, \infty) \subset \mathbb{R},\) let \(\delta = \epsilon/(1 + \epsilon + |a_0| + \|x_0\|) \in (0, 1) \subset \mathbb{R}\). \(\forall (a, x) \in B_{K \times X} ((a_0, x_0), \delta)\), we have
\[
\|ax - a_0x_0\| \leq \|ax - a_0x\| + \|a_0x - a_0x_0\|
= |a| \|x - x_0\| + \|x_0\| |a - a_0|
\leq \left(|a|^2 + \|x_0\|^2\right)^{1/2} \left(\|x - x_0\|^2 + |a - a_0|^2\right)^{1/2}
\leq (|a| + \|x_0\|) \delta \leq (1 + |a_0| + \|x_0\|) \delta < \epsilon
\]
where we have made use of Cauchy-Schwarz Inequality in the second inequality. Hence, \(\otimes_X\) is continuous at \((a_0, x_0)\). By the arbitrariness of \((a_0, x_0)\), we have that \(\otimes_X\) is continuous.

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

**Definition 7.24** Let \(X := (X, K, \|\cdot\|)\) and \(Y := (Y, K, \|\cdot\|)\) 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 $A : X \rightarrow Y$ 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 $K$, 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. \PassageSpacing $\forall \beta \in \Gamma$, 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 $E : \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \rightarrow \prod_{\alpha \in \Lambda} X_\alpha$ by, $\forall x \in X^{(\Gamma)}$, $\forall \alpha \in \Lambda$, $\exists! \beta_\alpha \in \Gamma \ni \alpha \in \Lambda_{\beta_\alpha}$, $\pi_\alpha(E(x)) = \pi_{\alpha}^{(\beta_\alpha)}(\pi^{(\Gamma)}_{\beta_\alpha}(x))$. 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. \PassageSpacing

### 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 $X$, $\sum_{n=1}^{\infty} x_n = \lim_{n \in \mathbb{N}} s_n \in X$.

"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 $\|x_n - x_{n_k}\| < 2^{-k}$, $\forall n, m \geq n_k$. 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 $\lim_{n \in \mathbb{N}} x_{n_k} = x_0 \in X$.

We will now show that $\lim_{n \in \mathbb{N}} x_n = x_0$. $\forall k \in \mathbb{N}$, $\exists N \in \mathbb{N}$ with $N \geq k$ such that $\|x_n| - x_0\| < 2^{-k}$, $\forall i \geq N$. $\forall n \geq n_N$, 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, $\lim_{n \in \mathbb{N}} x_n = x_0$. 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 $(\mathcal{M}, \mathbb{R}, \| \cdot \|)$ is incomplete. Take $(x_n)_{n=1}^{\infty} \subseteq \mathcal{M}$ 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 $x \in \mathcal{M}$ 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 \(x : [a, b] \to R\). 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 \(\lim_{n \to \infty} \|x_n - x\| = 0\). Hence, \(\lim_{n \in \mathbb{N}} x_n = x\). Hence, \(C([a, b])\) is a Banach space.

**Example 7.31** Let \(\mathcal{K}\) be a countably compact topological space, \(\mathcal{X}\) be a normed linear space over the field \(K\), \(Y := \{f : \mathcal{K} \to \mathcal{X}\}\). Define the usual vector addition \(\oplus\) and scalar multiplication \(\otimes\) and null vector \(\vartheta\) on \(Y\) as in Example 6.20. Then, \(\mathcal{Y} := (Y, \oplus, \otimes, \vartheta)\) 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, \(\mathcal{M} := (M, \oplus, \otimes, \vartheta)\) is a subspace of \((\mathcal{Y}, K)\). Define a function \(\|\| : M \to [0, \infty) \subset \mathbb{R}\) by \(\|f\| = \max \{\sup_{k \in \mathcal{K}} \|f(k)\|_\mathcal{X}, 0\}, \forall f \in M\). This function is well defined by Propositions 7.21, 3.12, and 5.29. We will show that \(\|\|\) defines a norm on \(\mathcal{M}\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\mathcal{K} = \emptyset\); Case 2: \(\mathcal{K} \neq \emptyset\). Case 1: \(\mathcal{K} = \emptyset\). Then, \(\mathcal{M}\) is a singleton set \(\{\emptyset\}\) and \(\|\emptyset\| = 0\). Clearly, \((\mathcal{M}, K, \|\|)\) is a normed linear space. Case 2: \(\mathcal{K} \neq \emptyset\). \(\forall f, g \in \mathcal{M}, \forall \alpha \in K\), \(\|f\| = \max_{k \in \mathcal{K}} \|f(k)\|_\mathcal{X}\) by Propositions 7.21, 3.12, and 5.29. \(\|f\| = 0 \iff \|f(k)\|_\mathcal{X} = 0, \forall k \in \mathcal{K} \iff f(k) = \vartheta_\mathcal{X}, \forall k \in \mathcal{K} \iff f = \vartheta\). \(\|f + g\| = \max_{k \in \mathcal{K}} \|f(k) + g(k)\|_\mathcal{X} \leq \max_{k \in \mathcal{K}} (\|f(k)\|_\mathcal{X} + \|g(k)\|_\mathcal{X}) \leq \max_{k \in \mathcal{K}} \|f(k)\|_\mathcal{X} + \max_{k \in \mathcal{K}} \|g(k)\|_\mathcal{X} = \|f\| + \|g\|\). \(\|\alpha f\| = \max_{k \in \mathcal{K}} |\alpha f(k)|_\mathcal{X} = \max_{k \in \mathcal{K}} |\alpha| \|f(k)\|_\mathcal{X} = |\alpha| \|f\|\). Hence, \((\mathcal{M}, K, \|\|)\) is a normed linear space. In both cases, we have shown that \(C(\mathcal{K}, \mathcal{X}) := (\mathcal{M}, K, \|\|)\) is a normed linear space.

**Example 7.32** Let \(\mathcal{K}\) be a countably compact topological space, \(\mathcal{X}\) be a Banach space over the field \(K\) (with norm \(\|\|_\mathcal{X}\)). Consider the normed linear space \(C(\mathcal{K}, \mathcal{X})\) (with norm \(\|\|\)) 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: \(\mathcal{K} = \emptyset\); Case 2: \(\mathcal{K} \neq \emptyset\). Case 1: \(\mathcal{K} = \emptyset\). Then, \(C(\mathcal{K}, \mathcal{X})\) is a singleton set. Hence, any Cauchy sequence must converge. Thus, \(C(\mathcal{K}, \mathcal{X})\) is a Banach space. Case 2: \(\mathcal{K} \neq \emptyset\). Take a Cauchy sequence \((x_n)_{n=1}^{\infty} \subseteq C(\mathcal{K}, \mathcal{X})\). \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N}\) such that \(\forall n, m \geq N, 0 \leq \|x_n(k) - x_m(k)\|_\mathcal{X} \leq \|x_n - x_m\| < \epsilon, \forall k \in \mathcal{K}\). This shows that, \(\forall k \in \mathcal{K}\), \((x_n(k))_{n=1}^{\infty} \subseteq \mathcal{X}\) is a Cauchy sequence, which converges to \(x(k) \in \mathcal{X}\) since \(\mathcal{X}\) is complete. This defines a function \(x : \mathcal{K} \to \mathcal{X}\). It is easy to show that \((x_n)_{n=1}^{\infty}\), viewed as a sequence of functions of \(\mathcal{K}\) to \(\mathcal{X}\), converges uniformly to \(x\). By Proposition 4.26, \(x\) is continuous. Hence, \(x \in C(\mathcal{K}, \mathcal{X})\). It is easy to see that \(\lim_{n \in \mathbb{N}} \|x_n - x\| = 0\). Hence, \(\lim_{n \in \mathbb{N}} x_n = x\). Hence, \(C(\mathcal{K}, \mathcal{X})\) is a Banach space. In both cases, we have shown that \(C(\mathcal{K}, \mathcal{X})\) is a Banach space when \(\mathcal{X}\) is a Banach space.

**Example 7.33** Let \(\mathcal{X}\) be a Banach space over the field \(K\). Consider the normed linear space \(l_p(\mathcal{X})\) (with norm \(\|\cdot\|_p\)) defined in Example 7.10,
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 \( x_n = (\xi_{n,k})_{k=1}^\infty \subseteq X, \forall n \in \mathbb{N} \). \( \forall \xi \in (0, \infty) \subset \mathbb{R}, \exists N_\xi \in \mathbb{N} \) such that \( \forall n, m \geq N_\xi \), 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 \( \lim_{n \to \infty} x_n = x \). Since \((x_n)_{n=1}^\infty \) is a Cauchy sequence, then it is bounded, that is \( \exists M \in [0, \infty) \subset \mathbb{R} \) 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 \( \sum_{i=1}^k \|\xi_{n,i}\|_X^p = \lim_{n \to \infty} \sum_{i=1}^k \|\xi_{n,i}\|_X^p \leq M^p \). Hence, we have \( \sum_{i=1}^\infty \|\xi_{n,i}\|_X^p \leq M^p \) and \( \|x\|_p = \left( \sum_{i=1}^\infty \|\xi_{n,i}\|_X^p \right)^{1/p} \leq M \). Hence, we have \( x \in l_p(X) \). \( \forall \xi \in (0, \infty) \subset \mathbb{R}, \forall n, m \in \mathbb{N} \) with \( n, m \geq N_\xi \), we have \( \|x_n - x_m\|_p < \epsilon \). Then, \( \forall k \in \mathbb{N}, \sum_{i=1}^k \|\xi_{n,i} - \xi_{m,i}\|_X^p < \epsilon^p \). Taking limit as \( m \to \infty \), by Propositions 7.21, 3.66, and 7.23, we have \( \sum_{i=1}^\infty \|\xi_{n,i} - \xi_i\|_X^p = \lim_{n \to \infty} \sum_{i=1}^\infty \|\xi_{n,i} - \xi_i\|_X^p \leq \epsilon^p, \forall k \in \mathbb{N} \). Hence, we have \( \|x_n - x\|_p = \left( \sum_{i=1}^\infty \|\xi_{n,i} - \xi_i\|_X^p \right)^{1/p} \leq \epsilon \). This shows that \( \lim_{n \to \infty} x_n = x \). 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 \( x_n = (\xi_{n,k})_{k=1}^\infty \subseteq X, \forall n \in \mathbb{N} \). \( \forall \xi \in (0, \infty) \subset \mathbb{R}, \exists N_\xi \in \mathbb{N} \) such that \( \forall n, m \geq N_\xi \), we have \( \|x_n - x_m\|_{\infty} < \epsilon \). \( \forall k \in \mathbb{N}, \|\xi_{n,k} - \xi_{m,k}\|_X \leq \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_{m,i}\|_X = \|x_n - x_m\|_{\infty} < \epsilon \). 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 \( x \in l_{\infty}(X) \) and \( \lim_{n \to \infty} x_n = x \). Since \((x_n)_{n=1}^\infty \) is a Cauchy sequence, then it is bounded, that is \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \|x_n\|_X \leq M, \forall n \in \mathbb{N} \). Then, we have \( \|\xi_n\|_X = \lim_{n \to \infty} \|\xi_n\|_X \leq M \), \( \forall k \in \mathbb{N} \). By Propositions 7.21 and 3.66, we have \( \|\xi_n\|_X = \lim_{n \to \infty} \|\xi_n\|_X \leq M \). Hence, \( \|x\|_X = \sup_{i \in \mathbb{N}} \|\xi_i\|_X \leq M \). Therefore, \( x \in l_{\infty}(X) \). \( \forall \xi \in (0, \infty) \subset \mathbb{R}, \forall n, m \in \mathbb{N} \) with \( n, m \geq N_\xi \), 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 \( \|\xi_{n,k} - \xi_k\|_X = \lim_{n \to \infty} \|\xi_{n,k} - \xi_{m,k}\|_X \leq \epsilon, \forall k \in \mathbb{N} \). Hence, we have \( \|x_n - x\|_{\infty} = \sup_{i \in \mathbb{N}} \|\xi_{n,i} - \xi_i\|_X \leq \epsilon \). This shows that \( \lim_{n \to \infty} x_n = x \). 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 $D := \{\sum_{i=1}^{n} \alpha_i y_i \mid n \in \mathbb{Z}_+, \alpha_i \in K_Q, y_i \in D, i = 1, \ldots, n\}$, which is a countable set. Clearly, $D \subseteq \text{span}(S_1)$ 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. □

**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 $\tilde{n} \in \mathbb{Z}_+$ be the dimension of $M$, which is well defined by Theorem 6.51. We will prove the theorem by mathematical induction on $\tilde{n}$.

1° $\tilde{n} = 0$. Then, $M = \{\emptyset\}$. Clearly, any Cauchy sequence in $M$ must converge to $\emptyset \in M$. Hence, $M$ is complete. $\tilde{n} = 1$. Let $\{e_1\} \subseteq M$ be a basis for $M$. Clearly, $e_1 \neq \emptyset$ 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 $\forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N}, \forall n, m \geq N, \text{we have } \|x_n - x_m\| < \varepsilon$. 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 $\tilde{n} = k - 1 \in \mathbb{N}$.

3° Consider the case $\tilde{n} = k \in \{2, 3, \ldots\} \subseteq \mathbb{N}$. 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 $\delta_i := \text{dist}(e_i, M_i)$, $i = 1, \ldots, k$. $\forall i = 1, \ldots, k$, we have $\delta_i \in [0, \infty) \subseteq \mathbb{R}$. $M_i$, 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. $\forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N}$ such that $\forall n, m \geq N, \text{we have } \|x_n - x_m\| < \varepsilon$. $\forall i \in \{1, \ldots, k\}$, we have $\varepsilon > \|x_n - x_m\| / \delta = \left(\sum_{j=1}^{k}(\lambda_{n,j} - \lambda_{m,j})e_j\right) / \delta > \sum_{j=1}^{k}(\lambda_{n,j} - \lambda_{m,j}) \geq |\lambda_{n,i} - \lambda_{m,i}|$. 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 $x := \sum_{i=1}^{k} \lambda_i e_i \in M$.
7.5. COMPACTNESS

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. \( \square \)

**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: \( n = 0 \). 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 \( K = \max \{ \delta, \delta^{-1} \} \in (0, \infty) \subset \mathbb{R} \). Then, \( \| x \|_1 \leq K \| x \|_2 \), \( \forall x \in \mathcal{X} \), \( \exists \alpha \in \mathbb{K} \) such that \( x = \alpha e_1 \). Hence, the two norms are equivalent.

Case 2: \( n = 1 \). 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.

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 6.39, \( M_i \) is closed (with respect to \( \| \cdot \|_i \)). By Proposition 4.39, \( M_i \) is closed (with respect to \( \| \cdot \|_i \)). 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, \( \| x \|_1 \leq \sum_{i=1}^n | \alpha_i | \| e_i \|_1 \leq \delta \left( \sum_{i=1}^n | \alpha_i | \right) \). 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 \( \| \cdot \|_i \)).

We have \( \| x \|_1 = \sum_{i=1}^n | \alpha_i | \| e_i \|_1 \geq | \alpha_i | \delta_i \geq | \alpha_i | \delta \), \( \forall i \in \{ 1, \ldots, n \} \). Hence, we have \( \| x \|_1 \geq (\delta / n) \left( \sum_{i=1}^n | \alpha_i | \right) \). Let \( K_1 = \max \{ \delta, \delta/n \} \). 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. \( \square \)

**7.5 Compactness**

**Definition 7.39** Let \( \mathcal{X} \) be a normed linear space and \( S \subseteq \mathcal{X} \). \( S \) 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) \(\vartheta_x + x = x\); (iv) \(\alpha (x + y) = \alpha x + \alpha y\); (v) \((\alpha + \beta)x = \alpha x + \beta x\); (vi) \((\alpha \beta)x = \alpha (\beta x)\); (vii) \(0 \cdot x = \vartheta_{\mathcal{X}}\) 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 \(\| x \| = 0 \iff x = \vartheta_{\mathcal{X}}\); (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, \(\text{id}_\mathcal{X} : \mathcal{X} \to \mathcal{X}_\mathbb{R}\) 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 and \(\mathcal{X}_\mathbb{R}\) 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, \(\mathcal{X} := (\mathbb{C}^n, \mathbb{R}, \| \cdot \|)\) 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 $M \subseteq \mathcal{X}$ be a subspace of a vector space $\mathcal{X}$ over a field $\mathcal{F} := (F, +, \cdot, 0, 1)$. $x_1, x_2 \in \mathcal{X}$ are said to be equivalent modulo $M$ if $x_1 - x_2 \in M$. This equivalence relationship (easy to show) partitions the space $\mathcal{X}$ 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 $M$. $\forall x \in \mathcal{X}$, there is a unique coset of $M$, $[x] := \{x + \alpha \mid \alpha \in M\}$, such that $x \in [x]$. The quotient of $\mathcal{X}$ modulo $M$ is defined to be the set $N$ of all cosets of $M$. Define vector addition and scalar multiplication on $N$ by, $\forall [x_1], [x_2] \in N$, $\exists \alpha \in \mathcal{F}$, $[x_1] + [x_2] := [x_1 + x_2]$ and $\alpha [x_1] := [\alpha x_1]$. Let the null vector on $N$ be $[\emptyset \mathcal{X}]$. Then, $N$ together with the vector addition, scalar multiplication, and the null vector form a vector space over $\mathcal{F}$. This vector space will be called the quotient space of $\mathcal{X}$ modulo $M$ and denoted by $\mathcal{X}/M$. Define a function $\phi : \mathcal{X} \to \mathcal{X}/M$ by $\phi(x) = [x]$, $\forall x \in \mathcal{X}$. Then, $\phi$ is a linear function and will be called the natural homomorphism.

Proof Fix any $[x_1], [x_2], [x_3] \in \mathcal{X}/M$ and any $\alpha, \beta \in \mathcal{F}$. We will first show that vector addition and scalar multiplication are uniquely defined. $\forall x_1 \in [x_1], \forall x_2 \in [x_2]$, we have $x_1 - y_1, x_2 - y_2 \in M$; and $(x_1 + x_2) - (y_1 + y_2) \in M$, since $M$ is a subspace. Then, $[x_1] + [x_2] = [x_1 + x_2] = [y_1 + y_2] = [y_1] + [y_2]$. Hence, the vector addition is uniquely defined. $\alpha x_1 - \alpha y_1 = \alpha (x_1 - y_1) \in M$. Then, $\alpha [x_1] = [\alpha x_1] = [\alpha y_1] = \alpha [y_1]$. Hence, the scalar multiplication is uniquely defined.

(i) $[x_1] + [x_2] = [x_1 + x_2] = [x_2 + x_1] = [x_2] + [x_1]$; (ii) $([x_1] + [x_2] + [x_3]) + [x_3] = ([x_1 + x_2 + x_3]) = [x_1] + ([x_2] + [x_3])$; (iii) $[x_1] + [\emptyset \mathcal{X}] = [x_1 + \emptyset \mathcal{X}] = [x_1]$; (iv) $\alpha ([x_1] + [x_2]) = [\alpha x_1 + \alpha x_2] = [\alpha x_2] + [\alpha x_1] = [\alpha x_1 + \alpha x_2]$; (v) $\alpha (\alpha + \beta) [x_1] = [\alpha x_1 + \beta x_1] = [\alpha x_1] + [\beta x_1] = [\alpha (\beta x_1) + \alpha x_1] = [\alpha (\beta x_1)] = \alpha (\beta x_1)$; (vi) $0 [x_1] = [0 x_1] = [\emptyset \mathcal{X}]$; 1 $[x_1] = [x_1]$. Hence, $\mathcal{X}/M$ is a vector space over $\mathcal{F}$. 

$n \in \mathbb{N}$. Case 1: $n = 0$. Then, $\mathcal{X}$ is a singleton set. Clearly $K$ is compact. Case 2: $n \in \mathbb{N}$. Let $S = \{e_1, \ldots, e_n\}$. $\forall x \in \mathcal{X}$, $x$ can be uniquely expressed as $\sum_{i=1}^{n} \alpha_i e_i$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$. This allows us to define a bijective mapping $\psi : \mathcal{X} \to \mathbb{K}^n$ by $\psi(x) = (\alpha_1, \ldots, \alpha_n)$, $\forall x = \sum_{i=1}^{n} \alpha_i e_i \in \mathcal{X}$. Define a alternative norm $\parallel \cdot \parallel_1$ on $\mathcal{X}$ by $\parallel x \parallel_1 = \parallel \sum_{i=1}^{n} \alpha_i e_i \parallel_1 = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}$, $\forall x \in \mathcal{X}$. It is easy to show that $\parallel \cdot \parallel_1$ is a norm. By Theorem 7.38, there exists $\xi \in [1, \infty) \subset \mathbb{R}$ such that $\parallel x \parallel_1 / \xi \leq \parallel x \parallel_1$, $\forall x \in \mathcal{X}$. Hence, $\psi$ is a homeomorphism. By Proposition 3.10, $\psi(K)$ is closed. By the equivalence of the two norms, $\psi(K)$ is bounded. By Proposition 5.40 or 7.41, $\psi(K)$ is compact. By Proposition 5.7, $K$ is compact. 

This completes the proof of the proposition. 

□
Clearly, $\phi$ is a linear function. This completes the proof of the proposition.

\begin{proposition}
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 $M$.

\begin{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 $\| \cdot \| : X/M \to [0, \infty) \subseteq \mathbb{R}$ 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) $\| [x_1] \| \in [0, \infty) \subseteq \mathbb{R}$, since $\| [x_1] \| \leq \| x_1 \|_X < +\infty$. (ii) $\| x_1 + x_2 \| = \inf_{m \in M} \| x_1 + x_2 - m \|_X$. Note that $\| [x_i] \| = \inf_{m \in M} \| x_i - m \|_X, i = 1, 2$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $i = 1, 2, \exists \delta_i \in M$ such that $\| x_i - \delta_i \|_X < \| [x_i] \| + \epsilon$. Then, we have $\| x_1 + x_2 - \delta_1 - \delta_2 \|_X \leq \| x_1 - \delta_1 \|_X + \| x_2 - \delta_2 \|_X \leq \| [x_1] \| + \| [x_2] \| + 2\epsilon$. 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) $\| \alpha [x_1] \| = \| [\alpha x_1] \| = \inf_{m \in M} \| \alpha x_1 - m \|_X$. 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: $\alpha \neq 0$. $\| \alpha [x_1] \| = \inf_{m \in M} \| \alpha x_1 - \alpha m \|_X = \inf_{m \in M} \| \alpha [x_1 - m] \|_X = \| [\alpha] \| [x_1] \|$. 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.

\end{proof}

\begin{proposition}
Let $X$ be a Banach space and $M \subseteq X$ be a closed subspace. Then, the quotient space $X/M$ is a Banach space.

\begin{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, $\sum_{n=1}^{\infty} \| [x_n] \| \in \mathbb{R}, \forall n \in \mathbb{N}, \exists y_n \in [x_n]$ such that $\| y_n \|_X < \| [x_n] \| + 2^{-n}$. Then, $\sum_{n=1}^{\infty} \| y_n \|_X < \sum_{n=1}^{\infty} (\| [x_n] \| + 2^{-n}) = \sum_{n=1}^{\infty} \| [x_n] \| + 1$. 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.

\end{proof}
Definition 7.46  Let \((\mathcal{X}, K)\) be a vector space and \(\| \cdot \| : \mathcal{X} \to [0, \infty) \subset \mathbb{R}\). Assume that \(\| \cdot \|\) satisfies (ii) and (iii) of Definition 7.1 and \(\| \vartheta \| = 0\), but not necessarily (i). Then, \(\| \cdot \|\) is called a pseudo-norm.

Proposition 7.47  Let \((\mathcal{X}, K)\) be a vector space and \(\| \cdot \| : \mathcal{X} \to [0, \infty) \subset \mathbb{R}\) be a pseudo-norm on \(\mathcal{X}\). Then, the set \(M := \{ x \in \mathcal{X} \mid \| x \| = 0 \}\) is a subspace of \((\mathcal{X}, K)\). On the quotient space \(\mathcal{X}/M\), define a norm \(\| \cdot \|_1 : \mathcal{X}/M \to [0, \infty) \subset \mathbb{R}\) by \(\| [x] \|_1 = \| x \|, \forall [x] \in \mathcal{X}/M\). Then, the space \((\mathcal{X}/M, K, \| \cdot \|_1)\) is a normed linear space, which will be called the quotient space of \((\mathcal{X}, K)\) modulo \(\| \cdot \|\).

Proof  We need the following claim.

Claim 7.47.1  \(\forall x \in \mathcal{X}, \forall m \in M, \text{we have } \| x \| = \| x + m \|\).

Proof of claim:  Note that \(\| x \| \leq \| x + m \| + \| -m \| = \| x + m \| \leq \| x \| + \| m \| = \| x \|\). This completes the proof of the claim. \(\Box\)

Clearly, \(\vartheta x \neq 0, \forall x \in M \neq \emptyset\). \(\forall m_1, m_2 \in M, \forall \alpha \in K\), 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 \((\mathcal{X}, K)\). By Proposition 7.43, \(\mathcal{X}/M\) is a vector space over \(K\).

By Claim 7.47.1, \(\| \cdot \|_1 : \mathcal{X}/M \to [0, \infty) \subset \mathbb{R}\) is uniquely defined.

Next, we show that \(\| \cdot \|_1\) is a norm on \(\mathcal{X}/M\). \(\forall [x_1], [x_2] \in \mathcal{X}/M, \forall \alpha \in K, \| [x_1] \|_1 = \| x_1 \| \in [0, \infty) \subset \mathbb{R}\). \([x_1] = [\vartheta x] \) implies that \(\| [x_1] \|_1 = \| [\vartheta x] \|_1 = \| \vartheta x \| = 0\). \(\| [x_1] \|_1 = 0\) implies that \(\| x_1 \| = 0\) and \(x_1 \in M\), which further implies that \(x_1 \in [\vartheta x]\) and \([x_1] = [\vartheta x]\). \(\| [x_1 + x_2] \|_1 = \| [x_1 + x_2] \| = \| x_1 + x_2 \| \leq \| x_1 \| + \| x_2 \| = \| [x_1] \|_1 + \| [x_2] \|_1\), \(\| \alpha [x_1] \|_1 = \| \alpha [x_1] \|_1 = \| \alpha x_1 \| = |\alpha| \| x_1 \| = |\alpha| \| [x_1] \|_1\). Hence, \(\| \cdot \|_1\) is a norm on \(\mathcal{X}/M\). Therefore, \((\mathcal{X}/M, K, \| \cdot \|_1)\) is a normed linear space. This completes the proof of the proposition. \(\Box\)

7.7 The Stone-Weierstrass Theorem

Definition 7.48  Let \(X\) be a set, \(f : X \to \mathbb{R}\), and \(g : X \to \mathbb{R}\). \(f \wedge g : X \to \mathbb{R}\) and \(f \vee g : X \to \mathbb{R}\) 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. \(\Box\)

Example 7.50  Let \(\mathcal{X}\) be a topological space and \(\mathcal{Y}\) be a normed linear space over the field \(K\). Let \((\mathcal{M}(\mathcal{X}, \mathcal{Y}), K)\) be the vector space defined in
Example 6.20. Let $V := \{ f \in \mathcal{M}(X, \mathbb{R}) \mid f \text{ is continuous} \}$. Then, by Propositions 3.12, 3.32, 6.25, and 7.23, $V$ is a subspace of $(\mathcal{M}(X, \mathbb{R}), \mathbb{R})$.

This subspace will be denoted by $C_v(X, \mathbb{R})$.

Definition 7.51 Let $X$ be a topological space, $C_v(X, \mathbb{R})$ be the vector space defined in Example 7.50. $L \subseteq C_v(X, \mathbb{R})$ is said to be a lattice if $L \not= \emptyset$ and $\forall f, g \in L$, $f \lor g \in L$ and $f \land g \in L$. A subspace $M \subseteq C_v(X, \mathbb{R})$ 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, $C(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq C(\mathcal{X}, \mathbb{R})$ be a lattice. Assume $h : \mathcal{X} \rightarrow \mathbb{R}$ defined by $h(x) = \inf_{f \in L} f(x)$, $\forall x \in \mathcal{X}$ is continuous. Then, $\forall 0 < \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists g \in L$ 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: $\mathcal{X} \not= \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $C(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = C(\mathcal{X}, \mathbb{R})$ since $L$ is a lattice and therefore nonempty. Then, the result holds. Case 2: $\mathcal{X} \not= \emptyset$. $\forall x \in (0, \infty) \subset \mathbb{R}$, $\exists x \in \mathcal{X}$, since $h(x) \in \mathbb{R}$, then $\exists f_x \in L$ such that $0 \leq f_x(x) - h(x) < \epsilon/3$. 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 $\mathcal{X} \not= \emptyset$ 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, $C(\mathcal{X}, \mathbb{R})$ be the Banach space as defined in Example 7.32, and $L \subseteq C(\mathcal{X}, \mathbb{R})$ be a lattice satisfying the following conditions:

(i) $L$ separates points, that is $\forall x, y \in \mathcal{X}$ with $x \not= y$, $\exists f \in L$ such that $f(x) \not= f(y)$;

(ii) $\forall f \in L$, $\forall c \in \mathbb{R}$, $c + f, cf \in L$.

Then, $\forall h \in C(\mathcal{X}, \mathbb{R})$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, 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: $\mathcal{X} \not= \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $C(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = C(\mathcal{X}, \mathbb{R})$ since $L$ is a lattice and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \not= \emptyset$. We need the following two results.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\mathcal{X} = \emptyset$; Case 2: $\mathcal{X} \not= \emptyset$. Case 1: $\mathcal{X} = \emptyset$. $C(\mathcal{X}, \mathbb{R})$ is a singleton set and $L = C(\mathcal{X}, \mathbb{R})$ since $L$ is a lattice and therefore nonempty. Then, the result holds.

Case 2: $\mathcal{X} \not= \emptyset$. We need the following two results.
Claim 7.53.1 \( \forall a, b \in \mathbb{R}, \forall x_1, x_2 \in X \text{ with } x_1 \neq x_2, \exists f \in L \text{ such that } f(x_1) = a \text{ and } f(x_2) = b. \)

**Proof of claim:** Let \( g \in L \) be such that \( g(x_1) \neq g(x_2) \), which exists by (i). Let
\[
f = \frac{a - b}{g(x_1) - g(x_2)}g + \frac{b g(x_1) - a g(x_2)}{g(x_1) - g(x_2)} \in L
\]
Then, \( f \) is the desired function.

Claim 7.53.2 \( \forall a, b \in \mathbb{R} \text{ with } a \leq b, \forall \text{ closed set } F \subseteq X, \forall x_0 \in X \text{ with } x_0 \notin F, \exists f \in L \text{ such that } f(x_0) = a, f(x) \geq a, \forall x \in X, \text{ and } f(x) > b, \forall x \in F. \)

**Proof of claim:** \( \forall x \in F, \) we have \( x \neq x_0. \) By Claim 7.53.1, \( \exists f_x \in L \) such that \( f_x(x_0) = a \) and \( f_x(x) = b + 1. \) Let \( O_x := \{ x \in X \mid f_x(x) > b \}. \) Then, \( O_x \in \mathcal{O} \) since \( f_x \) is continuous. Clearly, \( x \in O_x. \) \( F \subseteq \bigcup_{x \in F} O_x. \) By Proposition 5.5, \( F \) is compact. Then, there exists a finite set \( F_N \subseteq F \) such that \( F \subseteq \bigcup_{x \in F_N} O_x. \) Take \( g \in L \neq \emptyset. \) Then, \( a = 0 \) and \( a \in L \) by (ii). Let \( f := a \vee (\bigvee_{x \in F_N} f_x) \in L. \) Clearly, \( f(x_0) = a \) and \( f(x) \geq a, \forall x \in X. \) \( \forall x \in F, \exists x_0 \in F_N \) such that \( x \in O_{x_0}. \) Then, \( f(x) \geq f_{x_0}(x) > b. \) Hence, \( f \) is the desired function. This completes the proof of the claim. 

\( \forall h \in C(X, \mathbb{R}), \) let \( L := \{ f \in L \mid f(x) \geq h(x), \forall x \in X \}. \) By the continuity of \( h, \) the compactness of \( X, \) and Proposition 5.29, \( \exists b \in \mathbb{R} \) such that \( h(x) \leq b, \forall x \in X. \) Then, the constant function \( b \in L \neq \emptyset (b \in L \) by (iii)). \( \forall f_1, f_2 \in L, \) it is easy to show that \( f_1 \vee f_2 \in L \) and \( f_1 \wedge f_2 \in L. \) Hence, \( L \) is a lattice. We will show that \( h(x) = \inf_{f \in L} f(x), \forall x \in X. \) Then, the result follows from Proposition 7.52. \( \forall x_0 \in X, \forall \eta \in (0, \infty) \subseteq \mathbb{R}, \) let \( F_{x_0, \eta} := \{ x \in X \mid h(x) \geq h(x_0) + \eta \}. \) By the continuity of \( h \) and Proposition 3.10, \( F_{x_0, \eta} \) is closed. Clearly, \( x_0 \notin F_{x_0, \eta}. \) Now, by Claim 7.53.2, \( \exists f_{x_0, \eta} \in L \) such that \( f_{x_0, \eta}(x_0) = h(x_0) + \eta, f_{x_0, \eta}(x) \geq h(x_0) + \eta, \forall x \in X, \) and \( f_{x_0, \eta}(x) > b, \forall x \in F_{x_0, \eta}. \) It is clear that \( f_{x_0, \eta}(x) > h(x), \forall x \in X. \) Hence, \( f_{x_0, \eta} \in L. \) Then, \( \inf_{f \in L} f(x_0) \leq f_{x_0, \eta}(x_0) = h(x_0) + \eta. \) By the definition of \( L \) and the arbitrariness of \( \eta, \) we have \( h(x_0) \leq \inf_{f \in L} f(x_0) \leq h(x_0). \) Hence, we have \( h(x_0) = \inf_{f \in L} f(x_0). \)

This completes the proof of the proposition.

**Lemma 7.54** \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \) there exists a polynomial \( P(s) \) in one variable such that \( |P(s) - |s|| < \epsilon, \forall s \in [-1, 1] \subseteq \mathbb{R}. \)

**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 527.

**Lemma 7.55** Let \( X \) be a nonempty compact space and \( A \subseteq C(X, \mathbb{R}) \) be an algebra. Then, \( \overline{A} \) is an algebra.
Proof: Note that $\mathcal{A}$ is a subspace of $\mathcal{C}(\mathcal{X}, \mathbb{R})$ by Definition 7.51. Then, by Proposition 7.17, $\overline{\mathcal{A}}$ is a subspace of $\mathcal{C}(\mathcal{X}, \mathbb{R})$. \forall f, g \in \mathcal{A}, \forall \epsilon \in (0, 1) \subset \mathbb{R}, \exists \tilde{f}, \tilde{g} \in \mathcal{A}$ such that $\| f - \tilde{f} \| < \frac{\epsilon}{\alpha + \beta}$ and $\| g - \tilde{g} \| < \frac{\epsilon}{\alpha + \beta}$. Then, $\tilde{f} \tilde{g} \in \mathcal{A}$ since $\mathcal{A}$ is an algebra, and $\| \tilde{f} \tilde{g} - f \tilde{g} \| = \max_{x \in \mathcal{X}} | f(x)g(x) - \tilde{f}(x)\tilde{g}(x) | \leq \max_{x \in \mathcal{X}} ( | f(x)g(x) - f(x)\tilde{g}(x) | + | f(x)\tilde{g}(x) - \tilde{f}(x)\tilde{g}(x) | ) \leq \max_{x \in \mathcal{X}} | f(x) | | g(x) - \tilde{g}(x) | + \max_{x \in \mathcal{X}} | \tilde{g}(x) | | f(x) - \tilde{f}(x) | \leq \| f \| \| g - \tilde{g} \| + \| g \tilde{g} - g \| \| f - \tilde{f} \| < \frac{\epsilon}{\alpha + \beta} + \frac{\epsilon}{\alpha + \beta} < \epsilon$. Hence, $\tilde{f} \tilde{g} \in \overline{\mathcal{A}}$ by the arbitrariness of $\epsilon$. This shows that $\overline{\mathcal{A}}$ is an algebra. This completes the proof of the lemma. □

Theorem 7.56 (Stone-Weierstrass Theorem) Let $\mathcal{X} := (\mathcal{X}, \mathcal{O})$ be a compact space and $\mathcal{C}(\mathcal{X}, \mathbb{R})$ be the Banach space defined in Example 7.32. Assume that $\mathcal{A} \subseteq \mathcal{C}(\mathcal{X}, \mathbb{R})$ is an algebra and satisfies

(i) $\mathcal{A}$ separates points, that is, $\forall x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$, then $\exists f \in \mathcal{A}$ such that $f(x_1) \neq f(x_2)$;

(ii) $\mathcal{A}$ contains all constant functions.

Then, $\overline{\mathcal{A}} = \mathcal{C}(\mathcal{X}, \mathbb{R})$, 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, $\overline{\mathcal{A}}$ is an algebra. We need the following claim.

Claim 7.56.1 $\forall f \in \overline{\mathcal{A}}$, then $\| f \| \in \overline{\mathcal{A}}$.

Proof of claim: Fix $f \in \overline{\mathcal{A}}$. \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, $\exists \tilde{g} \in (0, \infty) \subset \mathbb{R}$ such that $\| g \| < N$. Then, $g/N \in \overline{\mathcal{A}}$ since $\mathcal{A}$ is an algebra. By Lemma 7.54, $\exists$ a polynomial $P(s)$ such that $| P(s) - g(s) \| < \frac{\epsilon}{\alpha + \beta}$, $\forall s \in [-1, 1] \subset \mathbb{R}$. Hence, we have $P\circ (g/N) \in \overline{\mathcal{A}}$ since $\mathcal{A}$ is an algebra. Furthermore, $| P(g(x)/N) - g(x) |/N \| < \frac{\epsilon}{\alpha + \beta}$, $\forall x \in \mathcal{X}$. Let $h = NP \circ (g/N) \in \overline{\mathcal{A}}$. Then, we have $\| h \| < \epsilon/2$. Note that $\| f - h \| \leq \| f - g \| + \| g \| - h \| \leq \| f - g \| + \| g \| - h \| < \epsilon$. Hence, $\| f \| \in \overline{\mathcal{A}}$, by the arbitrariness of $\epsilon$. This completes the proof of the claim. □

$\forall f, g \in \overline{\mathcal{A}}$, we have

\[
\begin{align*}
f \vee g &= \frac{1}{2}(f + g) + \frac{1}{2}| f - g | \in \overline{\mathcal{A}} \\
f \wedge g &= \frac{1}{2}(f + g) - \frac{1}{2}| f - g | \in \overline{\mathcal{A}}
\end{align*}
\]
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 \( h \in \mathcal{A} = \mathcal{A}, \) where the last equality follows from Proposition 3.3. Hence, we have \( \mathcal{A} = C(X, \mathbb{R}). \) This completes the proof of the theorem. \( \square \)

**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, \( \mathcal{A} = C(X, \mathbb{R}). \) This completes the proof of the corollary. \( \square \)

**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 \( \mathcal{M} = \{ g \in C(S_1, \mathbb{R}) \mid g(x) = \cos(nx), \forall x \in S_1 \text{ or } g(x) = \sin(nx), \forall x \in S_1, \text{ where } n \in \mathbb{Z}_+ \}. \) Let \( \mathcal{A} = \text{span}(\mathcal{M}) \subseteq C(S_1, \mathbb{R}). \) Then, \( f \in \mathcal{A}. \)

**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 \( \Phi \circ \Psi \in \mathcal{A}, \) 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, \( f \in \mathcal{A}, \) by the arbitrariness of \( \epsilon. \) This completes the proof of the corollary. \( \square \)
For ease of presentation below, we will define two functions: $\text{sqr} : \mathbb{R} \to \mathbb{R}$ by $\text{sqr}(x) = x^2$, $\forall x \in \mathbb{R}$, and $\sqrt{ } : [0, \infty) \to [0, \infty)$ by $\sqrt{x} = \sqrt{x}$, $\forall x \in [0, \infty)$.

**Corollary 7.59** Let $\mathcal{X} := (\mathcal{X}, \mathcal{O})$ 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) $\exists f_0 \in \mathcal{A}$ 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 $g_0 : \mathcal{X} \to [0, 1] \subseteq \mathbb{R}$.

By Corollary 7.57, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, there exists a polynomial $Q_\epsilon$ in one variable such that $|\sqrt{s} - Q_\epsilon(s)| < \epsilon$, $\forall s \in [0, 1] \subseteq \mathbb{R}$ and $Q_\epsilon(0) = 0$.

$\forall f : \mathcal{X} \to [0, 1] \subseteq \mathbb{R}$ with $f \in \overline{\mathcal{A}}$, $Q_\epsilon \circ f \in \overline{\mathcal{A}}$ since $\overline{\mathcal{A}}$ is an algebra and $Q_\epsilon(0) = 0$. 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 $g_0(x) \geq g_0(x_m) =: \gamma > 0$, $\forall x \in \mathcal{X}$, $\forall \epsilon \in (0, 1) \subseteq \mathbb{R}$, $\exists n_\epsilon \in \mathbb{N}$ such that $\left| 1 - \gamma^{2^{-n_\epsilon}} \right| < \epsilon$. Then, $\|1 - \text{sqr}^{n_\epsilon} \circ g_0\| < \epsilon$. 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.

**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 $\exists x_0 \in \mathcal{X}$ such that $\overline{\mathcal{A}} = \{ f \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid f(x_0) = 0 \}$.

**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: $1 \not\in \overline{\mathcal{A}}$. 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: $1 \not\in \overline{\mathcal{A}}$. 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.
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 \( X \) 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, \( g : X \to [-1, 1] \subset \mathbb{R} \), \( \forall x \in (0, \infty) \subset \mathbb{R} \), 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, \( |f| \in \mathcal{A} = \mathcal{A} \).

\[
\begin{align*}
    f \land g &= \frac{1}{2}(f + g) + \frac{1}{2}|f - g| \in \mathcal{A} \\
    f \lor g &= \frac{1}{2}(f + g) - \frac{1}{2}|f - g| \in \mathcal{A}
\end{align*}
\]

Hence, \( \mathcal{A} \) is a lattice. This completes the proof of the claim. \( \square \)

Claim 7.60.2 \( \exists x_0 \in X \) 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. \( \forall x \in X \), \( \exists f_x \in \mathcal{A} \) such that \( f_x(x) \neq 0 \). By Claim 7.60.1, \( |f_x| \in \mathcal{A} \) and \( |f_x(x)| > 0 \). Let \( O_x := \{ x \in X \mid |f_x(x)| > 0 \} \).

Since \( f_x \in \mathcal{A} \subseteq C(X, \mathbb{R}) \), then \( x \in O_x \in \mathcal{O}_X \). Hence, we have \( X \subseteq \bigcup_{x \in X} O_x \). By the compactness of \( X \), there exists a finite set \( X_N \subseteq X \) such that \( X \subseteq \bigcup_{x \in X_N} O_x \). Clearly, \( X_N \neq \emptyset \) since \( X \neq \emptyset \). Let \( f := \sum_{x \in X_N} |f_x| \in \mathcal{A} \). Hence, \( f(x) > 0 \), \( \forall x \in X \). 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 \( F \subseteq X \) with \( x_0 \notin F \), we have

\( \text{(i)} \) let \( A_F := \{ h \in C(F, \mathbb{R}) \mid \exists f \in \mathcal{A} \text{ such that } h = f|_F \} \), then \( \mathcal{A}_F = C(F, \mathbb{R}) \);

\( \text{(ii)} \) \( \forall \epsilon \in (0, 1) \subset \mathbb{R} \), \( \exists g \in \mathcal{A} \) such that \( g : X \to [0, 1] \subset \mathbb{R} \) and \( g(x) \geq 1 - \epsilon \), \( \forall x \in F \).

**Proof of claim:** Clearly, \( A_F \) is an algebra since \( \mathcal{A} \) is an algebra. Furthermore, \( A_F \) separate points on \( F \) since \( \mathcal{A} \) separates points on \( X \).

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 \( O_x := \{ x \in X \mid |f_x(x)| > 0 \} \).

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 \( \mathcal{A}_F = C(F, \mathbb{R}) \). Hence, (i) is true.
\[
\exists M \in (0, \infty) \subset \mathbb{R} \text{ such that } \|f\| \leq M. \text{ Then, } g_0 := f/M \in \mathcal{A} \text{ and } g_0 : \mathcal{X} \to [0, 1] \subset \mathbb{R}.
\]

By Corollary 7.57, \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), there exists a polynomial \(Q_\epsilon\) in one variable such that \(|\sqrt{s} - Q_\epsilon(s)| < \epsilon, \forall s \in [0, 1] \subset \mathbb{R}\), and \(Q_\epsilon(0) = 0\).

\[
\forall f : \mathcal{X} \to [0, 1] \subset \mathbb{R} \text{ with } \tilde{f} \in \mathcal{A}, Q_\epsilon \circ \tilde{f} \in \mathcal{A} \text{ since } \mathcal{A} \text{ is an algebra and } Q_\epsilon(0) = 0. \text{ Then, } \text{sqrt} \circ \tilde{f} \in \overline{\mathcal{A}} = \mathcal{A}. \text{ Recursively, we may conclude that } \text{sqrt}^n \circ g_0 \in \overline{\mathcal{A}}, \forall n \in \mathbb{N}.
\]

By the compactness of \(F\) and Proposition 5.29, \(\exists \gamma \in (0, 1) \subset \mathbb{R}\) such that \(g_0(x) \geq \gamma, \forall x \in F\). \(\forall \epsilon \in (0, 1) \subset \mathbb{R}\), \(\exists n_0 \in \mathbb{N}\) such that \(\gamma^{2^{-n_0}} \geq 1 - \epsilon\).

Then, \(g := \text{sqrt}^{n_0} \circ g_0 \in \overline{\mathcal{A}}, g : \mathcal{X} \to [0, 1] \subset \mathbb{R}\) and \(g(x) \geq 1 - \epsilon, \forall x \in F\).

Hence, (ii) is true.

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

Claim 7.60.4 \(\forall g \in \mathcal{C}(\mathcal{X}, \mathbb{R})\) with \(g(x_0) = 0\) and \(g(x) \geq 0, \forall x \in \mathcal{X}\), we have \(g \in \mathcal{A}\).

Proof of claim: \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), let \(O := \{x \in \mathcal{X} \mid g(x) < \epsilon/2\}\). Then, \(x_0 \in O \subset \mathcal{O}_{\mathcal{X}} \text{ since } g \text{ is continuous. Let } M := \|g\| \in [0, \infty) \subset \mathbb{R}.
\]

By Claim 7.60.3, \(\exists h_1 \in \mathcal{A}\) such that \(h_1 : \mathcal{X} \to [0, 1] \subset \mathbb{R}\) and \(h_1(x) \geq 2/3, \forall x \in \overline{O}\). Note that \(h_1(x_0) = 0\) by Claim 7.60.2. Let \(h_2 := 3Mh_1/2 \in \overline{\mathcal{A}}\). Then, we have \(0 \leq h_2(x) \leq 3M/2, \forall x \in \mathcal{X}\), and \(h_2(x) \geq M \geq g(x), \forall x \in \overline{O}\).

Let \(U := \{x \in O \mid h_2(x) < \epsilon/2\}\). Then, \(x_0 \in U \subset \mathcal{O}_{\mathcal{X}} \text{ since } O \in \mathcal{O}_{\mathcal{X}}\), \(h_2(x) = 0\) and \(h_3 \in \mathcal{C}(\overline{U}, \mathbb{R})\).

By Claim 7.60.3, \(\exists h_3 \in \mathcal{A}\) such that \(|h_3|_{\overline{U}}(x) - g|_{\overline{U}}(x)| < \epsilon, \forall x \in \overline{U}\). Define \(h_4 := (h_3 \lor 0) \land h_2\). Then, \(h_4 \in \mathcal{A}\) by Claim 7.60.1. \(\forall x \in \mathcal{X}\), we will show that \(|g(x) - h_4(x)| < \epsilon\) by distinguishing three exhaustive and mutually exclusive cases: Case A: \(x \in U\); Case B: \(x \in O \setminus U\); Case C: \(x \in \overline{O}\).

Case A: \(x \in U\). Then, \(0 \leq h_2(x) < \epsilon/2, 0 \leq g(x) < \epsilon/2\). Then, \(0 \leq h_4(x) = (h_3(x) \lor 0) \land h_2(x) \leq h_2(x) < \epsilon/2\). Therefore, \(|g(x) - h_4(x)| < \epsilon\).

Case B: \(x \in O \setminus U\). Then, \(0 \leq g(x) < \epsilon/2\). Hence, \(0 \leq h_4(x) \lor 0 < g(x) + \epsilon\). Then, \(0 \leq h_4(x) < g(x) + \epsilon\) and \(-\epsilon/2 < -g(x) \leq h_4(x) - g(x) < \epsilon\). Therefore, \(|g(x) - h_4(x)| < \epsilon\).

Case C: \(x \in \overline{O}\). Then, \(0 \leq g(x) \leq \epsilon\) and \(h_4(x) \geq M \geq g(x) \geq 0\). Hence, \(g(x) - \epsilon < h_4(x) < g(x) + \epsilon\). Then, \(0 \leq h_4(x) < g(x) + \epsilon\) and \(g(x) - \epsilon < h_4(x) < g(x) + \epsilon\). Therefore, \(|g(x) - h_4(x)| < \epsilon\).

Hence, in all three cases, we have \(|g(x) - h_4(x)| < \epsilon\). Hence \(||g - h_4|| < \epsilon\).

By the arbitrariiness of \(\epsilon\), we have \(g \in \overline{\mathcal{A}} = \mathcal{A}\). \(\square\)

\(\forall f \in \mathcal{C}(\mathcal{X}, \mathbb{R})\) with \(f(x_0) = 0\), let \(f^+ := f \lor 0\) and \(f^- := (f) \lor 0\). Clearly, \(f^+, f^- \in \mathcal{C}(\mathcal{X}, \mathbb{R}), f^+(x_0) = 0, f^-(x_0) = 0\), and \(f = f^+ - f^-\). By Claim 7.60.4, \(f^+, f^- \in \overline{\mathcal{A}}\). Then, \(f \in \overline{\mathcal{A}}\) since \(\mathcal{A}\) is an algebra. This coupled with Claim 7.60.2, we have that \(\mathcal{A} = \{f \in \mathcal{C}(\mathcal{X}, \mathbb{R}) \mid f(x_0) = 0\}\). Hence, the result holds in this case.
This completes the proof of the proposition. \(\square\)

### 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 \( \| Ax \|_y \leq M \| x \|_X \), \( \forall x \in \mathcal{X} \). 

Let \( \varepsilon \in (0, \infty) \subset \mathbb{R} \), let \( \delta = \varepsilon/(1 + M) \in (0, \infty) \subset \mathbb{R} \). \( \forall x_1, x_2 \in \mathcal{X} \) with \( \| x_1 - x_2 \|_X < \delta \), we have \( \| Ax_1 - Ax_2 \|_y = \| A(x_1 - x_2) \|_y \leq M \| x_1 - x_2 \|_X < \varepsilon \). Hence, \( A \) is uniformly continuous.

(ii). \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \), \( \forall x \in \mathcal{B}_x(x_0, \delta) \), we have \( \| Ax - Ax_0 \|_y < \varepsilon \). \( \forall \bar{x} \in \mathcal{B}_x(x_0, \delta) \), we have \( \| A\bar{x} \|_y = \| A(x_0 + \bar{x}) - Ax_0 \|_y < \varepsilon \). Hence, \( A \) is continuous at \( \vartheta_X \). Then, \( \exists \delta_0 \in (0, \infty) \subset \mathbb{R} \) such that \( \| Ax \|_y < 1 \), \( \forall x \in \mathcal{B}_x(\vartheta_X, \delta_0) \). Then, \( \forall x \in \mathcal{X} \), we have two possibilities: (a) \( x = \vartheta_X \), then \( \| Ax \|_y = 0 \leq \frac{1}{\delta_0} \| x \|_X \); (b) \( x \neq \vartheta_X \), then, \( \| Ax \|_y = \| A\left( \frac{\delta_0}{2\| x \|_X} - \frac{\delta_0}{\| x \|_X} \right) \|_y = \frac{2\| x \|_X}{\delta_0} \| A\left( \frac{\delta_0}{2\| x \|_X} - \frac{\delta_0}{\| x \|_X} \right) \|_y < \frac{2}{\delta_0} \| x \|_X \). Hence, \( A \) is bounded.

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

**Proposition 7.63** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over the field \( \mathbb{K} \). Let \( (\mathcal{M}(\mathcal{X}, \mathcal{Y}), \mathbb{K}) \) be the vector space defined in Example 6.20 with the null vector \( \vartheta \). Let \( P := \{ A \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid A \text{ is linear} \} \). Then, \( P \) is a subspace of \( \mathcal{M}(\mathcal{X}, \mathcal{Y}) \). 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

\[
\| A \| = \sup_{x \in \mathcal{X}, \| x \|_X \leq 1} \| Ax \|_y = \max \left\{ \sup_{x \in \mathcal{X}, \| x \|_X = 1} \frac{\| Ax \|_y}{\| x \|_X}, 0 \right\} = \max \left\{ \sup_{x \in \mathcal{X}, \| x \|_X = 1} \| Ax \|_y, 0 \right\}
\]

Let \( N := \{ A \in P \mid \| A \| < +\infty \} \). Then, \( (N, \mathbb{K}, \| \cdot \|) := (B(\mathcal{X}, \mathcal{Y})) \) is a normed linear space.
Proposition 7.65 Let $X$ and $Y$ be normed linear spaces over the field $K$ and $B(X,Y)$ be the normed linear space of bounded linear operators of $X$. Then, if $K = \mathbb{R}$ or $\mathbb{C}$, we have \( \| A \| \leq \| B \| \| A \| \| B \| \). Furthermore, if $X$ and $Y$ are real or complex vector spaces, we have \( \| A \| \leq \| B \| \| A \| \| B \| \). Therefore, \( \| A \| \leq \| B \| \| A \| \| B \| \) and \( \| A \| \leq \| B \| \| A \| \| B \| \). Hence, \( B(X,Y) \) is a normed linear space.

Proof This is straightforward, and is therefore omitted.

Proposition 7.66 Let $X$, $Y$, and $Z$ be normed linear spaces over the field $K$, $A \in B(Y,Z)$, and $B \in B(Z,X)$. Then, \( \forall x \in X \), 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.
7.8. LINEAR OPERATORS

Let $X$ and $Y$ be normed linear spaces over the field $K$. Let $B(X, Y)$ be the normed linear space defined in Proposition 7.63. If $Y$ is a Banach space, then $B(X, Y)$ is also a Banach space.

Proof

All we need to show is that $B(X, Y)$ is complete. Take a Cauchy sequence $(A_n)_{n=1}^\infty \subseteq B(X, Y)$. $\forall x \in X$, we have $\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \epsilon \|x\|$. Then, by Propositions 7.66, 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, $\|f(x)\| \leq \|f(x) - A_n x\| + \|A_n x\| \leq \epsilon \|x\| + \|A_n\| \|x\| = (\epsilon + \|A_n\|) \|x\|$. This implies that $f$ is bounded. Hence, $f \in B(X, Y)$. Note that $0 \leq \lim_{n \to \infty} \|A_n - f\| = \lim_{n \to \infty} \sup_{x \in X} (\|A_n - f\|) = \lim_{n \to \infty} \sup_{x \in X} (\|A_n - f\|) = \lim_{n \to \infty} \sup_{x \in X} (\|A_n - f\|) = 0$ and hence $\lim_{n \to \infty} A_n = f \in B(X, Y)$. Hence, $B(X, Y)$ is complete.

This completes the proof of the proposition.

□

Proposition 7.67

Let $X$ and $Y$ be normed linear spaces over the field $K$, $X$ be finite dimensional with dimension $n \in \mathbb{Z}_+$, and $A : X \to Y$ be a linear operator. Then, $A \in B(X, Y)$.

Proof

Let $X_N \subseteq X$ be a basis for $X$, 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, $X = \{0\}$. Then, by Proposition 7.63, we have $\|A\| = \sup_{x \in X} \|A x\| = 0$. Hence, $A \in B(X, Y)$.

Case 2: $n \in \mathbb{N}$. Let $X_N = \{x_1, \ldots, x_n\}$. $\forall x \in X$, 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 $X$ by $\|x\|_1 = \sqrt{\sum_{i=1}^n |\alpha_i|^2}$.
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,

\[
\| A \| = \sup_{x \in \mathcal{X}, \| x \| \leq 1} \| Ax \| = \sup_{\alpha_1, \ldots, \alpha_n \in \mathbb{R}, \| \sum_{i=1}^n \alpha_i x_i \| \leq 1} \left\| A \left( \sum_{i=1}^n \alpha_i x_i \right) \right\| \\
\leq \sup_{\alpha_1, \ldots, \alpha_n \in \mathbb{R}, \| \sum_{i=1}^n \alpha_i x_i \| \leq 1} \sum_{i=1}^n |\alpha_i| r_i \\
\leq \sup_{\alpha_1, \ldots, \alpha_n \in \mathbb{R}, \| \sum_{i=1}^n \alpha_i x_i \| \leq 1} \left( \sum_{i=1}^n r_i^2 \right)^{1/2} \left( \sum_{i=1}^n |\alpha_i|^2 \right)^{1/2} \\
= \sup_{x \in \mathcal{X}, \| x \| \leq 1} \left( \sum_{i=1}^n r_i^2 \right)^{1/2} \| x \|_1 \leq \sup_{x \in \mathcal{X}, \| x \| \leq 1} \left( \sum_{i=1}^n r_i^2 \right)^{1/2} M \| x \| \\
\leq M \left( \sum_{i=1}^n r_i^2 \right)^{1/2}
\]

where we have applied the Cauchy-Schwarz Inequality in the second inequality. Hence, \( A \in B(\mathcal{X}, \mathcal{Y}) \).

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 \( \{ \vartheta y \} \subseteq \mathcal{Y} \) is closed by Proposition 3.34. By Proposition 3.10, \( \mathcal{N}(A) = A_{\text{inv}}(\vartheta y) \) 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, \( \mathcal{R}(A) = \{ \vartheta y \} \). Hence, \( A = \vartheta B(\mathcal{X}, \mathcal{Y}) \), which is clearly bounded.

Case 2: \( n \in \mathbb{N} \). Then, \( \exists \mathcal{X}_N = \{ x_1, \ldots, x_n \} \subseteq \mathcal{X} \) with exactly \( n \) elements such that \( A(\mathcal{X}_N) \subseteq \mathcal{Y} \) 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. \( \forall x \in \mathcal{X}, Ax \in \mathcal{R}(A) = \text{span}(A(\mathcal{X}_N)) \). Then, by Definition 6.50 and Corollary 6.47, \( \exists! \alpha_1, \ldots, \alpha_n \in \mathbb{K} \) such that \( Ax = \sum_{i=1}^n \alpha_i Ax_i \). Then, we have \( A(x - \sum_{i=1}^n \alpha_i x_i) = \vartheta y \) 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 \( \forall \tilde{x} \in \{ x \} \), we have \( x - \tilde{x} \in \mathcal{N}(A) \) and \( Ax = A\tilde{x} \). Hence, \( A \) is uniquely defined. \( \forall \{ x_1 \}, \{ x_2 \} \in \mathcal{X}/\mathcal{N}(A), \forall \alpha, \beta \in \mathbb{K}, \) we have \( \tilde{A}(\alpha \{ x_1 \} + \beta \{ x_2 \}) = \tilde{A}(\alpha \{ x_1 \}) + \tilde{A}(\beta \{ x_2 \}) = \alpha Ax_1 + \beta Ax_2 = A(\alpha x_1 + \beta x_2) = \tilde{A} (\alpha \{ x_1 \} + \beta \{ x_2 \}) \). Hence, \( \tilde{A} \) is linear and bounded as a linear operator. Therefore, \( \tilde{A} \) is continuous. Hence, \( \mathcal{X}/\mathcal{N}(A) \) is closed. This completes the proof of the proposition. \( \square \)
be its dual. Then, the following statements hold.

**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 $\phi : X \rightarrow X/M$ 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.

**Proposition 7.70** Let $X$ and $Y$ be normed linear spaces over the field $\mathbb{K}$, $A \in B(X,Y)$, and $\phi : X \rightarrow X/\mathcal{N}(A)$ be the natural homomorphism. Then, $\exists B \in B(X/\mathcal{N}(A),Y)$ such that $A = B \circ \phi$, $B$ is injective, and $\|A\| = \|B\|$.

**Proof** By Proposition 7.68, $\mathcal{N}(A)$ is closed. Then, by Proposition 7.44, $X/\mathcal{N}(A)$ is a normed linear space. Define $B : X/\mathcal{N}(A) \rightarrow Y$ by $B([x]) = Ax$, $\forall [x] \in X/\mathcal{N}(A)$. Note that $\forall \bar{x} \in [x]$, we have $x - \bar{x} \in \mathcal{N}(A)$ and $Ax = A\bar{x}$. Hence, $B$ is uniquely defined. $\forall [x_1], [x_2] \in X/\mathcal{N}(A)$, $\forall \alpha, \beta \in \mathbb{K}$, we have $B(\alpha[x_1] + \beta[x_2]) = B(\alpha Ax_1 + \beta Ax_2) = A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \alpha B([x_1]) + \beta B([x_2])$. Hence, $B$ is linear. $\forall [x_1], [x_2] \in X/\mathcal{N}(A)$, $B[x_1] = B[x_2]$ implies that $Ax_1 = Ax_2$, $x_1 - x_2 \in \mathcal{N}(A)$, and hence, $[x_1] = [x_2]$. This shows that $B$ is injective.

$\forall x \in X$, $B \circ \phi(x) = B[x] = Ax$. Hence, $A = B \circ \phi$.

$\forall [x] \in X/\mathcal{N}(A)$, $\forall m \in \mathcal{N}(A)$, $\|B[x]\| = \|Ax\| = \|A(x - m)\| \leq \|A\| \|x - m\|$. This implies that $\|B[x]\| \leq \|A\| \|x\|$. Hence, $B \in B(X/\mathcal{N}(A),Y)$ 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.

## 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 $f : X \rightarrow \mathbb{K}$ 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 \( f : X \to K \), we have \( \|f\| = \inf \{ M \in [0, \infty) \subset R \mid |f(x)| \leq M \|x\|_X, \forall x \in X \} = \sup_{x \in X, \|x\|_X \leq 1} |f(x)| = \max \left\{ \sup_{x \in X, x \neq \theta_X} \frac{|f(x)|}{\|x\|_X} \right\} = \max \left\{ \sup_{x \in X, \|x\|_X = 1} |f(x)|, 0 \right\} \).

(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 \( f : X \to K \) 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 \( K \) 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 \( X := \{ \vartheta_X \}, K, \|\cdot\| \). Clearly, there is a single linear functional on \( X \) given by \( f : X \to K \) and \( f(\vartheta_X) = 0 \). Clearly, this linear functional is bounded with norm 0. Hence, \( X^* = (\{ f \}, K, \|\cdot\|_1) \), which is isometrically isomorphic to \( X \). ∗

**Example 7.74** Let us consider the dual of \( X := K^n, n \in N \), \( \forall x := (\xi_1, \ldots, \xi_n) \in K^n \), any functional of the form \( f(x) = \sum_{i=1}^n \eta_i \xi_i \) with \( \eta_1, \ldots, \eta_n \in K \) is clearly linear. By Cauchy-Schwarz Inequality, we have \( |f(x)| = |\sum_{i=1}^n \eta_i \xi_i| \leq \sum_{i=1}^n |\eta_i||\xi_i| \leq \sqrt{\sum_{i=1}^n |\eta_i|^2} \sqrt{\sum_{i=1}^n |\xi_i|^2} \). 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 \( \eta_i = f(e_i) \in K, i = 1, \ldots, n \). 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 \( \Psi : X^* \to K^n \) by \( \Psi(f) = (\eta_1, \ldots, \eta_n), \forall f \in X^* \). Clearly, \( \Psi \) is linear, bijective and norm preserving. Hence, the dual of \( K^n \) is (isometrically isomorphic to) \( K^n \). ∗
Lemma 7.75 Let $X$ be a normed linear space over $\mathbb{K}$, $X^*$ be its dual, and $x_* \in X^*$. Then, $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists x \in X$ with $\|x\| \leq 1$ such that $\langle\langle x_*, x \rangle\rangle \in \mathbb{R}$ and $\langle\langle x_*, x \rangle\rangle \geq (1 - \epsilon) \|x_*\|$. 

Proof $\forall \epsilon \in (0, 1) \subset \mathbb{R}$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|x_*\| = 0$; Case 2: $\|x_*\| > 0$. Case 1: $\|x_*\| = 0$. Take $x = \emptyset$ 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 $\langle\langle x_*, \bar{x} \rangle\rangle \geq (1 - \epsilon) \|x_*\|$. Take $x := \frac{\langle\langle x_*, \bar{x} \rangle\rangle}{\langle\langle x_*, \bar{x} \rangle\rangle} \bar{x} \in X$. Then, $\|x\| = \|\bar{x}\| \leq 1$ and $\langle\langle x_*, x \rangle\rangle = \langle\langle x_*, \frac{\langle\langle x_*, \bar{x} \rangle\rangle}{\langle\langle x_*, \bar{x} \rangle\rangle} \bar{x} \rangle\rangle = \langle\langle x_*, \bar{x} \rangle\rangle = \langle\langle x_*, \bar{x} \rangle\rangle \|x_*\| \leq \|x\| \leq 1$. This completes the proof of the lemma. $\square$

Example 7.76 Let $\mathcal{Y}_i$ be a normed linear space over the field $\mathbb{K}$, $\mathcal{Y}_i^*$ be its dual, $i = 1, \ldots, n$. $n \in \mathbb{N}$. We will study the dual of $X := \prod_{i=1}^n \mathcal{Y}_i$. We will show that $X^* = \prod_{i=1}^n \mathcal{Y}_i^*$ isometrically isomorphically.

Let $f \in X^*$. $\forall i = 1, \ldots, n$, define $A_i : \mathcal{Y}_i \to X$ by $x = (y_1, \ldots, y_n) = A_i y_i$, $\forall y_i \in \mathcal{Y}_i$, and $y_j = \delta_{y_i}$, $\forall j \in 1, \ldots, n$ with $j \neq i$, and $y_i = y$. Clearly, $A_i$ is well-defined and linear and bounded with $\|A_i\| \leq 1$. Define $f_i : \mathcal{Y}_i \to \mathbb{K}$ by $f_i = f \circ A_i$. Then, $f_i$ is a bounded linear functional on $\mathcal{Y}_i$ with $\|f_i\| \leq \|f\| \|A_i\| \leq \|f\|$. Hence, $f_i \in \mathcal{Y}_i^*$. Denote $f_i := y_{*i}$. $\forall x = (y_1, \ldots, y_n) \in X$, we have $y_i \in \mathcal{Y}_i$, $i = 1, \ldots, n$. By the linearity of $f$, we have

$$f(x) = f \left( \sum_{i=1}^n A_i y_i \right) = \sum_{i=1}^n f(A_i y_i) = \sum_{i=1}^n \langle\langle y_{*i}, y_i \rangle\rangle \tag{7.1}$$

Let $x_* := (y_{*1}, \ldots, y_{*n}) \in \prod_{i=1}^n \mathcal{Y}_i^*$. $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\forall i = 1, \ldots, n$, by Lemma 7.75, $\exists \bar{y}_i \in \mathcal{Y}_i$ with $\|\bar{y}_i\| \leq 1$ such that $\langle\langle y_{*i}, \bar{y}_i \rangle\rangle \geq (1 - \epsilon) \|y_{*i}\|$. Let $\tilde{y}_i := \frac{y_{*i}}{\|y_{*i}\|} \bar{y}_i$. Then, $\langle\langle y_{*i}, \tilde{y}_i \rangle\rangle \geq (1 - \epsilon) \|y_{*i}\|$. Then, we have $\|f(\sum_{i=1}^n A_i y_i)\| = \|\sum_{i=1}^n \langle\langle y_{*i}, \tilde{y}_i \rangle\rangle \| = \sum_{i=1}^n \langle\langle y_{*i}, \tilde{y}_i \rangle\rangle \geq (1 - \epsilon) \sum_{i=1}^n \|y_{*i}\|^2 = (1 - \epsilon) \|x_*\|^2$. Since $f \in X^*$, then, by Proposition 7.72, we have $(1 - \epsilon) \|x_*\|^2 \leq \|f\| \|\sum_{i=1}^n A_i \tilde{y}_i\| = \|f\| \left( \sum_{i=1}^n \|\tilde{y}_i\|^2 \right)^{1/2} \leq \|f\| \left( \sum_{i=1}^n \|y_{*i}\|^2 \right)^{1/2} = \|f\| \|x_*\|$. By the arbitrariness of $\epsilon$, we have $\|x_*\| \leq \|f\|$. 

Based on the preceding analysis, we may define a function $\psi : X^* \to \prod_{i=1}^n \mathcal{Y}_i^*$ 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 $f_1 \circ A_i = f_2 \circ A_i$, $\forall i \in \mathbb{N}$. $\forall x := (y_1, \ldots, y_n) \in X$, 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. $\forall x_* := (y_{*1}, \ldots, y_{*n}) \in \prod_{i=1}^n \mathcal{Y}_i^*$, define $\phi(x_*) : X \to \mathbb{K}$ 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

\[
\| \phi(x_*) \| = \sup_{x \in X, \|x\| \leq 1} |\phi(x_*)(x)| \leq \sup_{x = (y_1, \ldots, y_n) \in X, \|x\| \leq 1} \sum_{i=1}^{n} |\langle \phi(y_i), y_i \rangle| \\
\leq \sup_{x \in X, \|x\| \leq 1} \| x_* \| \| x \| \leq \| x_* \| \tag{7.2}
\]

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, \( \psi \circ \phi = \text{id}_{\prod_{i=1}^{n} Y} \). Hence, \( \psi \) is surjective. Therefore, \( \psi \) is bijective and admits inverse \( \psi_{\text{inv}} \). By Proposition 2.4, \( \phi = \psi_{\text{inv}} \).

**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 \( x = (y_1, y_2, \ldots) = A_i y, \forall y \in Y \), and \( y_j = \delta_{ij}, \forall j \in \mathbb{N} \) 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

\[
f(x) = \lim_{n \in \mathbb{N}} f \left( \sum_{i=1}^{n} A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f \left( A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} \langle y_i, y_i \rangle \tag{7.3}
\]

**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, \( \exists y_i \in Y \) with \( \| y_i \| \leq 1 \) such that \( \langle \langle y_i, y_i \rangle \rangle \in \mathbb{R} \) and \( \langle \langle y_i, y_i \rangle \rangle \geq (1 - \epsilon) \| y_i \| \rangle \). Let \( y_i = \| y_i \|^{q/p} y_i \), then \( \| y_i \| \leq \| y_i \|^{q/p} \), \( \langle \langle y_i, y_i \rangle \rangle \in \mathbb{R} \), and \( \langle \langle y_i, y_i \rangle \rangle \geq (1 - \epsilon) \| y_i \|^{q/p+1} = (1 - \epsilon) \| y_i \|^{q} \).

Then, \( \forall n \in \mathbb{N} \), \( |f(\sum_{i=1}^{n} A_i y_i)| = |\sum_{i=1}^{n} \langle \langle y_i, y_i \rangle \rangle| = \sum_{i=1}^{n} \langle \langle y_i, y_i \rangle \rangle \geq (1 - \epsilon) \sum_{i=1}^{n} \| y_i \|^{q} \leq \| f \| \sum_{i=1}^{n} \| A_i y_i \| \leq \| f \| \left( \sum_{i=1}^{n} \| y_i \|^{p} \right)^{1/p} \leq \| f \| \cdot \left( \sum_{i=1}^{n} \| y_i \|^{q} \right)^{1/p}. \)

Then, we have \( \left( \sum_{i=1}^{n} \| y_i \|^{q} \right)^{1/q} \leq \| f \|/(1 - \epsilon) \). 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 \((\sum_{i=1}^{\infty} \| y_{si} \|^q)^{1/q} \leq \| f \|\). 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, \( \exists y_{vi} \in \mathbb{Y} \) with \( \|y_{vi}\| \leq 1 \) such that \( \langle\langle y_{si}, y_{vi}\rangle\rangle \geq (1-\epsilon) \| y_{si} \| \). Then, \( (1-\epsilon) \| y_{si} \| \leq \| \langle\langle y_{si}, y_{vi}\rangle\rangle \| = \| f(A_i y_{vi}) \| \leq \| f \| \| A_i \| \| y_{vi} \| \leq \| f \| \). 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 \( \psi : X^* \to l_q(\mathbb{Y}^*) \) by \( \psi(f) = (f \circ A_1, f \circ A_2, \ldots) \), \( \forall f \in X^* \). Clearly, \( \psi \) is linear. \( \forall f_1, f_2 \in X^* \) with \( \psi(f_1) = \psi(f_2) \). Then, we have \( f_1 \circ A_i = f_2 \circ A_i, \forall i \in \mathbb{N} \).

\( \forall x := (y_1, y_2, \ldots) \in X \), by (7.3), we have \( f_1(x) = \lim_{n \to \infty} \sum_{i=1}^{n} f_1(A_i y_i) = \lim_{n \to \infty} \sum_{i=1}^{n} f_2(A_i y_i) = f_2(x) \). Then, we have \( f_1 = f_2 \). Hence, \( \psi \) is injective. \( \forall x := (y_1, y_2, y_3, \ldots) \in l_q(\mathbb{Y}^*) \), define \( \phi(x) : X \to \mathbb{K} \) by \( \phi(x)(x) = \sum_{i=1}^{\infty} \langle\langle y_{si}, y_{vi}\rangle\rangle \), \( \forall x := (y_1, y_2, \ldots) \in X \). Note that, \( \forall x := (y_1, y_2, \ldots) \in X \), \( \sum_{i=1}^{\infty} \| y_{si} \| \| y_{vi} \| \leq \| x \| \| x \| \), 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

\[
\| \phi(x) \| = \sup_{x \in X, \| x \| \leq 1} | \phi(x)(x) | \leq \sup_{x=(y_1, y_2, \ldots) \in X, \| x \| \leq 1} \sum_{i=1}^{\infty} |\langle\langle y_{si}, y_{vi}\rangle\rangle | \leq \sup_{x \in X, \| x \| \leq 1} \| x \| \| x \| \leq \| x \|. \tag{7.4}
\]

Hence, \( \phi(x) \in X^* \) and \( \psi(\phi(x)) = \phi(x) \circ A_1, \phi(x) \circ A_2, \ldots = (y_{s1}, y_{s2}, \ldots) = x \). Then, \( \psi \circ \phi = \text{id}_{l_q(\mathbb{Y}^*)} \). Hence, \( \psi \) is surjective. Therefore, \( \psi \) is bijective and admits inverse \( \psi_{\text{inv}} \). By Proposition 2.4, \( \phi = \psi_{\text{inv}} \).

\( \forall f \in X^* \), by Claim 7.77.1 and (7.4), we have \( \| f \| = \| \phi(\psi(f)) \| \leq \| \psi(f) \| \leq \| f \| \). Then, \( \psi(\phi(f)) = \| f \| \) and \( \psi \) is an isometry. Hence, \( \psi \) is an isometrical isomorphism.

**Example 7.78** Let \( \mathbb{Y} \) be a normed linear space over the field \( \mathbb{K} \), \( \mathbb{Y}^* \) be its dual, and \( M = \{ x = (y_1, y_2, \ldots) \in l_\infty(\mathbb{Y}) \mid \lim_{n \to \infty} y_n = \vartheta_y \} \). Clearly, \( M \) is a subspace of \( l_\infty(\mathbb{Y}) \). By Proposition 7.13, \( M \) is a normed linear space over \( \mathbb{K} \), which will be denoted by \( c_0(\mathbb{Y}) \). Next, we will study the dual of \( X := c_0(\mathbb{Y}) \). We will show that \( X^* = l_1(\mathbb{Y}^*) \) isometrical isomorphically.

Let \( f \in X^* \). \( \forall i \in \mathbb{N} \), define \( A_i : \mathbb{Y} \to X \) by \( x = (y_1, y_2, \ldots) = A_i y \), \( \forall y \in \mathbb{Y} \), 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 \( f_i : \mathbb{Y} \to \mathbb{K} \) by \( f_i = f \circ A_i \). Then, \( f_i \) is a bounded linear functional on \( \mathbb{Y} \) with \( \| f_i \| \leq \| f \| \| A_i \| \leq \| f \| \). Hence, \( f_i \in \mathbb{Y}^* \). Denote \( f_i := y_{si} \).

\( \forall x = (y_1, y_2, \ldots) \in X \), we have \( y_i \in \mathbb{Y} \), \( \forall i \in \mathbb{N} \), \( \lim_{n \to \infty} y_i = \vartheta_y \), and \( \lim_{n \to \infty} \| y_i \| = 0 \) (by Proposition 7.21). Then, \( \lim_{n \to \infty} \sup_{i \geq n+1} \| y_k \| = 0 \).

This implies that \( \lim_{n \to \infty} \| x - \sum_{i=1}^{n} A_i y_i \| = \lim_{n \to \infty} \sup_{k \geq n+1} \| y_k \| = 0 \).
\[ \lim_{n \in \mathbb{N}} \max_{k \geq n+1} \|y_k\| = 0. \] 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

\[
f(x) = \lim_{n \in \mathbb{N}} f \left( \sum_{i=1}^{n} A_i y_i \right) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f \left( A_i y_i \right) = \lim_{n \in \mathbb{R}} \sum_{i=1}^{n} \langle y_{s_i}, y_i \rangle \quad (7.5)\]

Let \( x_* := (y_{s_1}, y_{s_2}, \ldots, \forall \epsilon \in (0, 1) \subset \mathbb{R}, \forall i \in \mathbb{N} \), by Lemma 7.75, \( \exists \hat{y}_i \in Y \) with \( \|\hat{y}_i\| \leq 1 \) such that \( \langle \langle y_{s_i}, \hat{y}_i \rangle \rangle \in \mathbb{R} \) and \( \langle \langle y_{s_i}, \hat{y}_i \rangle \rangle \geq (1 - \epsilon)\|y_{s_i}\| \). Then, \( \forall n \in \mathbb{N}, (1 - \epsilon) \sum_{i=1}^{n} \|y_{s_i}\| \leq \|\sum_{i=1}^{n} \langle y_{s_i}, \hat{y}_i \rangle \| = \|\sum_{i=1}^{n} f(A_i y_i)\| = f(\sum_{i=1}^{n} A_i \hat{y}_i) \leq \|f\| \|\sum_{i=1}^{n} A_i \hat{y}_i\| \leq \|f\|. \) By the arbitrariness of \( n \), we have \( \sum_{i=1}^{\infty} \|y_{s_i}\| \leq \|f\|/(1 - \epsilon) \). By the arbitrariness of \( \epsilon \), we have

\[
\sum_{i=1}^{\infty} \|y_{s_i}\| \leq \|f\| \quad (7.6)
\]

Hence, \( x_* \in l_1(Y^*) \).

Based on the preceding analysis, we may define a function \( \psi : X^* \to l_1(Y^*) \) by \( \psi(f) = (f \circ A_1, f \circ A_2, \ldots), \forall f \in X^* \). Clearly, \( \psi \) is linear. \( \forall f_1, f_2 \in X^* \) with \( \psi(f_1) = \psi(f_2) \). Then, we have \( f_1 \circ A_1 = f_2 \circ A_1, \forall i \in \mathbb{N} \). \( \forall x := (y_1, y_2, \ldots) \in X \), by (7.5), we have \( f_1(x) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_i(A_i y_i) = \lim_{n \in \mathbb{N}} \sum_{i=1}^{n} f_2(A_i y_i) = f_2(x) \). Then, we have \( f_1 = f_2 \). Hence, \( \psi \) is injective. \( \forall x_* := (y_{s_1}, y_{s_2}, \ldots) \in l_1(Y^* \), define \( \phi(x_*) : X \to K \) by \( \phi(x_*) := \sum_{i=1}^{\infty} (\langle y_{s_i}, y_i \rangle), \forall x := (y_1, y_2, \ldots) \in X \). Note that, \( \forall x := (y_1, y_2, \ldots) \in X, \sum_{i=1}^{\infty} (\langle y_{s_i}, y_i \rangle) \leq \sum_{i=1}^{\infty} \|y_{s_i}\| \|y_i\| \leq \|x_*\| \|x\|, \) where we have made use of Proposition 7.72 in the first inequality and Hölder’s inequality in the second inequality. Hence, \( \phi(x_*) \) is well-defined. Hence, \( \phi(x_*) \) is well-defined. Clearly, \( \phi(x_*) \) is linear and bounded with

\[
\|\phi(x_*)\| = \sup_{x \in X, \|x\| \leq 1} \|\phi(x_*)\| \leq \sup_{x = (y_1, y_2, \ldots) \in X, \|x\| \leq 1} \sum_{i=1}^{\infty} \|\langle y_{s_i}, y_i \rangle\| \leq \sup_{x \in X, \|x\| \leq 1} \|x\| \|x\| \leq \|x_*\| \quad (7.7)
\]

Hence, \( \phi(x_*) \in X^* \) and \( \psi(\phi(x_*)) = (\phi(x_*) \circ A_1, \phi(x_*) \circ A_2, \ldots) = (y_{s_1}, y_{s_2}, \ldots) = x_* \). Then, \( \psi \circ \phi = \text{id}_{l_1(Y^*)} \). Hence, \( \psi \) is surjective. Therefore, \( \psi \) is bijective and admits inverse \( \psi_{\text{inv}} \). By Proposition 2.4, \( \phi = \psi_{\text{inv}} \). \( \forall f \in X^* \), by (7.6) and (7.7), we have \( \|f\| = \|\phi(f)\| \leq \|\psi(f)\| \leq \|f\| \). Then, \( \|\phi(f)\| = \|f\| \) and \( \psi \) is an isometry. Hence, \( \psi \) is an isometrical isomorphism. \( \diamond \)

### 7.9.3 Extension form of Hahn-Banach Theorem

**Definition 7.79** Let \( X \) be a vector space over \( K \). A sublinear functional is \( p : X \to \mathbb{R} \) satisfying, \( \forall x_1, x_2 \in X, \forall \alpha \in \mathbb{R} \) with \( \alpha \geq 0, \) (i) \( p(x_1 + x_2) \leq p(x_1) + p(x_2) \); and (ii) \( p(\alpha x_1) = \alpha p(x_1) \).
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 \( \vartheta_\mathcal{X} \), 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 \( (g_1, N_1) \succeq (g_2, N_2) \) 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, \( \vartheta_\mathcal{X} \in M \subseteq N_c \subseteq \mathcal{X} \). \( \forall x_1, x_2 \in N_c, \forall \alpha, \beta \in \mathbb{R}, \exists (g_1, N_1), (g_2, N_2) \in \mathcal{C} \) 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 \( (g_1, N_1) \succeq (g_2, N_2) \). 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, \( (g_1, N_1) \succeq (g_2, N_2) \). Then, \( x \in N_1 \subseteq N_2 \) and \( g_c(\alpha x_1 + \beta x_2) = g_2(\alpha x_1 + \beta x_2) = \alpha g_2(x_1) + \beta g_2(x_2) = \alpha g_c(x_1) + \beta g_c(x_2) \).

Hence, \( g_c \) is a linear functional. \( g_c(x_1) = g_2(x_1) \leq p(x_1), \forall x_1 \in M, g_c(m) = g_2(m) = f(m) \). 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 \( N_c = \{ x \in \mathcal{X} \mid \exists \alpha \in \mathbb{R}, \exists n \in N_M \setminus \{ x = \alpha x_0 + n \} \}. \) 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 \( g_\epsilon : N_c \to \mathbb{R} \) by \( g_\epsilon(x) = g_\epsilon(\alpha x_0 + n) = g_M(n) + \alpha g_\epsilon(x_0) \), \( \forall x = \alpha x_0 + n \in N_c, \alpha \in \mathbb{R} \), and \( n \in N_M \), where \( g_\epsilon(x_0) \in \mathbb{R} \) is a constant to be determined. Clearly, \( g_\epsilon \) is well-defined and is a linear functional on \( N_c \). \( \forall n \in N_M \), we have \( g_\epsilon(n) = g_M(n) \), and hence, \( g_\epsilon|_{N_M} = g_M \). Then, \( g_\epsilon|_M = f \).

We will show that \((g_\epsilon, N_c) \in \mathcal{E}\) by finding an admissible constant \( g_\epsilon(x_0) \) such that \( g_\epsilon(x) \leq p(x), \forall x \in N_c \). \( \forall n_1, n_2 \in N_M \), 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 \( \sup_{n \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq \inf_{n \in N_M} (p(n_1 + x_0) - g_M(n_1)) \).

Since \( \vartheta_x \in N_M \), then \( \exists c \in \mathbb{R} \) such that

\[
-\infty < \sup_{n \in N_M} (g_M(n_2) - p(n_2 - x_0)) \leq c
\]

\[
\leq \inf_{n \in N_M} (p(n_1 + x_0) - g_M(n_1)) < +\infty
\]

Let \( g_\epsilon(x_0) := c \).

\( \forall x \in N_c, x = \alpha x_0 + n \), 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: \( \alpha > 0 \). \( g_\epsilon(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(p(n/\alpha + x_0) - g_M(n/\alpha) + g_M(n/\alpha)) = \alpha p(x/\alpha) = p(x) \).

Case 2: \( \alpha = 0 \). Then, \( x \in N_M \) and \( g_\epsilon(x) = g_M(n) \leq p(n) = p(x) \).

Case 3: \( \alpha < 0 \). \( g_\epsilon(x) = \alpha c + g_M(n) = \alpha(c + g_M(n/\alpha)) \leq \alpha(g_M(-n/\alpha) - p(-n/\alpha - x_0) + g_M(n/\alpha)) = -\alpha p(-x/\alpha) = p(x) \). Hence, we have \( g_\epsilon(x) \leq p(x) \) in all three cases. Therefore, \((g_\epsilon, N_c) \in \mathcal{E}\). We have shown that \((g_M, N_M) \preceq (g_\epsilon, N_c) \) and \((g_M, N_M) \neq (g_\epsilon, N_c) \). This contradicts the fact that \((g_M, N_M) \) is maximal in \( \mathcal{E}\) 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, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( |p(x)| < \epsilon, \forall x \in \mathcal{B}(\vartheta_X, \delta) \).

\( \forall x \in \mathcal{B}(\vartheta_X, \delta), \) 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, \( 0 \leq F(x) \leq p(x) < \epsilon \) and \( |F(x)| < \epsilon \). Case 2: \( F(x) < 0 \). Then, \( |F(x)| = -F(x) = F(-x) \leq p(-x) < \epsilon \). Hence, in both cases, we have \( |F(x)| < \epsilon \). Then, \( F \) is continuous at \( \vartheta \). Thus, by Proposition 7.72, \( F \) is a bounded linear functional, and therefore continuous.

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 $X$ be a functional on $C\subset X$. Furthermore, if $\alpha : X \to \mathbb{R}$ is a linear functional on $(X, \mathbb{C})$, then $f$ is continuous. Hence, $g$ is a linear functional on $(X, \mathbb{C})$. Let $\alpha, \beta \in \mathbb{R}$, we have $g(\alpha x_1 + \beta x_2) = Re(f(\alpha x_1 + \beta x_2)) = Re(\alpha f(x_1) + \beta f(x_2)) = \alpha \Re(f(x_1)) + \beta \Re(f(x_2)) = \alpha g(x_1) + \beta g(x_2)$. Hence, $g$ is a linear functional on $(X, \mathbb{C})$, then $f(ix) = g(ix) + i\Im(f(ix)) = if(x) = -\Im(f(x)) + ig(x)$. Then, $\Im(f(x)) = -g(ix)$ and $f(x) = g(x) - ig(ix)$.

"Sufficiency" $\forall x_1, x_2 \in X$, $\forall \alpha := \alpha_r + i\alpha_i, \beta := \beta_r + i\beta_i \in \mathbb{C}$, where $\alpha_r, \alpha_i, \beta_r, \beta_i \in \mathbb{R}$, we have

$$f(\alpha x_1 + \beta x_2) = g(\alpha x_1 + \beta x_2) - ig(\alpha x_1 + i\beta x_2) = g(\alpha x_1 + \alpha_i x_1 + \beta_i x_2) - ig(\alpha_i x_1 + \alpha \beta x_2 - \beta_i x_2) = \alpha_r g(x_1) + \alpha_i g(ix_1) + \beta_r g(x_2) + \beta_i g(ix_2) - i\alpha_\beta g(ix_1) + i\alpha_i g(x_1) - i\beta g(x_2) = \alpha_r f(x_1) + \alpha_i f(x_1) + \beta_r f(x_2) + i\beta_i f(x_2) - i\alpha_\beta f(x_2) = \alpha f(x_1) + \beta f(x_2)$$

Hence, $f$ is a linear functional on $(X, \mathbb{C})$.

This completes the proof of the lemma.

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 $Re(f(m)) \leq p(m), \forall m \in M$. Then, there exists a linear functional $F : X \to \mathbb{C}$, such that $F|_M = f$ and $Re(F(x)) \leq p(x), \forall x \in X$. Furthermore, if $X$ is normed with $\| \cdot \|$ and $p$ is continuous at $\partial X$, 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 $g(m) = Re(f(m)), \forall m \in M$. 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, $g(m) = Re(f(m)) \leq p(m), \forall m \in M$. Then, by Hahn-Banach Theorem, Theorem 7.80, there 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 $F(x) = G(x) - G(ix), \forall x \in X$. 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, $F(m) = G(m) - G(im) = g(m) - ig(im) = f(m)$. Hence, we have $F|_M = f$. $\forall x \in X$, $Re(F(x)) = G(x) \leq p(x)$. Hence, $F$ is the functional we seek.

Furthermore, if $(X, \mathbb{C})$ is normed with $\| \cdot \|$ and $p$ is continuous at $\partial X$, then, by Lemma 7.40, $X_R := (X, \mathbb{R}, \| \cdot \|)$ is a normed linear space. By Theorem 7.80, $G$ is continuous. Then, $G : X_R \to \mathbb{C}$ is continuous. By Propositions 3.12, 3.32, and 7.23, $F : X_R \to \mathbb{C}$ 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 $F \in \mathcal{X}^*$ 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 $\|f\|_M = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\|$. On the other hand, we have, $\forall x \in \mathcal{X}$,

$-F(x) = F(-x) \leq p(-x) = \|f\|_M \|x\|$ 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, $\forall m \in M, \text{Re}(f(m)) \leq |\text{Re}(f(m))| \leq |f(m)| \leq p(m)$, 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 $\text{Re}(F(x)) \leq p(x), \forall x \in \mathcal{X}$. Furthermore, since $p$ is continuous, then $F$ is continuous. By Proposition 7.72, $F \in \mathcal{X}^*$. Note that $\|f\|_M = \sup_{m \in M, \|m\| \leq 1} |F(m)| \leq \sup_{x \in \mathcal{X}, \|x\| \leq 1} |F(x)| = \|F\|$.

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 $|F(x)| = \left(\frac{|F(x)|}{F(x)}\right) F(x) = F\left(\frac{|F(x)|}{F(x)} \cdot F\right)^* X = \|F\|_M \|x\|$. Thus, we have $|F(x)| \leq \|F\|_M \|x\|$, $\forall x \in \mathcal{X}$. Therefore, $\|F\| \leq \|f\|_M$. Hence, $\|F\| = \|f\|_M$.

This completes the proof of the theorem. \qed

**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) $\text{Re}(f(m)) \leq p(m), \forall m \in M$.

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 $\text{Re}(F(x)) \leq p(x), \forall x \in \mathcal{X}$.
Furthermore, if, in addition, $\mathcal{X}$ is normed with $\| \cdot \|$ and $p$ is continuous at $\vartheta_\mathcal{X}$, then $F$ is continuous.

**Proof** Define a functional $q : \mathcal{X} \to \mathbb{R}$ by

\[
q(x) = \inf_{A_1, \ldots, A_n \in \mathcal{G}} \frac{1}{n} p\left( \sum_{i=1}^{n} A_i x \right); \quad \forall x \in \mathcal{X}
\]

Since $p(\vartheta_\mathcal{X}) = 0$, then $q(\vartheta_\mathcal{X}) = 0$. Since $E \in \mathcal{G}$, then, $\forall x \in \mathcal{X}$, $q(x) \leq p(Ex) \leq p(x) < +\infty$. $\forall x_1, x_2 \in \mathcal{X}$, $\forall n, k \in \mathbb{N}$, $\forall A_1, \ldots, A_n \in \mathcal{G}$, $\forall B_1, \ldots, B_k \in \mathcal{G}$, 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 follows from the fact that $\mathcal{G}$ is an abelian semigroup of linear operators on $\mathcal{X}$, the second and third inequalities follows 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 $0 = q(\vartheta_\mathcal{X}) \leq q(x_1) + q(-x_1)$. Then, $q(x_1) > -\infty$. Hence, $q : \mathcal{X} \to \mathbb{R}$.
\( \forall \alpha \in [0, \infty) \subset \mathbb{R}, \forall x \in \mathcal{X}, \) we have

\[
q(\alpha x) = \inf_{\lambda_1, \ldots, \lambda_n \in \emptyset} \frac{1}{n} p \left( \sum_{i=1}^{n} A_i(\alpha x) \right) = \inf_{\lambda_1, \ldots, \lambda_n \in \emptyset} \frac{1}{n} p \left( \alpha \sum_{i=1}^{n} A_i x \right)
\]

\[
= \inf_{\lambda_1, \ldots, \lambda_n \in \emptyset} \left( \frac{1}{n} p \left( \sum_{i=1}^{n} A_i x \right) \right) = \alpha q(x)
\]

where the fourth equality follows from Proposition 3.81 and the fact that \( q(x) \in \mathbb{R} \). Hence, we have shown that \( q \) is a sublinear functional.

\[ \forall m \in M, \forall n \in \mathbb{N}, \forall A_1, \ldots, A_n \in \mathcal{G}, \] we have

\[
\text{Re} \left( f(m) \right) = \frac{1}{n} \sum_{i=1}^{n} \text{Re} \left( f(A_i m) \right) = \frac{1}{n} \text{Re} \left( f \left( \sum_{i=1}^{n} A_i m \right) \right)
\]

\[
\leq \frac{1}{n} p \left( \sum_{i=1}^{n} A_i m \right)
\]

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 \( \text{Re}(f(m)) \leq q(m) \), \( \forall m \in M \).

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 \( F : \mathcal{X} \to \mathbb{K} \) such that \( F|_M = f \) and \( \text{Re}(F(x)) \leq q(x) \), \( \forall x \in \mathcal{X} \).

\[ \forall x \in \mathcal{X}, \forall A \in \mathcal{G}, \forall n \in \mathbb{N}, \] we have

\[
q(x - Ax) \leq \frac{1}{n} p \left( E(x - Ax) + A(x - Ax) + \cdots + A^{n-1}(x - Ax) \right)
\]

\[
= \frac{1}{n} p(Ex - A^n x) \leq \frac{1}{n} (p(Ex) + p(A^n(-x))) \leq \frac{1}{n} (p(x) + p(-x))
\]

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 \( H : \mathcal{X} \to \mathbb{R} \) by \( H(x) = \text{Re}(F(x)) \), \( \forall x \in \mathcal{X} \). 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 \( F(x) = H(x) - iH(ix), \forall x \in \mathcal{X} \). Note that \( H(x) - H(Ax) = H(x - Ax) = \text{Re}(F(x - Ax)) \leq q(x - Ax) \leq 0 \). On the other hand, we have \( H(Ax) - H(x) = H(-x) - H(A(-x)) = \)
\[ H((-x) - A(-x)) = \text{Re} \left( F((-x) - A(-x)) \right) \leq q((-x) - A(-x)) \leq 0. \] Hence, \( H(x) = H(Ax) \). Then, \( F(Ax) = H(Ax) - iH(AiAx) = H(Ax) - iH(A(iz)) = H(x) - iH(iz) = F(x) \).

Furthermore, if, in addition, \( \mathcal{X} \) is normed with \( \| \cdot \| \) and \( p \) is continuous at \( \vartheta_X \), then \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall x \in B(\vartheta_X, \delta) \), we have \( |p(x) - p(\vartheta_X)| = |p(x)| < \varepsilon \). Note that \( q(x) \leq p(x) < \varepsilon \). Then, \( q(-x) \leq p(-x) < \varepsilon \). Since \( q \) is a sublinear functional, we have \( 0 = q(\vartheta_X) \leq q(x) + q(-x) \). Hence, we have \( q(x) \geq -q(-x) > -\varepsilon \). Therefore, \( |q(x) - q(\vartheta_X)| = |q(x)| < \varepsilon \). Hence, \( q \) is continuous at \( \vartheta_X \). Therefore, by the extension forms of Hahn-Banach Theorem, \( F \) is continuous.

This completes the proof of the corollary. \( \Box \)

**Proposition 7.85** Let \( \mathcal{X} \) be a normed linear space over the field \( \mathbb{K} \) and \( x_0 \in \mathcal{X} \) with \( x_0 \neq \vartheta_X \). Then, \( \exists x_\ast \in \mathcal{X}^* \) with \( \| x_\ast \| = 1 \) such that \( \langle x_\ast, x_0 \rangle = \| x_\ast \| \| x_0 \| \).

**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 \( x_\ast \in \mathcal{X}^* \) such that \( x_\ast |_M = f \) and \( \| x_\ast \| = \| f \|_M = 1 \). Then, \( \langle x_\ast, x_0 \rangle = f(x_0) = \| x_0 \| = \| x_\ast \| \| x_0 \| \).

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

**Example 7.86** Let \( \mathcal{Y} \) be a normed linear space over the field \( \mathbb{K} \), \( \mathcal{X} := l_\infty(\mathcal{Y}) \) be the normed linear space as defined in Example 7.10, and \( y_0 \in \mathcal{Y} \) with \( y_0 \neq \vartheta_Y \). We will show that \( \mathcal{X}^* \neq l_1(\mathcal{Y}^*) \) by Hahn-Banach Theorem.

Let \( M := \{ x := (\xi_1, \xi_2, \ldots) \in \mathcal{X} | \lim_{n \to \infty} \xi_n \in \mathcal{Y} \} \). Clearly, \( M \) is a subspace of \( \mathcal{X} \). By Proposition 7.85, there exists \( y_\ast \in \mathcal{Y}^* \) such that \( \| y_\ast \| = 1 \) and \( \langle y_\ast, y_0 \rangle = \| y_0 \| > 0 \). Define \( f : M \to \mathbb{K} \) by \( f(m) = \langle y_\ast, \lim_{n \to \infty} \xi_n \rangle , \forall m := (\xi_1, \xi_2, \ldots) \in M \). 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

\[
|f(m)| = \| \langle y_\ast, \lim_{n \to \infty} \xi_n \rangle \| \leq \| \lim_{n \to \infty} \xi_n \| = \lim_{n \to \infty} \| \xi_n \| \leq \sup_{n \geq 1} \| \xi_n \|
\]

\[
= \| m \| \leq 1
\]

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 \( |f(m_0)| = |\langle y_\ast, y_0 \rangle| = \| y_0 \| = \| m_0 \| > 0 \). Hence, by Proposition 7.72, \( \| f \|_M = 1 \). By the simple version of Hahn-Banach Theorem, there exists \( x_\ast \in \mathcal{X}^* \) such that \( x_\ast |_M = f \) and \( \| x_\ast \| = \| f \|_M = 1 \).

Clearly, there does not exist \( (\eta_1, \eta_2, \ldots) \in l_1(\mathcal{Y}^*) \) such that \( x_\ast \) is given by \( \langle x_\ast, x \rangle = \sum_{n=1}^{\infty} \langle \eta_n, \xi_n \rangle \), \( \forall x = (\xi_1, \xi_2, \ldots) \in \mathcal{X} \). Hence, \( \mathcal{X}^* \neq l_1(\mathcal{Y}^*) \). \( \diamond \)
7.9.4 Second dual space

Definition 7.87 Let $X$ be a normed linear space over the field $K$ and $X^*$ be its dual. The dual space of $X^*$ is called the second dual of $X$ and is denoted by $X^{**}$.

Remark 7.88 Let $X$ be a normed linear space over the field $K$, $X^*$ be its dual, and $X^{**}$ be its second dual. Then, $X$ is isometrically isomorphic to a dense subset of a Banach space, which can be taken as a subspace of $X^{**}$. This can be proved as follows.

By Proposition 7.72, $X^{**}$ is a Banach space over $K$. $\forall x \in X$, define a functional $f : X^* \to K$ by $f(x^*) = \langle x^*, x \rangle$, $\forall x^* \in X^*$. Clearly, $f$ is a linear functional on $X^*$. Note that $|f(x^*)| = |\langle x^*, x \rangle| \leq \|x\| \|x^*\|$, $\forall x^* \in X^*$, by Proposition 7.72. Then, $\|f\| = \|x\|$. When $x = \vartheta_X$, then $\|f\| = 0 = \|x\|$. When $x \neq \vartheta_X$, then, by Proposition 7.85, $\exists x_0 \in X$ with $\|x_0\| = 1$ such that $\langle x_0, x \rangle = \|x\|$. Then, $\|f\| = \sup_{x^*, \|x^*\| \leq 1} |f(x^*)| \geq |f(x_0)| = \|x\|$. Hence, we have $\|f\| = \|x\|$. Hence, $f \in X^{**}$. Thus, we may define a natural mapping $\phi : X \to X^{**}$ by $\phi(x) = f$, $\forall x \in X$.

$\forall x_1, x_2 \in X$, $\forall \alpha, \beta \in K$, $\forall x^* \in X^*$, we have $\phi(\alpha x_1 + \beta x_2)(x^*) = \langle x^*, \alpha x_1 + \beta x_2 \rangle = \alpha \langle x^*, x_1 \rangle + \beta \langle x^*, x_2 \rangle$. Hence, $\phi$ is a linear functional. $\forall x \in X$, we have $\|\phi(x)\| = \|x\|$. Hence, $\phi$ is an isometry and $\phi$ is injective. Therefore, $\phi$ is an isometrical isomorphism between $X$ and $\phi(X)$. Clearly, $\phi(X) \subseteq X^{**}$ is a subspace. Then, by Proposition 7.17, $\phi(X) \subseteq X^{**}$ is a closed subspace. By Proposition 4.39, $\phi(X)$ is complete. Hence, $\phi(X)$ is a Banach space. By Proposition 3.5, $\phi(X)$ is dense in $\phi(X)$. This completes the proof of the remark.

Definition 7.89 Let $X$ be a normed linear space over the field $K$, $X^{**}$ be its second dual, and $\phi : X \to X^{**}$ be the natural mapping as defined in Remark 7.88. $X$ is said to be reflexive if $\phi(X) = X^{**}$, that is $X$ and $X^{**}$ are isometrically isomorphic. Then, we may label $X^{**}$ such that $X^{**} = X$ and $\phi = \text{id}_X$.

A reflexive normed linear space is clearly a Banach space.

Proposition 7.90 Let $X$ be a Banach space over the field $K$ and $X^*$ be its dual. Then, $X$ is reflexive if, and only if, $X^{**}$ is reflexive.

Proof Let $X^{**}$ be the second dual of $X$ and $X^{***}$ be the dual of $X^{**}$.

"Necessity" Let $X$ be reflexive. Then, $X^{**} = X$ isometrically isomorphically. Then, $X^* = (X^{**})^* = X^{***}$ isometrically isomorphically. Hence, $X^*$ is reflexive.

"Sufficiency" Let $X^*$ be reflexive. Then, $X^{***} = \phi_*(X^*)$, where $\phi_* : X^* \to X^{***}$ is the natural mapping on $X^*$. We will show that $X$ is reflexive by an argument of contradiction. Suppose $X$ is not reflexive. Let $\phi : X \to X^{**}$ be the natural mapping. $\phi$ is an isometrical isomorphism between
X and \( \phi(X) \). Since \( X \) is complete, then \( \phi(X) \) is also complete. Then, by Proposition 4.39, \( \phi(X) \) is a closed subspace of \( X^{**} \). Since \( X \) is not reflexive, then there exists \( y_{**} \in X^{**} \setminus \phi(X) \). Then, \( y_{**} \neq \vartheta_{**} = \phi(\vartheta) \).

Let \( \delta := \inf_{m_{**} \in \phi(X)} \| y_{**} - m_{**} \| \). By Proposition 4.10, \( \delta \in (0, \infty) \subset \mathbb{R} \).

Consider the subspace \( N := \{ m_{**} + \alpha y_{**} \in X^{**} \mid \alpha \in \mathbb{K}, m_{**} \in \phi(X) \} \).

Since \( y_{**} \notin \phi(X) \), then \( \forall n_{**} \in N \), \( \exists \alpha \in \mathbb{K} \) and \( \exists m_{**} \in \phi(X) \) such that \( n_{**} = \alpha y_{**} + m_{**} \). Define a linear functional \( f : N \to \mathbb{K} \) by \( f(n_{**}) = \alpha \).

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

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 reflective. Let \( \mathcal{Y} \) be a reflexive Banach space. Then, \( l_p(\mathcal{Y}), p \in (1, \infty) \subset \mathbb{R} \), are reflexive. \( l_1(\mathcal{Y}) \) is not reflexive by Example 7.86. Then, by Proposition 7.90, \( l_\infty(\mathcal{Y}^*) \) is not reflexive either. \( \diamond \)

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 \( \phi : X \to X^{**} \) 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, \( X = (\{ \vartheta \}, \mathbb{K}, \| \cdot \|) \). By Example 7.73, \( X^* = (\{ \vartheta \}, \mathbb{K}, \| \cdot \|) \). Then \( X^{**} = (\{ \vartheta \}, \mathbb{K}, \| \cdot \|) \). Clearly, \( \phi(\vartheta) = \vartheta_{**} \). Hence, \( \phi(X) = X^{**} \) and \( X \) is reflexive.

Case 2: \( n \in \mathbb{N} \). Let \( \{ e_1, \ldots, e_n \} \subseteq X \) be a basis of \( X \). \( \forall x \in X \), by Corollary 6.47, \( x \) can be uniquely expressed as \( \sum_{j=1}^n \alpha_j e_j \) for some \( \alpha_1, \ldots, \alpha_n \in \mathbb{K} \). \( \forall i = 1, \ldots, n \), define \( f_i : X \to \mathbb{K} \) by \( f_i(x) = \alpha_i \), \( \forall x = \sum_{j=1}^n \alpha_j e_j \in X \). 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 \( \beta_i \in X^* \). \( \forall x_{**} \in X^{**}, \forall i = 1, \ldots, n \), let \( \beta_i = \langle \langle x_{**}, e_i \rangle \rangle \in \mathbb{K} \). \( \forall x = \sum_{j=1}^n \alpha_j e_j \in X \), we have

\[
\langle \langle x_{**}, x \rangle \rangle = \sum_{j=1}^n \alpha_j \langle \langle x_{**}, e_j \rangle \rangle = \sum_{j=1}^n \beta_j \alpha_j = \sum_{k=1}^n \beta_j \langle \langle e_{**}, x \rangle \rangle
\]

\[
= \langle \langle \sum_{k=1}^n \beta_j e_{**}, x \rangle \rangle
\]
Hence, \( x_\ast = \sum_{j=1}^n \beta_j e_{\ast j} \). Therefore, \( X^\ast = \text{span} (\{ e_{\ast 1}, \ldots, e_{\ast n} \}) \). \( \forall x_{\ast \ast} \in X^{\ast \ast} \), \( \forall i = 1, \ldots, n \), let \( \gamma_i = \langle \langle x_{\ast \ast}, e_{\ast i} \rangle \rangle \in \mathbb{K} \). \( \forall x_\ast = \sum_{j=1}^n \beta_j e_{\ast j} \in X^\ast \), we have

\[
\langle \langle x_{\ast \ast}, x_\ast \rangle \rangle = \sum_{j=1}^n \beta_j \langle \langle x_{\ast \ast}, e_{\ast j} \rangle \rangle = \sum_{j=1}^n \beta_j \gamma_j = \sum_{j=1}^n \beta_j \langle \langle e_{\ast j}, \sum_{k=1}^n \gamma_k e_k \rangle \rangle
\]

\[
= \langle \langle x_\ast, \sum_{k=1}^n \gamma_k e_k \rangle \rangle = \langle \langle \phi \left( \sum_{k=1}^n \gamma_k e_k \right), x_\ast \rangle \rangle
\]

Hence, \( x_{\ast \ast} = \phi \left( \sum_{k=1}^n \gamma_k e_k \right) \) and \( X^{\ast \ast} \subseteq \phi(X) \). 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, \( x \in X \), \( X^\ast \) be the dual, and \( x_\ast \in X^\ast \). We say that \( x_\ast \) is aligned with \( x \) if \( \langle \langle x_\ast, x \rangle \rangle = \| x_\ast \| \| x \| \).

Clearly, \( \vartheta \) aligns with any vector in the dual and \( \vartheta_\ast \) is aligned with any vector in \( X \). By Proposition 7.85, \( \forall x \in X \) with \( x \neq \vartheta \), there exists a \( x_\ast \in X^\ast \) with \( x_\ast \neq \vartheta_\ast \) that is aligned with it.

**Definition 7.94** Let \( X \) be a normed linear space, \( x \in X \), \( X^\ast \) be the dual, and \( x_\ast \in X^\ast \). We say that \( x \) and \( x_\ast \) are orthogonal if \( \langle \langle x_\ast, x \rangle \rangle = 0 \).

**Definition 7.95** Let \( X \) be a normed linear space, \( S \subseteq X \), and \( X^\ast \) be the dual. The orthogonal complement of \( S \), denoted by \( S^\perp \), consists of all \( x_\ast \in X^\ast \) that is orthogonal to every vector in \( S \), that is \( S^\perp := \{ x_\ast \in X^\ast \mid \langle \langle x_\ast, x \rangle \rangle = 0, \forall x \in S \} \).

**Definition 7.96** Let \( X \) be a normed linear space, \( X^\ast \) be the dual, and \( S \subseteq X^\ast \). The pre-orthogonal complement of \( S \), denoted by \( S^\perp \), consists of all \( x \in X \) that is orthogonal to every vector in \( S \), that is \( S^\perp := \{ x \in X \mid \langle \langle x_\ast, x \rangle \rangle = 0, \forall x_\ast \in S \} \).

**Proposition 7.97** Let \( X \) be a normed linear space over the field \( \mathbb{K} \), \( M \subseteq X \) be a subspace, and \( y \in X \). Then, the following statements hold.

1. \( \delta := \inf_{m \in M} \| y - m \| = \max_{x_\ast \in M, \| x_\ast \| \leq 1} \text{Re} (\langle \langle x_\ast, y \rangle \rangle) \), where the maximum is achieved at some \( x_{\ast 0} \in M^\perp \) with \( \| x_{\ast 0} \| \leq 1 \). If the infimum is achieved at \( m_0 \in M \) then \( y - m_0 \) is aligned with \( x_{\ast 0} \).

2. If \( \exists m_0 \in M \) and \( \exists x_{\ast 0} \in M^\perp \) with \( \| x_{\ast 0} \| = 1 \) such that \( y - m_0 \) is aligned with \( x_{\ast 0} \), then the infimum is achieved at \( m_0 \) and the maximum is achieved at \( x_{\ast 0} \), that is \( \delta = \| y - m_0 \| = \langle \langle x_{\ast 0}, y \rangle \rangle = \text{Re} (\langle \langle x_{\ast 0}, y \rangle \rangle) \).
7.9. DUAL SPACES

Proposition 7.98 Let \( X \) be a normed linear space over \( K \), \( A, C \subseteq X \), and \( B, D \subseteq X^* \). 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) \( A^\perp \subseteq X^* \) is a closed subspace.

(iii) \( B^\perp \subseteq X \) is a closed subspace.

(iv) \( \overline{\{A \}} = \text{span}(A) \).
\textbf{Proof} 

(i) This follows directly from Definition 7.95.

(ii) \( \theta_x \in A^\perp \neq \emptyset \). For every \( x_{s_1}, x_{s_2} \in A^\perp \), \( \forall \alpha, \beta \in K, \forall x \in A \), we have 
\[ \langle \alpha x_{s_1} + \beta x_{s_2}, x \rangle = \alpha \langle x_{s_1}, x \rangle + \beta \langle x_{s_2}, x \rangle = 0. \] Hence, \( \alpha x_{s_1} + \beta x_{s_2} \in A^\perp \) and \( A^\perp \) is a subspace. By Proposition 4.10, we have \( y_x \in A^\perp \) and \( A^\perp \) is closed, by Proposition 7.98, we have \( \inf_{n \in \mathbb{N}} \langle x_n, x \rangle = \langle x_\star, x \rangle = 0 \). Then, \( y_x \in A^\perp \) 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, \( A \subseteq \{ A^\perp \} \). Then, by (iii), we have \( M := \text{span}(A) \subseteq \{ A^\perp \} \). On the other hand, \( A \subseteq M \), then, by (i), we have \( A^\perp \supseteq M^\perp \) and \( \{ A^\perp \} \subseteq \{ M^\perp \} \). \( \forall x_0 \in \{ M^\perp \} \), by Proposition 7.97, we have 
\[ \inf_{m \in M} \| x_0 - m \| = \max_{x_0 \in M^\perp, \| x_0 \| \leq 1} \text{Re}(\langle x_0, x_0 \rangle) = 0 \]
Since \( M \) is closed, by Proposition 4.10, \( x_0 \in M \). Hence, we have \( \{ A^\perp \} \subseteq \{ M^\perp \} \subseteq M \). Hence, \( M = \{ A^\perp \} \). Hence, (iv) follows.

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

\textbf{Proposition 7.99} Let \( X \) be a normed linear space over the field \( K \), \( M \subseteq X \) be a subspace, and \( y_\star \in X^\star \). Then, the following statements holds.

(i) \( \delta := \min_{x_\star \in M^\perp} \| y_\star - x_\star \| = \sup_{m \in M, \| m \| \leq 1} \text{Re}(\langle y_\star, m \rangle) =: \delta \), where 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 \( m_0 \) is aligned with \( y_\star - x_{s_0} \).

(ii) If \( \exists m_0 \in M \) with \( \| m_0 \| = 1 \) and \( \exists x_{s_0} \in M^\perp \) such that \( y_\star - x_{s_0} \) is aligned with \( m_0 \), then the minimum is achieved at \( x_{s_0} \) and the supremum is achieved at \( m_0 \), that is \( \delta = \| y_\star - x_{s_0} \| = \text{Re}(\langle y_\star, m_0 \rangle) \).

\textbf{Proof} 

(i) \( \forall x_\star \in M^\perp, \forall m \in M \) with \( \| m \| \leq 1 \), we have 
\[ \text{Re}(\langle y_\star, m \rangle) = \text{Re}(\langle y_\star - x_\star, m \rangle) \leq \| y_\star - x_\star \| \| m \| \leq \| y_\star - x_\star \| \]
where we have applied Proposition 7.72 in the third inequality. Then, we have 
\[ \delta = \inf_{x_\star \in M^\perp} \| y_\star - x_\star \| \leq \sup_{m \in M, \| m \| \leq 1} \text{Re}(\langle y_\star, m \rangle) = \delta \]
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 \( y_\star \in M^\perp \) the minimum is achieved at a unique vector \( x_{s_0} := y_\star \). Take \( m_0 = \theta \) with \( \| m_0 \| = 0 \leq 1 \). Then, 
\[ \delta = \text{Re}(\langle y_\star, m_0 \rangle) \leq \sup_{m \in M, \| m \| \leq 1} \text{Re}(\langle y_\star, m \rangle) = \delta \leq 0. \] Hence,
7.9. DUAL SPACES

\[ \delta = \min_{x_* \in M^*} \| y_* - x_* \| = \sup_{m \in M, \|m\| \leq 1} \Re \left( \langle \langle y_*, m \rangle \rangle \right) = \tilde{\delta} \] and the minimum is achieved at \( x_{*0} = y_* \in M^* \). If the supremum is achieved at \( m_0 \in M \) with \( \| m_0 \| \leq 1 \). Then, clearly \( m_0 \) is aligned with \( y_* - x_{*0} \). This case is proved.

Case 2: \( \delta > 0 \). Then, \( y_* \notin M^* \). Consider the subspace \( M \). Then, we may define a functional \( f : M \to K \) by \( f(m) = \langle \langle y_*, m \rangle \rangle \). Clearly, \( f \) is a linear functional on \( M \). \( \forall m \in M \) with \( \| m \| \leq 1 \), we have either \( f(m) = 0 \), then \( |f(m)| = 0 \leq \sup_{m \in M, \|m\| \leq 1} \Re \left( \langle \langle y_*, m \rangle \rangle \right) = \delta \); or \( f(m) \neq 0 \), then \( |f(m)| = \frac{f(m)}{\|f(m)\|} \leq \delta \). Hence, \( \| f \|_M := \sup_{m \in M, \|m\| \leq 1} |f(m)| \leq \delta \leq \delta < +\infty \). By the simple version of Hahn-Banach Theorem, \( \exists y_{0*} \in X^* \) such that \( y_{0*} |_M = f \) and \( \| y_{0*} \| = \| f \|_M \leq \delta \). Let \( x_{0*} := y_* - y_{0*} \). \( \forall m \in M \), we have \( \langle \langle x_{0*}, m \rangle \rangle = \langle \langle y_*, m \rangle \rangle - \langle \langle y_{0*}, m \rangle \rangle = 0 \). Hence, \( x_{0*} \in M^* \). Note that \( \delta \geq \delta \geq \| f \|_M = \| y_* - x_* \| \geq \inf_{x_* \in M^*} \| y_* - x_* \| = \delta \).

Hence, \( \delta = \min_{x_* \in M^*} \| y_* - x_* \| = \sup_{m \in M, \|m\| \leq 1} \Re \left( \langle \langle y_*, m \rangle \rangle \right) = \delta \) and the minimum is achieved at some \( x_{0*} \in M^* \). If the supremum is achieved at \( m_0 \in M \) with \( \| m_0 \| \leq 1 \), then \( \delta = \| y_* - x_{0*} \| \geq \| y_* - x_{0*} \| \| m_0 \| \geq | \langle \langle y_* - x_{0*}, m_0 \rangle \rangle | = | \langle \langle y_*, m_0 \rangle \rangle | \geq \Re \left( \langle \langle y_*, m_0 \rangle \rangle \right) = \delta \), where the second inequality follows from Proposition 7.72. Hence, \( \langle \langle y_* - x_{0*}, m_0 \rangle \rangle = \langle \langle y_* - x_{0*}, m_0 \rangle \rangle = \langle \langle y_*, m_0 \rangle \rangle = \Re \left( \langle \langle y_*, m_0 \rangle \rangle \right) \leq \delta \). Hence, the result follows.

This completes the proof of the proposition.

\[ \Box \]

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) \( S^* \) is isometrically isomorphic to \( X^*/S^* \).

(ii) If \( X \) is reflexive and \( S \) is closed, then \( S \) is reflexive.

\[ \text{Proof} \quad \text{(i) Define a mapping } A : X^* \to S^* \text{ by } \langle \langle A(x_*), s \rangle \rangle = \langle \langle x_*, s \rangle \rangle, \forall x_* \in X^* \text{ and } \forall s \in S. \forall x_* \in X^*, \text{ we will show that } A(x_*) \in S^*. \text{ Clearly, } A(x_*) \text{ is a linear functional on } S. \langle \langle A(x_*), s \rangle \rangle := \sup_{s \in S, \|s\| \leq 1} | \langle \langle x_*, s \rangle \rangle | = \sup_{s \in S, \|s\| \leq 1} | \langle \langle x_*, s \rangle \rangle | \leq \sup_{s \in S, \|s\| \leq 1} \| x_* \| \| s \| \leq \| x_* \| < +\infty \). \text{ Hence, } A(x_*) \in S^* \text{ and } S \text{ is well-defined.} \]

\[ \forall x_{*1}, x_{*2} \in X^*, \forall \alpha, \beta \in K, \forall s \in S, \text{ we have} \]

\[ \langle \langle A(\alpha x_{*1} + \beta x_{*2}), s \rangle \rangle = \langle \langle \alpha x_{*1} + \beta x_{*2}, s \rangle \rangle = \alpha \langle \langle x_{*1}, s \rangle \rangle + \beta \langle \langle x_{*2}, s \rangle \rangle \]

\[ = \alpha \langle \langle A(x_{*1}), s \rangle \rangle + \beta \langle \langle A(x_{*2}), s \rangle \rangle = \langle \langle \alpha A(x_{*1}) + \beta A(x_{*2}), s \rangle \rangle \]

Hence, \( A \) is a linear function. Since \( \| A x_* \| \leq \| x_* \|, \forall x_* \in X^* \), then \( A \in B(X^*, S^*) \) with \( \| A \| \leq 1 \).
By Proposition 7.68, \( \mathcal{N}(A) \subseteq X^* \) is a closed subspace. \( \forall x \in \mathcal{N}(A) \), we have \( Ax = \partial S^* \). \( \forall s \in S, 0 = \langle Ax, s \rangle = \langle x, s \rangle \). Hence, \( x \in S^\perp \). Therefore, \( \mathcal{N}(A) \subseteq S^\perp \). On the other hand, \( \forall x \in S^\perp, \forall s \in S \), we have \( 0 = \langle x, s \rangle = \langle Ax, s \rangle \). Then, \( Ax = \partial S^* \). Hence, \( x \in \mathcal{N}(A) \). Thus, \( S^\perp \subseteq \mathcal{N}(A) \). In conclusion, \( S^\perp = \mathcal{N}(A) \).

By Proposition 7.45, \( X^*/S^\perp \) is a Banach space. Let \( \phi : X^* \to X^*/S^\perp \) be the natural homomorphism. By Proposition 7.70, there exists \( A \in B(X^*/S^\perp, S^*) \) such that \( A = A_D \circ \phi \), \( A_D \) is injective, and \( \|A_D\| = \|A\| \leq 1 \).

\( \forall s \in S^* \), by the simple version of Hahn-Banach Theorem, there exists \( x \in X^* \) such that \( x|_S = s \) and \( \|x\| = \|s\| \). \( \forall s \in S \), we have \( \langle s, s \rangle = \langle Ax, s \rangle = \langle x, s \rangle \). Hence, \( s = A_D(x) \). Then, \( A_D \) is surjective. Thus, \( A_D \) is bijective.

\( \forall [x] \in X^*/S^\perp \), \( \|x\| = \inf_{y \in S^\perp} \|x - y\| \), by Proposition 7.44. By Proposition 7.99, we have

\[
\|x\| = \min_{y \in S^\perp} \|x - y\| = \sup_{s \in S} \Re \left( \langle x, s \rangle \right) \\
= \sup_{s \in S} \Re \left( \langle Ax, s \rangle \right) = \sup_{s \in S} \Re \left( \langle A_D(x), s \rangle \right) \\
= \sup_{s \in S} \Re \left( \langle A_D(x), s \rangle \right) \leq \sup_{s \in S} \|A_D(x)\| = \|x\| \\
= \inf_{y \in \mathcal{N}(A)} \|x - y\| = \inf_{y \in S^\perp} \|x - y\| \\
\leq \inf_{y \in S^\perp} \|x\| \|x - y\| \leq \|x\|^2
\]

where we have applied Proposition 7.72 in the third inequality and Proposition 7.64 in the fifth inequality. Therefore, we have \( \|x\| = \|A_D(x)\| \) and \( A_D \) is an isometry. Thus, \( A_D \) is an isometrical isomorphism between \( X^*/S^\perp \) and \( S^* \). Hence, \( (i) \) is established.

(ii) Let \( X \) be reflexive and \( S \) be a closed subspace. Let \( \psi : S \to S^{**} \) be the natural mapping. All we need to show is that \( \psi(S) = S^{**} \). Fix a \( s \in S^{**} \). Define a functional \( \tau : X^* \to K \) by \( \tau(x) = \langle s, Ax \rangle \), \( \forall x \in X^* \). It is easy to show that \( \tau \) is a linear functional on \( X^* \). \( \forall x \in X^* \) with \( \|x\| \leq 1 \), we have \( |\tau(x) - \langle s, Ax \rangle| \leq \|s\| \|Ax\| \leq \|s\| \|x\| \leq \|s\| < +\infty \), where we have applied Propositions 7.72 and 7.64. Hence, \( \tau \in X^{**} \).

Since \( X \) is reflexive, then, by Remark 7.88 and Definition 7.89, \( \exists x_0 \in X \) such that \( \tau(x) = \langle x, x_0 \rangle \), \( \forall x \in X^* \). \( \forall y_0 \in S^\perp \), we have \( \langle y_0, x_0 \rangle = \tau(y_0) = \langle s, y_0 \rangle = \langle s, \partial S^* \rangle = 0 \), where the third equality follows from the fact that \( S^\perp = \mathcal{N}(A) \). Hence, \( x_0 \in S^\perp \) is \( S \), by Proposition 7.98. Hence, \( \forall s \in S^* \), \( \exists x \in X^* \) such that \( Ax = A_D(x) = s \).
since $A_D$ is an isometrical isomorphism. Then, we have $\langle(s_*, A)x_0\rangle = \langle(s_*, x_0)\rangle$. This implies that $s_* = \psi(x_0)$ and $\psi(S) = S^{**}$. Hence, (ii) is established.

This completes the proof of the proposition.

\section{The Open Mapping Theorem}

\textbf{Definition 7.101} Let $\mathcal{X} := (X, \mathcal{O}_X)$ and $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be topological spaces and $A : \mathcal{X} \to \mathcal{Y}$. $A$ 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.

\textbf{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 $\alpha \overline{S} = \overline{\alpha S}$.

(iii) $\overline{S + T} \subseteq \overline{S} + \overline{T}$.

(iv) If $\alpha \neq 0$, then $(\alpha S)^0 = \alpha S^0$.

(v) $S^0 + T^0 \subseteq (S + T)^0$.

\textbf{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 $\alpha \overline{S} = \emptyset = \alpha \overline{S}$. Case 2: $\alpha = 0$ and $S \neq \emptyset$. Then, $\overline{S} \neq \emptyset$ and $\alpha \overline{S} = \{0\} = \alpha \overline{S}$.

Case 3: $\alpha \neq 0$. $\forall x \in \overline{\alpha S}$, by Proposition 4.13, $\exists (x_n)_{n=1}^{\infty} \subseteq \alpha S$ such that $\lim_{n \to \infty} x_n = x$. Note that $(\alpha x_n)_{n=1}^{\infty} \subseteq \alpha S$ 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, $x \in \overline{\alpha S}$. This shows that $\overline{\alpha S} \subseteq \overline{\alpha S}$.

On the other hand, $\forall x \in \overline{\alpha S}$, then $\alpha^{-1} x \in \overline{S}$. By Proposition 4.13, $\exists (x_n)_{n=1}^{\infty} \subseteq \overline{S}$ 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, $\overline{\alpha S} \subseteq \overline{\overline{\alpha S}}$. Hence, $\overline{\alpha S} = \overline{s \overline{S}}$.

(ii) $\forall x \in \alpha \overline{S}$ if, and only if, $\alpha^{-1} x \in \overline{S}$ if, and only if, $x \in \alpha \overline{S}$. Hence, $\alpha \overline{S} = \overline{\alpha S}$.

(iii) $\forall x \in (\alpha S)^0$, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $B(x, \delta) \subseteq \alpha S$. Then, $B(\alpha^{-1} x, \delta/|\alpha|) = \alpha^{-1} B(x, \delta) \subseteq S$. Hence, $\alpha^{-1} x \in S^0$. Then, $x \in \alpha S^0$. This shows that $(\alpha S)^0 \subseteq \alpha S^0$.

On the other hand, $\forall x \in \alpha S^0$, we have $\alpha^{-1} x \in S^0$. $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $B(\alpha^{-1} x, \delta) \subseteq S$. Then, $B(x, |\alpha| \delta) = \alpha B(\alpha^{-1} x, \delta) \subseteq \alpha S$. Hence, $x \in (\alpha S)^0$. This shows that $\alpha S^0 \subseteq (\alpha S)^0$. Hence, we have $(\alpha S)^0 = \alpha S^0$. 


(iv) \( \forall x \in \overline{S + T}, \exists \tilde{s} \in \overline{S} \) and \( \exists \tilde{t} \in \overline{T} \) such that \( \bar{x} = \tilde{s} + \tilde{t} \). By Proposition 4.13, \( \exists (s_n)_{n=1}^{\infty} \subseteq S \) and \( \exists (t_n)_{n=1}^{\infty} \subseteq T \) such that \( \lim_{n \to \infty} s_n = \tilde{s} \) and \( \lim_{n \to \infty} t_n = \tilde{t} \). Then, \( (s_n + t_n)_{n=1}^{\infty} \subseteq \overline{S + T} \) and \( \lim_{n \to \infty} (s_n + t_n) = \tilde{s} + \tilde{t} = \bar{x} \), by Propositions 7.23, 3.66, and 3.67. By Proposition 4.13, \( \bar{x} \in \overline{S + T} \).

Hence, \( \overline{S + T} \subseteq \overline{S} \overline{+} \overline{T} \).

(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, \( \exists r_s, r_t \in (0, \infty) \subseteq \mathbb{R} \) such that \( B(s_0, r_s) \subseteq S \) and \( B(t_0, r_t) \subseteq T \). Thus, we have \( B(x, r_s + r_t) = B(s_0, r_s) + B(t_0, r_t) \subseteq S + T \) 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 \( A \in B(X, Y) \) be surjective. Then, \( A \) is an open mapping. Furthermore, if \( A \) is injective, then \( A_{\text{inv}} \in B(Y, X) \).

**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 \( B_{Y}(\vartheta y, \delta) \subseteq A(B_{X}(\vartheta x, 1)) \).

**Proof of claim:** Let \( S_n := B_{X}(\vartheta x, 2^{-n}) \), \( \forall n \in \mathbb{Z}_+ \). Since \( A \) is linear and surjective and \( \bigcup_{k=1}^{\infty} kS_1 = X \), 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, \( A(S_1) \subseteq (0, \infty) \subseteq \mathbb{R} \) such that \( B_{Y}(\vartheta y, \delta) \subseteq A(S_1) \). Then, we have \( B_{Y}(\vartheta y, \delta) = B_{Y}(\vartheta y, \delta - \vartheta y) \subseteq A(S_1) = A(S_1 - A(S_1)) \). 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 \( A \), \( B_{Y}(\vartheta y, 2^{-n}\delta) = 2^{-n}B_{Y}(\vartheta y, \delta) \subseteq 2^{-n}A(S_0) = A(2^{-n}S_0) = A(S_0) \subseteq Y, \forall n \in \mathbb{Z}_+ \).

Now, we proceed to show that \( B_{Y}(\vartheta y, \delta/2) \subseteq A(S_0) \). Fix an arbitrary vector \( y \in B_{Y}(\vartheta y, \delta/2) \). Then, \( y \in A(S_1) \) and \( \exists x_1 \in S_1 \), such that \( \| y - Ax_1 \| < 2^{-2}\delta \). This implies that \( y - Ax_1 \in B_{Y}(\vartheta y, 2^{-2}\delta) \subseteq A(S_2) \). 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 \( \| y - \sum_{k=1}^{n} Ax_k \| < 2^{-n}\delta \). This implies that \( y - \sum_{k=1}^{n} Ax_k \in B_{Y}(\vartheta y, 2^{-n}\delta) \subseteq A(S_{n+1}) \). Since \( x_n \in S_n \) and \( \| x_n \| < 2^{-n} \), \( \forall n \in \mathbb{N} \), then \( \sum_{k=1}^{n} \| x_k \| < 1 \). By Proposition 7.27, we have \( \sum_{k=1}^{n} x_k = x \in Y \). It is easy to show that \( x \in S_0 \). Then, \( y = \lim_{n \to \infty} \sum_{k=1}^{n} Ax_k = \lim_{n \to \infty} A \left( \sum_{k=1}^{n} x_k \right) = Ax \), where we have applied Proposition 3.66. Thus, \( y \in A(S_0) \) and \( B_{Y}(\vartheta y, \delta/2) \subseteq A(S_0) \).
This completes the proof of the claim. \[\square\]

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 \(r \in (0, \infty) \subseteq \mathbb{R}\) such that \(B_X(x, r) \subseteq O\). By Claim 7.103.1, \(\exists \delta \in (0, \infty) \subseteq \mathbb{R}\) such that \(B_Y(\vartheta y, \delta) \subseteq A(B_X(\vartheta x, r))\). By the linearity of \(A\), we have \(B_0(y, \delta) = y + B_0(\vartheta y, \delta) \subseteq Ax + A(B_X(\vartheta x, r)) = A(B_X(x, r)) \subseteq A(O)\). 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. \[\square\]

**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 \(\exists M \in [0, \infty) \subseteq \mathbb{R}\) 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, \(A \in B(X_1, X_2)\). By Open Mapping Theorem, \(A_{\text{inv}} \in B(X_2, X_1)\). Then, \(\exists M \in [0, \infty) \subseteq \mathbb{R}\) such that \(\| x \|_1 = \| Ax \|_1 \leq M \| x \|_2\), \(\forall x \in X\). Now, take \(K = \max \{ M, M \} + 1 \in (0, \infty) \subseteq \mathbb{R}\). Then, we have \(\| x \|_1 / K \leq \| x \|_2 \leq K \| x \|_1\), \(\forall x \in X\). Hence, the two norms are equivalent. This completes the proof of the proposition. \[\square\]

**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 \(\lim_{n \in \mathbb{N}} x_n \subseteq X\). \(\forall x \in X\). There exists \(x_0 \in X\) 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, \(\forall x \in X\), \(\exists x_0 \in X\), \(\exists y_0 \in Y\), we have \(\| x \| \leq \| x \|_1\). By Proposition 7.104, \(\| \cdot \|\) and \(\| \cdot \|_1\) are equivalent. Then, \(\exists M \in [0, \infty) \subseteq \mathbb{R}\) 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. \[\square\]

Recall that the graph of a function \(f : X \to Y\) is the set \(\{(x, y) \in X \times Y \mid x \in X, y = f(x)\}\). Then, we have the following alternative statement of the Closed Graph Theorem.
Theorem 7.106 (Closed Graph Theorem) Let \( \mathcal{X} \) and \( \mathcal{Y} \) be Banach spaces over the field \( \mathbb{K} \) and \( A : \mathcal{X} \rightarrow \mathcal{Y} \) be a linear operator. Then, \( A \in B(\mathcal{X}, \mathcal{Y}) \) if, and only if, the graph of \( A \) is a closed set in \( \mathcal{X} \times \mathcal{Y} \).

**Proof**  

"Sufficiency" Let the graph of \( A \) be graph \((A)\). Fix any sequence \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) with \( \lim_{n \in \mathbb{N}} x_n = x_0 \in \mathcal{X} \) and \( \lim_{n \in \mathbb{N}} Ax_n = y_0 \in \mathcal{Y} \). Note that \( ((x_n, Ax_n))_{n=1}^{\infty} \subseteq \text{graph}(A) \subseteq \mathcal{X} \times \mathcal{Y} \) and \( \lim_{n \in \mathbb{N}} (x_n, Ax_n) = (x_0, y_0) \), by Proposition 3.67. Then, by Proposition 4.13, we have \( (x_0, y_0) \in \text{graph}(A) \). By the Closed Graph Theorem, Theorem 7.105, we have \( A \in B(\mathcal{X}, \mathcal{Y}) \).

"Necessity" Fix any sequence \( (x_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) with \( \lim_{n \in \mathbb{N}} x_n = x_0 \in \mathcal{X} \) and \( \lim_{n \in \mathbb{N}} Ax_n = y_0 \). By Proposition 3.66, we have \( y_0 = \lim_{n \in \mathbb{N}} Ax_n = Ax_0 \). Hence, \( (x_0, y_0) \in \text{graph}(A) \). By Proposition 3.3, \( \text{graph}(A) \) is closed in \( \mathcal{X} \times \mathcal{Y} \).

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

**Proposition 7.107** Let \( \mathcal{X} \) be a Banach space over the field \( \mathbb{K} \), \( \mathcal{Y} \) be a normed linear space over the same field, and \( \mathcal{F} \subseteq B(\mathcal{X}, \mathcal{Y}) \). Assume that \( \forall x \in \mathcal{X}, \exists M_x \in [0, \infty) \subset \mathbb{R} \) such that \( \|Tx\| \leq M_x, \forall T \in \mathcal{F} \). Then, \( \exists M \in [0, \infty) \subset \mathbb{R}, \) such that \( \|T\| \leq M, \forall T \in \mathcal{F} \).

**Proof**  

By Baire Category Theorem, \( \mathcal{X} \) is second category everywhere. \( \forall T \in \mathcal{F} \), let \( f : \mathcal{X} \rightarrow \mathbb{R} \) be given by \( f(x) = \|Tx\|, \forall x \in \mathcal{X} \). By Propositions 7.21 and 3.12, \( f \) is a continuous real-valued function. By Uniform Boundedness Principle, there exist an open set \( O \subseteq \mathcal{X} \) with \( O \neq \emptyset \) and \( \bar{M} \in [0, \infty) \subset \mathbb{R} \) such that \( \|Tx\| \leq \bar{M}, \forall T \in \mathcal{F}, \forall x \in O \). Since \( O \) is nonempty and open, then \( \exists \mathcal{B}_\mathcal{X}(x_0, r) \subseteq O \) for some \( x_0 \in \mathcal{X} \) and some \( r \in (0, \infty) \subset \mathbb{R} \). \( \forall x \in \mathcal{X} \) with \( \|x\| < r, x + x_0 \in \mathcal{B}_\mathcal{X}(x_0, r) \subseteq O \) and \( \forall T \in \mathcal{F} \), we have

\[
\|Tx\| = \|T(x + x_0) - Tx_0\| \leq \|T(x + x_0)\| + \|Tx_0\| \leq \bar{M} + M_{x_0}
\]

Hence, \( \forall T \in \mathcal{F} \), we have, \( \forall \epsilon \in (0, r) \subset \mathbb{R}, \)

\[
\|T\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} \|Tx\| = \sup_{x \in \mathcal{X}, \|x\| \leq 1} (r - \epsilon)^{-1} \|T((r - \epsilon)x)\| \leq \frac{\bar{M} + M_{x_0}}{r - \epsilon}
\]

By the arbitrariness of \( \epsilon \), we have \( \|T\| \leq (\bar{M} + M_{x_0})/r =: M < +\infty \). This completes the proof of the proposition. \( \square \)

**Proposition 7.108** Let \( \mathcal{X} \) be a Banach space over the field \( \mathbb{K} \), \( \mathcal{Y} \) be a normed linear space over the same field, and \( (T_n)_{n=1}^{\infty} \subseteq B(\mathcal{X}, \mathcal{Y}) \). Assume that \( \forall x \in \mathcal{X}, \lim_{n \in \mathbb{N}} T_n x = T(x) \in \mathcal{Y} \). Then, \( T \in B(\mathcal{X}, \mathcal{Y}) \).
Proof \( \forall x_1, x_2 \in X, \forall \alpha, \beta \in \mathbb{K}, \) we have
\[
T(\alpha x_1 + \beta x_2) = \lim_{n \to \infty} T_n(\alpha x_1 + \beta x_2) = \lim_{n \to \infty} (\alpha T_n x_1 + \beta T_n x_2) = \alpha Tx_1 + \beta Tx_2
\]
by Propositions 7.23 and 3.66. Hence, \( T \) is linear. \( \forall x \in X, \) by Propositions 7.21 and 3.66, \( \lim_{n \to \infty} \| T_n x \| = \| Tx \| < +\infty. \) Then, \( \exists M_x \in [0, \infty) \subset \mathbb{R} \) such that \( \| T_n x \| \leq M_x, \forall n \in \mathbb{N}. \) By Proposition 7.107, \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \| T_n \| \leq M, \forall n \in \mathbb{N}. \) \( \forall x \in X \) 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, \( T \in \mathcal{B}(X, Y). \) This completes the proof of the proposition.

7.11 The Adjoints of Linear Operators

Proposition 7.109 Let \( X \) and \( Y \) be normed linear spaces over the field \( \mathbb{K} \) and \( A \in \mathcal{B}(X, Y). \) The adjoint operator of \( A \) is \( A' : Y^* \to X^* \) defined by
\[
\langle \langle A'(y_*), x \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle ; \quad \forall x \in X, \forall y_* \in Y^*
\]
Then, \( A' \in \mathcal{B}(Y^*, X^*) \) with \( \| A' \| = \| A \|. \)

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 \( A' \in \mathcal{B}(Y^*, X^*) \) and \( \| A' \| \leq \| A \|. \)

On the other hand, \( \forall x \in X, \) we have either \( Ax = 0, \) then \( \| Ax \| = 0 \leq \| A' \| \| x \|; \) or \( Ax \neq 0, \) 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.

Proposition 7.110 Let \( X, Y, \) and \( Z \) be normed linear spaces over the field \( \mathbb{K}. \) Then, the following statements hold.

(i) \( \text{id}_X' = \text{id}_X. \)

(ii) If \( A_1, A_2 \in \mathcal{B}(X, Y), \) then \( (A_1 + A_2)' = A'_1 + A'_2. \)
(iii) If \( A \in B(\mathcal{X}, \mathcal{Y}) \) and \( \alpha \in \mathbb{K} \), then \((\alpha A)' = \alpha A'\).

(iv) If \( A_1 \in B(\mathcal{X}, \mathcal{Y}) \) and \( A_2 \in B(\mathcal{Y}, \mathcal{Z}) \), then \((A_2A_1)' = A_1'A_2'\).

(v) If \( A \in B(\mathcal{X}, \mathcal{Y}) \) and \( A \) has a bounded inverse, then \((A^{-1})' = (A')^{-1}\).

**Proof**

(i) \( \forall x_*, \forall x \in \mathcal{X}, \langle \langle \text{id}_x'(x_*), x \rangle \rangle = \langle \langle \text{id}_x(x), x \rangle \rangle \). Hence, the result follows.

(ii) \( \forall y_*, \forall x \in \mathcal{X}, \langle \langle (A_1 + A_2)'y_*, x \rangle \rangle = \langle \langle y_*, (A_1 + A_2)x \rangle \rangle = \langle \langle y_*, A_1x \rangle \rangle + \langle \langle y_*, A_2x \rangle \rangle = \langle \langle A_1'y_*, x \rangle \rangle + \langle \langle A_2'y_*, x \rangle \rangle = \langle \langle A_1'y_*, A_2'y_*, x \rangle \rangle = \langle \langle (A_1 + A_2)'y_*, x \rangle \rangle \). Hence, the result follows.

(iii) \( \forall y_*, \forall x \in \mathcal{X}, \langle \langle (\alpha A)'y_*, x \rangle \rangle = \alpha \langle \langle A'y_*, x \rangle \rangle = \alpha \langle \langle A'y_*, x \rangle \rangle = \langle \langle (\alpha A)'y_*, x \rangle \rangle \). Hence, the result follows.

(iv) \( \forall z_* \in \mathcal{Z}^*, \forall x \in \mathcal{X}, \langle \langle (A_2A_1)'z_*, x \rangle \rangle = \langle \langle z_*, (A_2A_1)x \rangle \rangle = \langle \langle z_*, A_2(A_1)x \rangle \rangle = \langle \langle A_2z_*, x \rangle \rangle = \langle \langle A_2z_*, A_1x \rangle \rangle = \langle \langle A_1'(A_2z_*) \rangle \rangle = \langle \langle (A_1'(A_2z_*) \rangle \rangle \rangle \). Hence, the result follows.

(v) By (i) and (iv), we have \((A^{-1})'A' = (AA^{-1})' = \text{id}'_{\mathcal{Y}} = \text{id}'_{\mathcal{Y}} \) and \((A^{-1})' = (A')^{-1} = \text{id}'_{\mathcal{X}} = \text{id}'_{\mathcal{X}}\). By Proposition 2.4, we have \((A^{-1})' = (A')^{-1}\).

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

**Proposition 7.111** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over the field \( \mathbb{K} \), \( A \in B(\mathcal{X}, \mathcal{Y}), \phi_{\mathcal{X}} : \mathcal{X} \rightarrow \mathcal{X}^* \) and \( \phi_{\mathcal{Y}} : \mathcal{Y} \rightarrow \mathcal{Y}^* \) be the natural mappings on \( \mathcal{X} \) and \( \mathcal{Y} \), respectively, and \( A'' : \mathcal{X}^* \rightarrow \mathcal{Y}^* \) be the adjoint of the adjoint of \( A \). Then, we have \( A'' \circ \phi_{\mathcal{X}} = \phi_{\mathcal{Y}} \circ A \).

**Proof** \( \forall x \in \mathcal{X}, \forall y_* \in \mathcal{Y}^*, \) we have

\[
\langle \langle A''(\phi_{\mathcal{X}}(x)), y_* \rangle \rangle = \langle \langle \phi_{\mathcal{X}}(x), A'y_* \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle = \langle \langle \phi_{\mathcal{Y}}(Ax), y_* \rangle \rangle
\]

Hence, the desired result follows. This completes the proof of the proposition. \(\square\)

By Proposition 7.111 and Remark 7.88, \( \forall A, B \in B(\mathcal{X}, \mathcal{Y}) \), we have \( A' = B' \) if, and only if, \( A = B \).

**Proposition 7.112** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over the field \( \mathbb{K} \) and \( A \in B(\mathcal{X}, \mathcal{Y}) \). Then,

\[
(\mathcal{R}(A))^\perp = \mathcal{N}(A') \quad \text{and} \quad (\mathcal{R}(A'))^\perp = \mathcal{N}(A)
\]

**Proof** Fix a vector \( y_* \in \mathcal{N}(A') \). \( \forall y \in \mathcal{R}(A), \exists x \in \mathcal{X} \) such that \( y = Ax \). Then, \( \langle \langle y_*, y \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle = \langle \langle A'y_*, x \rangle \rangle = \langle \langle \partial_{\mathcal{X}^*}, x \rangle \rangle = 0 \).

Hence, \( y_* \in (\mathcal{R}(A))^\perp \). This shows that \( \mathcal{N}(A') \subseteq (\mathcal{R}(A))^\perp \).

On the other hand, fix a vector \( y_* \in (\mathcal{R}(A))^\perp \). \( \forall x \in \mathcal{X} \), we have \( \langle \langle A'y_*, x \rangle \rangle = \langle \langle y_*, Ax \rangle \rangle = 0 \). Hence, we have \( A'y_* = \partial_{\mathcal{X}^*} \). This shows that \( (\mathcal{R}(A))^\perp \subseteq \mathcal{N}(A') \).
7.11. THE ADJOINTS OF LINEAR OPERATORS

Hence, we have \((\mathcal{R}(A))^\perp = \mathcal{N}(A')\).

Fix a vector \(x \in \mathcal{N}(A)\). \(\forall x_* \in \mathcal{R}(A')\), \(\exists y_* \in Y^*\) such that \(x_* = A'y_*\).

Then, \(<x_*, x> = <y_*, Ax> = <y_*, \vartheta y> = 0\). 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 \(x \in \perp(\mathcal{R}(A'))\). \(\forall y_* \in Y^*\), we have \(0 = <A'y_*, x> = <y_*, Ax>\). 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 \(X\) and \(Y\) be Banach spaces over the field \(K\) and \(A \in B(X,Y)\). Assume that \(\mathcal{R}(A) \subseteq Y\) is closed. Then, there exists \(K \in [0,\infty) \subset \mathbb{R}\) such that, \(\forall y \in \mathcal{R}(A)\), there exists \(x \in X\) such that \(y = Ax\) and \(\|x\| \leq K \|y\|\).

**Proof** Since \(A\) is continuous, then \(\mathcal{N}(A) \subseteq X\) is closed. By Proposition 7.45, \(X/\mathcal{N}(A)\) is a Banach space. Let \(\phi : X \to X/\mathcal{N}(A)\) be the natural homomorphism, which is a bounded linear function by Proposition 7.69. By Proposition 7.70, there exists a bounded linear function \(A_D : X/\mathcal{N}(A) \to \mathcal{R}(A)\) 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, \(A_D : X/\mathcal{N}(A) \to \mathcal{R}(A)\) is a bijective bounded linear operator. By Open Mapping Theorem, \(A_D^{-1} \in B(\mathcal{R}(A), X/\mathcal{N}(A))\).

\(\forall y \in \mathcal{R}(A)\), Let \([x] := A_D^{-1}y \in X/\mathcal{N}(A)\). 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, \([x] = [\partial \chi]\). Take \(x = \partial \chi\), 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 \(X\) and \(Y\) be normed linear spaces over the field \(K\) and \(A \in B(X,Y)\). Assume that \(\mathcal{R}(A)\) is closed in \(Y\). Then,

\[\mathcal{R}(A) = \perp(\mathcal{N}(A'))\]

If, in addition, \(X\) and \(Y\) are Banach spaces, then

\[\mathcal{R}(A') = (\mathcal{N}(A))^\perp\]
Proof. Fix $y \in \mathcal{R}(A)$. Then $\exists x \in X$ such that $y = Ax$. $\forall y_* \in \mathcal{N}(A')$, we have $\langle y_* , y \rangle = \langle A'y_* , x \rangle = \langle \vartheta_{\mathcal{X}'}, x \rangle = 0$. 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, $\delta := \inf_{k \in \mathcal{R}(A)}\|y - k\| = \max_{y_* \in (\mathcal{R}(A))^{\perp}}\|y_*\| \leq 1$. By Proposition 7.112, we have $(\mathcal{R}(A))^{\perp} = \mathcal{N}(A) \ni y_0$. Then, $\delta = \max_{y_* \in (\mathcal{R}(A))^{\perp}}\|y_*\| = 0$. By Proposition 4.10, $y \in \mathcal{R}(A)$ since $\mathcal{R}(A) \subseteq \mathcal{N}(A')$ is closed. This shows that $\mathcal{R}(A') \subseteq \mathcal{R}(A)$.

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 $\langle x_* , x \rangle = \langle A'y_* , x \rangle = \langle y_* , Ax \rangle = \langle y_* , \vartheta_{\mathcal{X}} \rangle = 0$. Then, $x_* \in \mathcal{N}(A)$.

On the other hand, fix an $x_* \in (\mathcal{N}(A))^{\perp}$. Let $K \in [0, \infty] \subset \mathbb{R}$ 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 $f : \mathcal{R}(A) \to \mathbb{K}$ 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, $|f(y)| \leq \|x_*\| \|x\| \leq \|x_*\| K \|y\|$. 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. □

7.12 Weak Topology

Definition 7.115 Let $X$ be a normed linear space over the field $\mathbb{K}$. The weak topology on $X$, denoted by $\mathcal{O}_{\text{weak}}(X) \subseteq X'$, is the weak topology generated by $X^*$, that is the weakest topology on $X$ such that $x_* : X \to \mathbb{K}$, $\forall x_* \in X^*$, are continuous.

For a normed linear space $X$, let $\mathcal{O}_X$ be the natural topology induced by the norm on $X$. Then, $\mathcal{O}_{\text{weak}}(X) \subseteq \mathcal{O}_X$. We usually call the topology $\mathcal{O}_X$ the strong topology. A set $S \subseteq X$ is weakly open, that is $S \in \mathcal{O}_{\text{weak}}(X)$, 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

\[
\{ x \in X \mid \langle \langle x_i, x \rangle \rangle \in O_i, i = 1, \ldots, n \}
\]

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

\[
\{ x \in X \mid |\langle \langle x_i, x \rangle \rangle | < \epsilon, i = 1, \ldots, n \}
\]

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 (\(T_{\Delta+}\)).

(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, \(\lim_{k \in \mathbb{N}} \langle \langle x_s, x_k \rangle \rangle = \langle \langle x_s, x_0 \rangle \rangle\), \(\forall x_s \in X^*\). In this case, we will write \(\lim_{k \in \mathbb{N}} x_k = x_0\) 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 \(B = \{ x \in X \mid \langle \langle \vartheta_s, x \rangle \rangle \in \mathbb{K} \}\), which is of the form \((7.9)\) 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 \((7.9)\). 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, \(x_0 \in B \in \mathcal{O}_\text{weak}(X)\). \(\forall O \in \mathcal{O}_\text{weak}(X)\) with \(x_0 \in O\), by Definition 3.17, there exists a basis open set \(B_1 := \{ x \in X \mid \langle \langle x_{i_1}, x \rangle \rangle \in O_i, i = 1, \ldots, n \}\), 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 \(i = 1, \ldots, n\), \(c_i := \langle \langle x_{i_1}, x_0 \rangle \rangle \in O_i \in \mathcal{O}_\mathbb{K}\). Then, \(\exists \epsilon_i \in (0, \infty) \subset \mathbb{R}\) 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 \(B_2 := \{ x \in X \mid |\langle \langle x_{i_1}, x - x_0 \rangle \rangle | < \epsilon, i = 1, \ldots, n \}\). 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, \(\exists x_s \in X^*\) with \(\| x_s \| = 1\) such that \(\langle \langle x_s, x_1 - x_2 \rangle \rangle = \| x_1 - x_2 \| > 0\). Let \(O_1 := \{ a \in \mathbb{K} \mid \text{Re}(a) > \text{Re}(\langle \langle x_s, x_2 \rangle \rangle) + \| x_1 - x_2 \| / 2\}\) and \(O_2 := \{ a \in \mathbb{K} \mid \text{Re}(a) < \text{Re}(\langle \langle x_s, x_2 \rangle \rangle) + \| x_1 - x_2 \| / 2\}\). Then, \(O_1, O_2 \in \mathcal{O}_\mathbb{K}\) and \(O_1 \cap O_2 = \emptyset\). Let \(B_1 := \{ x \in X \mid \langle \langle x_s, x \rangle \rangle \in O_1 \}\) and \(B_2 := \{ x \in X \mid \langle \langle x_s, x \rangle \rangle \in O_2 \}\). Then, \(B_1, B_2 \in \mathcal{O}_\text{weak}(X), x_1 \in B_1, x_2 \in B_2,\) and \(B_1 \cap B_2 = \emptyset\). This shows that \(X_{\text{weak}}\) is Hausdorff.
Next, we show that $X_{\text{weak}}$ is completely regular. Fix any weakly closed set $F \subseteq X_{\text{weak}}$ and $x_0 \in \overline{F} \in O_{\text{weak}}(X)$. Then, there exists a basis open set $B = \{ x \in X \mid \langle\langle x_i, x \rangle\rangle \in O_i, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$ and some $x_i \in X^*$ and $O_i \in O_K$, $i = 1, \ldots, n$, 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 $p_0 := (\langle\langle x_{i1}, x_0 \rangle\rangle, \ldots, \langle\langle x_{in}, x_0 \rangle\rangle) \in \mathbb{K}^n$. 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 $g : X_{\text{weak}} \to [0, 1] \subset \mathbb{R}$ by $g(x) = f(\langle\langle x_{i1}, x \rangle\rangle, \ldots, \langle\langle x_{in}, x \rangle\rangle), \forall x \in X_{\text{weak}}$. By Propositions 3.12 and 3.32, $g$ is a continuous real-valued function on $X_{\text{weak}}$. $g(x_0) = f(p_0) = 1$, and $g|_F = 0$. Hence, $X_{\text{weak}}$ is completely regular. Thus, (ii) (iii)

" if $\forall x \in X^*, x : X \to \mathbb{K}$ is weakly continuous. By Proposition 3.66, we have $\lim_{x \in \mathbb{N}} \langle\langle x_i, x_k \rangle\rangle = \langle\langle x_i, x_0 \rangle\rangle$.

" if $\forall x \in X^*, x : X \to \mathbb{K}$ is weakly continuous. By Proposition 3.66, we have $\lim_{x \in \mathbb{N}} \langle\langle x_i, x_k \rangle\rangle = \langle\langle x_i, x_0 \rangle\rangle$.

Proof

Let $(x_k)_{k=1}^\infty$ satisfy that $\lim_{x \in \mathbb{N}} \langle\langle x_i, x_k \rangle\rangle = \langle\langle x_i, x_0 \rangle\rangle, \forall x \in X^*$. For any basis open set $B = \{ x \in X \mid \langle\langle x_i, x \rangle\rangle \in O_i, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$ and some $x_i \in X^*$ and $O_i \in O_K$, $i = 1, \ldots, n$, with $x_0 \in B, we have that $\forall i = 1, \ldots, n, \exists N_i \in \mathbb{N}$ such that $\langle\langle x_i, x_k \rangle\rangle \in O_i, \forall k \geq N_i$. 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. $\square$

Proposition 7.117 Let $X$ be a normed linear space over the field $\mathbb{K}$ and $X_{\text{weak}}$ be the topological space of $X$ endowed with the weak topology. Then, vector addition $\oplus : X_{\text{weak}} \times X_{\text{weak}} \to X_{\text{weak}}$ is continuous; and scalar multiplication $\odot : \mathbb{K} \times X_{\text{weak}} \to X_{\text{weak}}$ is continuous.

Proof

Fix $(x_0, y_0) \in X_{\text{weak}} \times X_{\text{weak}}$. 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 7.16, $B = \{ z \in X \mid \langle\langle x_i, z-x_0-y_0 \rangle\rangle < \epsilon, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_i \in X^*, i = 1, \ldots, n$.

Let $B_1 := \{ x \in X \mid \langle\langle x_i, x-x_0 \rangle\rangle < \epsilon/2, i = 1, \ldots, n \}$ and $B_2 := \{ y \in X \mid \langle\langle x_i, y-y_0 \rangle\rangle < \epsilon/2, i = 1, \ldots, n \}$. Clearly, $B_1, B_2 \subseteq O_{\text{weak}}(X)$, $B_1 \times B_2 \subseteq O_{\text{weak}} \times X_{\text{weak}}$, and $(x_0, y_0) \in B_1 \times B_2$. $\forall (x, y) \in B_1 \times B_2, \forall i = 1, \ldots, n, \langle\langle x_i, x+y-x_0-y_0 \rangle\rangle \leq \langle\langle x_i, x-x_0 \rangle\rangle + \langle\langle x_i, y-y_0 \rangle\rangle < \epsilon$. 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 $(\alpha_0, x_0) \in \mathbb{K} \times X_{\text{weak}}$. We will show that $\odot$ is continuous at $(\alpha_0, x_0)$. Fix a basis open set $B \in O_{\text{weak}}(X)$ with $\alpha_0 x_0 \in B$. Then, by Proposition 7.16, $B = \{ z \in X \mid \langle\langle x_i, z-\alpha_0 x_0 \rangle\rangle < \epsilon, i = 1, \ldots, n \}$, for some $n \in \mathbb{N}$, for some $\epsilon \in (0, \infty) \subset \mathbb{R}$, and for some $x_i \in X^*, i = 1, \ldots, n$. Let $M := \max_{1 \leq i \leq n} \| x_i \| \in (0, \infty) \subset \mathbb{R}$. Let $B_1 := \{ \alpha \in \mathbb{K} \mid \| \alpha - \alpha_0 \| < \epsilon/2M\|x_0\| \}$ and $B_2 := \{ x \in X \mid \langle\langle x_i, x-x_0 \rangle\rangle < \epsilon/2M\|x_0\|, i = 1, \ldots, n \}$. Clearly, $B_1 \in O_K$ and $B_2 \subseteq O_{\text{weak}}(X)$, $B_1 \times B_2 \subseteq O_{\text{weak}} \times X_{\text{weak}}$, 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_i, \alpha x-\alpha_0 x_0 \rangle\rangle \leq \langle\langle x_i, x-x_0 \rangle\rangle + \langle\langle x_i, \alpha x_0-\alpha_0 x_0 \rangle\rangle < \epsilon/2M\|x_0\| + \langle\langle x_i, \alpha x_0-\alpha_0 x_0 \rangle\rangle.
7.12. WEAK TOPOLOGY

| ⟨⟨x_1, αx − αx_0⟩⟩| + |⟨⟨x_1, αx_0 − α_0x_0⟩⟩| = |α| |⟨⟨x_1, x − x_0⟩⟩| + |α − α_0||⟨⟨x_1, x_0⟩⟩| ≤ (|α_0| + |α − α_0|)|⟨⟨x_1, x − x_0⟩⟩| + |α − α_0||x_1||x_0|| ≤ (|α_0| + 1)|⟨⟨x_1, x − x_0⟩⟩| + |α − α_0||x_0|| < e. Hence, αx ∈ B. This shows that ⊗ is continuous at (α_0, x_0). By the arbitrariness of (α_0, x_0) and Proposition 3.9, ⊗ is continuous.

This completes the proof of the proposition.

Proposition 7.118 Let X be a finite-dimensional normed linear space over the field K, O_X be the strong topology on X, and O_{weak} (X) be the weak topology on X. Then, O_X = O_{weak} (X).

Proof Clearly, O_{weak} (X) ⊆ O_X. Fix any basis open set O = B_X (x_0, r) ∈ O_X with x_0 ∈ X and r ∈ (0, ∞) ⊆ R. We will show that O ∈ O_{weak} (X). Let n ∈ Z_+ be the dimension of X. We will distinguish two exhaustive and mutually exclusive cases: Case 1: n = 0; Case 2: n ∈ N.

Case 1: n = 0. Then, X is a singleton set. O = X ∈ O_{weak} (X). This case is proved.

Case 2: n ∈ N. Let {e_1, ..., e_n} ⊆ X be a basis of X with ∥e_i∥ = 1, i = 1, ..., n. ∀i = 1, ..., n, let f_i : X → K be defined by f_i (x) = α_i, ∀x = ∑_{j=1}^{n} α_j e_j ∈ 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 ∈ X*. ∀x_1 ∈ O, let δ = (r − ∥x_1 − x_0∥)/n ∈ (0, ∞) ⊆ R. Let B := {x ∈ X | |⟨⟨e_i, x − x_1⟩⟩| < δ, i = 1, ..., n}. By Proposition 7.116, B ∈ O_{weak} (X). ∀x ∈ B, let x − x_1 = ∑_{j=1}^{n} α_j e_j. Then, ∀i = 1, ..., n, |α_i| = |⟨⟨e_i, x − x_1⟩⟩| < δ. Thus, ∥x − x_1∥ ≤ ∑_{j=1}^{n} |α_j| ∥e_j∥ < nδ = r − ∥x_1 − x_0∥. Then, we have ∥x − x_0∥ ≤ ∥x − x_1∥ + ∥x_1 − x_0∥ < r and x ∈ O. Hence, x_1 ∈ B ⊆ O. Therefore, O ∈ O_{weak} (X).

In both cases, we have shown that O ∈ O_{weak} (X). Then, O_X ⊆ O_{weak} (X). Hence, O_X = O_{weak} (X). This completes the proof of the proposition.

Given a normed linear space X, its dual X* is a Banach space. On X*, we can also talk about the notion of weak topology. The weak topology of X* is the weakest topology on X* such that all of the functional in X** are continuous. This topology turns out to be less useful than the weak topology for X* generated by X (or more precisely by φ(X), where φ : X → X** is the natural mapping). This leads us to the following definition.

Definition 7.119 Let X be a normed linear space over the field K, X* be its dual, and φ : X → X** be the natural mapping. The weak* topology on X*, denoted by O_{weak*} (X*) ∈ X**, is the weak topology generated by φ(X), that is the weakest topology on X* such that φ(x) : X* → K, ∀x ∈ X, are continuous.

For a normed linear space X, let O_X be the natural topology induced by the norm on X*. Then, O_{weak*} (X*) ⊆ O_{weak} (X*) ⊆ O_X. Therefore, the
weak* topology is weaker than the weak topology on \( X^* \), which is further weaker than the strong topology on \( X^* \), \( O_{X^*} \). Clearly, if \( X \) is reflexive, then the weak topology and the weak* topology coincide. We will denote the topological space \( (X^*, O_{weak^*}(X^*)) \) by \( X_{weak^*}^* \). 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 \( O_{weak^*}(X^*) \) consists of all sets of the form

\[
\{ x_* \in X^* \mid \langle \langle x_* , x_i \rangle \rangle = \langle \langle \phi(x_i) , x_* \rangle \rangle \in O_i , i = 1, \ldots, n \} \tag{7.10}
\]

where \( n \in \mathbb{N} \), \( O_i \in O_K \) and \( x_i \in X \), \( i = 1, \ldots, n \), and \( O_K \) is the natural topology on \( K \). A basis at \( x_{*0} \in X^* \) for \( O_{weak^*}(X^*) \) consists of all sets of the form

\[
\{ x_* \in X^* \mid | \langle \langle x_* - x_{*0} , x_i \rangle \rangle | = | \langle \langle \phi(x_i) , x_* - x_{*0} \rangle \rangle | < \epsilon , i = 1, \ldots, n \} \tag{7.11}
\]

where \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( n \in \mathbb{N} \), and \( x_i \in X \), \( i = 1, \ldots, n \).

(ii) \( X_{weak^*} \) is completely regular \((T_{3\frac{1}{2}})\).

(iii) For a sequence \( (x_{*k})_{k=1}^{\infty} \subseteq X_{weak^*} \), \( x_{*0} \in X^* \) 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, \( O_{weak^*}(X^*) \) 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 \( O_{weak^*}(X^*) \).

Let \( B \subseteq X^* \) be any set of the form \( (7.11) \). Clearly, \( x_{*0} \in B \in O_{weak^*}(X^*) \). \( \forall O \in O_{weak^*}(X^*) \) 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 \( O_i \in O_K \), \( i = 1, \ldots, n \), such that \( x_{*0} \in B_1 \subseteq O \). For each \( i = 1, \ldots, n \), \( c_i := \langle \langle x_{*0} , x_i \rangle \rangle \in O_i \in O_K \). Then, \( \exists \epsilon_i \in (0, \infty) \subseteq \mathbb{R} \) such that \( B_K(c_i, \epsilon_i) \subseteq O_i \). Take \( \epsilon = \min\{ \epsilon_i \mid i \leq n \} \epsilon_i \in (0, \infty) \subseteq \mathbb{R} \) 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 \( O_{weak^*}(X^*) \). 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
7.12. WEAK TOPOLOGY

⟨⟨x₁ − x₂, x⟩⟩ ≥ ∥x₁ − x₂∥/2 > 0. Let O₁ := \{ a ∈ K | Re(a) > Re(⟨⟨x₁ − x₂, x⟩⟩) + ∥x₁ − x₂∥/4 \} and O₂ := \{ a ∈ K | Re(a) < Re(⟨⟨x₁ − x₂, x⟩⟩) + ∥x₁ − x₂∥/4 \}. Then, O₁, O₂ ∈ O_K and O₁ ∩ O₂ = ∅. Let B₁ := \{ x ∈ X* | ⟨⟨x, x⟩⟩ ∈ O₁ \} and B₂ := \{ x ∈ X* | ⟨⟨x, x⟩⟩ ∈ O₂ \}. Then, B₁, B₂ ∈ O_{weak*}(X*), x₁ ∈ B₁, x₂ ∈ B₂, and B₁ ∩ B₂ = ∅. This shows that X_{weak*} is Hausdorff.

Next, we show that X_{weak*} is completely regular. Fix any weak* closed set F ⊆ X_{weak*} and x₀ ∈ F ∈ O_{weak*}(X*). Then, there exists a basis open set B = \{ x ∈ X* | ⟨⟨x, x_i⟩⟩ ∈ O_i, i = 1, ..., n \}, for some n ∈ N and some xᵢ ∈ X and Oᵢ ∈ O_K, i = 1, ..., n, such that x₀ ∈ B ⊆ F. Let O := \bigcap_{i=1}^{n} Oᵢ ⊆ K^n which is open. Let p₀ := (⟨⟨x₀, x_1⟩⟩, ..., ⟨⟨x₀, x_n⟩⟩) ∈ K^n. Then, p₀ ∈ O. By Propositions 4.11 and 3.61, K^n is normal and therefore completely regular. Then, there exists a continuous function f : K^n → [0, 1] ⊆ R such that f|₀ = 0 and f(p₀) = 1. Define g : X_{weak*} → [0, 1] ⊆ R by g(x₀) = f(⟨⟨x₀, x_1⟩⟩, ..., ⟨⟨x₀, x_n⟩⟩), ∀x₀ ∈ X_{weak*}. By Propositions 3.12 and 3.32, g is a continuous real-valued function on X_{weak*}, g(x₀) = f(p₀) = 1, and |g|_F = 0. Hence, X_{weak*} is completely regular. Thus, (ii) holds.

(iii) “Only if” ∀x ∈ X, φ(x) : X* → K is weak* continuous. By Proposition 3.66, we have \lim_{k∈N} ⟨⟨x, x_k⟩⟩ = ⟨⟨x, x⟩⟩ = \lim_{k∈N} ⟨⟨φ(x), x_k⟩⟩ = ⟨⟨φ(x), x⟩⟩.

“If” Let (xₖ)_{k=1}^{∞} satisfy that \lim_{k∈N} ⟨⟨x, x_k⟩⟩ = ⟨⟨x, x⟩⟩, ∀x ∈ X. For any basis open set B = \{ x ∈ X* | ⟨⟨x, xᵢ⟩⟩ ∈ Oᵢ, i = 1, ..., n \}, for some n ∈ N and some xᵢ ∈ X and Oᵢ ∈ O_K, i = 1, ..., n, with x₀ ∈ B, we have that ∀i = 1, ..., n, ∃Nᵢ ∈ N such that ⟨⟨xᵢ, xᵢ⟩⟩ ∈ Oᵢ, ∀k ≥ Nᵢ. Take N = max_{1≤i≤n} Nᵢ ∈ N. ∀k ≥ N, xᵢ ∈ B. This shows that (xₖ)_{k=1}^{∞} converges weak* to x₀. Hence, (iii) holds.

This completes the proof of the proposition.

□

Proposition 7.121 Let X be a normed linear space over the field K and X_{weak*} be the topological space of X* endowed with the weak* topology. Then, vector addition ⊕ : X_{weak*} × X_{weak*} → X_{weak*} is continuous; and scalar multiplication ⊗ : K × X_{weak*} → X_{weak*} is continuous.

Proof Fix (x₀, y₀) ∈ X_{weak*} × X_{weak*}. We will show that ⊕ is continuous at (x₀, y₀). Fix a basis open set B ∈ O_{weak*}(X*) with x₀ + y₀ ∈ B. Then, by Proposition 7.120, B = \{ z ∈ X* | |⟨⟨z, -x₀ - y₀, xᵢ⟩⟩| < ε, i = 1, ..., n \}, for some n ∈ N, for some ε ∈ (0, ∞) ⊆ R, and for some xᵢ ∈ X, i = 1, ..., n. Let B₁ := \{ x ∈ X* | |⟨⟨x, -x₀, xᵢ⟩⟩| < ε/2, i = 1, ..., n \} and B₂ := \{ y ∈ X* | |⟨⟨y, -y₀, xᵢ⟩⟩| < ε/2, i = 1, ..., n \}. Clearly, B₁, B₂ ∈ O_{weak*}(X*), B₁ × B₂ ∈ O_{weak*} × X_{weak*}, and (x₀, y₀) ∈ B₁ × B₂. ∀(xᵢ, yᵢ) ∈ B₁ × B₂, ∀i = 1, ..., n, |⟨⟨xᵢ, yᵢ⟩⟩| ≤ |⟨⟨xᵢ, -x₀, xᵢ⟩⟩| + |⟨⟨yᵢ, -y₀, xᵢ⟩⟩| < ε. Hence, xᵢ + yᵢ ∈ B. This shows that ⊕ is continuous at (x₀, y₀). By the arbitrariness of (x₀, y₀) and Proposition 3.9, ⊗ is continuous.
Theorem 7.122 (Alaoglu Theorem) Let $\mathcal{X}$ be a normed linear space over the field $\mathbb{K}$ and $S = \overline{B}_\mathcal{X}(0, r)$, for some $r \in \mathbb{Z}_+$. Then, $S$ is weak* compact.

Proof Let the weak* topology on $S$ be $\mathcal{O}_\mathcal{X}^*(S)$. Denote the topological space $(S, \mathcal{O}_\mathcal{X}^*(S))$ by $S$. \forall x \in S, \forall x \in S, by Proposition 7.72, we have $|\langle\langle x, x \rangle\rangle| \leq \|x\| \leq r \|x\|$. Let $I_x := \overline{B}_K(0, r \|x\|)$. Then, $\langle\langle x, x \rangle\rangle \in I_x$. Let $I_S \subseteq \mathbb{K}$ be endowed with the subset topology $\mathcal{O}_S$. Denote the topological space $(I_S, \mathcal{O}_S)$ by $I_x$. By Proposition 5.40 or 7.41, $I_x$ is compact. Let $\mathcal{P} := \prod_{x \in S} I_x$, whose topology is denoted by $\mathcal{O}_\mathcal{P}$. By Tychonoff Theorem, $\mathcal{P}$ is compact. Define the equivalence map $E : S \to \mathcal{P}$ by $\pi_x(E(x)) = \langle\langle x, x \rangle\rangle$, $\forall x \in S, \forall x \in S$. Let $\mathcal{E} := \mathcal{O}_\mathcal{P}$ be the homeomorphism.

Claim 7.122.1 $E : S \to E(S) \subseteq \mathcal{P}$ is a homeomorphism.

Proof of claim: Clearly, $E : S \to E(S)$ is surjective. \forall x_1, x_2 \in S with $x_1 \neq x_2$, then $\exists x \in S$ such that $\pi_x(E(x_1)) = \langle\langle x_1, x \rangle\rangle \neq \langle\langle x_2, x \rangle\rangle = \pi_x(E(x_2))$. Then, $E(x_1) \neq E(x_2)$. Therefore, $E : S \to E(S)$ is bijective and admits inverse $E_{inv} : E(S) \to S$.

Fix any $x_0 \in S$. Fix any basis open set $O \in \mathcal{O}_S$ with $E(x_0) \in O$. By Proposition 3.25, $O = \bigcup_{x \in X} O_x$, where $O_x \in \mathcal{O}_S, \forall x \in S$, 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 $x_1 = \emptyset$ and $O_1 = \mathbb{K}$. Let $B := \{x_1 \in S \mid \langle\langle x_1, x_1 \rangle\rangle \in O_1\} = S$. Clearly, $x_0 \in B \in \mathcal{O}_\mathcal{X}^*(S)$ and $\forall x \in B, \forall x \in S, \pi_x(E(x_1)) = \langle\langle x_1, x \rangle\rangle \in I_x = O_x$. Then, $E(x_1) \in O$ and $E(B) \subseteq O$. Case 2: $X_N \neq \emptyset$. $\forall x \in X_N$, by Proposition 3.4, let $O_x \in \mathcal{O}_S$ be such that $O_x \cap I_x = O_x$. Let $B := \{x \in S \mid \langle\langle x, x \rangle\rangle \in O_x, \forall x \in X_N\}$. Clearly, $x_0 \in B \in \mathcal{O}_\mathcal{X}^*(S)$. $\forall x \in B, \forall x \in X_N, \pi_x(E(x_1)) = \langle\langle x, x \rangle\rangle \in I_x = O_x$. Therefore, $\langle\langle x, x \rangle\rangle \in I_x = O_x$, for all $x$'s.
7.12. WEAK TOPOLOGY

$\bar{O}_x \cap I_x = O_x$. Hence, $E(x_*) \in O$ and $E(B) \subseteq O$. In both cases, we have shown that $fB \in \mathcal{O}_{\text{weak}^*(S)}$ 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 $x_0 = E_{\text{inv}}(p_0) \in S$. Fix any basis open set $B \in \mathcal{O}_{\text{weak}^*(S)}$ with $x_0 \in B$. By Proposition 7.120, $B = \{ x_\ast \in S \mid \langle \langle x_\ast, x_i \rangle \rangle \in O_i, i = 1, \ldots, n \}$, 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 $O_i = O_i$, $i = 1, \ldots, n - 2$, 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 $O_x = O_i \cap I_x \in \mathcal{O}_Z$; otherwise, let $O_x = I_x \in \mathcal{O}_X$. Let $O := \prod_{x \in X} O_x$. Clearly, $p_0 \in O$. $\forall p \in O \cap E(S)$, let $x_\ast = E_{\text{inv}}(p) \in S$. $\forall i = 1, \ldots, n$, $\langle \langle x_\ast, x_i \rangle \rangle = \pi_{x_\ast}(E(x_\ast)) \in O_{x_i} = O_i \cap I_x$. Then, $x_\ast \in B$ and $E_{\text{inv}}(O \cap E(S)) \subseteq B$. This shows that $E_{\text{inv}}$ is continuous at $p_0$. By the arbitrariness of $p_0$ and Proposition 3.9, $E_{\text{inv}}$ 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. $\forall p_0 \in \overline{E(S)} \subseteq \mathcal{P}$, 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 $z = \alpha x_1 + \beta x_2$, $\forall \epsilon \in (0, \infty)$, $\forall x \in X$, if $x \in \{ x_1, x_2, z \}$, let $O_x = B_K(\pi_x(p_0), \epsilon) \cap I_x \in \mathcal{O}_Z$, otherwise, let $O_x = I_x \in \mathcal{O}_Z$. Let $O := \prod_{x \in X} O_x \in \mathcal{O}_P$. Clearly, $p_0 \in O$. By Proposition 3.3, $\exists p_1 \in E(S) \cap O$. Let $x_\ast = E_{\text{inv}}(p_1) \in S$. $\forall x \in \{ x_1, x_2, z \}$, $\langle \langle x_\ast, x \rangle \rangle = \pi_{x_\ast}(p_1) \in O_{x} = B_K(\pi_x(p_0), \epsilon) \cap I_x \subseteq B_K(f_0(x), \epsilon)$. Hence, we have $\| \langle \langle x_\ast, x_1 \rangle \rangle - f_0(x_1) \| < \epsilon, \| \langle \langle x_\ast, x_2 \rangle \rangle - f_0(x_2) \| < \epsilon$, and $\| \langle \langle x_\ast, z \rangle \rangle - f_0(z) \| < \epsilon$. This implies that

$$
| f_0(z) - \alpha f_0(x_1) - \beta f_0(x_2) | = | f_0(z) - \langle \langle x_\ast, z \rangle \rangle - \alpha f_0(x_1) + \alpha \langle \langle x_\ast, x_1 \rangle \rangle - \beta f_0(x_2) + \beta \langle \langle x_\ast, x_2 \rangle \rangle | < \epsilon + | \alpha | \epsilon + | \beta | \epsilon
$$

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. $\forall x \in X$, $f_0(x) = \pi_x(p_0)$, then $\| f_0(x) \| = R \| x \|$. Therefore, $f_0 \in X^\ast$ and $\| f_0 \| \leq R$. $\forall x \in X$, $\pi_X(E(f_0)) = f_0(x) = \pi_x(p_0)$. 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 $\mathcal{P}$, we have $E(S)$ is compact. By Proposition 5.7, $S$ is compact. This completes the proof of the theorem.

\textbf{Proposition 7.123} Let $X$ and $Y$ be normed linear spaces over $K$, $A \in B(X,Y)$, and $X_{\text{weak}}$ and $Y_{\text{weak}}$ be the topological spaces of $X$ and $Y$ endowed with the weak topology. Then, $A : X_{\text{weak}} \rightarrow Y_{\text{weak}}$ is continuous.

\textbf{Proof} Fix any basis open set $O_Y \in \mathcal{O}_{\text{weak}^*(Y)}$, we will show that $A_{\text{inv}}(O_Y) \in \mathcal{O}_{\text{weak}^*(X)}$. 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 $n \in \mathbb{N}$, $O_i \in \mathcal{O}_K$, $y_i \in Y^\ast$, $i = 1, \ldots, n$. Then, $A_{\text{inv}}(O_Y) = \{ x \in X \mid \langle \langle y_i, Ax \rangle \rangle \in O_i, i = 1, \ldots, n \} \in \mathcal{O}_X$. \qed
\[ \{ \langle A^* y_i, x \rangle \in O_i \mid i = 1, \ldots, n \} \in \mathcal{O}_{\text{weak}}(X). \] Hence, \( A : X_{\text{weak}} \to Y_{\text{weak}} \) is continuous. This completes the proof of the proposition. \( \square \)

**Definition 7.124** Let \( X \) be a normed linear space and \( D \subseteq X \). \( D \) is said to be locally convex at \( x_0 \in D \) if \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \), such that \( D_0 := D \cap B_X(x_0, \delta_0) \) 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 \( B \subseteq \mathcal{O} \) at \( x_0 \) for the subset topological space \( (D, \mathcal{O}) \) such that each \( B \in B \) is convex. Furthermore, \( D \subseteq X \) is locally convex if, and only if, there exists a basis \( B \subseteq \mathcal{O} \) for the subset topological space \( (D, \mathcal{O}) \) such that each \( B \in B \) is convex.

**Proposition 7.125** Let \( X \) be a Banach space over \( \mathbb{K} \). Then, \( X \) is \( \sigma \)-compact if, and only if, \( X \) is finite dimensional.

**Proof** “Sufficiency” Assume that \( X \) is finite dimensional. Let \( K_n := \overline{B}_X(\vartheta_X, n), \forall n \in \mathbb{N} \). 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, \( \overline{X} = \bigcup_{n=1}^{\infty} K_n \), where \( K_n \subseteq \overline{X} \) 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 \( F : \overline{X} \to \overline{X} \) by \( F(x) = \frac{1}{\delta}(x - x_0), \forall x \in \overline{X} \). 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, \( \forall x \in K_n, x \in \overline{K_n}, \forall n \in \mathbb{N} \), since \( x \) is not an interior point of \( K_n \). Then, \( \overline{X} \supseteq \overline{K_n} \supseteq K_n \cup \overline{K_n} = X \). This implies that \( \overline{X} = \overline{K_n} = \overline{K_n} \). Hence, \( K_n \) is a nowhere dense subset in \( \overline{X} \). This then leads to the conclusion that \( \overline{X} = \bigcup_{n=1}^{\infty} K_n \) is of first category. By Baire Category Theorem 4.40, \( \overline{X} \) is second category everywhere. Then, \( \overline{X} \neq \emptyset \) 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 \( x_1 := y_1 \in K, z_n \in \text{argmin}_{z} \text{span}(\{x_1, \ldots, x_{n-1}\}) \|y_n - z\|, \) 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 $X$. By $E$ being a linearly independent set, we have $\| y_n - z_n \| > 0$, $\forall n \in \mathbb{N}$. Hence, $x_n \in X$ 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 $(x_n)_1^{\infty}$ of $(x_n)_n^{\infty}$ such that $\lim_{n \in \mathbb{N}} x_n = x_0$.

Since $x_n \in \text{span}(E)$, $\forall i \in \mathbb{N}$, 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 $x_0 = \lim_{i \in \mathbb{N}} x_n = \lim_{i \in \mathbb{N}} \sum_{j=1}^{n_i} a_j x_j$ and thus $\lim_{i \in \mathbb{N}} (x_n - \sum_{j=1}^{n_i} a_j x_j) = \varnothing_X$. But, $\forall i \in \mathbb{N}$, $\| x_n - \sum_{j=1}^{n_i} a_j x_j \| = \frac{1}{\| y_{n_i} - z_{n_i} \|} \| y_{n_i} - \sum_{j=1}^{n_i} (\| y_{n_i} - z_{n_i} \|) a_j x_j - z_{n_i} \|$.

The vector $\sum_{j=1}^{n_i-1} (\| y_{n_i} - z_{n_i} \|) a_j x_j - z_{n_i} \in \text{span}(\{x_1, \ldots, x_{n_i-1}\})$. By the definition of $z_{n_i}$, we have $\| x_n - \sum_{j=1}^{n_i-1} a_j x_j \| = \frac{1}{\| y_{n_i} - z_{n_i} \|} \| y_{n_i} - \sum_{j=1}^{n_i-1} (\| y_{n_i} - z_{n_i} \|) a_j x_j - z_{n_i} \| \geq 1$. This contradicts with $\lim_{i \in \mathbb{N}} (x_n - \sum_{j=1}^{n_i-1} a_j x_j) = \varnothing_X$. Hence, the hypothesis does not hold, and $X$ must be finite dimensional.

This completes the proof of the proposition. □

**Proposition 7.126** Let $X$ be a separable normed linear space, $Y$ be a Banach space, $D \subseteq X$, and $f : D \rightarrow Y$ be continuous. Then, $\mathcal{W} := \text{span}(f(X)) \subseteq Y$ is a separable normed linear subspace of $Y$, and $\overline{\mathcal{W}} \subseteq Y$ is a separable Banach subspace of $Y$.

**Proof** $\mathcal{W}$ is a subspace of $Y$ and hence a normed linear subspace of $Y$. $\overline{\mathcal{W}}$ is a closed subspace of $Y$ and hence a complete normed linear subspace of $Y$, by Proposition 4.39. Then, $\overline{\mathcal{W}}$ is a Banach subspace of $Y$. All we need to show is that they are separable. By Proposition 4.38, $D$ is separable as a metric subspace of $X$. 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 $f(D) = f(X)$. Then, $f(X)$ is a separable subset of $Y$. By Proposition 7.35, $\mathcal{W}$ and $\overline{\mathcal{W}}$ are separable subsets of $Y$. This completes the proof of the proposition. □
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 $V \supseteq H$ is a linear variety, then either $V = H$ or $V = X$.

Proposition 8.2 Let $X$ be a real vector space. $H \subseteq X$ 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. $\forall x \in \mathcal{X}$, $x$ can be uniquely written as $\alpha x_0 + m$, where $\alpha \in \mathbb{R}$ and $m \in M$. Define $f : \mathcal{X} \to \mathbb{R}$ by $f(x) = f(\alpha x_0 + m) = \alpha$, $\forall x \in \mathcal{X}$. Clearly, $f$ is a linear functional and is not identically equal to zero. It is straightforward to verify that $H = \{ x \in \mathcal{X} \mid f(x) = 1 \}$. Case 2: $x_0 \in M$. Then, $H = M$. By Definition 8.1, $\exists x_1 \in \mathcal{X} \setminus H$. Let $\bar{M} := \text{span}(M \cup \{x_1\})$. Clearly, $\bar{M} \supset H$ and is a linear variety. Then, $\bar{M} = \mathcal{X}$ by Definition 8.1. $\forall x \in \mathcal{X}$, $x$ can be uniquely written as $\alpha x_1 + m$, where $\alpha \in \mathbb{R}$ and $m \in M$. Define $f : \mathcal{X} \to \mathbb{R}$...
by $f(x) = f(\alpha x_1 + m) = \alpha, \forall x \in \mathcal{X}$. Clearly, $f$ is a linear functional and is not identically equal to zero. It is straightforward to verify that $H = \{ x \in \mathcal{X} \mid f(x) = 0 \}$.

"Sufficiency" Let $H = \{ x \in \mathcal{X} \mid f(x) = c \}$, where $f : \mathcal{X} \rightarrow \mathbb{R}$ 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 $\mathcal{X}$. Since $f$ is not identically equal to zero, then $\exists x_0 \in \mathcal{X} \setminus M$ such that $f(x_0) = 1$, $\forall x \in \mathcal{X}$, $f(x - f(x)x_0) = 0$. Then, $x - f(x)x_0 \in M$ and $x \in \text{span} (M \cup \{x_0\})$. Hence, $\mathcal{X} = \text{span} (M \cup \{x_0\})$ and $M$ is a maximal proper subspace. Then, $H = cx_0 + M$ is a hyperplane.

This completes the proof of the proposition. □

Consider a hyperplane $H$ in a real normed linear space $\mathcal{X}$. By Proposition 7.17, $\mathcal{P}$ is a linear variety. By Definition 8.1, $\mathcal{P} = \mathcal{X}$ or $\mathcal{P} = H$. Thus, a hyperplane in a real normed linear space must either be dense or closed.

**Proposition 8.3** Let $\mathcal{X}$ be a real normed linear space and $H \subseteq \mathcal{X}$. $H$ is a closed hyperplane if, and only if, there exist $x_\ast \in \mathcal{X}^\ast$ and $c \in \mathbb{R}$ with $x_\ast \neq 0$, such that $H = \{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle = c \}$.

**Proof** "Necessity" Let $H$ be a closed hyperplane. By Proposition 8.2, there exists a linear functional $f : \mathcal{X} \rightarrow \mathbb{R}$ and $c \in \mathbb{R}$ with $f$ being not identically equal to zero such that $H = \{ x \in \mathcal{X} \mid f(x) = c \}$. All we need to show is that $f \in \mathcal{X}^\ast$. 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, $f \in \mathcal{X}^\ast$.

"Sufficiency" By Proposition 8.2, $H$ is a hyperplane. By the continuity of $x_\ast$ and Proposition 3.10, $H$ is closed. Hence, $H$ is a closed hyperplane.

This completes the proof of the proposition. □

For a real normed linear space $\mathcal{X}$ and a closed hyperplane $H \subseteq \mathcal{X}$. Then, $H = \{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle = c \}$, where $x_\ast \neq x_\ast \in \mathcal{X}^\ast$ and $c \in \mathbb{R}$. We associate four sets with $H$: (a) $\{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle \leq c \}$; (b) $\{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle < c \}$; (c) $\{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle \geq c \}$; (d) $\{ x \in \mathcal{X} \mid \langle x_\ast, x \rangle > c \}$, 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 $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be convex with $\vartheta \in K^\circ$. The Minkowski functional $p : \mathcal{X} \rightarrow \mathbb{R}$ of $K$ is defined by $p(x) = \inf \{ r \in \mathbb{R} \mid r^{-1}x \in K, r > 0 \}$, $\forall x \in \mathcal{X}$.

**Proposition 8.5** Let $\mathcal{X}$ be a real normed linear space and $K \subseteq \mathcal{X}$ be convex with $\vartheta \in K^\circ$. Then, the Minkowski functional $p : \mathcal{X} \rightarrow \mathbb{R}$ of $K$ satisfies

(i) $0 \leq p(x) < +\infty$, $\forall x \in \mathcal{X}$;

(ii) $p(\alpha x) = \alpha p(x)$, $\forall x \in \mathcal{X}$, $\forall \alpha \in [0, \infty) \subseteq \mathbb{R}$;
(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) \( \mathcal{K} = \{ x \in \mathcal{X} \mid p(x) \leq 1 \}; \ K^\circ = \{ x \in \mathcal{X} \mid p(x) < 1 \}. \)

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) \( \forall x \in \mathcal{X}, \forall \alpha \in [0, \infty) \subseteq \mathbb{R}, \) 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 \( p(ax) = p(\vartheta) = 0 = \alpha p(x) \).

Case 2: \( x \neq \vartheta \) and \( \alpha = 0 \). Then, we have \( p(ax) = p(\vartheta) = 0 = \alpha p(x) \).

Case 3: \( x \neq \vartheta \) and \( \alpha > 0 \). \( \forall r \in \mathbb{R} \mid r^{-1} x \in K, r > 0 \}, \) we have \( \alpha x > 0 \) and \( \alpha x \in \mathbb{R} \mid r^{-1} (ax) \in K, r > 0 \}. \) Hence, \( \alpha p(x) \geq p(ax) \).

On the other hand, \( \forall r \in \mathbb{R} \mid r^{-1} (ax) \in K, r > 0 \}, \) we have \( r/\alpha \in \{ r \in \mathbb{R} \mid r^{-1} x \in K, r > 0 \} \). Hence, \( \alpha^{-1} p(x) \geq p(x) \. \) Therefore, we have \( \alpha p(x) = p(\alpha x) \).

(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 \( r_1 + r_2 \geq p(x_1 + x_2) \) and \( p(x_1) + p(x_2) \geq p(x_1 + x_2) \).

(iv) \( \forall x \in \mathcal{X}, \exists \in (0, \infty) \subset \mathbb{R} \) such that \( B(x, \epsilon) \subseteq K \), \( \forall x_1, x_2 \in \mathcal{X} \) with \( \| x_1 - x_2 \| < \delta \), we have \( p(x_1) \leq p(x_2) + p(x_1 - x_2) \leq p(x_2) + \| x_1 - x_2 \| / \epsilon_0 < p(x_2) + \epsilon \) and \( p(x_2) \leq p(x_1) + p(x_2 - x_1) \leq p(x_1) + \| x_2 - x_1 \| / \epsilon_0 < p(x_1) + \epsilon \). 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 \( B(x, \epsilon) \subseteq K \), 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 \( p(y) < 1, \forall y \in B(x, \delta), \forall y \in B(y, \delta), y \in K \) by the convexity of \( K \), \( p(y) < 1 \), and the fact that \( \vartheta \in K \). Then, \( B(y, \delta) \subseteq K \). 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 $K^\circ \neq \emptyset$. Then, $\overline{K} = \overline{K^\circ}$.

Proof Clearly, $K^\circ \subseteq K$. Then, $\overline{K^\circ} \subseteq \overline{K}$.

To show that $\overline{K} \subseteq \overline{K^\circ}$, we will distinguish two exhaustive and mutually exclusive cases: Case 1: $\emptyset \in K^\circ$; Case 2: $\emptyset \notin K^\circ$. 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, $K^\circ = \{x \in X \mid p(x) < 1\}$, and $\overline{K} = \{x \in X \mid p(x) \leq 1\}$. $\forall x \in \overline{K}$, $\forall x \in (0, \infty) \subseteq \mathbb{R}$, $p(x) \leq 1$. By the convexity of $K$, $y := \frac{p(x)}{p(x) - 1} x \in K$. Then, $p(y) = \frac{\|x\|}{\|x\| - 1} p(x) < 1$. Then, $y \in K^\circ$.

Case 1: $\emptyset \notin K^\circ$. Let $x_0 \in K^\circ \neq \emptyset$. Let $K_1 = K - x_0$. By Propositions 7.16 and 6.39, $K_1^\circ = K^\circ - x_0 \neq \emptyset$ and $K_1$ is convex. By Case 1, $\overline{K_1} \subseteq \overline{K_1}^\circ$. By Proposition 7.16, $\overline{K} = x_0 + K_1 = x_0 + K_1 \subseteq x_0 + \overline{K_1}^\circ = x_0 + \overline{K_1} = (x_0 + K_1)^\circ = K^\circ$. This case is proved.

Hence, $\overline{K} = \overline{K^\circ}$. 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 $V \cap K^\circ = \emptyset$. 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 $x_* \neq \emptyset$, such that $\langle x_*, v \rangle = c$, $\forall v \in V$, $\langle x_*, k \rangle < c$, $\forall k \in K^\circ$, and $\langle x_*, k \rangle \leq c$, $\forall k \in K$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\emptyset \in K^\circ$; Case 2: $\emptyset \notin K^\circ$. 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 $\emptyset \notin V$, then $\exists x_1 \in V$ such that $V - x_1 =: M$ is a subspace in $M$ and $x_1 \notin M$. $\forall m \in M$, $\exists \alpha \in \mathbb{R}$ and $\exists \bar{m} \in M$ 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 $v \notin K^\circ$ and $p(v) \geq 1 = f(v)$. $\forall m = \alpha x_1 + \bar{m} \in M$, we have either $\alpha > 0$, then $f(m) = \alpha f(x_1 + \alpha^{-1} \bar{m}) = \alpha \leq \alpha p(x_1 + \alpha^{-1} \bar{m}) = p(m)$, 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 $\langle \langle x_*, x \rangle \rangle \leq p(x)$, $\forall x \in X$. Clearly, $x_* \neq \varnothing$. $\forall v \in V$, $\langle \langle x_*, v \rangle \rangle = f(v) = 1 =: c \in \mathbb{R}$. $\forall k \in K$, we have $\langle \langle x_*, k \rangle \rangle \leq p(k) \leq 1$. $\forall k \in K^\circ$, we have $\langle \langle x_*, k \rangle \rangle \leq p(k) < 1$. Hence, the closed hyperplane $H := \{ x \in X \mid \langle \langle x_*, x \rangle \rangle = c \}$ is the one we seek.

Case 2: $\varnothing \notin K^\circ$. Let $x_0 \in K^\circ$. 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, $K_1^\circ = K^\circ - x_0$. Then, $\varnothing \notin K_1^\circ$ and $V_1 \cap K_1^\circ = \emptyset$. By Case 1, there exist $c_1 \in \mathbb{R}$ and $x_* \in X^*$ with $x_* \neq \varnothing$, such that $\langle \langle x_*, v_1 \rangle \rangle = c_1$, $\forall v_1 \in V_1$, $\langle \langle x_*, k_1 \rangle \rangle < c_1$, $\forall k_1 \in K_1^\circ$ and $\langle \langle x_*, k_1 \rangle \rangle \leq c_1$, $\forall k_1 \in K_1^\circ$. Let $c := c_1 + \langle \langle x_*, x_0 \rangle \rangle \in \mathbb{R}$. Then, $\forall v \in V$, $\langle \langle x_*, v \rangle \rangle = c$, $\forall k \in K^\circ$, $\langle \langle x_*, k \rangle \rangle < c$. By Proposition 7.16, $K_1^\circ = K - x_0$. Then, $\forall k \in K_1^\circ$, $\langle \langle x_*, k \rangle \rangle \leq c$. Thus, the closed hyperplane $H := \{ x \in X \mid \langle \langle x_*, x \rangle \rangle = c \}$ is the one we seek.

This completes the proof of the theorem. □

**Definition 8.8** Let $X$ be a real normed linear space. A closed hyperplane $H := \{ x \in X \mid \langle \langle x_*, x \rangle \rangle = c \}$ 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 $\inf_{k \in K} \langle \langle x_*, k \rangle \rangle = c$ or $\sup_{k \in K} \langle \langle x_*, k \rangle \rangle = c$.

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 $K_1^\circ \neq \emptyset$ and $K_1^\circ \cap K_2 = \emptyset$. 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

$$\sup_{k_1 \in K_1} \langle \langle x_*, k_1 \rangle \rangle \leq c \leq \inf_{k_2 \in K_2} \langle \langle x_*, k_2 \rangle \rangle$$

**Proof** Let $K := K_1^\circ - K_2$. By Propositions 7.15 and 6.39, $K$ is convex. Clearly, $K^\circ \neq \emptyset$ since $K_1^\circ \neq \emptyset$ and $K_2 \neq \emptyset$. Since $K_1^\circ \cap K_2 = \emptyset$, then $\varnothing \notin K$. Let $V = \{ \varnothing \}$. Then, $V$ is a linear variety and $V \cap K^\circ = \emptyset$. By Mazur’s Theorem, $\exists x_* \in X^*$ with $x_* \neq \varnothing$ such that $\langle \langle x_*, k \rangle \rangle \leq \langle \langle x_*, v \rangle \rangle = 0$, $\forall k \in K$. $\forall k_1 \in K_1^\circ$, $\forall k_2 \in K_2$, $k_1 - k_2 \in K$ and $\langle \langle x_*, k_1 - k_2 \rangle \rangle \leq 0$. Then, $\langle \langle x_*, k_1 \rangle \rangle \leq \langle \langle x_*, k_2 \rangle \rangle$. Hence, $-\infty < \sup_{k_1 \in K_1^\circ} \langle \langle x_*, k_1 \rangle \rangle \leq c \leq \inf_{k_2 \in K_2} \langle \langle x_*, k_2 \rangle \rangle < +\infty$ for some $c \in \mathbb{R}$. By Proposition 8.6, $\overline{K_1} = K_1^\circ$. By Proposition 4.13, $\forall k_1 \in \overline{K_1}$, $\langle \langle x_*, k_1 \rangle \rangle \leq \sup_{k_1 \in K_1^\circ} \langle \langle x_*, k_1 \rangle \rangle \leq c$. Hence, we have

$$\sup_{k_1 \in K_1} \langle \langle x_*, k_1 \rangle \rangle = \sup_{k_1 \in K_1^\circ} \langle \langle x_*, k_1 \rangle \rangle \leq c \leq \inf_{k_2 \in K_2} \langle \langle x_*, k_2 \rangle \rangle$$

Thus, the closed hyperplane $H := \{ x \in X \mid \langle \langle x_*, x \rangle \rangle = c \}$ is the one we seek. This completes the proof of the theorem. □
Proof. We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( K = \emptyset \); Case 2: \( K \neq \emptyset \). Case 1: \( K = \emptyset \). Let \( x_* = \varnothing \). Then, \( \langle x_*, x_0 \rangle = 0 < +\infty = \inf_{k \in K} \langle x_*, k \rangle \). Hence, \( K \) is weakly closed. This case is proved.

Case 2: \( K \neq \emptyset \). \( \forall x_0 \in X \setminus K \), by Proposition 4.10, \( \inf_{k \in K} \| x_0 - k \| =: d \in (0, \infty) \subset \mathbb{R} \). Then, \( K_1 := B_X(x_0, d) \) is a convex set with nonempty interior and \( K_1 \cap K = \emptyset \). By Eidelheit Separation Theorem, there exists \( x_* \in X^* \) with \( x_* \neq \varnothing \) and \( \exists c \in \mathbb{R} \) such that \( \sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \leq \inf_{k \in K} \langle x_*, k \rangle \). By Lemma 7.75, \( \langle x_*, x_0 \rangle = \sup_{k_1 \in K_1} \langle x_*, k_1 \rangle \leq c \). Then, \( \langle x_*, x_0 \rangle < \inf_{k \in K} \langle x_*, k \rangle \).

Let \( \bar{K} \) be the closure of \( K \) in the weak topology \( O_{\text{weak}}(X) \). \( x_0 \in \{ x \in X \mid \langle x_*, x \rangle < c \} =: O \). Clearly, \( O \) is a weakly open set and \( O \cap K = \emptyset \). Then, by Proposition 3.3, \( x_0 \notin \bar{K} \). Thus, we have shown that \( \forall x_0 \in \bar{K}, x_0 \notin \bar{K} \). 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 \( X \) be a reflexive real normed linear space and \( K \subseteq X^* \) 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 \( K \subseteq \mathcal{B}_{X^*}(\varnothing, n) := S \). By Alaoglu Theorem, \( S \) is weak* compact. Since \( K \) is closed and convex, then, by Proposition 8.10, \( K \) is weakly closed. Since \( X \) is reflexive, then the weak topology and the weak* topology on \( X^* \) coinide. 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 \( X \) be a real normed linear space and \( K \subseteq X \) be a nonempty convex set. The support of \( K \) is the set \( K_{\text{supp}} := \{ x_* \in X^* \mid \sup_{k \in K} \langle x_*, k \rangle < +\infty \} \).

Proposition 8.13 Let \( X \) be a real normed linear space and \( K, G \subseteq X \) be nonempty convex sets. Then, the following statements hold.

(i) \( K_{\text{supp}} \subseteq X^* \) is a convex cone.

(ii) If \( K \) is closed, then \( K = \bigcap_{x_* \in K_{\text{supp}}} \{ x \in X \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle \} \)

\[ = \bigcap_{x_* \in K_{\text{supp}}, x_* \neq \varnothing} \{ x \in X \mid \langle x_*, x \rangle \leq \sup_{k \in K} \langle x_*, k \rangle \}, \text{ that is, } K \]

equals to the intersection of all closed half-spaces containing \( K \).
(iii) \((K + G)_{\text{supp}} = K_{\text{supp}} \cap G_{\text{supp}}\).

(iv) \((K \cap G)_{\text{supp}} \supseteq K_{\text{supp}} + G_{\text{supp}}\).

**Proof**

(i) Clearly, \(\vartheta_* \in K_{\text{supp}}\). \(\forall x_1, x_2 \in K_{\text{supp}}, \forall \alpha \in [0, \infty) \subseteq \mathbb{R},\) we have either \(\alpha = 0\), then \(\alpha x_1 = \vartheta_* \in K_{\text{supp}}\), or \(\alpha > 0\), then 

\[
\sup_{k \in K} \langle \alpha x_1, k \rangle = \sup_{k \in K} \alpha \langle x_1, k \rangle = \alpha \sup_{k \in K} \langle x_1, k \rangle < +\infty,
\]

where we have applied Proposition 3.8.1, which further implies that \(\alpha x_1 \in K_{\text{supp}}\).

This shows that \(K_{\text{supp}}\) is a cone. Note that \(\sup_{k \in K} \langle x_1 + x_2, k \rangle = \sup_{k \in K} \langle x_1, k \rangle + \sup_{k \in K} \langle x_2, k \rangle \leq +\infty\), where we have applied Proposition 3.8.1. Then, \(x_1 + x_2 \in K_{\text{supp}}\).

This coupled with the fact that \(K_{\text{supp}}\) is a cone implies that \(K_{\text{supp}}\) is a convex cone.

(ii) Clearly \(K \subseteq \bigcap_{x \in K_{\text{supp}}} \{ x \in X \mid \langle x, x \rangle \leq \sup_{k \in K} \langle x, k \rangle \} =: \hat{K}. \) \(\forall x \in X \setminus K,\) by Proposition 8.10, \(\exists x_0 \in X^*_+\) such that \(\langle x_0, x \rangle < \inf_{k \in K} \langle x_0, k \rangle\). Then, \(\sup_{k \in K} \langle -x_0, k \rangle < \langle -x_0, x \rangle < +\infty\). Hence, \(-x_0 \in K_{\text{supp}}\) and \(x \in X \setminus \hat{K}.\) This shows that \(X \setminus K \subseteq X \setminus \hat{K}\) and \(\hat{K} \subseteq K.\) Hence, \(K = \hat{K}.\) Note that, for \(\vartheta_* \in K_{\text{supp}}, \{ x \in X \mid \langle \vartheta_* , x \rangle \leq \sup_{k \in K} \langle \vartheta_* , k \rangle \} \subseteq \hat{K}.\) Then, \(K = \bigcap_{x \in K_{\text{supp}}} \{ x \in X \mid \langle \vartheta_* , x \rangle \leq \sup_{k \in K} \langle \vartheta_* , k \rangle \} \subseteq \hat{K}.\) Clearly, \(K \subseteq \hat{K}.\) Any closed half-space containing \(K\) can be expressed as \(\{ x \in X \mid \langle x, x \rangle \leq c \}\) for some \(x \in X^*\) with \(x \neq \vartheta_*\) and for some \(c \in \mathbb{R}.\) Then, \(c \geq \sup_{k \in K} \langle x, k \rangle = \mathbb{R}.\) Hence, \(x \in K_{\text{supp}}\) and \(\hat{K} \subseteq \bigcap_{x \in K_{\text{supp}}} \{ x \neq \vartheta_* \mid \sup_{k \in K} \langle x, k \rangle \} \subseteq \bigcap_{x \in K_{\text{supp}}} \{ x \in X \mid \langle x, x \rangle \leq \sup_{k \in K} \langle x, k \rangle \} = K.\) Hence, \(K = \hat{K}.\)

(iii) \(\forall x \in K_{\text{supp}} \cap G_{\text{supp}},\) 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 \(x \in (K + G)_{\text{supp}}.\) Hence, \(K_{\text{supp}} \cap G_{\text{supp}} \subseteq (K + G)_{\text{supp}}.\)

On the other hand, \(\forall x \in (K + G)_{\text{supp}},\) \(\sup_{x \in K + G} \langle x, x \rangle =: c < +\infty.\) Fix \(k_0 \in K\) 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, \(x \in K_{\text{supp}} \cap G_{\text{supp}}.\) This shows that \((K + G)_{\text{supp}} \subseteq K_{\text{supp}} \cap G_{\text{supp}}.\) Therefore, we have \((K + G)_{\text{supp}} = K_{\text{supp}} \cap G_{\text{supp}}.\)

(iv) \(\forall x \in K_{\text{supp}} + G_{\text{supp}},\) let \(x = x_1 + x_2\) with \(x_1 \in K_{\text{supp}}\) and \(x_2 \in G_{\text{supp}}.\) \(\forall x \in K \cap G,\) 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, \(x \in (K \cap G)_{\text{supp}}.\) Hence, we have \(K_{\text{supp}} + G_{\text{supp}} \subseteq (K \cap G)_{\text{supp}}.\)

This completes the proof of the proposition.

\(\square\)

**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 \(h : K_{\text{supp}} \to \mathbb{R}\) given by \(h(x) = \sup_{k \in K} \langle x, k \rangle, \forall x \in K_{\text{supp}}.\)
In the above definition, \( h \) takes value in \( \mathbb{R} \) since \( K \neq \emptyset \) and \( x_0 \in \text{supp}(K) \).

**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) \( \delta := \inf_{k \in K} \| y - k \| = \max_{x_0 \in K_{\text{supp}} \text{ and } \| k \| \leq 1} (\langle x_0, y \rangle - h(x_0)) \), where the infimum is achieved at some \( x_0 \in K_{\text{supp}} \) with \( x_0 \| \leq 1 \). If the inequality is equivalent to \( \| y - k_0 \| = \langle x_0, y \rangle - h(x_0) \).

(ii) If \( \exists k_0 \in K \) and \( \exists x_0 \in K_{\text{supp}} \) with \( \| x_0 \| = 1 \) 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 \( \delta = \| y - k_0 \| = \langle x_0, y \rangle - h(x_0) \).

**Proof**

(i) \( \forall x_0 \in K_{\text{supp}} \text{ with } \| x_0 \| \leq 1 \), \( \forall k \in K \), we have \( \| y - k \| \geq \| x_0 \| \| y - k \| \geq \langle x_0, y - k \rangle = \langle x_0, y \rangle - \langle x_0, k \rangle \geq \langle x_0, y \rangle - h(x_0) \). Hence, \( \delta = \inf_{k \in K} \| y - k \| \geq \sup_{x_0 \in K_{\text{supp}} \text{ and } \| x_0 \| \leq 1} (\langle x_0, y \rangle - h(x_0)) \).

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \delta = 0 \); Case 2: \( \delta > 0 \). Case 1: \( \delta = 0 \). Take \( x_0 = 0 \in K_{\text{supp}} \) and \( \| x_0 \| = 0 \leq 1 \). Then, \( \delta = 0 = \langle x_0, y \rangle - h(x_0) \). Then, \( \delta := \inf_{k \in K} \| y - k \| = \max_{x_0 \in K_{\text{supp}} \text{ and } \| x_0 \| \leq 1} (\langle x_0, y \rangle - h(x_0)) \). 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 \( K_1 = B_X(0, \delta) \). Then, \( K_1 \) is a convex set with nonempty interior, \( K \cap K_1 = \emptyset \). By Eidelheit Separation Theorem, there exists \( x_0 \in K_1^\circ \) with \( x_0 \neq 0 \) and \( \exists c \in \mathbb{R} \) such that \( \sup_{k \in K_1} \langle x_0, k \rangle \leq c \leq \inf_{k \in K} \langle x_0, k \rangle \). Take \( x_0 = -\| x_0 \| ^{-1} x_0 \). Then, by Proposition 3.81, we have \( \inf_{k \in K_1} \langle x_0, k \rangle \geq -c/\| x_0 \| \geq \sup_{k \in K} \langle x_0, k \rangle \). This shows, \( x_0 \in K_{\text{supp}} \) and \( \| x_0 \| = 1 \). The above inequality is equivalent to \( \langle x_0, y \rangle + \inf_{x \in B_X(0, \delta)} \langle x_0, x \rangle \geq h(x_0) \).

By Lemma 7.75, we have \( \langle x_0, y \rangle - h(x_0) \geq \delta \). Hence, we have \( \delta = \langle x_0, y \rangle - h(x_0) = \max_{x_0 \in K_{\text{supp}} \text{ and } \| x_0 \| \leq 1} (\langle x_0, y \rangle - h(x_0)) \). If the infimum is achieved at some \( k_0 \in K \), then \( \delta = \| y - k_0 \| = \langle x_0, y - k_0 \rangle - h(x_0) \leq \langle x_0, y \rangle - \langle x_0, k_0 \rangle = \| x_0 \| \| y - k_0 \| \leq \| y - k_0 \| \), where the second inequality follows from Proposition 7.72. Hence, \( \langle x_0, y - k_0 \rangle = \| x_0 \| \| y - k_0 \| \), \( x_0 \) is aligned with \( y - k_0 \), and \( \langle x_0, k_0 \rangle = h(x_0) \). This case is proved.

(ii) Note that \( \delta \leq \| y - k_0 \| = \| x_0 \| \| y - k_0 \| = \langle x_0, y - k_0 \rangle = \langle x_0, y \rangle - h(x_0) \leq \delta \). Then, the result follows.

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

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 8.17 Let \( X \) be a reflexive real normed linear space, \( x_* \in X^* \), and \( K \subseteq X^* \) be a nonempty closed convex set. Then, \( \delta = \min_{k \in K} \| x_* - k_* \| \) and the minimum is achieved at some \( k_{s0} \in K \).

Proof Fix a \( k_{s1} \in K \neq \emptyset \). Let \( \mu := \| x_* - k_{s1} \| \in [0, \infty) \subset \mathbb{R} \) and \( d = \mu + \| x_* \| + 1 \in (0, \infty) \subset \mathbb{R} \). Then, \( \forall k_* \in K \) with \( \| k_* \| > d \), we have \( \| x_* - k_* \| \geq \| k_* \| - \| x_* \| > \mu + 1 \). Thus, \( \delta = \inf_{k_* \in K} \| x_* - k_* \| \) and any \( k_{s0} \) achieving the infimum for the original problem must be in the set \( K \cap \overline{B}_{X^*}(\varnothing_* ,d) =: K_1 \). By Proposition 6.40, \( K_1 \) is bounded closed and convex. Note that \( k_{s1} \in K_1 \neq \emptyset \). Then, by Proposition 8.11, \( K_1 \) is weak* compact. By Propositions 7.121, 8.16, and 3.16, we have \( f : X^* \to \mathbb{R} \) given by \( f(k_*) = \| x_* - k_* \| \), \( \forall k_* \in X^* \), is lower semicontinuous. By Proposition 5.30 and Definition 3.14, there exists \( k_{s0} \in K_1 \) that achieves the infimum on \( K_1 \). Such \( k_{s0} \) achieves the infimum on \( K \). This completes the proof of the proposition. \( \square \)

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.
Proposition 8.21 Let $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ 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, $(f(x_0),x_0) \in [f,C]$. For all $(r,x) \in v([f,C]) = (f(x_0),x_0) + \text{span}([f,C] - (f(x_0),x_0))$, where the equality follows from Proposition 6.37, there exist $n \in \mathbb{Z}_+$, $r_1, x_1, \ldots, r_n, x_n \in [f,C]$, and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ 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, $x \in v(C) = x_0 + \text{span}(C - x_0)$ and $(r, x) \in v([f,C])$. Hence, $v([f,C]) \subseteq \mathbb{R} \times v(C)$.

On the other hand, $(r, x) \in \mathbb{R} \times v(C)$, then, by Proposition 6.37, $x \in v(C) = x_0 + \text{span}(C - x_0)$. Then, $\exists n \in \mathbb{Z}_+, \exists x_1, \ldots, x_n \in C$, and $\exists \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that $x - x_0 = \sum_{i=1}^n \alpha_i (x_i - x_0)$. $\exists r_0 \in (0, \infty) \subseteq \mathbb{R}$ 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 $(r - f(x_0), x - x_0) = \sum_{i=1}^n \alpha_i (f(x_i) - f(x_0), x_i - x_0) + \alpha_0 (r_0 - f(x_0), x_0 - x_0)$. 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 $\mathcal{X}$ be a real normed linear space, $C \subseteq \mathcal{X}$ 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 $r_0 \in (f(x_0), +\infty) \subseteq \mathbb{R}$.

Proof “Sufficiency” Fix $x_0 \in \partial C$ and $r_0 \in (f(x_0), +\infty) \subseteq \mathbb{R}$. Let $f$ be continuous at $x_0$. Then, for $\epsilon = (r_0 - f(x_0))/2 \in (0, \infty) \subseteq \mathbb{R}$, $\exists \delta \in (0, r_0 - f(x_0) - \epsilon) \subseteq \mathbb{R}$ such that $\forall x \in B(x_0, \delta) \cap V(C)$, we have $x \in C$ and $|f(x) - f(x_0)| < \epsilon$. Clearly, $(r_0, x_0) \in [f,C]$. $\forall (r, x) \in \overline{B}_{\mathbb{R} \times \mathcal{X}}((r_0, x_0), \delta) \cap V([f,C]) = \overline{B}_{\mathbb{R} \times \mathcal{X}}((r_0, x_0), \delta) \cap (\mathbb{R} \times V(C))$, we have $x \in B(x_0, \delta) \cap V(C)$ 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 $(r_0, x_0) \in \partial [f,C]$. □
Proposition 8.23 Let \((r_0, x_0) \in \circ [f, C]\). 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, \(B_X(x_0, \epsilon_0) \cap V(C) \subseteq C\), \(x_0 \in \circ C\), \(f(x) \leq r_0 - \epsilon_0\), \(\forall x \in B_X(x_0, \epsilon_0) \cap V(C)\), 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

\[
\begin{align*}
\delta &\leq (1 - \delta)f(x_0) + \delta f(x_0) < \epsilon, \\
\delta &\leq \frac{1}{1 + \delta} f(x) + \frac{\delta}{1 + \delta} f(x_0), \\
\Rightarrow &\quad (1 + \delta)f(x_0) \leq f(x) + (r_0 - \epsilon_0) \delta, \\
\Rightarrow &\quad f(x) \geq f(x_0) - (r_0 - f(x_0)) - (r_0 - f(x_0)) \delta > f(x_0) - \epsilon.
\end{align*}
\]

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, \(x_0 \in \circ C \neq \emptyset\), and \(f : C \rightarrow \mathbb{R}\) be convex. If \(f\) is continuous at \(x_0\), then \(f\) is continuous at \(x\), \(\forall x \in \circ C\).

**Proof** Fix any \(x \in \circ C\). 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 \(x_0 \in \circ C\), \(\exists \delta \in (0, \infty) \subset \mathbb{R}\) such that \(|f(y) - f(x_0)| < \epsilon\), \(\forall y \in B_X(x_0, \delta) \cap V(C) \subseteq C\). Since \(x \in \circ C\), then \(\exists \delta_1 \in (0, \delta] \subset \mathbb{R}\) such that \(B_X(x, \delta_1) \cap V(C) \subseteq C\). Take \(\beta = (1 - \beta)/\beta \in (1, \infty) \subset \mathbb{R}\).

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 \(\beta/\beta - 1\) \((y - x) + x_0 \in B_X(x, \delta_1) \cap V(C) \subseteq C\). 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, \(B_\mathbb{R}(r_1 + (\beta - 1)\delta_1/\beta) \cap V([f, C]) = B_\mathbb{R}(r, (\beta - 1)\delta_1/\beta) \cap (\mathbb{R} \times V(C)) \subseteq [f, C]\). Hence, \((r, x) \in \circ [f, C]\). By Proposition 8.22, \(f\) is continuous at \(x\). This completes the proof of the proposition. \(\square\)
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, $\forall x_0 \in \partial C$, $f$ is continuous at $x_0$.

Proof We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\partial C = \emptyset$; Case 2: $\partial C \neq \emptyset$. Case 1: $\partial C = \emptyset$. This is trivial.

Case 2: $\partial C \neq \emptyset$. Fix any $x_0 \in \partial C$. Then, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $K := f(B_X(x_0, \delta) \cap V(C)) \subseteq C$. 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: $n \in \mathbb{N}$. Case 2b: $n = 0$. 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 $\alpha_i, \beta_i \in [0, 1] \subseteq \mathbb{R}, i = \ldots, n$, 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, $\exists \xi \in [1, \infty) \subseteq \mathbb{R}$ such that $p(m)/\xi \leq \|m\| \leq \xi p(m)$, $\forall m \in M. \forall x \in B_X(x_0, \delta/e) \cap V(C)$, we have $x = x_0 + \sum_{i=1}^n \alpha_i e_i$ for some $\alpha_1, \ldots, \alpha_n \in R$. 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, $B_X(x_0, \delta/\xi) \cap V(C) \subseteq K \subseteq C$. Take $r = \delta/\xi + r_1$. It is easy to show that $B_{R \times X}(r, x_0), \delta/\xi) \cap V([f, C]) \subseteq [f, C]$, by Proposition 8.21. Then, $(r, x_0) \in \partial [f, C]$. By Proposition 8.22, $f$ is continuous at $x_0$.

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

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

$$\{x \in C \mid f(x) < a\} = C \setminus \{x \in C \mid f(x) \leq -a\} = C \setminus T_{-a}$$
8.5. **Conjugate Convex Functionals**

Clearly, \( C \setminus T_{-\alpha} \) 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 \( \lim_{n \in \mathbb{N}} r_n = r_0 \) and \( \lim_{n \in \mathbb{N}} x_n = x_0 \). 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, \( -f(x_0) \geq \limsup_{x \rightarrow x_0} (-f(x)) = -\liminf_{x \rightarrow x_0} f(x) \). By Proposition 3.87, we have \( f(x_0) \leq \liminf_{x \rightarrow x_0} f(x) \leq \liminf_{n \in \mathbb{N}} f(x_n) \leq \liminf_{n \in \mathbb{N}} r_n = r_0 \). 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.

### Proposition 8.26

Let \( X \) be a real normed linear space, \( C \subseteq X \) be convex, and \( f : C \rightarrow \mathbb{R} \) be convex. Assume that \([f, C] \) is closed. Then, \( f \) is weakly lower semicontinuous.

#### Proof

\( \forall a \in \mathbb{R}, V_a := \{(a, x) \in \mathbb{R} \times X \mid x \in X\} \) 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. Since \( f \) is convex, then \( T_a \) is convex. By Proposition 8.10, \( T_a \) is weakly closed. Note that

\[ \{x \in C \mid -f(x) < a\} = C \setminus \{x \in C \mid f(x) \leq -a\} = C \setminus T_{-\alpha} \]

Clearly, \( C \setminus T_{-\alpha} \) is weakly upper semicontinuous and \( f \) is weakly lower semicontinuous. This completes the proof of the proposition.

### 8.5 Conjugate Convex Functionals

#### Definition 8.27

Let \( X \) be a real normed linear space, \( C \subseteq X \) be a nonempty convex set, and \( f : C \rightarrow \mathbb{R} \) be convex. The conjugate set \( C_{\text{conj}} \) is defined as

\[ C_{\text{conj}} := \{x^* \in X^* \mid \sup_{x \in C} (\langle x^*, x \rangle - f(x)) < +\infty\} \]

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 \(x_\ast \in C_{\text{conj}}\) and \(C \neq \emptyset\).

**Proposition 8.28** Let \(X\) be a real normed linear space, \(C \subseteq X\) be a nonempty convex set, \(f : C \rightarrow R\) be convex. Then, \(C_{\text{conj}} \subseteq X^*\) is convex, \(f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}^*\) is convex, and \([f_{\text{conj}},C_{\text{conj}}]\) \(\subseteq \mathbb{R} \times X^*\) is a closed convex set.

**Proof** \(\forall x_1, x_2 \in C_{\text{conj}}, \forall \alpha \in [0, 1] \subseteq \mathbb{R}\), let \(M_i := f_{\text{conj}}(x_i) = \sup_{x \in C} (\langle x_i, x \rangle - f(x)) \in \mathbb{R}, i = 1, 2\). \(\forall x \in C\), we have \(\langle \alpha x_1 + (1 - \alpha) x_2, x \rangle - f(x) = \alpha (\langle x_1, x \rangle - f(x)) + (1 - \alpha) (\langle x_2, x \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha) M_2 + \infty\). \(\forall \alpha \in [0, 1]\), \(\sup_{x \in C} (\langle \alpha x_1 + (1 - \alpha) x_2, x \rangle - f(x)) \leq \alpha M_1 + (1 - \alpha) M_2\). Hence, \(f_{\text{conj}}(x_1) + (1 - \alpha)f_{\text{conj}}(x_2)\). 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 \((s_k, x_k)_{k=1}^\infty \subseteq [f_{\text{conj}},C_{\text{conj}}]\) such that \(\lim_{k \in \mathbb{N}} (s_k, x_k) = (s_0, x_0)\). By Proposition 3.67, \(\lim_{k \in \mathbb{N}} s_k = s_0\) and \(\lim_{k \in \mathbb{N}} x_k = x_0\). \(\forall x \in C\), \(\forall k \in \mathbb{N}\), since \((s_k, x_k) \in [f_{\text{conj}},C_{\text{conj}}]\), \(\langle x_k, x \rangle - f(x) \leq f_{\text{conj}}(x_k) \leq s_k\). By Propositions 7.72 and 3.66, \(\langle x_0, x \rangle - f(x) \leq \lim_{k \in \mathbb{N}} \langle x_k, x \rangle - f(x) \leq \lim_{k \in \mathbb{N}} s_k = s_0 + \infty\). 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.

Geometric interpretation of \(f_{\text{conj}} : C_{\text{conj}} \rightarrow \mathbb{R}\). \(\forall x_\ast \in C_{\text{conj}}, \forall (r,x) \in [f,C]\), we have \(f(x) \leq r\). Then, \(\langle ((-1, x_\ast), (r,x) \rangle = \langle (x_\ast, x) \rangle - r \leq f_{\text{conj}}(x_\ast)\). It is easy to recognize that \(\sup_{(r,x) \in [f,C]} \langle ((-1, x_\ast), (r,x) \rangle = f_{\text{conj}}(x_\ast)\). Hence, the closed hyperplane \(H := \{(r,x) \in \mathbb{R} \times X | \langle ((-1, x_\ast), (r,x) \rangle = f_{\text{conj}}(x_\ast)\}\) 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 \(f : C \rightarrow \mathbb{R}\) be a convex functional. Assume that \([f,C] =: K \subseteq \mathbb{R} \times X\) is closed. Then, \(\exists x_0 \in X^*\) such that

\[
\sup_{(r,x) \in K} \langle ((-1, x_\ast), (r,x) \rangle < +\infty
\]

Then, \(x_0 \in C_{\text{conj}} \neq \emptyset\).

**Proof** \(\forall x_\ast \in C \neq \emptyset\). 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, \(\langle (r_0, x_0) \rangle \notin K\). By Proposition 8.10 and Example 7.76, \(\exists (\bar{s}_0, \bar{x}_0) \in \mathbb{R} \times X^*\) such that

\[
\langle (\bar{s}_0, \bar{x}_0), (r_0, x_0) \rangle < \inf_{(r,x) \in K} \langle (\bar{s}_0, \bar{x}_0), (r,x) \rangle
\]  (8.1)
Since \((f(x_0), x_0) \in K\), then \(\bar{s}_0 r_0 + \langle \bar{x}_0, x_0 \rangle < \bar{s}_0 f(x_0) + \langle \bar{x}_0, x_0 \rangle\), 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

\[
\langle\langle (-1, x_0), (r_0, x_0) \rangle\rangle > \sup_{(r, x) \in K} \langle\langle (-1, x_0), (r, x) \rangle\rangle
\]

The above implies that \(\langle\langle (-1, x_0), (r_0, x_0) \rangle\rangle > \sup_{x \in C} \langle\langle (x_0, x) \rangle\rangle - f(x)\).

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 \(K \subseteq \mathbb{R} \times X =: W\) 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 \(\exists (s_1, x_1) \in W = \mathbb{R} \times X^*\) with \(s_1 \neq 0\) such that \(\sup_{(r, x) \in K} \langle\langle (s_1, x_1), (r, x) \rangle\rangle = +\infty\).

Then, \(\forall (r_0, x_0) \in W \setminus K\), there exists a nonvertical hyperplane separating \((r_0, x_0)\) and \(K\), that is, \(\exists \bar{x}_0 \in X^*\) such that either

\[
\sup_{(r, x) \in K} \langle\langle (-1, x_0), (r, x) \rangle\rangle < \langle\langle (-1, x_0), (r_0, x_0) \rangle\rangle \quad (8.2a)
\]

or

\[
\inf_{(r, x) \in K} \langle\langle (-1, x_0), (r, x) \rangle\rangle > \langle\langle (-1, x_0), (r_0, x_0) \rangle\rangle \quad (8.2b)
\]

**Proof** Fix \((r_0, x_0) \in W \setminus K\). By Proposition 8.10, \(\exists (\bar{s}_0, \bar{x}_0) \in \mathbb{R} \times X^*\) such that

\[
\langle\langle (\bar{s}_0, \bar{x}_0), (r_0, x_0) \rangle\rangle < \inf_{(r, x) \in K} \langle\langle (\bar{s}_0, \bar{x}_0), (r, x) \rangle\rangle \quad (8.3)
\]

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

\[
-r_0 + \langle\langle x_0, x_0 \rangle\rangle > \sup_{(r, x) \in K} (-r + \langle\langle x_0, x \rangle\rangle)
\]

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

\[
-r_0 + \langle\langle x_0, x_0 \rangle\rangle < \inf_{(r, x) \in K} (-r + \langle\langle x_0, x \rangle\rangle)
\]

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 \(x_0 = \emptyset X^*\). Then,

\[
-r_0 > -\infty = \sup_{(r, x) \in K} \langle\langle (-1, x_0), (r, x) \rangle\rangle
\]
Hence, (8.2a) holds.

Case 3b: \( K \neq \emptyset \). Define

\[
M_1 := \sup_{(r,x) \in K} \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle - \langle\langle (s_1, x_{s1}), (r_0, x_0) \rangle\rangle \in \mathbb{R}
\]

\[
M_2 := \inf_{(r,x) \in K} \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r, x) \rangle\rangle - \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r_0, x_0) \rangle\rangle \in (0, \infty) \subset \mathbb{R}
\]

Let \( \delta := \frac{M_2}{1 + |M_1|} \in (0, \infty) \subset \mathbb{R} \) and \( (\tilde{s}_0, \tilde{x}_{s0}) = (\tilde{s}_0, \tilde{x}_{s0}) - \delta(s_1, x_{s1}) \in \mathbb{R} \times \mathbb{X}^* \). Then, we have

\[
\inf_{(r,x) \in K} \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r, x) \rangle\rangle - \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
= \inf_{(r,x) \in K} \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r, x) \rangle\rangle - \delta \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
\geq \inf_{(r,x) \in K} \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r, x) \rangle\rangle + \inf_{(r,x) \in K} (\delta \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
= \inf_{(r,x) \in K} \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r, x) \rangle\rangle - \delta \sup_{(r,x) \in K} \langle\langle (s_1, x_{s1}), (r, x) \rangle\rangle
\]

\[
- \langle\langle (\tilde{s}_0, \tilde{x}_{s0}), (r_0, x_0) \rangle\rangle
\]

\[
= M_2 - \delta M_1 > 0
\]

where we have applied Proposition 3.81 in the above. Hence, we have obtained an alternative \((\tilde{s}_0, \tilde{x}_{s0}) := (\tilde{s}_0, \tilde{x}_{s0})\) such that (8.3) holds. For this alternative pair, \( \tilde{s}_0 = -\delta s_1 \neq 0 \). Hence, this case can be solved by Case 1 or Case 2 with the alternative pair.

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

**Definition 8.31** Let \( \mathbb{X} \) be a real normed linear space, \( \Gamma \subseteq \mathbb{X}^* \) be a nonempty convex set, and \( \varphi : \Gamma \rightarrow \mathbb{R} \) be convex. The pre-conjugate set \( \text{conj} \Gamma \) is defined as

\[
\text{conj} \Gamma := \{ x \in \mathbb{X} \mid \sup_{x_* \in \Gamma} (\langle\langle x_*, x \rangle\rangle - \varphi(x_*) < +\infty \}
\]

and the functional \( \text{conj} \varphi : \text{conj} \Gamma \rightarrow \mathbb{R} \) pre-conjugate to \( \varphi \) is defined by

\[
\text{conj} \varphi (x) = \sup_{x_* \in \Gamma} (\langle\langle x_*, x \rangle\rangle - \varphi(x_*)), \quad \forall x \in \text{conj} \Gamma
\]

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, \([f, C] = \text{conj}[[f, C]_{\text{conj}}]\). Therefore, \( \forall x_0 \in C, f(x_0) = \sup_{x \in C_{\text{conj}}}((\langle x_*, x_0 \rangle) - f_{\text{conj}}(x_*)). \)

**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, \( \text{conj}(C_{\text{conj}}) \) and \( \text{conj}(f_{\text{conj}}) \) are well defined. Thus, \( \text{conj}[[f, C]_{\text{conj}}] \) is a well-defined set.

We will first show that \([f, C] \subseteq \text{conj}[[f, C]_{\text{conj}}] \). Fix any \((r, x) \in [f, C]\). \( \forall x_0 \in C_{\text{conj}}, \) we have \( f_{\text{conj}}(x_0) \geq \langle \langle x_0, x \rangle \rangle - f(x) \). Hence, we have \( r \geq f(x) \geq \langle \langle x_0, x \rangle \rangle - f_{\text{conj}}(x_0) \). Thus, we have

\[
\sup_{x_0 \in C_{\text{conj}}} \langle \langle x_0, x \rangle \rangle - f_{\text{conj}}(x_0) = \text{conj}(f_{\text{conj}})(x)
\]

Hence, \((r, x) \in \text{conj}[[f, C]_{\text{conj}}] \) and \([f, C] \subseteq \text{conj}[[f, C]_{\text{conj}}] \).

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

\[
\sup_{(r, x) \in [f, C]} \langle \langle -1, x_0 \rangle \rangle, (r, x) \rangle < +\infty
\]

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

\[
\sup_{(r, x) \in K} \langle \langle -1, x_0 \rangle \rangle, (r, x) \rangle < \langle \langle -1, x_0 \rangle \rangle, (r_0, x_0) \rangle \tag{8.4a}
\]

or

\[
\inf_{(r, x) \in K} \langle \langle -1, x_0 \rangle \rangle, (r, x) \rangle > \langle \langle -1, x_0 \rangle \rangle, (r_0, x_0) \rangle \tag{8.4b}
\]

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, \( c := \sup_{(r, x) \in K} \langle \langle -1, x_0 \rangle \rangle, (r, x) \rangle = \sup_{x \in C}((\langle x_0, x \rangle) - f(x)) = f_{\text{conj}}(x_0) \in \mathbb{R} \) and \( x_0 \in C_{\text{conj}} \). Then, \( (c, x_0) \in [f_{\text{conj}}, C_{\text{conj}}] \). But, \( c < -r_0 + \langle \langle x_0, x_0 \rangle \rangle \) implies that \( r_0 - f_{\text{conj}}(x_0) \leq \sup_{x_0 \in C_{\text{conj}}} \langle \langle x_0, x_0 \rangle \rangle - f_{\text{conj}}(x_0) \). Therefore, \( (r_0, x_0) \notin \text{conj}[[f, C]_{\text{conj}}] \) and \( [f, C] \subseteq \text{conj}[[f, C]_{\text{conj}}] \).

This yields \([f, C] = \text{conj}[[f, C]_{\text{conj}}] \), which implies that \( f \) and \( \text{conj}(f_{\text{conj}}) \) admit the same domain of definition and equal to each other on \( C \). Therefore, \( \forall x_0 \in C, f(x_0) = \text{conj}(f_{\text{conj}})(x_0) = \sup_{x_0 \in C_{\text{conj}}}((\langle x_0, x_0 \rangle) - f_{\text{conj}}(x_0)) \).}

**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 semi-continuous 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, \( f(x_0) = \sup_{x_0 \in C_{\text{conj}}}((\langle x_0, x \rangle) - f_{\text{conj}}(x_0)). \)
Proof: \( \forall x_0 \in C_{\text{conj}}, \) we have \( \langle \langle x_0, x_0 \rangle \rangle - f_{\text{conj}}(x_0) = \langle \langle x_0, x_0 \rangle \rangle + \inf_{x \in C}(\langle \langle x_0, x \rangle \rangle + f(x)) \leq f(x_0). \) Then,

\[
\begin{align*}
f(x_0) & \geq \sup_{x_0 \in \mathcal{K}} (\langle \langle x_0, x_0 \rangle \rangle - f_{\text{conj}}(x_0)) \\
& \geq \sup_{r \in C} (\langle \langle x_0, x_0 \rangle \rangle - f_{\text{conj}}(x_0)) \quad \forall x_0 \in C_{\text{conj}}, \end{align*}
\]

Let \( r < f(x_0) - \epsilon + \delta \leq f(x_0) - \epsilon/2 \) and \( r \geq f(x) \). This leads to \( f(x) < f(x_0) - \epsilon/2 \), which is a contradiction. Hence, \( K^*_1 \cap K_2 = 0 \). This completes the proof of the claim.

By Eidelheit Separation Theorem and Example 8.76, \( \exists c \in \mathbb{R} \) and \( \exists (s_0, \bar{x}_0) \in \mathcal{W}^* = \mathbb{R} \times \mathcal{X}^* \) with \( (s_0, \bar{x}_0) \neq (0, \vartheta_{\mathcal{X}^*}) \) such that

\[
\sup_{(r,x) \in K_1} \langle \langle s_0, x_0 \rangle, (r,x) \rangle \leq c \leq \inf_{(r,x) \in K_2} \langle \langle s_0, x_0 \rangle, (r,x) \rangle \quad (8.5)
\]

The second inequality in (8.5) implies that \( -\infty < c \leq \inf_{(r,x) \in K_2} \langle \langle s_0, x_0 \rangle, (r,x) \rangle \). Since \( K_2 = [f, C] \), then \( s_0 \geq 0 \).

Claim 8.33.2 \( \bar{s}_0 > 0 \).

Proof of claim: Suppose \( \bar{s}_0 = 0 \). By the fact that \( (s_0, \bar{x}_0) \neq (0, \vartheta_{\mathcal{X}^*}) \), we have \( \bar{x}_0 \neq \vartheta_{\mathcal{X}^*} \). Then, (8.5) implies that \( \langle \langle \bar{x}_0, x_0 \rangle \rangle \geq c = \sup_{(r,x) \in K_1} \langle \langle \bar{x}_0, x \rangle \rangle = \sup_{x \in B_{\mathcal{X}}(x_0, \delta)} \langle \langle \bar{x}_0, x \rangle \rangle \). This is impossible since \( \bar{x}_0 \neq \vartheta_{\mathcal{X}^*} \). Therefore, \( s_0 < 0 \). This completes the proof of the claim.

Let \( x_0 = -s_0^{-1} \bar{x}_0 \in \mathcal{X}^* \). Then, (8.5) is equivalent to

\[
\inf_{(r,x) \in K_1} (\langle \langle x_0, x_0 \rangle \rangle - r) \geq -\bar{s}_0^{-1}c \geq \sup_{(r,x) \in K_2} (\langle \langle x_0, x \rangle \rangle - r) = \sup_{x \in C} (\langle \langle x_0, x \rangle \rangle - f(x))
\]

Hence, \( x_0 \in C_{\text{conj}} \) and \( f_{\text{conj}}(x_0) \leq \langle \langle x_0, x_0 \rangle \rangle - f(x_0) + \epsilon \). Therefore, we have

\[
(\langle \langle x_0, x_0 \rangle \rangle - f_{\text{conj}}(x_0)) \geq f(x_0) - \epsilon
\]

By the arbitrariness of \( \epsilon \), \( f(x_0) = \sup_{x \in C_{\text{conj}}} (\langle \langle x_0, x \rangle \rangle - f_{\text{conj}}(x_0)) \).

This completes the proof of the proposition.

Proposition 8.34 Let \( \mathcal{X} \) be a real normed linear space, \( C \subseteq \mathcal{X} \) be a nonempty convex set, \( C^\circ \neq \emptyset \), \( f : C \to \mathbb{R} \) be a convex functional, \( f \) be continuous at \( \bar{x} \in C^\circ \), \( C_{\text{conj}} \subseteq \mathcal{X}^* \) be the conjugate set, and \( f_{\text{conj}} : C_{\text{conj}} \to \mathbb{R} \) be the conjugate functional. Then, \( \forall x_0 \in C^\circ \),

\[
f(x_0) = \max_{x \in C_{\text{conj}}} (\langle \langle x_0, x \rangle \rangle - f_{\text{conj}}(x_0)).
\]
8.5. **CONJUGATE CONVEX FUNCTIONALS**

\[ \sup x \in C_{\text{conj}} \text{, we have } \langle \langle x_*, x_0 \rangle \rangle - f_{\text{conj}}(x_*) = \langle \langle x_*, x_0 \rangle \rangle + \inf_{x \in C} \langle \langle x_*, x \rangle \rangle + f(x) \leq f(x_0). \text{ Then, } \]

\[ f(x_0) \geq \sup_{x_* \in C_{\text{conj}}} \left( \langle \langle x_*, x_0 \rangle \rangle - f_{\text{conj}}(x_*) \right) \]

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 \((\bar{r}, \bar{x})\) for some \( \bar{r} > f(\bar{x}) \). By Proposition 8.21, \( V([f,C]) = \mathbb{R} \times V(C) = \mathbb{R} \times X = W \) since \( C^o \neq \emptyset \). Hence, \((\bar{r}, \bar{x}) \in K^o \neq \emptyset \). Note that \((f(x_0) - \delta, x_0) \notin [f, C], \forall \delta \in (0, \infty) \subset \mathbb{R}\). Then, \((f(x_0), x_0) \notin K^o\). Hence, \( V \cap K^o = \emptyset \). By Mazur’s Theorem and Example 7.76, \( \exists x \in \mathbb{R} \) and \( \exists (s_0, \bar{x}_s) \in W^* = \mathbb{R} \times X^* \) with \( (s_0, \bar{x}_s) \neq (0, \emptyset X) \) such that

\[ \langle \langle (s_0, \bar{x}_s), (f(x_0), x_0) \rangle \rangle = c \geq \sup_{(r, x) \in K} \langle \langle (s_0, \bar{x}_s), (r, x) \rangle \rangle \]

which is equivalent to

\[ s_0f(x_0) + \langle \langle \bar{x}_s, x_0 \rangle \rangle \geq \sup_{(r, x) \in K} (s_0r + \langle \langle \bar{x}_s, x \rangle \rangle) \tag{8.6} \]

Since \( K = [f, C] \), then \( s_0 \leq 0 \), otherwise the right-hand-side of (8.6) equals to \(+\infty\).

**Claim 8.34.1** \( s_0 < 0 \).

**Proof of claim:** Suppose \( s_0 = 0 \). By the fact that \((s_0, \bar{x}_s) \neq (0, \emptyset X^*)\), we have \( \bar{x}_s \neq \emptyset X^* \). Then, (8.6) implies that

\[ \langle \langle \bar{x}_s, x_0 \rangle \rangle \geq \sup_{x \in C} \langle \langle \bar{x}_s, x \rangle \rangle \]

Note that \( x_0 \in C^o \) and \( \bar{x}_s \neq \emptyset X^* \) implies that

\[ \sup_{x \in C} \langle \langle \bar{x}_s, x \rangle \rangle > \langle \langle \bar{x}_s, x_0 \rangle \rangle \]

This is a contradiction. Therefore, \( s_0 < 0 \). This completes the proof of the claim. \( \square \)

Let \( x_{s_0} := [s_0]^{-1} \bar{x}_s \in X^* \). Then, (8.6) is equivalent to

\[ -f(x_0) + \langle \langle x_{s_0}, x_0 \rangle \rangle \geq \sup_{(r, x) \in K} (-r + \langle \langle x_{s_0}, x \rangle \rangle) = \sup_{x \in C} (\langle \langle x_{s_0}, x \rangle \rangle - f(x)) \]

Hence, \( x_{s_0} \in C_{\text{conj}} \) and \(-f(x_0) + \langle \langle x_{s_0}, x_0 \rangle \rangle \geq f_{\text{conj}}(x_{s_0})\).

Therefore, we have

\[ f(x_0) \leq \langle \langle x_{s_0}, x_0 \rangle \rangle - f_{\text{conj}}(x_{s_0}) \leq \sup_{x_* \in C_{\text{conj}}} (\langle \langle x_*, x_0 \rangle \rangle - f_{\text{conj}}(x_*)) \leq f(x_0) \]

Hence, \( f(x_0) = \max_{x_* \in C_{\text{conj}}} (\langle \langle x_*, x_0 \rangle \rangle - f_{\text{conj}}(x_*)) \), where the maximum is achieved at \( x_{s_0} \). \( \square \)
Figure 8.1: Fenchel duality.

8.6 Fenchel Duality Theorem

Let $\mathcal{X}$ be a real normed linear space, $C_1, C_2 \subseteq \mathcal{X}$ 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 $f_{1\text{conj}}, f_{1\text{conj}}, f_{2\text{conj}}, f_{2\text{conj}}$ 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

$$\mu = \inf_{(r_1, x) \in [f_1, C_1], (r_2, x) \in [f_2, C_2]} r_1 + r_2$$

The idea of Fenchel Duality Theorem can be illustrated in Figure 8.1.

**Theorem 8.35 (Fenchel Duality Theorem)** Let $\mathcal{X}$ be a real normed linear space, $C_1, C_2 \subseteq \mathcal{X}$ be nonempty convex sets, $f_1 : C_1 \to \mathbb{R}$ and $f_2 : C_2 \to \mathbb{R}$ be convex functionals, and $C_1 \cap C_2 \neq \emptyset$. Assume that that $f_1$ is continuous at $\bar{x} \in C_2$ 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) $\mu = \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \max_{x^* \in C_{1\text{conj}} \cap (-C_{2\text{conj}})} (-f_{1\text{conj}}(x^*) - f_{2\text{conj}}(-x^*))$, where the maximum is achieved at some $x_{*0} \in C_{1\text{conj}} \cap$
8.6. Fenchel Duality Theorem

\(-C_{2\text{conj}}\). If the infimum is achieved by some \(x_0 \in C_1 \cap C_2\), then, we have

\[
\begin{align*}
 f_{1\text{conj}}(x_0) &= \langle \langle x_0, x_0 \rangle \rangle - f_1(x_0) \quad (8.7a) \\
 f_{2\text{conj}}(-x_0) &= \langle \langle -x_0, x_0 \rangle \rangle - f_2(x_0) \quad (8.7b)
\end{align*}
\]

(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 the infimum is achieved at \(x_0\) and the maximum is achieved at \(x_0\).

Proof \(\forall x_0 \in C_{1\text{conj}} \cap (-C_{2\text{conj}}), we have\)

\[
\begin{align*}
 -f_{1\text{conj}}(x_0) &= -\sup_{x \in C_1} (\langle \langle x, x \rangle \rangle - f_1(x)) = \inf_{x \in C_1} (-\langle \langle x, x \rangle \rangle + f_1(x)) \\
 &\leq \inf_{x \in C_1 \cap C_2} (-\langle \langle x, x \rangle \rangle + f_1(x)) \\
 -f_{2\text{conj}}(-x_0) &= -\sup_{x \in C_2} (\langle \langle -x, x \rangle \rangle - f_2(x)) \\
 &= \inf_{x \in C_2} (\langle \langle x, x \rangle \rangle + f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (\langle \langle x, x \rangle \rangle + f_2(x))
\end{align*}
\]

Then, we have

\[
- f_{1\text{conj}}(x_0) - f_{2\text{conj}}(-x_0) \leq \inf_{x \in C_1 \cap C_2} (-\langle \langle x, x \rangle \rangle + f_1(x))
\]

+ \[
\inf_{x \in C_1 \cap C_2} (\langle \langle x, x \rangle \rangle + f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) = \mu (8.8)
\]

(i) Consider the sets \(K_1 := \{f_1 - \mu, C_1\} \) and \(K_2 := \{(r, x) \in \mathbb{R} \times X \mid x \in C_2, r \leq -f_2(x)\} = \{(r, x) \in \mathbb{R} \times X \mid (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, \(V(\{f_1 - \mu, C_1\}) = \mathbb{R} \times V(C_1) = \mathbb{R} \times X\). Then, \((\bar{r}, \bar{x}) \in K_1^\circ \neq \emptyset\).

We will show that \(K_1^\circ \cap K_2 = \emptyset\) by an argument of contradiction. Suppose \((r, x) \in K_1^\circ \cap K_2 \neq \emptyset\). Then, we have \((r, x) \in K_1^\circ\), 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, \(K_1^\circ \cap K_2 = \emptyset\).

By Eidelheit Separation Theorem and Example 7.76, \(\exists (\bar{s}_0, \bar{x}_0) \in \mathbb{R} \times X^*\) with \((\bar{s}_0, \bar{x}_0) \notin \partial_{\mathbb{R} \times X_\ast}\) such that

\[
-\infty < \sup_{(r, x) \in K_1} (\langle \langle \bar{x}_0, x \rangle \rangle + r\bar{s}_0) \leq \inf_{(r, x) \in K_2} (\langle \langle \bar{x}_0, x \rangle \rangle + r\bar{s}_0) < +\infty (8.9)
\]

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, \(\sup_{(r, x) \in K_1} (\langle \langle \bar{x}_0, x \rangle \rangle + r\bar{s}_0) = +\infty\), which contradicts (8.9). Case 2: \(\bar{s}_0 = 0\). Then, \(\bar{x}_0 \notin \partial_{X_\ast}\). By (8.9), we have

\[
-\infty < \sup_{x \in C_1} (\langle \langle \bar{x}_0, x \rangle \rangle \leq \inf_{x \in C_2} (\langle \langle \bar{x}_0, x \rangle \rangle < +\infty
\]
Let \( x \in C_1^0 \cap C_2 \neq \emptyset \). Then, \( \sup_{x \in C_1} \langle x_0, x \rangle \leq \langle x_0, \hat{x} \rangle \). This is not possible since \( \hat{x} \in C_1^0 \) and \( \hat{x} \neq \partial \mathcal{X} \). Hence, we have a contradiction in both cases. Therefore, \( \bar{s}_0 < 0 \).

Let \( x_0 = |\bar{s}_0|^{-1}x_0 \). Then, (8.9) is equivalent to

\[
-\infty < \sup_{(r,x) \in K_1} \langle x_0, x \rangle - r \leq \inf_{(r,x) \in K_2} \langle x_0, x \rangle - r < +\infty \quad (8.10)
\]

Note that

\[
+\infty > \sup_{(r,x) \in K_2} \langle x_0, x \rangle - r = \sup_{x \in C_1} \langle x_0, x \rangle - f_1(x) + \mu \\
\geq \sup_{x \in C_1 \cap C_2} \langle x_0, x \rangle - f_1(x) + \mu =: d_1 > -\infty
\]

Hence, \( x_0 \in C_{1\text{conj}} \). Note also that,

\[
-\infty < \inf_{(r,x) \in K_2} \langle x_0, x \rangle - r = \inf_{x \in C_2} \langle x_0, x \rangle + f_2(x) \\
= -\sup_{x \in C_2} (\langle -x_0, x \rangle - f_2(x)) \leq \inf_{x \in C_1 \cap C_2} (\langle x_0, x \rangle + f_2(x)) \\
=: d_2 < +\infty
\]

Hence, \( x_0 \in -C_{2\text{conj}} \). By (8.10), \( d_1 \leq d_2 \). On the other hand, we have

\[
d_2 - d_1 = \inf_{x \in C_1 \cap C_2} (\langle x_0, x \rangle + f_2(x)) + \inf_{x \in C_1 \cap C_2} (-\langle x_0, x \rangle + f_1(x)) - \mu \\
\leq \inf_{x \in C_1 \cap C_2} (f_1(x) + f_2(x)) - \mu = 0
\]

Hence, we have

\[
d_1 = \sup_{x \in C_1} (\langle x_0, x \rangle + f_1(x) + \mu) = -\sup_{x \in C_2} (\langle -x_0, x \rangle - f_2(x)) = d_2
\]

This implies that \( d_1 = f_{1\text{conj}}(x_0) + \mu = -f_{2\text{conj}}(-x_0) = d_2 \). Therefore, we have \( \mu = -f_{1\text{conj}}(x_0) - f_{2\text{conj}}(-x_0) \) and \( x_0 \in C_{1\text{conj}} \cap (-C_{2\text{conj}}) \).

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 \( x_0 \in C_{1\text{conj}} \cap (-C_{2\text{conj}}) \).

If the infimum is achieved at \( x_0 \in C_1 \cap C_2 \), then we have

\[
-f_{1\text{conj}}(x_0) \leq \inf_{x \in C_1 \cap C_2} (\langle x_0, x \rangle + f_1(x)) \leq -\langle x_0, x_0 \rangle + f_1(x_0) \\
-f_{2\text{conj}}(-x_0) \leq \inf_{x \in C_1 \cap C_2} (\langle x_0, x \rangle + f_2(x)) \leq \langle x_0, x_0 \rangle + f_2(x_0)
\]
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

\[ \mu \geq -f_{1\text{conj}}(x_0) - f_{2\text{conj}}(-x_0) = -\langle \langle x_0, x_0 \rangle \rangle + f_1(x_0) + f_2(x_0) = f_1(x_0) + f_2(x_0) \leq \mu \]

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, \( J : \Omega \to \mathbb{R}_e \), \( \Lambda \) be an index set. \( A : \Lambda \to \Omega^2 \). Assume \( \bigcup_{\lambda \in \Lambda} A(\lambda) = \Omega \). Then,

\[ \inf_{\omega \in \Omega} J(\omega) = \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega); \quad \sup_{\omega \in \Omega} J(\omega) = \sup_{\lambda \in \Lambda} \sup_{\omega \in A(\lambda)} J(\omega) \]

**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,

\[ \inf_{\omega \in \Omega} J(\omega) \leq \inf_{\lambda \in \Lambda} \inf_{\omega \in A(\lambda)} J(\omega) \]

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, \( g_1 : X \to \mathbb{R}_e, g_2 : Y \to \mathbb{R}_e \). Assume that, \( \forall x \in X, \forall y \in Y, g_1(x) + g_2(y) \in \mathbb{R}_e \) 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.
Hence, we have
\[ W \]

By Proposition 3.81, we have
\[ X \]
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: \( \mu_r < +\infty \). Then, \( +\infty \geq \mu_l \geq +\infty \). Hence, the desired equality holds.

Case 2b: \( \mu_r < +\infty \). \( \forall x \in X \) and \( \exists y \in Y \) such that \( g_1(x_0) + g_2(y_0) < m \). Then, again by Proposition 3.81, we have \( \mu_l < \mu_r \). 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 \( V : \mathcal{X} \times Y \rightarrow \mathbb{R} \). 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 \( W : \mathcal{X} \rightarrow \mathbb{R} \).

(ii) \( \forall y \in Y, \) define the function \( f_y : \mathcal{X} \rightarrow \mathbb{R} \) 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, \( \forall \bar{x} \in U, \forall y \in Y \), we have \( |f_y(\bar{x}) - f_y(x_0)| = |V(\bar{x}, y) - V(x_0, y)| < \epsilon \).

By (i), \( \exists \bar{y}, y_0 \in Y \) such that \( |W(\bar{x}) - V(\bar{x}, \bar{y})| < \epsilon \) and \( |W(x_0) - V(x_0, y_0)| < \epsilon \). Note that
\[
-2\epsilon < -|W(\bar{x}) - V(\bar{x}, \bar{y})| - |V(\bar{x}, \bar{y}) - V(x_0, \bar{y})| \leq W(\bar{x}) - V(x_0, \bar{y}) \\
\leq W(\bar{x}) - V(x_0, \bar{y}) + V(x_0, \bar{y}) - W(x_0) = W(\bar{x}) - W(x_0) \\
= W(\bar{x}) - V(x_0, y_0) + V(x_0, y_0) - W(x_0) \leq V(x_0, y_0) - W(x_0) \\
\leq |V(x_0, y_0) - V(x_0, y_0)| + |V(x_0, y_0) - W(x_0)| < 2\epsilon
\]

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.

Next, we present a result on game theory.
Proposition 8.39  Let $\mathcal{X}$ be a reflexive real normed linear space and $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{X}^*$ be nonempty bounded closed convex sets. Then,

$$\min_{x \in A} \max_{x_* \in B} \langle \langle x_* , x \rangle \rangle = \max_{x_* \in B} \min_{x \in A} \langle \langle x_* , x \rangle \rangle$$

Proof  Let $\mu := \inf_{x \in A} \sup_{x_* \in B} \langle \langle x_* , x \rangle \rangle$. Let $h : \mathcal{X} \times A \to \mathbb{R}$ be given by $h(x,x_*) = \langle \langle x_* , x \rangle \rangle$, $\forall x \in \mathcal{X}$, $\forall x_* \in B$. By Proposition 7.72, $h$ is continuous. Define $f_1 : \mathcal{X} \to \mathbb{R}$ by $f_1(x) = \sup_{x_* \in B} h(x,x_*)$, $\forall x \in \mathcal{X}$, define $h_x : \mathcal{X}^* \to \mathbb{R}$ by $h_x(x_*) = h(x,x_*) = \langle \langle x_* , x \rangle \rangle$, $\forall x_* \in \mathcal{X}^*$. 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, $f_1 : \mathcal{X} \to \mathbb{R}$ takes value in $\mathbb{R}$.

Since $B$ is bounded, then $\exists M_B \in [0, \infty) \subset \mathbb{R}$ such that $\| x_* \| \leq M_B$, $\forall x_* \in B$. $\forall x \in \mathcal{X}$, define $h_{x_*} : \mathcal{X} \to \mathbb{R}$ by $h_{x_*}(x) = h(x,x_*)$, $\forall x \in \mathcal{X}$, $\forall x_0 \in \mathcal{X}$, $\forall \alpha \in (0, \infty) \subset \mathbb{R}$, let $\delta = \epsilon/(1 + M_B) \in (0, \infty) \subset \mathbb{R}$. $\forall x \in B \times (x_0, \delta)$, $\forall x_* \in B$, we have $| h_{x_*}(x) - h_{x_*}(x_0) | = | h(x_*) - h(x_0) | = \langle \langle x_* , x - x_0 \rangle \rangle \leq \| x_* \| \| x - x_0 \| \leq M_B \delta < \epsilon$, where we have applied Proposition 7.72. Hence, $\{ h_{x_*} \mid \| x_* \| \leq M_B \}$ is equicontinuous. By Proposition 8.38, $f_1 : \mathcal{X} \to \mathbb{R}$ is continuous.

$\forall x_1 , x_2 \in \mathcal{X}$, $\forall \alpha \in [0, 1] \subset \mathbb{R}$, we have

$$f_1(\alpha x_1 + (1 - \alpha) x_2) = \max_{x_* \in B} \langle \langle x_* , \alpha x_1 + (1 - \alpha) x_2 \rangle \rangle$$

$$\leq \sup_{x_* \in B} \alpha \langle \langle x_* , x_1 \rangle \rangle + (1 - \alpha) \langle \langle x_* , x_2 \rangle \rangle$$

$$= \alpha \max_{x_* \in B} \langle \langle x_* , x_1 \rangle \rangle + (1 - \alpha) \max_{x_* \in B} \langle \langle x_* , x_2 \rangle \rangle$$

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 $\| x \| \leq M_A$, $\forall x \in A$. $\forall x \in A$, $\forall x_* \in B$, by Proposition 7.72, we have $\langle \langle x_* , x \rangle \rangle \geq -\| x_* \| \| x \| \geq -M_A M_B$. 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,

$$\mu = \max_{x_* \in X_{\text{conj}}(\mathcal{X}_{\text{conj}})} ( - f_{1\text{conj}}(x_*) - f_{2\text{conj}}(-x_*) )$$

Claim 8.39.1 $X_{\text{conj}} = B$ and $f_{1\text{conj}}(x_*) = 0$, $\forall x_* \in B$.

Proof of claim: $X_{\text{conj}} = \{ x_* \in \mathcal{X}^* \mid \sup_{x \in \mathcal{X}} (\langle \langle x_* , x \rangle \rangle - f_1(x)) < +\infty \}.

\forall x_0 \in B$, $\forall x \in \mathcal{X}$, we have $f_1(x) = \max_{x_* \in B} \langle \langle x_* , x \rangle \rangle \geq \langle \langle x_* , x \rangle \rangle$. Then, $\langle \langle x_* , x \rangle \rangle - f_1(x) \leq 0$. Hence, $x_0 \in X_{\text{conj}}$. Hence, $B \subseteq X_{\text{conj}}$. The above also implies that $f_{1\text{conj}}(x_0) = \sup_{x \in \mathcal{X}} (\langle \langle x_0 , x \rangle \rangle - f_1(x)) = 0$. 


CHAPTER 8. GLOBAL THEORY OF OPTIMIZATION

On the other hand, \( \forall x_0 \in X^* \setminus B, \) by Proposition 8.10, \( \exists x_{*0} \in X^{**} \) such that \( \langle \langle x_{*0}, x_0 \rangle \rangle < \inf_{x_0 \in B} \langle \langle x_{*0}, x_0 \rangle \rangle. \) Since \( X \) is reflexive, then \( x_{*0} = \phi(x_0) \) for some \( x_0 \in X, \) where \( \phi : X \to X^{**} \) is the natural mapping. Then, \( \langle \langle x_0, x_0 \rangle \rangle < \inf_{x_0 \in B} \langle \langle x_0, x_0 \rangle \rangle \) for some \( x_0 \in X \). By Proposition 7.90, \( X \) is weak compact. Then, \( \langle \langle x_0, x_0 \rangle \rangle = \inf_{x_0 \in X} \langle \langle x_0, x_0 \rangle \rangle = \langle \langle x_{*0}, x_0 \rangle \rangle. \) This completes the proof of the claim.

This completes the proof of the claim.

Claim 8.39.2 \( A_{\text{conj}} = X^* \) and \( f_{2\text{conj}}(x_*) = \max_{x \in A} \langle \langle x, x \rangle \rangle, \forall x_* \in X^*. \)

Proof of claim: \( A_{\text{conj}} = \{ x_* \in X^* \mid \sup_{x \in A} \langle \langle x, x \rangle \rangle < +\infty \}. \)

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. \( \forall x_* \in X^*, \) \( h_{x_*} \circ \phi_{\text{inv}} : X^{**} \to \mathbb{R} \) is weak* continuous. By Proposition 5.29, we have

\[
\sup_{x \in A} \langle \langle x, x \rangle \rangle = \sup_{x_* \in \phi(A)} \langle \langle x_*, x_* \rangle \rangle = \langle \langle x_{*0}, x_0 \rangle \rangle = \langle \langle x_{*0}, \phi_{\text{inv}}(x_{*0}) \rangle \rangle 
\]

for some \( x_{*0} \in \phi(A) \) and \( x_0 = \phi_{\text{inv}}(x_{*0}) \in A. \) Then, \( x_* \in A_{\text{conj}} \) and \( A_{\text{conj}} = X^*. \) Clearly, \( f_{2\text{conj}}(x_*) = \langle \langle x_*, x_0 \rangle \rangle = \max_{x \in A} \langle \langle x, x \rangle \rangle. \) This completes the proof of the claim.

Then,

\[
\mu = \inf_{x_* \in A_{\text{conj}}} \max_{x \in B} \langle \langle x, x \rangle \rangle = \max_{x_* \in B \setminus X^*} \left( -\max_{x \in A} \langle \langle -x, x \rangle \rangle \right) = \max_{x \in B} \langle \langle x, x \rangle \rangle.
\]

To show that the infimum in the above equation is actually achieved, we note that

\[
-\mu = \min_{x_* \in \phi(-A)} \max_{x \in (A)} \langle \langle x, x \rangle \rangle = \max_{x_* \in B} \min_{x \in \phi(-A)} \langle \langle x_*, x_* \rangle \rangle
\]

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

\[
-\mu = \max_{x_* \in \phi(-A)} \min_{x \in (A)} \langle \langle x_*, x_* \rangle \rangle = \max_{x \in (A)} \min_{x_* \in B} \langle \langle x_*, x_* \rangle \rangle
\]

which is equivalent to \( \mu = \min_{x \in A} \max_{x_* \in B} \langle \langle x, x \rangle \rangle. \)

This completes the proof of the proposition.
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 $x \gtrless y$ (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 $y \lesssim x$ if $y - x \in N$. We will write $x \rhd y$ ($y \rhd x$) if $x - y \in P^\circ$ ($y - x \in N^\circ = -P^\circ$).

It is easy to check that relations $\lesssim$ and $\rhd$ 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 $-\alpha x_1 \lesssim \vartheta$;

(ii) $x_1 \lesssim x_2$;

(iii) $x_1 \lesssim x_2$ and $x_2 \lesssim x_3$ implies $x_1 \lesssim x_3$;

(iv) $x_1 \lesssim x_2$ and $x_3 \lesssim x_4$ implies $x_1 + x_3 \lesssim x_2 + x_4$;

(v) $x_1 \lesssim x_2$ implies $\alpha x_1 \lesssim \alpha x_2$;

(vi) $x_1 \rhd \vartheta$ and $x_2 \rhd \vartheta$ implies $x_1 + x_2 \rhd \vartheta$;

(vii) $x_1 \lesssim x_2$ and $\alpha > 0$ implies $\alpha x_1 \lesssim \alpha x_2$.

Proof This is straightforward. \(\square\)

Definition 8.42 Let $X$ be a real normed linear space and $S \subseteq X$. The set $S^\oplus := \{ x \in X^* | \langle x, x \rangle \geq 0, \forall x \in S \}$ is called the positive conjugate cone of $S$. The set $S^\ominus := \{ x \in X^* | \langle x, x \rangle \leq 0, \forall x \in S \}$ 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, $\vartheta_x \in S^\oplus$. \(\forall x \in S^\oplus, \forall \alpha \in [0, \infty) \subset \mathbb{R}, \forall x \in S\), we have $\langle \alpha x, x \rangle = \alpha \langle x, x \rangle \geq 0$. 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 $\langle x_1 + x_2, x \rangle = \langle x_1, x \rangle + \langle x_2, x \rangle \geq 0$. Hence, $x_1 + x_2 \in S^\oplus$. Therefore, $S^\oplus$ is a convex cone.

\(\forall x_\ast \in S^\ominus\), by Proposition 4.13, \(\exists (x_{n_\ast})_{n=1}^\infty \subseteq S^\ominus\) such that \(\lim_{n \to \infty} x_{n_\ast} = x_\ast\). \(\forall x \in S\), by Propositions 7.72 and 3.66, $\langle x_{n_\ast}, x \rangle = \lim_{n \to \infty} \langle x_{n_\ast}, x \rangle \geq 0$. Hence, $x_\ast \in S^\oplus$ and $S^\ominus \subseteq S^\oplus$. By Proposition 3.3, $S^\ominus$ is closed.

(ii) This is straightforward.

This completes the proof of the proposition. \(\square\)
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 \hspace{1cm} \forall y_s \in (A(S))^{\oplus} \subseteq Y^*, \forall s \in S$, we have $\langle\langle A'y_s, s \rangle\rangle = \langle\langle y_s, As \rangle\rangle \geq 0$. Then, $A'y_s \in S^{\oplus}$ and $y_s \in A'_\text{inv}(S^{\oplus})$. Hence, $(A(S))^{\oplus} \subseteq A'_\text{inv}(S^{\oplus})$.

On the other hand, $\forall y_s \in A'_\text{inv}(S^{\oplus})$, $A'y_s \in S^{\oplus}$. $\forall s \in S$, we have $\langle\langle y_s, As \rangle\rangle = \langle\langle A'y_s, s \rangle\rangle \geq 0$. Therefore, we have $y_s \in (A(S))^{\oplus}$ and $A'_\text{inv}(S^{\oplus}) \subseteq (A(S))^{\oplus}$.

Hence, $(A(S))^{\oplus} = A'_\text{inv}(S^{\oplus})$. This completes the proof of the proposition. \qed

Definition 8.45 Let $X$ be a real normed linear space and $P \subseteq X$ be the positive cone. We will define $P^{\oplus} \subseteq X^*$ 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 $\langle\langle x_*, x_0 \rangle\rangle \geq 0$, $\forall x_* \in P^{\oplus}$ (or $x_* \nleq \vartheta_*), then x_0 \in P$.

Proof \hspace{1cm} 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 $\vartheta_* \in P$). Hence, $x_* \in P^{\oplus}$ and $\langle\langle x_*, x_0 \rangle\rangle < 0$. This contradicts with the assumption. Therefore, we must have $x_0 \in P$. This completes the proof of the proposition. \qed

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) $x \nleq \vartheta$ and $x_* \nleq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle \geq 0$;

(ii) $x \leq \vartheta$ and $x_* \nleq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle \leq 0$;

(iii) $x > \vartheta$, $x_* \nleq \vartheta_*$, and $x_* \neq \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle > 0$;

(iv) $x \nleq \vartheta$, $x \neq \vartheta$, and $x_* > \vartheta_*$ implies $\langle\langle x_*, x \rangle\rangle > 0$.

Proof \hspace{1cm} This is straightforward. \qed

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 $G : \Omega \to Z$ is said to be convex if $\Omega$ is convex and $\forall x_1, x_2 \in \Omega$, $\forall \alpha \in [0, 1] \subseteq \mathbb{R}$, 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 $P$. 
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 $G: \Omega \rightarrow \mathcal{Z}$ 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 $z \in \mathcal{Z}$. $\forall x_1, x_2 \in \Omega_z$, $\forall \alpha \in [0, 1] \subseteq \mathbb{R}$, 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.

### 8.8 Lagrange Multipliers

The basic problem to be considered in this section is

$$\mu_0 := \inf_{x \in \Omega, \ f(x) \preceq \partial_z} f(x)$$  \hfill (8.11)

where $\mathcal{X}$ is a real vector space, $\Omega \subseteq \mathcal{X}$ is a nonempty convex set, $f: \Omega \rightarrow \mathbb{R}$ is a convex functional, $\mathcal{Z}$ is a real normed linear space with positive cone $P \subseteq \mathcal{Z}$, and $G: \Omega \rightarrow \mathcal{Z}$ is a convex mapping.

Toward a solution to the above problem, we consider a class of problems:

$$\Gamma := \{ z \in \mathcal{Z} \mid \exists x \in \Omega, \ \exists \gamma G(x) \preceq z \}$$  \hfill (8.12a)

$$\omega(z) := \inf_{x \in \Omega, \ G(x) \preceq \partial_z} f(x); \ \forall z \in \Gamma$$  \hfill (8.12b)

where $\omega: \Gamma \rightarrow \mathbb{R}_+$ is the *primal functional*. To guarantee that $\omega$ is real-valued, we make the following assumption.

**Assumption 8.50** $\partial_z \in \Gamma$ and (i) $\exists \tilde{z} \in \gamma \Gamma$ such that $\omega(\tilde{z}) \in \mathbb{R}$ or (ii) $\mu := \inf_{x \in \Omega} f(x) > -\infty$ holds.

**Fact 8.51** $\Gamma \subseteq \mathcal{Z}$ is convex.

**Proof** $\forall z_1, z_2 \in \Gamma$, $\forall \alpha \in [0, 1] \subseteq \mathbb{R}$, 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.

**Fact 8.52** Under Assumption 8.50, $\omega: \Gamma \rightarrow \mathbb{R}$ is real-valued, convex, and nonincreasing, that is, $\forall z_1, z_2 \in \Gamma$ with $z_1 \preceq z_2$, we have $\omega(z_1) \geq \omega(z_2)$.
Proof: 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 \( \mathcal{B}_z(\delta, \delta) \cap V(\Gamma) \subseteq \Gamma \). Let \( \tilde{z}_1 := \bar{z} + \frac{\delta}{\|\bar{z} - z\|}(\bar{z} - z) \in \mathcal{B}_z(\delta, \delta) \cap V(\Gamma) \subseteq \Gamma \) and \( \tilde{\alpha} = \frac{\|\bar{z} - z\|}{\|\bar{z} - z\| + \frac{\delta}{2}} \in (0, 1) \subset \mathbb{R} \). It is easy to verify that \( \bar{z} = \tilde{\alpha}\tilde{z}_1 + (1 - \tilde{\alpha})z \). Then, we have

\[
-\infty < \omega(\bar{z}) = \inf_{\tilde{x} \in \Omega, \, G(\tilde{x}) \leq \tilde{\alpha}\tilde{z}_1 + (1 - \tilde{\alpha})z} f(\tilde{x}) \leq \inf_{\tilde{x} = \tilde{\alpha}\tilde{z}_1 + (1 - \tilde{\alpha})z, \, \tilde{x}_1 \in \Omega, \, G(\tilde{x}_1) \leq \tilde{z}_1, \, x \in \Omega, \, G(x) \leq \tilde{z}} f(\tilde{x}) \\
\leq \inf_{\tilde{x}_1 \in \Omega, \, G(\tilde{x}_1) \leq \tilde{z}_1, \, x \in \Omega, \, G(x) \leq \tilde{z}} (\tilde{\alpha}f(\tilde{x}_1) + (1 - \tilde{\alpha})f(x)) \\
= \inf_{\tilde{x}_1 \in \Omega, \, G(\tilde{x}_1) \leq \tilde{z}_1, \, x \in \Omega, \, G(x) \leq \tilde{z}} \tilde{\alpha}f(\tilde{x}_1) + \inf_{\tilde{x}_1 \in \Omega, \, G(\tilde{x}_1) \leq \tilde{z}_1, \, x \in \Omega, \, G(x) \leq \tilde{z}} (1 - \tilde{\alpha})f(x) \\
= \tilde{\alpha} \omega(\tilde{z}_1) + (1 - \tilde{\alpha})\omega(z)
\]

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, \( \omega : \Gamma \to \mathbb{R} \) 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 \( \Gamma \), we have \( \omega(z) < +\infty \). Hence, \( \omega(z) \in \mathbb{R} \).

Thus, under Assumption 8.50, \( \omega : \Gamma \to \mathbb{R} \) is real-valued.

\[
\forall z_1, z_2 \in \Gamma, \forall \alpha \in [0, 1] \subset \mathbb{R}, \text{ we have}
\]

\[
\omega(\alpha z_1 + (1 - \alpha)z_2) = \inf_{x \in \Omega, \, G(x) \leq \alpha z_1 + (1 - \alpha)z_2} f(x) \\
\leq \inf_{x = \alpha x_1 + (1 - \alpha)x_2, \, x_1 \in \Omega, \, G(x_1) \leq z_1, \, x_2 \in \Omega, \, G(x_2) \leq z_2} f(x) \\
\leq \inf_{x_1 \in \Omega, \, G(x_1) \leq z_1, \, x_2 \in \Omega, \, G(x_2) \leq z_2} (\alpha f(x_1) + (1 - \alpha)f(x_2)) \\
= \inf_{x_1 \in \Omega, \, G(x_1) \leq z_1, \, x_2 \in \Omega, \, G(x_2) \leq z_2} \alpha f(x_1) + \inf_{x_2 \in \Omega, \, G(x_2) \leq z_2} (1 - \alpha)f(x_2) \\
= \alpha \inf_{x_1 \in \Omega, \, G(x_1) \leq z_1} f(x_1) + (1 - \alpha) \inf_{x_2 \in \Omega, \, G(x_2) \leq z_2} f(x_2) \\
= \alpha \omega(z_1) + (1 - \alpha)\omega(z_2)
\]

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. □

Fact 8.53 Under Assumption 8.50, we have $\Gamma_{\text{conj}} \subseteq P^\ominus$ and $-\Gamma_{\text{conj}} \subseteq P^\oplus$.

Proof Since $\vartheta_z \in \Gamma$ implies that $\exists x_1 \in \Omega$ such that $G(x_1) \subseteq \vartheta_z$, then $\Gamma \supseteq G(x_1) + P$. ∀$z_*$ ∈ $\Gamma_{\text{conj}}$, by Definition 8.27, we have

$$+\infty > \sup_{z \in \Gamma}((z_*, z)) - \omega(z) \geq \sup_{z \in G(x_1) + P}((z_*, z)) - \omega(G(x_1))$$

$$= \sup_{z \in P}((z_*, z)) + \omega(G(x_1)) - \omega(G(x_1))$$

where the third inequality follows from Fact 8.52, and the equality follows from Proposition 8.37. Hence, $\sup_{z \in \vartheta_z}((z_*, z)) < +\infty$. Since $P$ is a cone, then $\sup_{z \in \vartheta_z}((z_*, z)) = 0$. This implies that $z_* \in P^\ominus$, $\Gamma_{\text{conj}} \subseteq P^\ominus$, and $-\Gamma_{\text{conj}} \subseteq P^\oplus$. This completes the proof of the fact. □

Fact 8.54 Let Assumption 8.50 hold. Define $\tilde{\omega} : P^\ominus \rightarrow \mathbb{R}_e$ by

$$\tilde{\omega}(z_*) = \sup_{z \in \Gamma}((z_*, z)) - \omega(z); \quad \forall z_* \in P^\ominus$$

Then, $\sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in P^\ominus} -\tilde{\omega}(z_*)$. ∀$z_* \in P^\ominus$, we have that

$$-\tilde{\omega}(-z_*) = \inf_{x \in \Omega} (f(x) + ((z_*, G(x)))) = \varphi(z_*)$$

where $\varphi : P^\oplus \rightarrow \mathbb{R}_e$ is called the dual functional.

Proof Clearly, $\forall z_* \in \Gamma_{\text{conj}}, \tilde{\omega}(z_*) = \omega_{\text{conj}}(z_*)$ and, by the definition of $\Gamma_{\text{conj}}$, $\forall z_* \in P^\ominus \setminus \Gamma_{\text{conj}}, \tilde{\omega}(z_*) = +\infty$. Hence, $\sup_{z_* \in \Gamma_{\text{conj}}} -\omega_{\text{conj}}(z_*) = \sup_{z_* \in P^\ominus} -\tilde{\omega}(z_*)$.

$$\forall z_* \in P^\ominus, -z_* = \vartheta_{z_*},$$

$$-\tilde{\omega}(-z_*) = -\sup_{z \in \Gamma}((z_*, z)) + \omega(z) = \inf_{z \in \Gamma}((z_*, z)) + \omega(z)$$

$$= \inf_{z \in \Gamma}((z_*, z)) + \inf_{x \in \Omega} f(x) = \inf_{z \in \Gamma} \inf_{x \in \Omega} ((z_*, z)) + f(x)$$

$$= \inf_{x \in \Omega} \inf_{z \in \Gamma} ((z_*, z)) + f(x) = \inf_{x \in \Omega} f(x) + \inf_{z \in \Gamma}((z_*, z))$$

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
\[ \omega \text{ is lower semicontinuous at } \vartheta_Z. \]
To apply Proposition 8.34, we assume that Assumption 8.56
\[ \exists x_1 \in \Omega \text{ such that } G(x_1) \leq \vartheta_Z. \]

Now, we state the two Lagrange duality results.

**Proposition 8.57** 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, \( Z \) be a real normed linear space with the positive cone \( P \subseteq Z \), \( G : \Omega \to Z \) be a convex mapping, and \( \mu_0 \) be defined as in (8.11). Let Assumptions 8.50 and 8.55 hold. Then,
\[ \mu_0 = \sup_{z^* \geq \vartheta_Z} \inf_{x \in \Omega} (f(x) + \langle \langle z^*, G(x) \rangle \rangle) = \sup_{z^* \geq \vartheta_Z} \varphi(z^*) \]}
(8.13)
Furthermore, if the supremum in (8.13) is achieved at \( z_{*0} \in P^\oplus \), that is,
\[ \mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_{*0}, G(x) \rangle \rangle) \]}
(8.14)
then the following statement holds: the infimum in (8.11) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \leq \vartheta_Z \) if, and only if, the infimum in (8.14) is achieved at \( x_0 \in \Omega \) with \( G(x_0) \leq \vartheta_Z \) 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
\[ \mu_0 = \sup_{z^* \in P^\oplus} -\tilde{\omega}(z^*) = \sup_{z^* \geq \vartheta_Z} -\tilde{\omega}(-z^*) = \sup_{z^* \geq \vartheta_Z} \varphi(z^*) \]
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 \( G(x_0) \leq \vartheta_Z \), 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 (8.11) 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 \( G(x_0) \leq \vartheta_Z \) 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 $P \subseteq \mathcal{Z}$, $P^0 \neq \emptyset$, $G : \Omega \to \mathcal{Z}$ be a convex mapping. Let Assumptions 8.50 and 8.56 hold. Then,

$$
\mu_0 = \max_{z \in \partial \vartheta_z} \inf_{x \in \Omega} (f(x) + \langle \langle z_*, G(x) \rangle \rangle) = \max_{z \in \partial \vartheta_z} \varphi(z_*) \tag{8.15}
$$

where the maximum is achieved at $z_0 \in P^\circ$, that is,

$$
\mu_0 = \inf_{x \in \Omega} (f(x) + \langle \langle z_0, G(x) \rangle \rangle) \tag{8.16}
$$

Furthermore, the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_z$, if, and only if, the infimum in (8.16) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_z$ and $\langle \langle z_0, G(x_0) \rangle \rangle = 0$.

Proof By Assumption 8.56, $G(x_1) \leq \vartheta_z$. Then, $-G(x_1) \in P^0$ and $G(x_1) + P \subseteq \Gamma$. By Proposition 7.16, $\Gamma^0 \supseteq G(x_1) + P^0$ and $\vartheta_z \in \Gamma^0$. For every $z \in G(x_1) + P$, we have $z \supseteq G(x_1)$ and, by Fact 8.52, $\omega(z) \leq \omega(G(x_1))$. Since $G(x_1) \leq \vartheta_z$, then $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $\mathcal{B}_\vartheta (\vartheta_z, \delta) \subseteq G(x_1) + P$. Take $r_0 := \omega(G(x_1)) + \delta \in \mathbb{R}$. It is easy to check that $\mathcal{B}_\vartheta (\vartheta_z, \delta) \subseteq [\omega, \Gamma]$. Hence, $(r_0, \vartheta_z) \in [\omega, \Gamma]^\circ$. By Proposition 8.22, $\omega$ is continuous at $\vartheta_z$. By Facts 8.51 and 8.52 and Proposition 8.34, we have

$$
\mu_0 = \omega(\vartheta_z) = \max_{z \in \Gamma^\circ} -\omega_{\text{conj}}(z_*)
$$

Then, by Facts 8.53 and 8.54, we have

$$
\mu_0 = \max_{z \in P^0} -\omega(z_*) = \max_{z \in \partial \vartheta_z} -\omega(-z_*) = \max_{z \in \partial \vartheta_z} \varphi(z_*)
$$

Therefore, (8.15) holds and the maximum is achieved at $z_0 \in P^\circ$.

If the infimum in (8.11) is achieved at $x_0 \in \Omega$ with $G(x_0) \leq \vartheta_z$, 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 (8.16) 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 $G(x_0) \leq \vartheta_z$ 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. □

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 $G(x_0) \equiv \partial \mathcal{Z}$ achieves the infimum in (8.11). Then, there exists $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \equiv \partial \mathcal{Z}^*$ such that the Lagrangian $L : \Omega \times P^\circ \to \mathbb{R}$ defined by

$$L(x, z_*) := f(x) + \langle (z_*, G(x)) \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

admits a saddle point at $(x_0, z_{*0})$, i.e.

$$L(x_0, z_*) \leq L(x_0, z_{*0}) \leq L(x, z_{*0}); \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

Proof By Proposition 8.58, there exists $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \equiv \partial \mathcal{Z}^*$ such that $L(x_0, z_{*0}) \leq L(x, z_{*0})$, $\forall x \in \Omega$ and $\langle (z_{*0}, G(x_0)) \rangle = 0$. Then, $\forall z_* \in P^\circ$, we have $L(x_0, z_*) - L(x_0, z_{*0}) = \langle (z_*, G(x_0)) \rangle - \langle (z_{*0}, G(x_0)) \rangle = 0$. 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 $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \equiv \partial \mathcal{Z}^*$ such that

$$f(x_0) + \langle (z_{*0}, G(x_0)) \rangle \leq f(x) + \langle (z_{*0}, G(x)) \rangle; \quad \forall x \in \Omega$$

Then,

$$f(x_0) = \inf_{x \in \Omega, G(x) \equiv G(x_0)} f(x)$$

Proof $\forall x \in \Omega$ with $G(x) \equiv G(x_0)$, we have $\langle (z_{*0}, G(x_0)) \rangle \leq \langle (z_{*0}, G(x_0)) \rangle$, since $z_{*0} \in P^\circ$. By the assumption of the proposition, $f(x_0) + \langle (z_{*0}, G(x_0)) \rangle \leq f(x) + \langle (z_{*0}, G(x)) \rangle$. 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 $z_{*0} \in \mathcal{Z}^*$ with $z_{*0} \equiv \partial \mathcal{Z}^*$ such that the Lagrangian $L : \Omega \times P^\circ \to \mathbb{R}$ given by

$$L(x, z_*) = f(x) + \langle (z_*, G(x)) \rangle; \quad \forall x \in \Omega, \forall z_* \in P^\circ$$

admits a saddle point at $(x_0, z_{*0})$, i.e.

$$L(x_0, z_*) \leq L(x_0, z_{*0}) \leq L(x, z_{*0}); \quad \forall x \in \Omega, \forall z_* \in P^\circ$$
Then, \( G(x_0) \subseteq \partial_Z \) and
\[
f(x_0) = L(x_0, z_{*0}) = \inf_{x \in \Omega, \; G(x) \subseteq \partial_Z} f(x)
\]

**Proof** By the first inequality in the saddle-point condition, we have
\[
\langle\langle z_{*}, G(x_0) \rangle\rangle \leq \langle\langle z_{*0}, G(x_0) \rangle\rangle; \quad \forall z_{*} \in P^0
\]
\[
\forall z_{*} \in P^0, \text{we have } z_{*} \not\subseteq \partial_Z \text{ and } z_{*} + z_{*0} \geq \partial_Z. \text{ Then, } \langle\langle z_{*} + z_{*0}, G(x_0) \rangle\rangle \leq \langle\langle z_{*0}, G(x_0) \rangle\rangle \text{ and } \langle\langle z_{*}, G(x_0) \rangle\rangle \leq 0. \text{ By Proposition 8.46, } G(x_0) \in (-P) \text{ and } G(x_0) \subseteq \partial_Z. \text{ Furthermore, } 0 = \langle\langle \partial_Z, G(x_0) \rangle\rangle \leq \langle\langle z_{*0}, G(x_0) \rangle\rangle \leq 0 \implies \langle\langle z_{*0}, G(x_0) \rangle\rangle = 0.
\]
\[
\forall x \in \Omega \text{ with } G(x) \subseteq \partial_Z, \text{ we have } f(x) \geq f(x) + \langle\langle z_{*0}, G(x) \rangle\rangle \geq f(x) + \langle\langle z_{*0}, G(x_0) \rangle\rangle = f(x_0). \text{ This completes the proof of the proposition.} \quad \square
\]

Next, we present a result on the sensitivity of the infimization problem.

**Proposition 8.62** Let \( \mathcal{X} \) be a real vector space, \( \Omega \subseteq \mathcal{X} \) be nonempty, \( f : \Omega \to \mathbb{R} \), \( Z \) be a real normed linear space with the positive cone \( P \subseteq Z \), \( G : \Omega \to Z \).

Let \( z_i \in Z \), \( \mu_i = \inf_{x \in \Omega, \; G(x) \subseteq z_i} f(x) \), \( i = 0, 1 \). Let \( z_{*0} \in P^0 \subseteq Z^* \) be the Lagrange multiplier associated with \( \mu_0 \), that is
\[
\mu_0 = \inf_{x \in \Omega} (f(x) + \langle\langle z_{*0}, G(x) - z_0 \rangle\rangle)
\]
Assume that \( \mu_0 \in \mathbb{R} \). Then, we have
\[
\mu_1 - \mu_0 \geq \langle\langle - z_{*0}, z_1 - z_0 \rangle\rangle
\]

**Proof** \( \forall x \in \Omega \text{ with } G(x) \subseteq z_1, \text{ we have } \mu_0 \leq f(x) + \langle\langle z_{*0}, G(x) - z_0 \rangle\rangle \), which implies that
\[
f(x) \geq \mu_0 - \langle\langle z_{*0}, G(x) - z_0 \rangle\rangle \geq \mu_0 - \langle\langle z_{*0}, z_1 - z_0 \rangle\rangle
\]
Hence, \( \mu_1 \geq \mu_0 + \langle\langle - z_{*0}, z_1 - z_0 \rangle\rangle \). This completes the proof of the proposition. \( \square \)
CHAPTER 8. GLOBAL THEORY OF OPTIMIZATION
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 \( X \) be a normed linear space over the field \( \mathbb{K} \), \( D \subseteq X \), and \( x_0 \in D \). \( u \in X \) is said to be an admissible deviation in \( D \) at \( x_0 \) if
\[
\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \text{ we have } \{ x_0 + r \bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \bar{u} \in B(u, \epsilon) \} \cap D \neq \emptyset.
\]

Let \( A_D(x_0) \) be the set of admissible deviations in \( D \) at \( x_0 \).

Clearly, if \( x_0 \in D^\circ \), then \( A_D(x_0) = X \). Another fact is that when \( D_1 \subseteq D_2 \subseteq X \) and \( x_0 \in D_1 \), then \( A_{D_1}(x_0) \subseteq A_{D_2}(x_0) \). Yet another fact is that when \( x_0 \in \overline{D_1}, D_1, D_2 \subseteq X \) 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 \( X \) be a normed linear space over the field \( \mathbb{K} \), \( D \subseteq X \), and \( x_0 \in \overline{D} \). Then, \( A_D(x_0) \subseteq X \) 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 \( D \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \bar{x} \in (B(x_0, \epsilon^2) \cap D) \setminus \{ x_0 \} \). Let \( v := \bar{x} - x_0 \). Then, \( \delta := \sqrt{\|v\|} \in (0, \epsilon) \subset \mathbb{R} \). Let \( \bar{v} := \delta^{-1} v \). Then, \( \bar{v} \in B(\vartheta, \epsilon) \) and \( \bar{x} = x_0 + \delta \bar{v} \in \{ x_0 + r \bar{u} \mid r \in (0, \epsilon) \subset \mathbb{R}, \bar{u} \in B(\vartheta, \epsilon) \} \cap D \neq \emptyset \). 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: \( \alpha > 0 \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), let \( \bar{c} = \min \{ \alpha \epsilon, \epsilon / \alpha \} \in (0, \infty) \subset \mathbb{R} \).
By \( u \in A_D(x_0), \exists \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \varepsilon) \subset \mathbb{R}, \bar{u} \in B(u, \varepsilon) \} \cap D. \)
Hence, \( \bar{x} \in D \) and \( \bar{x} = x_0 + r\bar{u} = x_0 + (r/\alpha) (\alpha \bar{u}) \) with \( r \in (0, \varepsilon) \subset \mathbb{R} \)
and \( \bar{u} \in B(u, \varepsilon). \) Then, \( r/\alpha \in (0, \varepsilon) \subset \mathbb{R} \) and \( \alpha \bar{u} \in B(\alpha u, \varepsilon). \) Hence,
\( \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \varepsilon) \subset \mathbb{R}, \bar{u} \in B(\alpha u, \varepsilon) \} \cap D \neq \emptyset. \) This implies that
\( \alpha u \in A_D(x_0). \) Hence, \( A_D(x_0) \) is a cone.

\( \forall u \in A_D(x_0), \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \) by Proposition 3.3, \( \exists u_1 \in A_D(x_0) \cap B(u, \varepsilon/2). \) By \( u_1 \in A_D(x_0), \exists \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \varepsilon/2) \subset \mathbb{R}, \bar{u} \in B(u_1, \varepsilon/2) \} \cap D. \) Then, \( \bar{x} \in \{ x_0 + r\bar{u} \mid r \in (0, \varepsilon) \subset \mathbb{R}, \bar{u} \in B(u_1, \varepsilon/2) \} \cap D \neq \emptyset. \) Hence, \( u \in A_D(x_0). \) Then, \( A_D(x_0) \subseteq A_D(x_0) \) and \( A_D(x_0) \) is closed.
This completes the proof of the proposition. \( \square \)

**Definition 9.3** Let \( X \) and \( Y \) be normed linear spaces over \( \mathbb{K}, D \subseteq X, f : D \rightarrow Y, \) and \( x_0 \in D. \) Assume that \( \text{span} (A_D(x_0)) = X. \) Let \( L \in B(X, Y) \)
be such that \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall x \in D \cap B_X(x_0, \delta), \) we have
\[
\| f(x) - f(x_0) - L(x - x_0) \| \leq \varepsilon \| x - x_0 \|
\]
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, \( Df : D \rightarrow B(X, Y) \) or \( f^{(1)} : D \rightarrow B(X, Y). \)

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 \( \mathbb{K}, D \subseteq X, f : D \rightarrow Y, x_0 \in D, \) 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 \( \hat{g} : M_D(x_0) \rightarrow Y \)
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 \( \mathbb{K}, D \subseteq X, f : D \rightarrow Y, x_0 \in D, \) \( \text{span} (A_D(x_0)) = X, L \in B(X, Y) \) be the Fréchet derivative of \( f \) at \( x_0, \) and \( u \in A_D(x_0). \) Then, \( Df(x_0; u) = Lu. \)

**Proof** \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \) let \( \bar{\varepsilon} = \varepsilon/(2 + 2 \| u \|) \in (0, \infty) \subset \mathbb{R}. \) By
\( Df(x_0) = L, \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that
\[
\| f(x) - f(x_0) - L(x - x_0) \| \leq \bar{\varepsilon} \| x - x_0 \|, \forall x \in B_X(x_0, \delta) \cap D.
\]
Let \( M_D(x_0) \) and \( \hat{g} : M_D(x_0) \rightarrow Y \) be as defined in Definition 9.4. Let \( \delta = \min \{ \varepsilon/(2 + 2 \| L \|), \delta/(1 + \| u \|), 1 \} \in \mathbb{R}. \)
(0, \infty) \subset \mathbb{R}$. \forall (r, \bar{u}) \in (M_D(x_0) \cap \mathbb{B}_{\mathbb{R} \times \mathbb{X}} ((0, u), \bar{\delta})) \setminus \{(0, u)\}, we have $\bar{x} := x_0 + r \bar{u} \in D \cap \mathbb{B}_X (x_0, \delta)$. This implies that

\[
\|\bar{g}(r, \bar{u}) - Lu\| = \|f(x_0 + r \bar{u}) - f(x_0) - rLu\| / r
\]

\[
= \|f(\bar{x}) - f(x_0) - L(\bar{x} - x_0) + rL(\bar{u} - u)\| / r
\]

\[
\leq \|f(\bar{x}) - f(x_0) - L(\bar{x} - x_0)\| / r + \|L(\bar{u} - u)\|
\]

\[
\leq \epsilon \|\bar{x} - x_0\| / r + \|L\| \|\bar{u} - u\| = \epsilon \|\bar{u}\| + \|L\| \epsilon \|\bar{u} - u\| = \epsilon \| u \| + \|L\| \epsilon \|\bar{u} - u\| = \epsilon \| u \| + \|L\| \epsilon \|\bar{u} - u\|
\]

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. \hfill \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 $L_1 u_0 \neq L_2 u_0$. By Proposition 9.5, $Df(x_0; u_0) = L_1 u_0 = L_2 u_0$. This is a contradiction. Hence, the result holds. This completes the proof of the proposition. \hfill \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. \hfill \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 \((x_1, \ldots, x_n) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n\), 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 \(\mathbb{K}\), \(m \in \mathbb{N}\). \(\forall L \in \mathcal{B}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_n, y_1 \times \cdots \times y_m)\) can be equivalently written as
\[
\begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & \ddots & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix}
\]
where \(L_{ij} \in \mathcal{B}(\mathcal{X}_j, y_i)\). Then, \((y_1, \ldots, y_m) = L(x_1, \ldots, x_n)\) can be equivalently expressed as
\[
\begin{bmatrix}
y_1 \\
\vdots \\
y_m
\end{bmatrix} =
\begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & \ddots & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
\vdots \\
x_n
\end{bmatrix} =
\begin{bmatrix}
\sum_{j=1}^n L_{1j}x_j \\
\vdots \\
\sum_{j=1}^n L_{mj}x_j
\end{bmatrix}
\]

Let \(\mathcal{Z}_1, \ldots, \mathcal{Z}_p\) be normed linear spaces over \(\mathbb{K}\), \(p \in \mathbb{N}\), and \(\bar{L} \in \mathcal{B}(y_1 \times \cdots \times y_m, \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_p)\) be given by
\[
\begin{bmatrix}
\bar{L}_{11} & \cdots & \bar{L}_{1m} \\
\vdots & \ddots & \vdots \\
\bar{L}_{p1} & \cdots & \bar{L}_{pm}
\end{bmatrix}
\]
Then, \(\bar{L}L \in \mathcal{B}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_n, \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_p)\) and
\[
\bar{L}L =
\begin{bmatrix}
\bar{L}_{11} & \cdots & \bar{L}_{1m} \\
\vdots & \ddots & \vdots \\
\bar{L}_{p1} & \cdots & \bar{L}_{pm}
\end{bmatrix}
\begin{bmatrix}
L_{11} & \cdots & L_{1n} \\
\vdots & \ddots & \vdots \\
L_{m1} & \cdots & L_{mn}
\end{bmatrix}
\]
\[
= 
\begin{bmatrix}
\sum_{j=1}^m \bar{L}_{1j}L_{j1} & \cdots & \sum_{j=1}^m \bar{L}_{1j}L_{jn} \\
\vdots & \ddots & \vdots \\
\sum_{j=1}^m \bar{L}_{pj}L_{j1} & \cdots & \sum_{j=1}^m \bar{L}_{pj}L_{jn}
\end{bmatrix}
\]
It is easy to show that the adjoint of \(L\) is
\[
L' =
\begin{bmatrix}
L'_{11} & \cdots & L'_{1n} \\
\vdots & \ddots & \vdots \\
L'_{m1} & \cdots & L'_{mn}
\end{bmatrix}
\]

**Proposition 9.9** Let \(\mathcal{X}, \mathcal{Y}, \mathcal{Z}\) 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 \(Df(x_0, y_0) \in \mathcal{B}(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})\) exists. Let \(D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}\). Assume that \(\text{span}\left(A_{D_{x_0}}(y_0)\right) = \)
Then, $f$ is partial differentiable with respect to $y$ at $(x_0, y_0)$ and $f_y(x_0, y_0) \in \mathbb{B}(y, \mathcal{Z})$ is given by $f_y(x_0, y_0)(k) = Df(x_0, y_0)(\partial_x, k), \forall k \in \mathcal{Y}.$

Let $D_{y} := \{ x \in \mathcal{X} \mid (x, y_0) \in D \}$. Assume that \text{span} \( \mathcal{A}_{D_{y}}(x_0) \) = $\mathcal{X}$. By symmetry, $f$ is partial differentiable with respect to $x$ at $(x_0, y_0)$ and $f_x(x_0, y_0) \in \mathbb{B}(X, \mathcal{Z})$ is given by $f_x(x_0, y_0)(h) = Df(x_0, y_0)(h, \partial_y), \forall h \in X.$

Hence, we have $Df(x_0, y_0) = \left[ \begin{array}{ccc} f_y(x_0, y_0) & f_x(x_0, y_0) \end{array} \right]$ in “matrix” notation.

**Proof**

Let $\tilde{g} : D_{x_0} \to \mathcal{Z}$ 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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta)$, we have $\| f(x, y) - f(x_0, y_0) - Df(x_0, y_0)(x - x_0, y - y_0) \| \leq \epsilon \| (x - x_0, y - y_0) \|$. Then, $\forall y \in D_{y} \cap B_{y}(y_0, \delta)$, we have $(x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta)$ and $\epsilon \| y - y_0 \| \geq \| f(x, y) - f(x_0, y_0) - Df(x_0, y_0)(\partial_x, y - y_0) \| = \| \tilde{g}(y) - \tilde{g}(y_0) - Df(x_0, y_0)(\partial_x, y - y_0) \|$. Let $L : \mathcal{Y} \to \mathcal{Z}$ be given by $L(k) = Df(x_0, y_0)(\partial_x, k), \forall k \in \mathcal{Y}$. Clearly, $L$ is a linear operator. Note that

$$||L|| = \sup_{k \in \mathcal{Y}, \|k\| \leq 1} ||Lk|| = \sup_{k \in \mathcal{Y}, \|k\| \leq 1} ||Df(x_0, y_0)(\partial_x, k)||,$$

$$\leq \sup_{k \in \mathcal{Y}, \|k\| \leq 1} \|Df(x_0, y_0)||||\|\partial_x, k|| \leq ||Df(x_0, y_0)|| < +\infty$$

where the first inequality follows from Proposition 7.64 and the last inequality follows from the fact that $Df(x_0, y_0) \in \mathbb{B}(X \times \mathcal{Y}, \mathcal{Z})$. Hence, $L \in B(\mathcal{Y}, \mathcal{Z})$. Then, $\| \tilde{g}(y) - \tilde{g}(y_0) - L(y - y_0) \| \leq \epsilon \| y - y_0 \|, \forall y \in D_{x_0} \cap B_{y}(y_0, \delta)$. This implies that $D\tilde{g}(y_0) = L$. By Definition 9.8, $\frac{\partial \tilde{g}}{\partial y}(x_0, y_0) = L$.

By symmetry, $f$ is partial differentiable with respect to $x$ at $(x_0, y_0)$ and $\frac{\partial f}{\partial x}(x_0, y_0) \in \mathbb{B}(X, \mathcal{Z})$ is given by $\frac{\partial f}{\partial x}(x_0, y_0)(h) = Df(x_0, y_0)(h, \partial_y), \forall h \in \mathcal{X}.$

Note that, $\forall h \in \mathcal{X}$ and $\forall k \in \mathcal{Y},$

$$Df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) + \frac{\partial f}{\partial y}(x_0, y_0)(k)$$

Then,

$$Df(x_0, y_0) = \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}$$

Hence, the desired “matrix” notation follows. This completes the proof of the proposition. \hfill \Box

9.2 The Derivatives of Some Common Functions

**Proposition 9.10** Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \to \mathcal{Y}$, $x_0 \in D$, and $\text{span} \( \mathcal{A}_{D}(x_0) \) = \mathcal{X}$. Assume that $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$
\( \mathbb{R} \) and \( \exists y_0 \in \mathcal{Y} \) such that \( f(x) = y_0, \forall x \in D \cap B_\mathcal{X}(x_0, \delta_0) \). Then, \( f \) is Fréchet differentiable at \( x_0 \) and \( Df(x_0) = \psi_{B(\mathcal{X}, \mathcal{Y})} \).

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

**Proposition 9.11** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), \( x_0 \in D_2 \subseteq D_1 \subseteq \mathcal{X} \), and \( f : D_1 \to \mathcal{Y} \). If \( f \) is Fréchet differentiable at \( x_0 \) and \( \text{span} (\mathcal{A}_{D_2}(x_0)) = \mathcal{X} \), then \( g := f|_{D_2} \) 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 \( D_1 \cap B_\mathcal{X}(x_0, \delta_0) = D_2 \cap B_\mathcal{X}(x_0, \delta_0) \), 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 \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), and \( f : \mathcal{X} \to \mathcal{X} \) be given by \( f = \text{id}_\mathcal{X} \), that is \( f(x) = x, \forall x \in \mathcal{X} \). Then, \( f \) is Fréchet differentiable and \( Df(x)(h) = h, \forall x \in \mathcal{X} \) and \( h \in \mathcal{X} \). Then, \( Df(x) = \text{id}_\mathcal{X}, \forall x \in \mathcal{X} \).

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

**Proposition 9.13** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), and \( f : \mathcal{X} \times \mathcal{Y} \to \mathcal{X} \) be given by \( f = \pi_\mathcal{X} \), that is \( f(x, y) = x, \forall (x, y) \in \mathcal{X} \times \mathcal{Y} \). Then, \( f \) is Fréchet differentiable and \( Df(x, y)(h, k) = h, \forall (x, y) \in \mathcal{X} \times \mathcal{Y} \) and \( \forall (h, k) \in \mathcal{X} \times \mathcal{Y} \). In “matrix” notation, we have \( Df(x, y) = \begin{bmatrix} \text{id}_\mathcal{X} & \psi_{B(\mathcal{Y}, \mathcal{X})} \end{bmatrix}, \forall (x, y) \in \mathcal{X} \times \mathcal{Y} \).

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

**Proposition 9.14** Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( f : \mathcal{X} \times \mathcal{X} \to \mathcal{X} \) be given by \( f(x_1, x_2) = x_1 + x_2, \forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X} \). Then, \( f \) is Fréchet differentiable and \( Df : \mathcal{X} \times \mathcal{X} \to B(\mathcal{X} \times \mathcal{X}, \mathcal{X}) \) is given by \( \forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X}, \forall (h_1, h_2) \in \mathcal{X} \times \mathcal{X}, Df(x_1, x_2)(h_1, h_2) = h_1 + h_2 \). In “matrix” notation, we have \( Df(x_1, x_2) = \begin{bmatrix} \text{id}_\mathcal{X} & \text{id}_\mathcal{X} \end{bmatrix} \).

**Proof** \( \forall (x_1, x_2) \in \mathcal{X} \times \mathcal{X} \), let \( L : \mathcal{X} \times \mathcal{X} \to \mathcal{X} \) be given by \( L(h_1, h_2) = h_1 + h_2, \forall (h_1, h_2) \in \mathcal{X} \times \mathcal{X} \). Clearly, \( L \) is a linear operator. Note that

\[
\| L \| = \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \| L(h_1, h_2) \| \\
\leq \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \| h_1 \| + \| h_2 \| \\
\leq \sup_{(h_1, h_2) \in \mathcal{X} \times \mathcal{X}, \|(h_1, h_2)\| \leq 1} \sqrt{2} \left( \| h_1 \|^2 + \| h_2 \|^2 \right)^{1/2} \leq \sqrt{2}
\]

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, \( L \in B(\mathcal{X} \times \mathcal{X}, \mathcal{X}) \).
Clearly, \( A_{X\times X} (x_1, x_2) = X \times X \) since \((x_1, x_2) \in (X \times X)^\circ \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) set \( \delta = 1 \in \mathbb{R}, \) \( \forall (\bar{x}_1, \bar{x}_2) \in (X \times X) \cap B_{X \times X} ((x_1, x_2), \delta), \) we have \( \| f(\bar{x}_1, \bar{x}_2) - f(x_1, x_2) - L(x_1 - x_1, \bar{x}_2 - x_2) \| = 0 \leq \epsilon \| (\bar{x}_1 - x_1, \bar{x}_2 - x_2) \|. \) By Definition 9.3, \( Df(x_1, x_2) = L. \) This completes the proof of the proposition. \( \square \)

**Proposition 9.15** Let \( X \) and \( Y \) be normed linear spaces over \( \mathbb{K}, D \subseteq X, f_1 : D \to Y, f_2 : D \to Y, x_0 \in D, \alpha_1, \alpha_2 \in \mathbb{K}, \) 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

\[
\| g(x) - g(x_0) - L(x - x_0) \| = \| \alpha_1 (f_1(x) - f_1(x_0) - f_1^{(1)}(x_0)(x - x_0)) + \alpha_2 (f_2(x) - f_2(x_0) - f_2^{(1)}(x_0)(x - x_0)) \| \leq (|\alpha_1| + |\alpha_2|) \epsilon \| x - x_0 \|
\]

Hence, \( g^{(1)}(x_0) = L. \) This completes the proof of the proposition. \( \square \)

**Proposition 9.16** Let \( X \) be a normed linear space over \( \mathbb{K}, f : \mathbb{K} \times X \to X \) be given by \( f(\alpha, x) = \alpha x, \forall (\alpha, x) \in \mathbb{K} \times X. \) Then, \( f \) is Fréchet differentiable and \( Df : \mathbb{K} \times X \to B(\mathbb{K} \times X, X) \) is given by \( \forall (\alpha, x) \in \mathbb{K} \times X, \forall (d, h) \in \mathbb{K} \times X, \) \( Df(\alpha, x)(d, h) = \alpha h + dx. \) Thus, \( Df(\alpha, x) = [x \text{ aid } x] \) in “matrix” notation.

**Proof** \( \forall (\alpha, x) \in \mathbb{K} \times X, \) let \( L : \mathbb{K} \times X \to X \) be given by \( L(d, h) = \alpha h + dx, \forall (d, h) \in \mathbb{K} \times X. \) Clearly, \( L \) is a linear operator. Note that

\[
\| L \| = \sup_{(d, h) \in \mathbb{K} \times X, \| (d, h) \| \leq 1} \| L(d, h) \| \\
\leq \sup_{(d, h) \in \mathbb{K} \times X, \| (d, h) \| \leq 1} |\alpha| \| h \| + |d| \| x \|
\leq \sup_{(d, h) \in \mathbb{K} \times X, \| (d, h) \| \leq 1} \left( |\alpha|^2 + \| x \|^2 \right)^{1/2} \left( |d|^2 + \| h \|^2 \right)^{1/2}
\leq \| (\alpha, x) \| < +\infty
\]

where the second inequality follows from Cauchy-Schwarz Inequality. Hence, \( L \in B(\mathbb{K} \times X, X). \)
Clearly, $A_{\mathbb{K} \times \mathcal{X}}(\alpha, x) = \mathbb{K} \times \mathcal{X}$ since $(\alpha, x) \in (\mathbb{K} \times \mathcal{X})^\circ$. \forall \epsilon \in (0, \infty) \subset \mathbb{R}$, set $\delta = 2\epsilon \in (0, \infty) \subset \mathbb{R}$, \forall $(\bar{\alpha}, \bar{x}) \in (\mathbb{K} \times \mathcal{X}) \cap \mathcal{B}_{\mathbb{K} \times \mathcal{X}}((\alpha, x), \delta)$, we have

$$\|f(\alpha, x) - L(\bar{\alpha} - \alpha, \bar{x} - x)\| = \|\bar{\alpha} - \alpha x - \alpha(\bar{x} - x) - (\bar{\alpha} - \alpha)x\| = \|\bar{\alpha} - \alpha\|\|\bar{x} - x\| = \frac{1}{2}(\|\bar{\alpha} - \alpha\|^2 + \|\bar{x} - x\|^2) = \frac{1}{2}(\|\bar{\alpha} - \alpha\|, \|\bar{x} - x\|)^2$$

$$\leq \epsilon \|\bar{\alpha} - \alpha, \bar{x} - x\|$$

By Definition 9.3, $Df(\alpha, x) = L$. This completes the proof of the proposition. \hfill \Box

In the previous proposition, we have abuse the notation using the “matrix” notation that we identify $x$ as $dx$.

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 $g(A) : B(\mathcal{Y}, \mathcal{Z}) \to B(\mathcal{X}, \mathcal{Z})$ 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 $g : B(\mathcal{X}, \mathcal{Y}) \to B(B(\mathcal{Y}, \mathcal{Z}), B(\mathcal{X}, \mathcal{Z}))$ 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 $\text{ro}$ notation 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 $\mathbb{K}$, $f : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \to \mathcal{Y}$ 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 $Df : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \to B(B(\mathcal{Y}, \mathcal{Z}), B(\mathcal{X}, \mathcal{Z}))$ 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 $L : B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X} \to \mathcal{Y}$ 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

$$\|L\| = \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|\Delta, h\| \leq 1} \|L(\Delta, h)\|$$

$$\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|\Delta, h\| \leq 1} \|A\| \|\Delta\| \|h\| + \|\Delta\| \|x\|$$

$$\leq \sup_{(\Delta, h) \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}, \|\Delta, h\| \leq 1} \left(\|A\|^2 + \|x\|^2\right)^{1/2} \left(\|\Delta\|^2 + \|h\|^2\right)^{1/2}$$

$$\leq \|\text{ro}(x)\| < +\infty$$

where the first inequality follows from Proposition 7.64, and the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(\mathcal{X}, \mathcal{Y}) \times \mathcal{X}$. 

9.3. Chain Rule and Mean Value Theorem

Clearly, \( A_{B(X, Y) \times X} (A, x) = B(X, Y) \times X \) since \((A, x) \in (B(X, Y) \times X)^c = B(X, Y) \times X \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), set \( \delta = 2\epsilon \in (0, \infty) \subset \mathbb{R} \), \( \forall (A, \bar{x}) \in (B(X, Y) \times X) \cap B_{B(X, Y) \times X} ((A, x), \delta) \), we have

\[
\begin{align*}
\| f(\bar{x}) - f(x) - L(\bar{x} - A, \bar{x} - x) \| &= \| \bar{x} - Ax - A(\bar{x} - x) -(\bar{x} - A)x \| = \| (\bar{x} - A)(\bar{x} - x) \| \\
&\leq \frac{1}{2} \| \| A - A \| ^2 + \| \bar{x} - x \|^2 \| = \frac{1}{2} \| (\bar{x} - A, \bar{x} - x) \|^2 \\
&\leq \epsilon \| (\bar{x} - A, \bar{x} - x) \|
\end{align*}
\]

where the first inequality follows from Proposition 7.64. By Definition 9.3, \( Df(A, x) = L \). This completes the proof of the proposition.

9.3 Chain Rule and Mean Value Theorem

**Theorem 9.18 (Chain Rule)** Let \( X, Y, \) and \( Z \) be normed linear spaces over \( \mathbb{R} \), \( D_x \subseteq X, D_y \subseteq Y, f : D_x \rightarrow D_y, g : D_y \rightarrow Z, x_0 \in D_x, y_0 := f(x_0) \in D_y, \) and \( h := g \circ f : D_x \rightarrow Z \). 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 \( L : X \rightarrow Z \) 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, \( L \in B(X, Z) \).

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), let \( \tilde{\epsilon} = (\sqrt{4\epsilon} + (\| Df(x_0) \| + \| Dg(y_0) \|)^2 - \| Df(x_0) \| - \| Dg(y_0) \|)/2 \in (0, \infty) \subset \mathbb{R} \). 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

\[
\| g(y) - g(y_0) - Dg(y_0)(y - y_0) \| \leq \tilde{\epsilon} \| y - y_0 \|
\]

By the fact that \( Df(x_0) \in B(X, Y) \), we have \( \text{span}(A_{D_x}(x_0)) = X \) and \( \exists \delta_2 \in (0, \delta_1/(\tilde{\epsilon} + \| Df(x_0) \|)) \subset \mathbb{R} \) such that \( \forall x \in D_x \cap B_{\mathbb{R}}(x_0, \delta_2), \) we have

\[
\| f(x) - f(x_0) - Df(x_0)(x - x_0) \| \leq \tilde{\epsilon} \| x - x_0 \|
\]

Then, by Proposition 7.64, \( \| f(x) - y_0 \| = \| f(x) - f(x_0) \| \leq \tilde{\epsilon} \| x - x_0 \| + \| Df(x_0) \| \| x - x_0 \| \leq (\tilde{\epsilon} + \| Df(x_0) \|) \| x - x_0 \| < \delta_1 \). This implies that \( f(x) \in D_y \cap B_y(y_0, \delta_1). \) Then, we have

\[
\begin{align*}
\| h(x) - h(x_0) - L(x - x_0) \| &= \| g(f(x)) - g(y_0) - L(x - x_0) \| \\
&\leq \| g(f(x)) - g(y_0) - Dg(y_0)(f(x) - y_0) \| \\
&\quad + \| Dg(y_0) (f(x) - f(x_0) - Df(x_0)(x - x_0)) \|
\end{align*}
\]
Proposition 9.19 Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f_1 : D \rightarrow \mathcal{Y}$, $f_2 : D \rightarrow \mathcal{Z}$, $x_0 \in D$, and $g : D \rightarrow \mathcal{Y} \times \mathcal{Z}$ 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, $Dg(x_0)(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$, $\forall h \in \mathcal{X}$. 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 $\operatorname{span}(A_D(x_0)) = \mathcal{X}$. Define $L : \mathcal{X} \rightarrow \mathcal{Y} \times \mathcal{Z}$ by, $\forall h \in \mathcal{X}$, $L(h) = (Df_1(x_0)(h), Df_2(x_0)(h))$. Clearly, $L$ is a linear operator. Note that

$$
\| L \| = \sup_{h \in \mathcal{X}, \| h \| \leq 1} \| L(h) \|
$$

$$
= \sup_{h \in \mathcal{X}, \| h \| \leq 1} \sqrt{\| Df_1(x_0)(h) \|^2 + \| Df_2(x_0)(h) \|^2}
$$

$$
\leq \sup_{h \in \mathcal{X}, \| h \| \leq 1} \sqrt{\| Df_1(x_0) \|^2 \| h \|^2 + \| Df_2(x_0) \|^2 \| h \|^2}
$$

$$
\leq \sqrt{\| Df_1(x_0) \|^2 + \| Df_2(x_0) \|^2} < +\infty
$$

where the first inequality follows from Proposition 7.64 and the last inequality follows from the fact $Df_1(x_0) \in B(\mathcal{X}, \mathcal{Y})$ and $Df_2(x_0) \in B(\mathcal{X}, \mathcal{Z})$. Hence, $L \in B(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$.

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 $\forall x \in D \cap B_{\mathcal{X}}(x_0, \delta_1)$, 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 $\forall x \in D \cap B_{\mathcal{X}}(x_0, \delta_2)$, we have $\| f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0) \| \leq \epsilon / \sqrt{2} \| x - x_0 \|$. Let $\delta = \min \{ \delta_1, \delta_2 \} \in (0, \infty) \subset \mathbb{R}$. $\forall x \in D \cap B_{\mathcal{X}}(x_0, \delta)$, we have

$$
\| g(x) - g(x_0) - L(x - x_0) \|
$$

$$
= \| (f_1(x), f_2(x)) - (f_1(x_0), f_2(x_0)) - (Df_1(x_0)(x - x_0), Df_2(x_0)(x - x_0)) \|
$$

$$
= \| (f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0), f_2(x)) - (f_2(x) - Df_2(x_0)(x - x_0)) \|
$$

$$
= \| (f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0), f_2(x)) - (f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0), f_2(x)) \|
$$

$$
+ \| f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0) \|^2
$$

$$
= \left( \| f_1(x) - f_1(x_0) - Df_1(x_0)(x - x_0) \|^2 + \| f_2(x) - f_2(x_0) - Df_2(x_0)(x - x_0) \|^2 \right)^{1/2}
$$
9.3. Chain Rule and Mean Value Theorem

\[ \leq \epsilon \| x - x_0 \| \]

Hence, \( g \) is differentiable at \( x_0 \) and \( D g(x_0) = L \).

“Necessity” By the differentiability of \( g \) at \( x_0 \), we have \( \text{span} \{ A_D(x_0) \} \subseteq \mathcal{X} \). Note that \( f_1 = \pi_y \circ g \). By Chain Rule and Proposition 9.13, \( f_1 \) is Fréchet differentiable at \( x_0 \) and

\[ D f_1(x_0) = \begin{bmatrix} \text{id}_y & \partial_{B(z,y)} \end{bmatrix} D g(x_0) \]

By symmetry, \( f_2 \) is Fréchet differentiable at \( x_0 \) and

\[ D f_2(x_0) = \begin{bmatrix} \partial_{B(y,z)} & \text{id}_x \end{bmatrix} D g(x_0) \]

Then, \( D g(x_0) = \begin{bmatrix} D f_1(x_0) & D f_2(x_0) \end{bmatrix} \).

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 \( \varphi(t) \), \( \forall t \in I^\circ \). Then, there exists \( t_0 \in I^\circ \) such that

\[ f(x_1) - f(x_2) = D f(\varphi(t_0))(x_1 - x_2) \]

**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 \( t \), \( \forall t \in I^\circ \) and

\[ D g(t)(h) = D f(\varphi(t))(D \varphi(t)(h)) = h D f(\varphi(t))(x_1 - x_2), \forall h \in \mathbb{R}, \forall t \in I^\circ \]

Then, \( D g(t) = D f(\varphi(t))(x_1 - x_2), \forall t \in I^\circ \). By Proposition 3.12, \( g \) is continuous. By Mean Value Theorem (Bartle, 1976, see Theorem 27.6), we have

\[ f(x_1) - f(x_2) = g(1) - g(0) = D g(t_0)(1 - 0) = D f(t_0 x_1 + (1 - t_0)x_2)(x_1 - x_2) \]

for some \( t_0 \in I^\circ \). 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 \( D := \{ a + i0 \mid a \in I := [0, 1] \subseteq \mathbb{R} \} \subseteq \mathbb{C} \), \( x_0 := x_r + i x_i \in D \), where \( x_r \in I \) and \( x_i = 0 \), \( f : D \to \mathbb{C} \) be Fréchet differentiable at \( x_0 \), \( D f(x_0) = d_r + i d_i \in \mathbb{C} \), 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 \( D g(x_r) = d_r = \text{Re} ( D f(x_0) ) \in \mathbb{R} \).
Proof  Note that span (A_t (x_r)) = \mathbb{R}. \forall \epsilon \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} = \bar{x} + i\bar{\epsilon}, \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}_r \in I and \bar{x}_r = 0. Then, the above implies that |(Re (f(\bar{x}_r) - Re (f(x_r))) - (\bar{x}_r - x_r)d_r) + i(Im (f(\bar{x}_r)) - Im (f(x_r))) - (\bar{x}_r - x_r)d_r)| \leq \epsilon |\bar{x}_r - x_r|. This further implies that |Re (f(\bar{x}_r) - Re (f(x_r))) - (\bar{x}_r - x_r)d_r| \leq \epsilon |\bar{x}_r - x_r|. Then, \forall t \in I \cap B_{\mathbb{R}} (x_0, \delta), \bar{x} := t + i\epsilon \in D \cap B_{\mathbb{C}} (x_0, \delta) and |g(t) - g(x_r) - (t - x_r)d_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 + i0 \mid a \in I := [0, 1] \subset \mathbb{R}\} \subset \mathbb{C}, f : D \rightarrow \mathbb{R} be continuous, f be Fréchet differentiable at a + i0, \forall a \in I^o. Then, \exists t_0 \in I^o such that Re (f(1) - f(0)) = Re (Df(t_0)).

Proof  Let g : I \rightarrow \mathbb{R} be given by g(t) = Re (f(t + i0)), \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 + i0)), \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 + i0)). Then, Re (f(1) - f(0)) = g(1) - g(0) = Re (Df(t_0 + i0)) = 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 \subseteq X, f : D \rightarrow 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 \rightarrow 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 \rightarrow \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 = \langle y_*, Df(\varphi(t))(d(x_1 - x_2)) \rangle = \langle y_*, Df(\varphi(t))(x_1 - x_2) \rangle d, \forall 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)\| = \|Df(\varphi(t_0))(x_1 - x_2)\| \leq \|Df(\varphi(t_0))(x_1 - x_2)\|, where the last inequality follows from Proposition 7.72. The desired result follows.
Case 3: $f(x_1) \neq f(x_2)$ and $K = \mathbb{C}$. By Proposition 7.85, $\exists y_* \in \mathbb{Y}^*$ with $\|y_*\| = 1$ such that $\langle y_*, f(x_1) - f(x_2) \rangle = \|f(x_1) - f(x_2)\|$. Let $D := \{a + it \mid a \in I\} \subset \mathbb{C}$. Define $\varphi : D \rightarrow D$ by $\varphi(t) = tx_1 + (1 - t)x_2$, $\forall t \in D$. 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 $\forall t \in D$. Define $g : D \rightarrow \mathbb{C}$ by $g(t) = \langle (y_*, f(\varphi(t))) \rangle$, $\forall t \in D$. 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 $Dg(a + i0)(d) = \langle (y_*, Df(\varphi(a) + i0))(d(x_1 - x_2)) \rangle$, $\forall d \in \mathbb{C}$ and $\forall a \in I^\circ$. Hence, $Dg(a + i0) = \langle (y_*, Df(\varphi(a))(x_1 - x_2)) \rangle$, $\forall a \in I^\circ$. Note that $g(1) - g(0) = \langle (y_*, f(x_1)) \rangle - \langle (y_*, f(x_2)) \rangle = \| f(x_1) - f(x_2) \| \in \mathbb{R}$. By Lemma 9.22, there exists $t_0 \in I^\circ$ such that $\left| g(1) - g(0) \right| = \text{Re}(Dg(t_0))$. Then, we have $\| f(x_1) - f(x_2) \| = |g(1) - g(0)| \leq \| Dg(t_0) \| = \langle (y_*, Df(\varphi(t_0))(x_1 - x_2)) \rangle \leq \| Df(\varphi(t_0))(x_1 - x_2) \|$, where the last inequality follows from Proposition 7.72. The desired result follows.

This completes the proof of the theorem. \hfill \square

**Proposition 9.24** Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X} \times \mathcal{Y}$, $f : D \rightarrow \mathcal{Z}$, $(x_0, y_0) \in D$. Assume that the following conditions hold.

(i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$, $\forall (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_0)$, we have $(x, ty + (1 - t)y_0) \in D$, $\forall t \in I := [0, 1] \subset \mathbb{R}$.

(ii) $f$ is partial differentiable with respect to $x$ at $(x_0, y_0)$ and $\frac{\partial f}{\partial x}(x_0, y_0) \in B(\mathcal{X}, \mathcal{Z})$.

(iii) $\forall (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_0)$, $f$ is partial differentiable with respect to $y$ at $(x, y)$ and $\frac{\partial f}{\partial y}$ is continuous at $(x_0, y_0)$.

Then, $f$ is Fréchet differentiable at $(x_0, y_0)$ and $Df(x_0, y_0) \in B(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$ is given by $Df(x_0, y_0)(h, k) = \frac{\partial f}{\partial x}(x_0, y_0)(h) \frac{\partial f}{\partial y}(x_0, y_0)(k)$, $\forall (h, k) \in \mathcal{X} \times \mathcal{Y}$.

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 $\text{span}(A_{D}(x_0, y_0)) = \mathcal{X} \times \mathcal{Y}$. Define $D_{x_0} := \{y \in \mathcal{Y} \mid (x_0, y) \in D\}$ and $D_{y_0} := \{x \in \mathcal{X} \mid (x, y_0) \in D\}$. By the partial differentiability of $f$ with respect to $x$ at $(x_0, y_0)$, we have $\text{span}(A_{D_{x_0}}(x_0)) = \mathcal{X}$. By the partial differentiability of $f$ with respect to $y$ at $(x_0, y_0)$, we have $\text{span}(A_{D_{y_0}}(y_0)) = \mathcal{Y}$. $\forall u \in A_{D_{x_0}}(x_0)$, $\forall v \in (0, \infty) \subset \mathbb{R}$, $\exists \tilde{u} := x_0 + r\tilde{u} \in D_{y_0}$ with $0 < r < \epsilon$ and $\tilde{u} \in B_{\mathcal{X}}(u, \epsilon)$. Then, $(\tilde{x}, y_0) = (x_0, y_0) + (r\tilde{u}, \vartheta_y) \in D$ and $(\tilde{u}, \vartheta_y) \in B_{\mathcal{X} \times \mathcal{Y}}((u, \vartheta_y), \epsilon)$. Hence, $(u, \vartheta_y) \in A_{D}(x_0, y_0)$. Then, $A_{D_{y_0}}(x_0) \times \{\vartheta_y\} \subseteq A_{D}(x_0, y_0)$, which implies that $\mathcal{X} \times \{\vartheta_y\} \subseteq \text{span}(A_{D}(x_0, y_0))$. By symmetry, we have $\{\vartheta_x\} \times \mathcal{Y} \subseteq \text{span}(A_{D}(x_0, y_0))$. By Proposition 7.17, $\text{span}(A_{D}(x_0, y_0)) = \mathcal{X} \times \mathcal{Y}$.
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
\[
\| L \| = \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \|(h, k)\| \leq 1} \| L(h, k) \|
\leq \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \|(h, k)\| \leq 1} \left( \| \frac{\partial f}{\partial x}(x_0, y_0) \| \| h \| + \| \frac{\partial f}{\partial y}(x_0, y_0) \| \| k \| \right)
\leq \sup_{(h, k) \in \mathcal{X} \times \mathcal{Y}, \|(h, k)\| \leq 1} \sqrt{\| \frac{\partial f}{\partial x}(x_0, y_0) \|^2 + \| \frac{\partial f}{\partial y}(x_0, y_0) \|^2}
\leq \sqrt{\| x \|^2 + \| y \|^2} \leq \sqrt{\| x \|^2 + \| y \|^2} \leq +\infty
\]
where the first inequality follows from Proposition 7.64 and the second inequality follows from Cauchy-Schwarz Inequality. Hence, $L \in B(\mathcal{X} \times \mathcal{Y}, \mathcal{Z})$.

$\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, by the partial differentiability of $f$ with respect to $x$ at $(x_0, y_0)$, there exists $\exists \delta_1 \in (0, \delta_0) \subset \mathbb{R}$ such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_1)$, we have $\| f(x_0, y_0) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \| \leq \epsilon/\sqrt{2} \| x - x_0 \|$. By the continuity of $\frac{\partial f}{\partial y}$ at $(x_0, y_0)$, there exists $\exists \delta_2 \in (0, \delta_1) \subset \mathbb{R}$, such that $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$, we have $\| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y_0) \| < \epsilon/\sqrt{2}$.

Claim 9.24.1 $\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \| \leq \epsilon/\sqrt{2} \| y - y_0 \|$, $\forall (x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$.

Proof of claim: Fix any $(x, y) \in D \cap \mathcal{B}_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta_2)$. Let $D = I$ if $\mathbb{K} = \mathbb{R}$; or $D = \{ a + i \theta \mid a \in I \} \subset \mathbb{C}$ if $\mathbb{K} = \mathbb{C}$. Define $\psi : D \to \mathcal{Z}$ by $\psi(t) = f(x, ty + (1 - t)y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(t(y - y_0))$, $\forall t \in D$. 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, $\exists \theta_0 \in I^2$,
\[
\| f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \| = \| \psi(1) - \psi(0) \| \leq \| D\psi(t_0) \|
\leq \| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0)y_0)(y - y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \|
\leq \| \frac{\partial f}{\partial y}(x, t_0 y + (1 - t_0)y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \| \| y - y_0 \| \leq \epsilon/\sqrt{2} \| y - y_0 \|
\]
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
\[
\| f(x, y) - f(x_0, y_0) - L(x - x_0, y - y_0) \|
\]
\[
\leq \left\| f(x, y) - f(x, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right\| \\
+ \left\| f(x, y_0) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \right\|
\]
\[
\leq \frac{\epsilon}{\sqrt{2}} \| y - y_0 \| + \frac{\epsilon}{\sqrt{2}} \| x - x_0 \| \leq \epsilon \| (x - x_0, y - y_0) \|
\]

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 \( K \). 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 multi-linear \( \mathcal{Y} \)-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 \( K \).

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 \( K \), \( D \subseteq \mathcal{X} \), \( f : D \to \mathcal{Y} \), 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 \( D^2 f(x_0) := f^{(2)}(x_0) \in B(\mathcal{X}, B(\mathcal{X}, \mathcal{Y})) = B_2(\mathcal{X}, \mathcal{Y}) \). \( D^2 f \) 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 $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \to \mathcal{Y}$, and $x_0 \in D$. Assume that $\exists \delta_0 \in (0, \infty) \subseteq \mathbb{R}$ such that $f$ is $k$-times Fréchet differentiable at $x$, $\forall x \in D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0)$, where $k \in \mathbb{N}$, that is $D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0) \subseteq \text{dom}(f^{(k)})$, 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 $\forall x \in D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0)$, $\forall k \in \mathbb{N}$, 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 $C_k$ at $x_0$ since $\text{dom}(f^{(k)})$ may not contain $D \cap \mathcal{B}_\mathcal{X}(x_0, \delta)$, $\forall \delta \in (0, \infty) \subseteq \mathbb{R}$. When $\text{dom}(f^{(k)}) \supseteq D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subseteq \mathbb{R}$, and $f$ is $(k+1)$-times differentiable at $x_0$, then $f$ is $C_k$ at $x_0$. In particular, if $f$ is $C_{k+1}$ at $x_0$, then $f$ is $C_k$ at $x$, $\forall x \in D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0)$, for some $\delta_0 \in (0, \infty) \subseteq \mathbb{R}$. If $f$ is infinitely many times differentiable at $x$, $\forall x \in D$, then $f$ is $C_\infty$.

Proposition 9.27 Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces over $\mathbb{K}$, $D \subseteq \mathcal{X}$, $f : D \to \mathcal{Y}$ 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 $(h_1, \ldots, h_n) \in \mathbb{R}^n$. Define the function $g : \text{dom}(f^{(n)}) \to \mathcal{Y}$ 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 $g^{(m)}(x_0) \in \mathcal{B}_m(\mathcal{X}, \mathcal{Y})$ is given by $g^{(m)}(x_0)(h_{n+m})\cdots(h_{n+1}) = f^{(n+m)}(x_0)(h_{n+m})\cdots(h_1)$, $\forall (h_{n+1}, \ldots, h_{n+m}) \in \mathbb{R}^m$.

(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 $\exists \delta_0 \in (0, \infty) \subseteq \mathbb{R}$ such that $D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0) \subseteq D_n$. By the $(n+1)$-times differentiability of $f$ at $x_0$, we have $\text{span}(A_{D_n}(x_0)) = \mathcal{X}$. Define $L : \mathcal{X} \to \mathcal{Y}$ by, $\forall h \in \mathcal{X}$, $L(h) = f^{(n+1)}(x_0)(h)(h_n)\cdots(h_1)$. Clearly, $L$ is a linear operator. Note that

$$
\|L\| = \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|L(h)\| = \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|f^{(n+1)}(x_0)(h)(h_n)\cdots(h_1)\|
\leq \sup_{h \in \mathcal{X}, \|h\| \leq 1} \|f^{(n+1)}(x_0)\| \|h\| \|h_n\| \cdots \|h_1\|
\leq \|f^{(n+1)}(x_0)\| \|h_n\| \cdots \|h_1\| < +\infty
$$

where the first inequality follows from Proposition 7.64. Hence, $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

$\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, by the differentiability of $f^{(n)}$ at $x_0$, $\exists \delta \in (0, \delta_0) \subseteq \mathbb{R}$ such that $\forall x \in D_n \cap \mathcal{B}_\mathcal{X}(x_0, \delta)$, we have $\|f^{(n)}(x) - f^{(n)}(x_0)\| \leq \epsilon/(1 + \|h_n\| \cdots \|h_1\|) \|x - x_0\|$. Then, we have

$$
\|g(x) - g(x_0) - L(x - x_0)\|
$$
\[ = \| f^{(n)}(x)(h_n) \cdots (h_1) - f^{(n)}(x_0)(h_n) \cdots (h_1) \\
- f^{(n+1)}(x_0)(x-x_0)(h_n) \cdots (h_1) \| \\
= \| (f^{(n)}(x) - f^{(n)}(x_0)) - f^{(n+1)}(x_0)(x-x_0)(h_n) \cdots (h_1) \| \\
\leq \| f^{(n)}(x) - f^{(n)}(x_0) - f^{(n+1)}(x_0)(x-x_0) \| \| h_n \| \cdots \| h_1 \|
\leq \epsilon \| x-x_0 \|
\]

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 \( C_{n+k} \) 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 \( C_{n+k-1} \) at \( \bar{x} \), \( \forall \bar{x} \in D \cap B_X(x_0, \delta_0) \). By inductive assumption, \( g \) is \( k \)-times Fréchet differentiable at \( \bar{x} \) and \( g^{(k)}(\bar{x})(h_{n+k}) \cdots (h_1) = f^{(n+k)}(\bar{x})(h_{n+k}) \cdots (h_1) \), \( \forall (h_{n+1}, \ldots, h_{n+k}) \in X^k \). Hence, \( \bar{x} \in D_{n+k} := \text{dom} (g^{(k)}) \). 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 \( (A_{D_{n+k}}(x_0)) = \text{span}(A_{D}(x_0)) = X \), since \( f \) is differentiable at \( x_0 \).

Define \( L : X \to B_k(X, Y) \) by, \( \forall h \in X, \forall (h_{n+1}, \ldots, h_{n+k}) \in X^k \), \( L(h)(h_{n+k}) \cdots (h_{n+1}) = f^{(n+k+1)}(x_0)(h)(h_{n+k}) \cdots (h_1) \). Clearly, \( L \) is a linear operator. Note that

\[
\| L \| = \sup_{h \in X, \| h \| \leq 1} \| L(h) \| \\
= \sup_{h \in X, \| h \| \leq 1, h_{n+1} \in X, \| h_{n+1} \| \leq 1, \ldots, k} \| L(h)(h_{n+k}) \cdots (h_{n+1}) \|
\]

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 \( \forall x \in \text{dom} (f^{(n+k)}) \cap B_X(x_0, \delta) = D \cap B_X(x_0, \delta) \), we have

\[
\| f^{(n+k)}(x) - f^{(n+k)}(x_0) - f^{(n+k+1)}(x_0)(x-x_0) \| \leq \epsilon/(1 + \| h_n \| \cdots \| h_1 \|) \| x-x_0 \|
\]

Then, \( \forall x \in D_{n+k} \cap B_X(x_0, \delta) = D \cap B_X(x_0, \delta) \),

\[
\| g^{(k)}(x) - g^{(k)}(x_0) - L(x-x_0) \| \\
= \sup_{h_{n+1} \in X, \| h_{n+1} \| \leq 1, \ldots, k} \| (g^{(k)}(x) - g^{(k)}(x_0) - L(x-x_0))(h_{n+1}) \| \\
= \sup_{h_{n+1} \in X, \| h_{n+1} \| \leq 1, \ldots, k} \| f^{(n+k)}(x)(h_{n+k}) \cdots (h_1) - f^{(n+k)}(x_0)(h_{n+k}) \cdots (h_1) \| \\
- f^{(n+k)}(x)(h_{n+k}) \cdots (h_1) - f^{(n+k+1)}(x_0)(x-x_0)(h_{n+k}) \cdots (h_1) \|
\]

where the first inequality follows from Proposition 7.64. Hence, \( g^{(k+1)}(x_0) = Dg^{(k)}(x_0) = L \). 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 \( D \cap \mathcal{B}_X(x_0, \delta_0) \subseteq D_{n+m} := \text{dom}(f^{(n+m)}) \subseteq D \) and \( f^{(n+m)} \) is continuous at \( x_0 \) \( \forall x \in D_{n+m} \cap \mathcal{B}_X(x_0, \delta_0) = D \cap \mathcal{B}_X(x_0, \delta_0) \), \( f \) is \( C_{n+m-1} \) at \( x_0 \) and \((n+m)\)-times differentiable at \( x \). By (i), \( g \) is \( m \)-times Fréchet differentiable at \( x \) and \( \forall(h_{n+1}, \ldots, h_{n+m}) \in \mathbb{X}^m \), we have
\[
\frac{\partial^{(m)}(x)(h_{n+1}) \cdots (h_{n+m})}{\partial h_{n+i}}(x_0) = f^{(n+m)}(x)(h_{n+1}) \cdots (h_{n+i}) + O(h_i)
\]
Therefore, \( f^{(n+m)} \) is \((n+m)\)-times differentiable at \( x \).

This completes the proof of the proposition.

**Proposition 9.28** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \), \( x_0 \in D \), and \( f : D \to \mathcal{Y} \) be \( C_n \) at \( x_0 \), where \( n \in \mathbb{N} \). Assume that \( D \) is locally convex at \( x_0 \). Then, \( f^{(n)}(x_0) \in \mathcal{B}_\mathcal{S}_n(\mathcal{X}, \mathcal{Y}) \).

**Proof** By the assumption, \( \exists \delta_0 \in (0, \infty) \subset \mathbb{R} \) such that \( D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0) \) is convex. Without loss of generality, assume \( f \) is \( n \)-times differentiable at \( x \), \( \forall x \in D \cap \mathcal{B}_\mathcal{X}(x_0, \delta_0) \). We will prove the proposition by mathematical induction on \( n \).

1° \( n = 1 \). Clearly \( f^{(1)}(x_0) \in \mathcal{B}_\mathcal{X}(x_0) = \mathcal{B}_1(\mathcal{X}, \mathcal{Y}) = \mathcal{B}_\mathcal{S}_1(\mathcal{X}, \mathcal{Y}) \).

Next, we consider \( n = 2 \). We will prove this case using an argument of contradiction. Suppose \( f^{(2)}(x_0) \) is not symmetric. Then, \( \exists h_0, l_0 \in \mathcal{X} \) 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 \( \text{span}(A_D(x_0)) = \mathcal{X} \). Then, \( f^{(2)}(x_0) \in \mathcal{B}_2(\mathcal{X}, \mathcal{Y}) \) and Proposition 3.56, \( \exists h_0, l_0 \in \text{span}(A_D(x_0)) \) such that \( f^{(2)}(x_0)(h_0)(l_0) \neq f^{(2)}(x_0)(l_0)(h_0) \). Since \( f^{(3)}(x_0) \) 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 \( f^{(2)}(x_0) \),
\[ \exists \epsilon_1 \in (0, \infty) \subset \mathbb{R} \text{ such that } \forall h \in \mathcal{B}_X \left( \hat{h}_0, \epsilon_1 \right), \forall \ell \in \mathcal{B}_X \left( \hat{\ell}_0, \epsilon_1 \right), \text{ we have } f^{(2)}(x_0)(\hat{h}_0(\hat{h}))(\hat{h}_0) \neq f^{(2)}(x_0)(\hat{\ell}(\hat{h}_0)) \hat{h}_0 \neq f^{(2)}(x_0)(\hat{\ell}_0(\hat{h}_0)) \hat{h}_0. \]

By \( \hat{h}_0, \hat{\ell}_0 \in A_D(x_0) \), \( \exists h, \ell \in (0, \epsilon_1) \subset \mathbb{R}, \exists h_0 \in \mathcal{B}_X \left( \hat{h}_0, \epsilon_1 \right), \exists \ell_0 \in \mathcal{B}_X \left( \hat{\ell}_0, \epsilon_1 \right) \), such that \( x_0 + r_h \hat{h}_0, x_0 + r_{\ell} \hat{\ell}_0 \in D \cap \mathcal{B}_X(x_0, \delta_0) \). Clearly, we have \( f^{(2)}(x_0)(r_h \hat{h}_0)(r_{\ell} \hat{\ell}_0) = r_h r_{\ell} f^{(2)}(x_0)(\hat{h}_0)(\hat{\ell}_0) \neq r_h r_{\ell} f^{(2)}(x_0)(\hat{h}_0)(\hat{\ell}_0) = f^{(2)}(x_0)(r_h \hat{h}_0)(r_{\ell} \hat{\ell}_0). \) Let \( h_0 := r_h \hat{h}_0 \) and \( \ell_0 := r_{\ell} \hat{\ell}_0 \). Then, by the convexity of the set \( D \cap \mathcal{B}_X(x_0, \delta_0) \), we have \( x_0, x_0 + h_0, x_0 + \ell_0, x_0 + h_0 + \ell_0 \in D \cap \mathcal{B}_X(x_0, \delta_0) \) and \( f^{(2)}(x_0)(h_0)(\ell_0) \neq f^{(2)}(x_0)(\hat{h}_0)(\hat{\ell}_0). \)

Clearly, \( h_0 \neq \hat{h}_0 \) and \( \ell_0 \neq \hat{\ell}_0 \). Let \( \epsilon_0 := \| f^{(2)}(x_0)(h_0)(\ell_0) - f^{(2)}(x_0)(h_0)(\ell_0) \| \| (\| h_0 \| \| \ell_0 \| ) \in (0, \infty) \subset \mathbb{R}. \) Since \( f \) is \( C_2 \) at \( x_0 \), then \( \exists \delta_1 \in (0, \delta_0) \subset \mathbb{R} \text{ such that } \forall x \in D \cap \mathcal{B}_X(x_0, \delta_1), \text{ we have } \| f^{(2)}(x) - f^{(2)}(x_0) \| < \epsilon_0/2. \) By proper scaling of \( h_0 \) and \( \ell_0 \), we may assume that \( \forall t_1, t_2 \in I := [0, 1] \subset \mathbb{R}, x_0 + t_1 h_0 + t_2 \ell_0 \in D \cap \mathcal{B}_X(x_0, \delta_0) \).

In summary, \( \exists h_0, l_0 \in \mathcal{X} \setminus \{ \emptyset \} \subset \mathbb{R} \) such that \( \epsilon_0 := \| f^{(2)}(x_0)(h_0)(\ell_0) - f^{(2)}(x_0)(h_0)(\ell_0) \| \| (\| h_0 \| \| \ell_0 \| ) \in (0, \infty) \subset \mathbb{R}, \forall t_1, t_2 \in I, f \) is twice differentiable at \( x_0 + t_1 h_0 + t_2 \ell_0 \in D \cap \mathcal{B}_X(x_0, \delta_0) \) and \( \| f^{(2)}(x_0 + t_1 h_0 + t_2 \ell_0) - f^{(2)}(x_0) \| < \epsilon_0/2. \)

Let \( D := I, \) if \( \mathcal{K} = \mathbb{R}; \) or \( D := \{ a + i0 \mid a \in I \} \subset \mathbb{C}, \) if \( \mathcal{K} = \mathbb{C}. \) \( \forall t_1 \in I, \) define \( \psi_{t_1} : D \rightarrow B(\mathcal{X}, \mathcal{Y}) \) by \( \psi_{t_1}(t_2) = f^{(1)}(x_0 + t_1 h_0 + t_2 \ell_0) - f^{(1)}(x_0 + t_1 h_0) - f^{(2)}(x_0)(t_2 \ell_0), \forall t_2 \in D. \) By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of \( \psi_{t_1} \) is Fréchet differentiable. Then, by Proposition 9.15, \( \psi_{t_1} \) is Fréchet differentiable at \( t_2, \forall t_2 \in D. \) By Mean Value Theorem, \( \exists \ell_2 \in I^0 \text{ such that } \)

\[
\begin{align*}
&\| f^{(1)}(x_0 + t_1 h_0 + t_2 \ell_0) - f^{(1)}(x_0) - f^{(2)}(x_0)(\ell_0) \| \\
&= \| \psi_{t_1}(1) - \psi_{t_1}(0) \| \leq \| D\psi_{t_1}(t_2) \| \\
&= \| f^{(2)}(x_0 + t_1 h_0 + t_2 \ell_0)(\ell_0) - f^{(2)}(x_0)(\ell_0) \| \\
&\leq \| f^{(2)}(x_0 + t_1 h_0 + t_2 \ell_0) - f^{(2)}(x_0) \| \| \ell_0 \| \leq \epsilon_0 \| \ell_0 \| /2
\end{align*}
\]

where the second inequality follows from Proposition 7.64.

Define \( \gamma : D \rightarrow \mathcal{Y} \) by \( \gamma(t_1) = f(x_0 + t_1 h_0) - f(x_0 + t_1 h_0) - f^{(2)}(x_0)(t_1 h_0), \forall t_1 \in D. \) 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, \( \exists \ell_1 \in I^0 \text{ such that } \)

\[
\begin{align*}
&\| f(x_0 + t_1 h_0 + \ell_1 h_0) - f(x_0 + h_0) - f(x_0 + \ell_1 h_0) + f(x_0) - f^{(2)}(x_0)(h_0) \| \\
&= \| \gamma(1) - \gamma(0) \| \leq \| D\gamma(t_1) \| = \| f^{(1)}(x_0 + \ell_1 h_0)(h_0) \\
&- f^{(1)}(x_0 + t_1 h_0)(h_0) - f^{(2)}(x_0)(h_0) \| \\
&\leq \| f^{(1)}(x_0 + \ell_1 h_0)(h_0) - f^{(1)}(x_0 + t_1 h_0)(h_0) - f^{(2)}(x_0)(h_0) \| \| h_0 \| \\
&= \| \psi_{t_1}(1) - \psi_{t_1}(0) \| \| h_0 \| \leq \epsilon_0 \| h_0 \| /2
\end{align*}
\]

where the second inequality follows from Proposition 7.64.
By symmetry, we have \( \| f(x_0 + h_0 + l_0) - f(x_0 + l_0) - f(x_0 + h_0) + f(x_0) - f^{(2)}(x_0)(h_0)(l_0) \| < \epsilon_0 \| h_0 \| \| l_0 \| \). Then, \( \| f^{(2)}(x_0)(h_0)(l_0) \| < \epsilon_0 \| h_0 \| \| l_0 \| \). This leads to the contradiction \( \epsilon_0 := \| f^{(2)}(x_0)(h_0)(l_0) \| / (\| h_0 \| \| l_0 \| ) < \epsilon_0 \).

Hence, \( f^{(2)}(x_0) \) must be symmetric and \( f^{(2)}(x_0) \in B_{S_2}(\mathcal{X}, \mathcal{Y}) \).

2° Assume that the result holds for \( n \leq k, k \in \{2, 3, \ldots\} \).

3° Consider the case \( n = k + 1 \). \( \forall (h_1, \ldots, h_{k+1}) \in \mathcal{X}^{k+1} \), let \( (v_1, \ldots, v_{k+1}) \) be a permutation of \( (h_1, \ldots, h_{k+1}) \). We need to show that

\[
\begin{align*}
& f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1) = f^{(k+1)}(x_0)(v_{k+1}) \cdots (v_1) \\
& \text{Since any permutation can be arrived at in finite number of steps by interchanging two consecutive elements, then, all we need to show is that,} \\
& \forall i = 1, \ldots, k,
\end{align*}
\]

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( 1 \leq i < k \); Case 2: \( i = k \). Define \( g : \text{dom}(f^{(k)}) \to \mathcal{Y} \) 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 \( \text{dom}(f^{(k)}) \supseteq D \cap B_X(x_0, \delta_0) \). Assume that \( f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_{i+1})(h_i) \cdots (h_1) = f^{(k+1)}(x_0)(v_{k+1}) \cdots (v_{i+1})(v_i) \cdots (v_1) \). This case is proved.

Case 2: \( i = k \). Define \( g : \text{dom}(f^{(k)}) \to \mathcal{Y} \) by \( \forall x \in \text{dom}(f^{(k)}) \), \( g(x) = f^{(k)}(x)(h_{k-1}) \cdots (h_1) \). Then, \( g \) is \( C_2 \) at \( x_0 \) and \( g^{(2)}(x_0)(u)(v) = f^{(k+1)}(x_0)(u)(v)(h_{k-1}) \cdots (h_1) \), \( \forall u, v \in \mathcal{X} \). Note that \( \text{dom}(f^{(k+1)}) \cap B_X(x_0, \delta_0) = D \cap B_X(x_0, \delta_0) \) is convex. By the case of \( n = 2 \), we have \( f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1) = f^{(k+1)}(x_0)(h_{k+1}) \cdots (h_1) \). This case is also proved.

Hence, \( f^{(k+1)}(x_0) \in B_{S_{k+1}}(\mathcal{X}, \mathcal{Y}) \).

This completes the induction process and the proof of the proposition. \( \square \)

**Definition 9.29** Let \( \mathcal{X}_1, \ldots, \mathcal{X}_p \), and \( \mathcal{Y} \) be normed linear spaces over \( \mathbb{K} \), where \( p \in \{2, 3, \ldots\} \), \( D \subseteq \mathcal{X} := \prod_{i=1}^p \mathcal{X}_i \), \( f : D \to \mathcal{Y} \), and \( x_0 := (x_{i_0}, \ldots, x_{i_0}) \in D \). Assume that \( \frac{\partial f}{\partial x_{i_1}} : \text{dom} \left( \frac{\partial f}{\partial x_{i_1}} \right) \to B(\mathcal{X}_i, \mathcal{Y}) \) is partial differentiable with respect to \( x_{i_2} \) at \( x_0 \in \text{dom} \left( \frac{\partial f}{\partial x_{i_1}} \right) \), where \( i_1, i_2 \in \{1, \ldots, p\} \). Then, this partial derivative is denoted by \( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(x_0) \in B(\mathcal{X}_{i_2}, B(\mathcal{X}_{i_1}, \mathcal{Y})) \), which is one of the second order partial derivatives of
f. We will let \( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}} \) denote the \( B(X_{i_2}, B(X_{i_1}, y)) \)-valued function whose domain of definition is \( \text{dom} \left( \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}} \right) := \left\{ x \in \text{dom} \left( \frac{\partial f}{\partial x_{i_1}} \right) \mid \frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}(x) \in B(X_{i_2}, B(X_{i_1}, y)) \right\} \). There are a total of \( p^2 \) second order partial derivatives of \( f \). Recursively, we may define kth-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 \( \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}} \), where \( i_1, \ldots, i_k \in \{1, \ldots, p\} \).

Let \( X \) be a normed linear space over \( \mathbb{K} \) and \( D \subseteq X \). Define \( \text{Sm} \left( D \right) := \bigcup_{n \in \mathbb{K}} nD \). Clearly, \( \text{Sm} \left( D \right) \subseteq \text{span} \left( D \right) \).

**Proposition 9.30** Let \( X \) be a normed linear space over \( \mathbb{K} \), \( D \subseteq X \), and \( x_0 \in \overline{D} \). Then, \( \text{span} \left( A_D \left( x_0 \right) \right) \subseteq \text{span} \left( \left( D \cap B(x_0, \delta) \right) - x_0 \right) \) and \( A_D \left( x_0 \right) \subseteq \text{Sm} \left( \left( D \cap B(x_0, \delta) \right) - x_0 \right) \), \( \forall \delta \in (0, \infty) \subseteq \mathbb{R} \).

**Proof** \( \forall u \in A_D \left( x_0 \right) \), we will show that \( u \in \text{Sm} \left( D - x_0 \right) \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), by \( u \in A_D \left( x_0 \right) \), we have \( \exists \tilde{r} := x_0 + r\tilde{u} \in D \) such that \( 0 < r < \epsilon \) and \( \tilde{u} \in B(u, \epsilon) \). Then, \( \tilde{u} \in \text{Sm} \left( D - x_0 \right) \cap \tilde{B}(u, \epsilon) \neq \emptyset \). Hence, by Proposition 3.3, \( u \in \text{Sm} \left( D - x_0 \right) \). By the arbitrariness of \( u \), we have \( A_D \left( x_0 \right) \subseteq \text{Sm} \left( D - x_0 \right) \subseteq \text{span} \left( D - x_0 \right) \). By Proposition 7.17, \( \text{span} \left( A_D \left( x_0 \right) \right) \subseteq \text{span} \left( D - x_0 \right) \). By Definition 3.2, \( \text{span} \left( A_D \left( x_0 \right) \right) \subseteq \text{span} \left( D - x_0 \right) \). \( \forall \delta \in (0, \infty) \subseteq \mathbb{R} \), we have \( x_0 \in \overline{D \cap B(x_0, \delta)} \) and \( A_D \left( x_0 \right) = A_{D \cap B(x_0, \delta)} \). Then, \( A_D \left( x_0 \right) \subseteq \text{Sm} \left( \left( D \cap B(x_0, \delta) \right) - x_0 \right) \) and \( \text{span} \left( A_D \left( x_0 \right) \right) \subseteq \text{span} \left( \left( D \cap B(x_0, \delta) \right) - x_0 \right) \). This completes the proof of the proposition.

### 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 \( D = \text{dom} \left( f \right) \). 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 \( X, Y, \) and \( Z \) be normed linear spaces over \( \mathbb{K} \), \( D \subseteq X \times Y \), \( f : D \to Z \), and \((x_0, y_0) \in D \). Assume that the following conditions are satisfied.

1. \( D \) is locally convex at \((x_0, y_0)\), that is, \( \exists \delta_0 \in (0, \infty) \subseteq \mathbb{R} \) such that \( D \cap B_{X \times Y} \left( (x_0, y_0), \delta_0 \right) =: \hat{D} \) is convex. Furthermore, \( \forall (x, y) \in \hat{D} \), we have \((x, y_0), (x_0, y) \in \hat{D} \).
2. \( \frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y), \) 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 \( \forall (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta) =: \bar{D}_\delta \), we have \( \| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \| \leq \varepsilon \| x - x_0 \| \| y - y_0 \|. \)

Furthermore, if, in addition, the following condition is satisfied.

(iii) \( \exists M \in (0, \infty) \subset \mathbb{R}, \forall h \in \mathcal{X}, \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \forall (x, y) \in \bar{D}, \exists \delta \in (0, \delta_0) \subset \mathbb{R}, \exists h_1, h_2 \in \text{Sm} \left( (D_y \cap B_{\mathcal{X}}(x_0, \delta)) - x_0 \right) \) such that \( h - h_1 + h_2 \leq \varepsilon \| h \|, \| h_1 \| \leq M \| h \|, \) and \( \| h_2 \| \leq M \| h \|, \) where \( D_y := \{ x \in \mathcal{X} \mid (x, y) \in D \}, \forall y \in \mathcal{Y}. \)

Then, \( \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \) exists and is given by

\[
\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k); \quad \forall h \in \mathcal{X}, \forall k \in \mathcal{Y}
\]

**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 \( \forall (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}}((x_0, y_0), \delta) =: \bar{D}_\delta \subseteq \bar{D}, \) 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), \( \bar{D}_\delta \) is convex and, \( \forall (x, y) \in \bar{D}_\delta, \) we have \((x_0, y), (x, y) \in \bar{D}_\delta. \)

**Claim 9.31.1** \( \forall (x, y) \in \bar{D}_\delta, \) \( \left\| f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0)(y - y_0) \right\| \leq \varepsilon /3 \| x - x_0 \| \| y - y_0 \|. \)

**Proof of claim:** FIX \((x, y) \in \bar{D}_\delta. \) Let \( I := [0, 1] \subset \mathbb{R}. \) Define \( \bar{D} := I \) if \( \mathbb{K} = \mathbb{R}; \) or \( \bar{D} := \{ a + i0 \mid a \in I \} \subseteq \mathbb{C} \) if \( \mathbb{K} = \mathbb{C}. \)

\( \forall t \in I, \) define \( \psi_{t_2} : \bar{D} \rightarrow B(\mathcal{Y}, Y) \) by \( \psi_{t_2}(t_1) = \frac{\partial f}{\partial y}(t_1 x + (1 - t_1)x_0, t_2 y + (1 - t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(t_1(x - x_0)), \forall t_1 \in \bar{D}. \)

By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of \( \psi_{t_2} \) is differentiable. By Proposition 9.15, \( \psi_{t_2} \) is differentiable. By Mean Value Theorem, \( \exists \bar{t}_1 \in I^0 \) such that

\[
\left\| \frac{\partial f}{\partial y}(t_1 x + (1 - t_1)x_0, t_2 y + (1 - t_2)y_0) - \frac{\partial f}{\partial y}(x_0, y_0)(x - x_0) \right\|
\]

\[
= \left\| \psi_{t_2}(1) - \psi_{t_2}(0) \right\| \leq \left\| D\psi_{t_2}(\bar{t}_1) \right\|
\]

\[
= \left\| \frac{\partial^2 f}{\partial x \partial y}(t_1 x + (1 - t_1)x_0, t_2 y + (1 - t_2)y_0)(x - x_0)ight.
\]

\[
- \left. \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x - x_0) \right\| \leq \varepsilon /3 \| x - x_0 \|
\]

where the last inequality follows from Proposition 7.64.
Define $\gamma : \bar{D} \rightarrow \mathbb{Z}$ by $\gamma(t_2) = f(x,t_2y + (1-t_2)y_0) - f(x_0, t_2y + (1-t_2)y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x-x_0)(y-y_0)$, $\forall t_2 \in \bar{D}$. By Propositions 9.19, 9.16, and 9.15 and Chain Rule, each term in the definition of $\gamma$ is differentiable. By Proposition 9.15, $\gamma$ is differentiable. By Mean Value Theorem, $\exists \tilde{t}_2 \in I^0$ such that

$$\| f(x,y) - f(x_0, y_0) - f(x_0, y_0) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(x-x_0)(y-y_0) \| = \| \gamma(1) - \gamma(0) \| \leq \| D\gamma(\tilde{t}_2) \|$$

where the second inequality follows from Proposition 7.64. This completes the proof of the claim.

Define $D_{x_0} := \{ y \in \mathbb{Y} | (x_0,y) \in D \}$; and $\hat{D} := \{ y \in \mathbb{Y} | (x_0,y) \in \text{dom} \left( \frac{\partial f}{\partial x} \right) \}$. By (ii), we have span $\{ A_{D_{x_0}}(y_0) \} = \mathbb{Y}$ and span $\{ A_{D_x}(x) \} = \mathbb{X}$, $\forall (x,y) \in \hat{D}$. Clearly, we have $\hat{D} \cap \mathcal{B}_y(y_0, \delta_0) = \mathcal{D}_{x_0} \cap \mathcal{B}_y(y_0, \delta_0)$. This implies that span $\{ A_{\hat{D}}(y_0) \} = \mathbb{Y}$.

$\forall k \in \mathbb{Y}$, define $L_k : \mathbb{X} \rightarrow \mathbb{Z}$ by $L_k(h) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(h)(k)$, $\forall h \in \mathbb{X}$. Clearly, $L_k$ is a linear operator. Note that $\| L_k \| = \sup_{h \in \mathbb{X}, \| h \| \leq 1} \| L_k(h) \| \leq \left\| \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| \| k \| < +\infty$, where the first inequality follows from Proposition 7.64. Hence, $L_k \in \mathbb{B}(\mathbb{X}, \mathbb{Z})$.

Define $L : \mathbb{Y} \rightarrow \mathbb{B}(\mathbb{X}, \mathbb{Z})$ by $L(k) = L_k$, $\forall k \in \mathbb{Y}$. Clearly, $L$ is a linear operator. Note that $\| L \| = \sup_{k \in \mathbb{Y}, \| k \| \leq 1} \| L(k) \| \leq \left\| \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right\| < +\infty$. Hence, $L \in \mathbb{B}(\mathbb{Y}, \mathbb{B}(\mathbb{X}, \mathbb{Z}))$.

Now, fix any $y \in \hat{D} \cap \mathcal{B}_y(y_0, \delta/\sqrt{2})$, we will show that $\| \Delta_y \| \leq (2M + 1)\varepsilon \| y - y_0 \|$, where $\Delta_y := \frac{\partial^2 f}{\partial x \partial y}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - L(y-y_0)$ $\in \mathbb{B}(\mathbb{X}, \mathbb{Z})$. This immediately implies that $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = L$ 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 $\frac{\partial f}{\partial x}(x_0, y)$, $\exists \delta_{y_1} \in (0, \delta/\sqrt{2}) \subset \mathbb{R}$ such that $\forall x \in D_y \cap \mathcal{B}_{x}(x_0, \delta_{y_1})$, we have $\left\| f(x, y) - f(x_0, y) - \frac{\partial f}{\partial x}(x_0, y)(x-x_0) \right\| \leq \left\| \Delta_y \right\| < (2M + 1)\varepsilon \| y - y_0 \|$.
Proposition 9.32

Let \( D \) be a subset of \( \mathbb{R}^n \). Then, we have

\[ \forall \epsilon > 0, \exists \delta > 0 \text{ such that } \forall (x, y) \in D \times \mathbb{R}^n, \| f(y) - f(x) - \nabla f(x)(y-x) \| \leq \epsilon \| y-x \|. \]

where \( f \) is a function defined on \( D \). By the existence of \( \frac{\partial f}{\partial x}(x_0, y_0) \), \( \exists \delta_2 \in (0, \delta/\sqrt{2}] \subset \mathbb{R} \) such that \( \forall x \in D \cap B_X(x_0, \delta_2) \), we have

\[ \| f(x, y) - f(x_0, y_0) - \left( \frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0) \right)(x-x_0) \| \leq \epsilon/3 \| y-y_0 \| \| x-x_0 \|. \]

Let \( \delta_y = \min \{ \delta_1, \delta_2 \} \in (0, \delta/\sqrt{2}] \subset \mathbb{R} \). \( \forall x \in D \cap B_X(x_0, \delta_y) \), we have \( (x, y), (x, y_0), (x_0, y) \in \hat{D}_y \). Then, \( x \in D \cap B_X(x_0, \delta_y) \). By Claim 9.31.1 and the preceding argument, we have

\[ \| \Delta_y(x-x_0) \| \]

\[ \leq \epsilon \| x-x_0 \| \| y-y_0 \| \]

\( \forall h \in \mathbb{R} \), by (iii) and the continuity of \( \Delta_y \), \( \exists h_1, h_2 \in \text{Sm} ((D \cap B_X(x_0, \delta_y)) - x_0) \), such that \( \| \Delta_y(h_1 + h_2) \| \leq \epsilon \| h \| \| y-y_0 \| \), \( \| h_1 \| \leq M \| h \| \), and \( \| h_2 \| \leq M \| h \| \). Then, \( \exists \alpha_1, \alpha_2 \in \mathbb{K} \) and \( \exists x_1, x_2 \in D \cap B_X(x_0, \delta_y) \) such that \( h_1 = \alpha_i (x_i - x_0) \), \( i = 1, 2 \). Then, we have

\[ \| \Delta_y(h) \| \]

\[ \leq \| \Delta_y(h_1 - h_2) \| + \epsilon \| h \| \| y-y_0 \| \]

\[ = \| \alpha_1 \Delta_y(x_1 - x_0) - \alpha_2 \Delta_y(x_2 - x_0) \| + \epsilon \| h \| \| y-y_0 \| \]

\[ \leq \| \alpha_1 \Delta_y(x_1 - x_0) \| + \| \alpha_2 \| \| \Delta_y(x_2 - x_0) \| + \epsilon \| h \| \| y-y_0 \| \]

\[ \leq \| \alpha_1 \| \| x_1 - x_0 \| + \| \alpha_2 \| \| x_2 - x_0 \| + \epsilon \| h \| \| y-y_0 \| \]

\[ = \| \alpha_1 + \| h_1 \| + \| h \| \| y-y_0 \| \]

This completes the proof of the proposition.

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 \( \mathbb{K} \), \( D \subseteq X \times Y \), \( f : D \to Z \), 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 \( D \cap B_X \times B_Y \left( (x_0, y_0), \delta_0 \right) =: \hat{D} \) is convex. Furthermore, \( \forall (x, y) \in \hat{D} \), we have \((x, y) \in \hat{D} \).

(ii) \( \frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y), \frac{\partial^2 f}{\partial x^2}(x, y), \frac{\partial^2 f}{\partial y^2}(x, y), \frac{\partial^2 f}{\partial x \partial y}(x, y), \) and \( \frac{\partial^2 f}{\partial y \partial x}(x, y) \) exist, \( \forall (x, y) \in \hat{D} \), and \( \frac{\partial^2 f}{\partial x \partial y} \) and \( \frac{\partial^2 f}{\partial y \partial x} \) are continuous at \((x_0, y_0)\).
Then, \( \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0)(k)(h) = \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(h)(k) \), \( \forall h \in \mathcal{X}, \forall k \in \mathcal{Y} \).

**Proof** Define \( D_{x_0} := \{ y \in \mathcal{Y} \mid (x_0, y) \in D \} \) and \( D_{y_0} := \{ x \in \mathcal{X} \mid (x, y_0) \in D \} \). By (ii), we have \( \text{span} \left( A_{D_{x_0}} (x_0) \right) = \mathcal{Y} \) and \( \text{span} \left( A_{D_{y_0}} (y_0) \right) = \mathcal{X} \).

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by Proposition 9.31, \( \exists \delta_1 \in (0, \delta_0) \subset \mathbb{R} \) such that \( (x, y) \in D \cap B_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_1) \) implies \( f(x, y) - f(x_0, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(x-x_0)(y-y_0) \leq \epsilon/2 \| x-x_0 \| \| y-y_0 \| \).

By the multi-linearity of \( f \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( (x, y) \in B_{\mathcal{X} \times \mathcal{Y}} ((x_0, y_0), \delta_2) \) and \( D := \mathcal{D}_{\delta_2} \subseteq \mathcal{D} \), we have \( f(x, y) - f(x_0, y_0) - f(x_0, y) + f(x_0, y_0) - \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0)(y-y_0)(x-x_0) \leq \epsilon/2 \| x-x_0 \| \| y-y_0 \| \).

Therefore, \( \exists \delta := \min \{ \delta_1, \delta_2 \} \in (0, \delta_0) \subset \mathbb{R} \). By the continuity of \( f \) and \( \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0) \), there exists \( \exists \delta_0 \in \mathbb{R} \) such that \( (x_0, y_0)(k_0)(\hat{h_0}) \neq \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(\hat{h_0})(k_0) \).

By the multi-linearity of these two second-order partial derivatives, \( \exists \delta_0 \in A_{D_{x_0}} (x_0) \) and \( \exists \delta_0 \in A_{D_{y_0}} (y_0) \) such that \( \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(\hat{h_0})(k_0) \neq \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0)(\hat{h_0})(k_0) \).

Therefore, \( \exists \delta_0 \in A_{D_{x_0}} (x_0) \) and \( \exists \delta_0 \in A_{D_{y_0}} (y_0) \) such that \( \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(\hat{h_0})(k_0) \neq \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0)(\hat{h_0})(k_0) \). By Proposition 9.30, we have

\[
A_{D_{x_0}} (x_0) \subseteq \text{Sm} \left( (D_{y_0} \cap B_{\hat{y}} (x_0, \delta_0)) - x_0 \right)
\]

\[
A_{D_{y_0}} (y_0) \subseteq \text{Sm} \left( (D_{x_0} \cap B_{\hat{x}} (y_0, \delta_0)) - y_0 \right)
\]

This implies that, by the continuity of these two second-order partial derivatives, \( \exists \delta_0 \in \text{Sm} ((D_{y_0} \cap B_{\hat{y}} (x_0, \delta_0)) - x_0) \) and \( \exists \delta_0 \in \text{Sm} ((D_{x_0} \cap B_{\hat{x}} (y_0, \delta_0)) - y_0) \). Then, \( \exists \tilde{x} \in D_{y_0} \cap B_{\hat{y}} (x_0, \delta_0), \exists \tilde{y} \in D_{x_0} \cap B_{\hat{x}} (y_0, \delta_0), \text{ and } \exists \alpha_h, \alpha_k \in \mathbb{R} \) such that \( \hat{h_0} = \alpha_h (\tilde{x} - x_0) \) and \( \hat{k_0} = \alpha_k (\tilde{y} - y_0) \). By the multi-linearity of these two second-order partial derivatives, \( \frac{\partial^2 f}{\partial y \partial x} (x_0, y_0)(\tilde{y} - y_0)(\tilde{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y} (x_0, y_0)(\tilde{x} - x_0)(\tilde{y} - y_0) \). Clearly, \( (\tilde{x}, y_0)(x_0, \tilde{y} \in \mathcal{D}_{\delta} \). By the convexity of \( \mathcal{D}_{\delta_0} \), we have \( (\tilde{x}, \tilde{y}) := 0.5 (\tilde{x}, y_0) + 0.5 (x_0, \tilde{y} \in \mathcal{D}_{\delta_0} \). By the multi-linearity of these two second-order partial derivatives, we have
\[ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)(\dot{y} - y_0)(\dot{x} - x_0) \neq \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)(\dot{x} - x_0)(\dot{y} - y_0). \] 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 \( \exists y_0 \in Y \) 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 \( f^{(i)}(x) = \partial_{B_X,Y}(x,y) \), \( \forall x \in D \cap B_X(x_0, \delta_0), \forall i \in \mathbb{N} \).

**Proof** \( \forall x \in D \cap B_X(x_0, \delta_0) \), let \( \delta_\xi := \delta_0 - \| x - x_0 \| > 0 \). Then, \( \forall \xi \in D \cap B_X(x, \delta_\xi) \subseteq D \cap B_X(x_0, \delta_0) \), we have \( \dot{f}(x) = y_0 \). Note that \( \text{span}(A_D(x)) = X \). By Proposition 9.10, \( f^{(1)}(x) = \partial_{B_X,Y} \). 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 \( f^{(i)}(x) = \partial_{B_X,Y}, \forall x \in D \cap B_X(x_0, \delta_0), \forall i \in \mathbb{N} \). This completes the proof of the proposition. \( \Box \)

**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 \( C_\infty \) 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 \( C_\infty \) at \( x_0 \).

(iv) If \( g \) is \( C_\infty \) 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 \( C_\infty \) at \( x_0 \).

**Proof** (i) We will use mathematical induction on \( k \) to prove this statement.

1. Assume \( k = 1 \). \( \forall x \in D_2 \), 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 \( k \leq k \in \mathbb{N} \).

3. Consider the case \( k = k + 1 \). By inductive assumption, \( g^{(k)}(x) \) exists and \( g^{(i)}(x) = f^{(i)}(x), \forall x \in D_2, \forall i \in \{1, \ldots, k\} \). Then, \( \partial_{B_X,Y}(g^{(k)})(x) = \partial_{B_X,Y}(f^{(k)})(x) \). This completes the proof of the proposition.
9.4. HIGHER ORDER DERIVATIVES

\( D_2 \subseteq \text{dom}\left( f^{(k)} \right) =: \bar{D}_1. \ \forall x \in D_2, \) by the assumption, \( f^{(k)} \) is differentiable at \( x \) and \( \text{span}\left( \text{span}(A_{D_2}(x)) \right) = X. \) By Proposition 9.11, \( g^{(k)} \) is Fréchet differentiable at \( x \) and \( g^{(k+1)}(x) = Dg^{(k)}(x) = Df^{(k)}(x) = f^{(k+1)}(x). \) This completes the induction process.

(ii) We will use mathematical induction on \( k \) to prove this statement.

1° \( k = 1. \ \forall x \in D_2, \) by Proposition 9.11, \( f^{(1)}(x) \) exists and \( f^{(1)}(x) = g^{(1)}(x). \) Hence, the result holds.

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, \( f^{(k)}(x) \) exists and \( f^{(k)}(x) = g^{(k)}(x), \) \( \forall x \in D_2, \forall i \in \{1, \ldots, \bar{k}\}. \) Then, \( \text{dom}\left( g^{(k)} \right) = D_2 \subseteq \text{dom}\left( f^{(k)} \right) =: \bar{D}_1 \subseteq D_1. \ \forall x \in D_2, \) by the assumption, \( g^{(k)} \) is differentiable at \( x \) and \( \delta x \in (0, \infty) \subset \mathbb{R} \) such that \( \bar{D}_1 \cap B_X(x, \delta x) = D_2 \cap B_X(x, \delta x). \) Then, by Proposition 9.11, \( f^{(k)} \) is Fréchet differentiable at \( x \) and \( f^{(k+1)}(x) = Df^{(k)}(x) = Dg^{(k)}(x) = g^{(k+1)}(x) \). This completes the induction process.

(iii) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: \( n \in \mathbb{N}; \) Case 2: \( n = \infty. \)

Case 1: \( n \in \mathbb{N}. \) Without loss of generality, assume \( f \) is \( n \)-times differentiable at \( x, \forall x \in D_1 \cap B_X(x_0, \delta). \) Let \( D_2 := D_2 \cap B_X(x_0, \delta) \) and \( \bar{g} := f|_{D_2}, \) \( \forall x \in D_2, f^{(n)}(x) \) exists. Let \( \delta x := \delta - \|x - x_0\| \in (0, \infty) \subset \mathbb{R}. \) Then, \( D_2 \cap B_X(x, \delta x) = D_2 \cap B_X(x, \delta x). \) Hence, \( \text{span}\left( \text{span}(A_{D_2}(x)) \right) = \mathbb{X}. \) Then, by (i), \( g^{(n)}(x) \) exists and \( g^{(n)}(x) = \bar{g}^{(n)}(x), \) \( \forall x \in D_2, \forall i \in \{1, \ldots, n\}. \) By (ii), \( g^{(n)}(x) \) is Fréchet differentiable at \( x, \delta x \in (0, \infty) \subset \mathbb{R} \) such that \( \bar{D}_1 \cap B_X(x, \delta x) = D_2 \cap B_X(x, \delta x). \) Then, by Proposition 9.11, \( f^{(n)} \) is Fréchet differentiable at \( x \) and \( f^{(n+1)}(x) = Df^{(n)}(x) = Dg^{(n)}(x) = g^{(n+1)}(x). \) Hence, \( g^{(n)} \) is continuous at \( x_0. \) Therefore, \( \bar{g} \in \mathcal{C}_n \) at \( x_0. \)

Case 2: \( n = \infty. \) By the assumption, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( f \) is \( \mathcal{C}_\delta \) at \( x, \forall x \in D \cap B_X(x_0, \delta), \forall i \in \mathbb{N}. \) By Case 1, \( g \) is \( \mathcal{C}_\delta \) at \( x, \forall x \in D \cap B_X(x_0, \delta), \forall i \in \mathbb{N}. \) Hence, \( g \) is \( \mathcal{C}_\infty \) at \( x_0. \)

(iv) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: \( n \in \mathbb{N}; \) Case 2: \( n = \infty. \)

Case 1: \( n \in \mathbb{N}. \) Without loss of generality, assume \( g \) is \( n \)-times differentiable at \( x, \forall x \in D_2 \cap B_X(x_0, \delta) =: \bar{D}_2. \) Let \( \bar{g} := g|_{\bar{D}_2}, \) \( \forall x \in \bar{D}_2, g^{(n)}(x) \) exists. Let \( \delta x := \delta - \|x - x_0\| \in (0, \infty) \subset \mathbb{R}. \) Then, \( D_2 \cap B_X(x, \delta x) = D_2 \cap B_X(x, \delta x) = D_1 \cap B_X(x, \delta x). \) Hence, \( \text{span}\left( \text{span}(A_{\bar{D}_2}(x)) \right) = \text{span}\left( \text{span}(A_{\bar{D}_2}(x)) \right) = X. \) Then, by (i), \( g^{(n)}(x) \) is Fréchet differentiable at \( x, \forall x \in \bar{D}_2, \forall i \in \{1, \ldots, n\}. \) By (ii), \( g^{(n)}(x) \) is Fréchet differentiable at \( x, \forall x \in \bar{D}_2, \forall i \in \{1, \ldots, n\}. \) By the continuity of \( g^{(n)} \) at \( x_0, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \delta) \subset \mathbb{R}, \forall x \in \text{dom}\left( g^{(n)} \right) \cap B_X(x_0, \delta) = D_2 \cap B_X(x_0, \delta) = D_1 \cap B_X(x_0, \delta), \) we have
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, \( \forall x \in D_1, f(1)(\bar{x}) = f(1)(1)x \), \( \forall i \in \mathbb{N} \) 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. \( \square \)

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 \( C_{\infty} \), \( f(1)(x) = \text{id}_X \), and \( f(1)(x) = \vartheta_{B_{\infty+1}(X, X)}, \forall x \in X, \forall i \in \mathbb{N} \).

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

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 \( C_{\infty} \), \( f(1)(x, y) = [ \text{id}_X \vartheta_{B_{\infty+1}(X, X)} ] \) and \( f(1)(x, y) = \vartheta_{B_{\infty+1}(X \times Y, X)} \), \( \forall (x, y) \in X \times Y, \forall i \in \mathbb{N} \).

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

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 \( f(1)(x_1, x_2) = \vartheta_{B_{\infty+1}(X \times X, X)} \), \( \forall (x_1, x_2) \in X \times X, \forall i \in \mathbb{N} \).

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

Proposition 9.39 Let \( X \) be a normed linear space over \( \mathbb{K} \), \( f : \mathbb{K} \times X \to X \) be given by \( f(\alpha, x) = \alpha x \), \( \forall (\alpha, x) \in \mathbb{K} \times X \). Then, \( f \) is \( C_{\infty} \), \( f(1)(\alpha, x) = [ x \text{id}_X ] \), \( f(2)(\alpha, x)(d_2, h_2)(d_1, h_1) = d_1 h_2 + d_2 h_1 \), and \( f(2)(\alpha, x) = \vartheta_{B_{\infty+2}(\mathbb{K} \times X, X)} \), \( \forall (\alpha, x) \in \mathbb{K} \times X, \forall i \in \mathbb{N}, \forall (d_1, h_1) \in \mathbb{K} \times X, \forall (d_2, h_2) \in \mathbb{K} \times X \).
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 $K$, $D \subseteq X$, $f_1 : D \rightarrow Y$, $f_2 : D \rightarrow Y$, $x_0 \in D$, $\alpha_1, \alpha_2 \in K$, $n \in N$, $k \in N \cup \{ \infty \}$, and $g : D \rightarrow Y$ 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 $g^{(i)}(x) = \alpha_1 f_1^{(i)}(x) + \alpha_2 f_2^{(i)}(x)$, $\forall x \in D \cap B_X(x_0, \delta_0)$, $\forall i \in N$ with $i \leq k$, for some $\delta_0 \in (0, \infty) \subseteq R$.

Proof  This is straightforward, and is therefore omitted.

Proposition 9.41 Let $X$ and $Y$ be normed linear spaces over $K$, $f : B(X, Y) \times X \rightarrow Y$ be given by $f(A, x) = Ax$, $\forall (A, x) \in B(X, Y) \times X$. Then, $f$ is $C_\infty$, $f^{(1)}(A, x) = \left[ \begin{array}{c} \rho(A) \end{array} \right]$, $f^{(2)}(A, x)(\Delta_2, h_2)(\Delta_1, h_1) = \Delta_1 h_2 + \Delta_2 h_1$, and $f^{(i+2)}(A, x) = \partial_{B_{gi+2}(B(\Delta_1, h_1))} B(\Delta_2, h_2)$, $\forall (A, x) \in B(X, Y) \times X$, $\forall i \in N$, $\forall (\Delta_1, h_1) \in B(X, Y) \times X$, $\forall (\Delta_2, h_2) \in B(X, Y) \times X$.

Proof  This is straightforward, and is therefore omitted.

Proposition 9.42 Let $X$, $Y$, and $Z$ be normed linear spaces over $K$, $f : B(Y, Z) \times B(X, Y) \rightarrow B(X, Z)$ be given by $f(A_2, A_1) = A_2 A_1$, $\forall (A_2, A_1) \in B(Y, Z) \times B(X, Y)$. Then, $f$ is $C_\infty$, $f^{(1)}(A_2, A_1) = \left[ \begin{array}{c} \rho(A_2) \\ \rho(A_1) \end{array} \right]$, $f^{(2)}(A_2, A_1)(\Delta_{xy_1}, \Delta_{xy_2}, \Delta_{x_1y_1}) = \Delta_{x_1y_2} \Delta_{x_2y_1} + \Delta_{x_2y_2} \Delta_{x_1y_1}$, and $f^{(i+2)}(A_2, A_1) = \partial_{B_{gi+2}(B(Y, Z) \times B(X, Y))} B(Y, Z) \times B(X, Y)$, $\forall (A_2, A_1) \in B(Y, Z) \times B(X, Y)$, $\forall (\Delta_{xy_1}, \Delta_{xy_2}) \in B(Y, Z) \times B(X, Y)$.

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 $p, m \in N$, $Z := B(X_i, Y_j)$, $i = 1, \ldots, p$, $j = 1, \ldots, m$, $Z := \Pi_{j=1}^m \Pi_{i=1}^p Z_{ij}$, and $f : Z \rightarrow B(\Pi_{j=1}^m X_i, \Pi_{i=1}^p Y_j)$ be given by $f(A_1, \ldots, A_m) = \left[ \begin{array}{c} A_{11} \\ \vdots \\ A_{mp} \end{array} \right]$, $\forall (A_1, \ldots, A_m) \in Z$.

Then, $f$ is $C_\infty$ and

$f^{(1)}(A_1, \ldots, A_m) = \left[ \begin{array}{c} \Delta_{11} \\ \vdots \\ \Delta_{1p} \\ \Delta_{m1} \\ \vdots \\ \Delta_{mp} \end{array} \right]$

$f^{(i+1)}(A_1, \ldots, A_m) = \partial_{B_{gi+1}(Z, B(\Pi_{j=1}^m X_i, \Pi_{i=1}^p Y_j))} (\Pi_{j=1}^m X_i, \Pi_{i=1}^p Y_j)$

$\forall (A_1, \ldots, A_m) \in Z$, $\forall i \in N$, $\forall (\Delta_{11}, \ldots, \Delta_{mp}) \in Z$.

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.

1. $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ 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 $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_X(x_0, \delta_0)$, $\forall i \in \{1, \ldots, k\}$.

2. 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 $k = \tilde{k} \in \mathbb{N}$.

3. Consider the case $k = \tilde{k} + 1$. “Necessity” Let $f_1^{(\tilde{k}+1)}(x)$ and $f_2^{(\tilde{k}+1)}(x)$ exist, $\forall x \in D \cap B_X(x_0, \delta_0)$. By inductive assumption, $g^{(\tilde{k})}(x)$ exists and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_X(x_0, \delta_0)$, $\forall i \in \{1, \ldots, \tilde{k}\}$. Let

$$\tilde{D}_1 := \text{dom} \left( f_1^{(\tilde{k})} \right), \quad \tilde{D}_2 := \text{dom} \left( f_2^{(\tilde{k})} \right) \text{ and } \tilde{D} := \text{dom} \left( g^{(\tilde{k})} \right).$$

Then, $\tilde{D}_1 \subseteq D$, $\tilde{D}_2 \subseteq D$, $\tilde{D} \subseteq D \cap \text{dom} \left( f_1^{(\tilde{k})} \right)$, and $\tilde{D} \cap \text{dom} \left( f_1^{(\tilde{k})} \right) \subseteq \tilde{D}_1 \cap \tilde{D}_2$. This implies that $\tilde{D} \cap B_X(x_0, \delta_0) = D \cap B_X(x_0, \delta_0) =: \tilde{D}$. By the assumption, $f_1^{(\tilde{k})}$ and $f_2^{(\tilde{k})}$ are differentiable at $x$, $\forall x \in \tilde{D}$. By Propositions 9.19 and 9.11, $\forall x \in \tilde{D}$, $g^{(\tilde{k})}$ is differentiable at $x$ and $g^{(\tilde{k}+1)}(x) = D \left. g^{(\tilde{k})} \right|_D (x) = \begin{bmatrix} D \left. f_1^{(\tilde{k})} \right|_D (x) \\ D \left. f_2^{(\tilde{k})} \right|_D (x) \end{bmatrix} = \begin{bmatrix} f_1^{(\tilde{k}+1)}(x) \\ f_2^{(\tilde{k}+1)}(x) \end{bmatrix}$.

“Sufficiency” Let $g^{(\tilde{k}+1)}(x)$ exist, $\forall x \in D \cap B_X(x_0, \delta_0)$. By inductive assumption, $f_1^{(\tilde{k})}(x)$ and $f_2^{(\tilde{k})}(x)$ exist and $g^{(i)}(x) = \begin{bmatrix} f_1^{(i)}(x) \\ f_2^{(i)}(x) \end{bmatrix}$, $\forall x \in D \cap B_X(x_0, \delta_0)$, $\forall i \in \{1, \ldots, \tilde{k}\}$. Let $\tilde{D}_1 := \text{dom} \left( f_1^{(\tilde{k})} \right)$, $\tilde{D}_2 := \text{dom} \left( f_2^{(\tilde{k})} \right)$ and $\tilde{D} := \text{dom} \left( g^{(\tilde{k})} \right)$. Then, $\tilde{D}_1 \subseteq D$, $\tilde{D}_2 \subseteq D$, $\tilde{D} \subseteq D \cap \text{dom} \left( f_1^{(\tilde{k})} \right)$, and $\tilde{D} \cap \text{dom} \left( f_1^{(\tilde{k})} \right) \subseteq \tilde{D}_1 \cap \tilde{D}_2$. This implies that $\tilde{D} \cap B_X(x_0, \delta_0) = D \cap B_X(x_0, \delta_0) =: \tilde{D}$. Then, $\text{span} \left( A_D(x) \right) = \text{span} \left( A_D(x) \right) = X$, $\forall x \in \tilde{D}$. By Propositions 9.11 and 9.19, we have, $\forall x \in \tilde{D}$, $f_1^{(\tilde{k})} \left|_D \right.$ and $f_2^{(\tilde{k})} \left|_D \right.$ are differen-
9.4. **HIGHER ORDER DERIVATIVES**

Differentiable at \( x \) and \( g^{(k+1)}(x) = Dg^{(k)}(x) \) \( \big|_{\tilde{D}} \). By Proposition 9.11, we have \( Df_1^{(k)} \bigg|_{\tilde{D}}(x) = Df_2^{(k)}(x) \) and \( Df_2^{(k)}(x) = Df_2^{(k+1)}(x) \), \( \forall x \in \tilde{D} \). Then, \( g^{(k+1)}(x) = \begin{bmatrix} f_1^{(k+1)}(x) \\ f_2^{(k+1)}(x) \end{bmatrix}, \forall x \in \tilde{D} \).

This completes the induction process.

(ii) We will distinguish two exhaustive and mutually exclusive cases:

Case 1: \( n \in \mathbb{N} \). "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 \( x, \forall x \in D \cap B_\delta(x_0, \delta_0) =: \tilde{D} \). By (i), \( f_1 \) and \( f_2 \) are \( n \)-times differentiable at \( x, \forall x \in \tilde{D} \) and \( g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix} \), \( \forall x \in \tilde{D} \).

"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 \( x, \forall x \in D \cap B_\delta(x_0, \delta_0) =: \tilde{D} \). By (i), \( g \) is \( n \)-times differentiable at \( x, \forall x \in \tilde{D} \) and \( g^{(n)}(x) = \begin{bmatrix} f_1^{(n)}(x) \\ f_2^{(n)}(x) \end{bmatrix} \), \( \forall x \in \tilde{D} \).

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 \( C_i \) at \( x, \forall x \in D \cap B_\delta(x_0, \delta), \forall i \in \mathbb{N} \). By Case 1, \( f_1 \) and \( f_2 \) are \( C_i \) at \( \forall x \in D \cap B_\delta(x_0, \delta), \forall i \in \mathbb{N} \). 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 \( \forall x \in D \cap B_\delta(x_0, \delta), \forall i \in \mathbb{N} \). By Case 1, \( g \) is \( C_i \) at \( \forall x \in D \cap B_\delta(x_0, \delta), \forall i \in \mathbb{N} \). 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, \( \forall y \in D_2 \cap B_y(y_0, \delta_1) \), and \( g^{(1)} \) is continuous at \( y_0 \). By
CHAPTER 9. DIFFERENTIATION IN BANACH SPACES

If $f$ being $C_1$ at $x_0$, then, by Proposition 9.7, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f(x) \in D_2 \cap B_\mathbb{R} (y_0, \delta_1)$ and $f^{(1)}(x)$ exists, $\forall x \in D_1 \cap B_\mathbb{R} (x_0, \delta)$, and $f^{(1)}$ is continuous at $x_0$. $\forall x \in D_1 \cap B_\mathbb{R} (x_0, \delta)$, by Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$. By Propositions 3.12, 9.7, 3.32, and 9.42, $h^{(1)}$ is continuous at $x_0$. Hence, $h$ is $C_1$ at $x_0$.

2° Assume that the result holds for $k \leq \tilde{k} \in \mathbb{N}$.

3° Consider the case $k = \tilde{k} + 1$. By $g$ being $C_{k+1}$ at $y_0$, then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$ such that $g^{(1)}(y)$ exists, $\forall y \in D_2 \cap B_\mathbb{R} (y_0, \delta_1) =: \bar{D}$, and $g^{(1)}$ is $C_k$ at $y_0$. By $f$ being $C_{k+1}$ at $x_0$, then, by Proposition 9.7, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $f(x) \in \bar{D}$ and $f^{(1)}(x)$ exists, $\forall x \in D_1 \cap B_\mathbb{R} (x_0, \delta) =: \bar{D}$, and $f^{(1)}$ is $C_k$ at $x_0$. $\forall x \in \bar{D}$, by Chain Rule, $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$. Then, $h^{(1)}|_{\bar{D}} : \bar{D} \rightarrow B(\mathbb{R}, \mathbb{Z})$ is given by $h^{(1)}|_{\bar{D}} (x) = g^{(1)}|_{\bar{D}} (f^{(1)}|_{D}(x)) f^{(1)}|_{\bar{D}} (x), \forall x \in \bar{D}$. By Proposition 9.34, $g^{(1)}|_{\bar{D}}$ is $C_k$ at $y_0$ and $f^{(1)}|_{\bar{D}}$ and $f^{(1)}|_{\bar{D}}$ are $C_k$ at $x_0$. By inductive assumption and Propositions 9.44, 9.42, and 3.32, $h^{(1)}|_{\bar{D}}$ is $C_k$ at $x_0$. By Proposition 9.34, $h^{(1)}$ is $C_k$ at $x_0$. Hence, $h$ is $C_{k+1}$ at $x_0$.

This completes the induction process.

When $k = \infty$, then $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $g$ is $C_1$ at $y$, $\forall y \in D_2 \cap B_\mathbb{R} (y_0, \delta)$, and $f$ is $C_1$ at $x$, $\forall x \in D_1 \cap B_\mathbb{R} (x_0, \delta), \forall i \in \mathbb{N}$, which further implies by the induction conclusion, $h$ is $C_1$ at $x$, $\forall x \in D_1 \cap B_\mathbb{R} (x_0, \delta), \forall i \in \mathbb{N}$. Hence, $h = C_\infty$ at $x_0$.

(ii) We will first use mathematical induction on $k$ to show that the result holds.

1° $k = 1$. By Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x)$, $\forall x \in D_1$. Hence, the result holds.

2° Assume that the result holds for $k \leq \tilde{k} \in \mathbb{N}$.

3° Consider the case $k = \tilde{k} + 1$. By Chain Rule, $h^{(1)}(x)$ exists and $h^{(1)}(x) = g^{(1)}(f(x))f^{(1)}(x), \forall x \in D_1$. By inductive assumption and Propositions 9.44 and 9.42, $h^{(1)}$ is $k$-times differentiable. Hence, $h$ is $(k+1)$-times differentiable. This completes the induction process and the proof of the proposition. □

**Proposition 9.46** Let $X$, $Y$, and $Z$ be normed linear spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $f : D \rightarrow Z$ 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 $(x_0, y_0)$.


9.4. HIGHER ORDER DERIVATIVES

Proof

By Proposition 9.9, \( f^{(1)}(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{bmatrix} \), \( \forall (x, y) \in D \). Define \( g : X \times Y \to B(X, X \times Y) \) by \( g(x, y) = \begin{bmatrix} \text{id}_X & \theta_{B(X, Y)} \end{bmatrix} \).

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. \( \square \)

Proposition 9.47 Let \( X_1, \ldots, X_p \), and \( Y \) be normed linear spaces over \( \mathbb{K} \), where \( p \in \{2, 3, \ldots\}, D \subseteq X := \prod_{i=1}^p X_i, f : D \to Y, x_0 \in D^\circ \), 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 \( x \), \( \forall x \in \tilde{D} := B_X(x_0, \delta_0) \subseteq D \). Then, \( f \) is \( C_k \) at \( x \), \( \forall x \in \tilde{D} \).

Proof

We will prove this using mathematical induction on \( k \).

1° \( k = 1 \). \( \forall x \in \tilde{D} \), by repeated application of Proposition 9.24, we have \( f^{(1)}(x) \) exists and \( f^{(1)}(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_p}(x) \end{bmatrix} \).

2° Assume that the result holds for \( k = \bar{k} \in \mathbb{N} \).

3° Consider the case \( k = \bar{k} + 1 \). \( \forall x \in \tilde{D} \), by the case \( k = 1 \), \( f^{(1)}(x) \) exists and \( f^{(1)}(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_p}(x) \end{bmatrix} \). \( \forall i \in \{1, \ldots, p\} \), by the assumption, all partial derivatives of \( \frac{\partial f}{\partial x_i} \) up to \( \bar{k} \)th order exist and are continuous at \( x \). Then, by inductive assumption, \( \frac{\partial f}{\partial x_i} \) is \( C_{\bar{k}} \) at \( x \). Define the function \( g : \tilde{D} \to \prod_{i=1}^p B(X_i, Y) \) by \( g(x) = (\frac{\partial f}{\partial x_i}(x), \ldots, \frac{\partial f}{\partial x_p}(x)) \), \( \forall x \in \tilde{D} := \bigcap_{i=1}^p \text{dom} \left( \frac{\partial f}{\partial x_i} \right) \). Clearly, \( \tilde{D} \subseteq \hat{D} \subseteq D \). Then, by Proposition 9.44,

\( g \) is \( \bar{k} \)-times differentiable at \( x \) and \( g^{(j)}(x) = \begin{bmatrix} D^{j} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ D^{j} \frac{\partial f}{\partial x_p}(x) \end{bmatrix}, \forall j = 1, \ldots, \bar{k} \).

By Proposition 3.32, \( g \) is \( C_{\bar{k}} \) at \( x \). By Propositions 9.45 and 9.43, \( f^{(1)} \) is \( C_{\bar{k}} \) at \( x \). Therefore, \( f \) is \( C_{\bar{k}+1} \) 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 $\bar{D} := I^o$ if $K = \mathbb{R}$ or $\bar{D} := \{ a + i0 \mid a \in I \} \subset \mathbb{C}$ and $\bar{D} := \{ a + i0 \mid a \in I^o \} \subset \mathbb{C}$ if $K = \mathbb{C}$. Let $\varphi : \bar{D} \to D$ be given by $\varphi(t) = tx_1 + (1 - t)x_0$, $\forall t \in \bar{D}$. Assume that $\text{dom}(f(n)) \supseteq \varphi(\bar{D})$, $\text{dom}(f^{(n+1)}) \supseteq \varphi(\bar{D})$, and $f^{(n)}$ is continuous at $x = \varphi(t)$, $\forall t \in \bar{D}$. Let $R_n \in Y$ be given by

$$R_n := f(x_1) - \left(f(x_0) + \frac{1}{1!} f^{(1)}(x_0)(x_1 - x_0) + \cdots + \frac{1}{n!} f^{(n)}(x_0)(x_1 - x_0)\right) \cdot \cdot \cdot (x_1 - x_0)$$

Then, the following statements hold.

(i) If $Y = \mathbb{R}$ and $K = \mathbb{R}$, then $\exists \bar{t}_0 \in I^o$ such that

$$R_n = \frac{1}{(n + 1)!} f^{(n+1)}(\varphi(\bar{t}_0)) (x_1 - x_0) \cdots (x_1 - x_0)$$

(ii) $\exists \bar{t}_0 \in I^o$ such that

$$\| R_n \| \leq \frac{1}{(n + 1)!} \| f^{(n+1)}(\varphi(\bar{t}_0)) \| \| x_1 - x_0 \|^{n+1}$$

Proof

(i) Let $Y = \mathbb{R}$ and $K = \mathbb{R}$. Define $F : I \to \mathbb{R}$ by

$$F(t) = f(\varphi(1)) - \left(f(\varphi(1 - t)) + \frac{t}{1!} f^{(1)}(\varphi(1 - t))(x_1 - x_0) + \cdots + \frac{t^n}{n!} f^{(n)}(\varphi(1 - t))(x_1 - x_0) \right) \cdot \cdot \cdot (x_1 - x_0) + R_n t^{n+1}; \quad \forall t \in I$$

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 $t$, $\forall t \in I^o$. Clearly, $F(0) = F(1) = 0$. By Mean Value Theorem 9.20, $\exists \bar{t}_0 \in I^o$ such that $0 = F(1) - F(0) = DF(t_0)$. Then, we have

$$0 = -f^{(1)}(\varphi(1 - t_0))(x_1 - x_0) + f^{(1)}(\varphi(1 - t_0))(x_1 - x_0)$$

$$-\frac{t_0}{1!} f^{(2)}(\varphi(1 - t_0))(x_1 - x_0)(x_1 - x_0) + \cdots$$

$$+ \frac{t_0^{n-1}}{(n - 1)!} f^{(n)}(\varphi(1 - t_0))(x_1 - x_0) \cdots (x_1 - x_0)$$
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,

\[R_n = \frac{1}{(n+1)!} f^{(n+1)}(\varphi(\tilde{t}_0)) (x_1-x_0) \cdots (x_1-x_0)\]

where \(\tilde{t}_0 = 1 - t_0 \in I^0\).

(ii) By Proposition 7.85, \(\exists y_0 \in \mathbb{Y}^*\) with \(\| y_0 \| \leq 1\) such that \(\| R_n \| = \langle \langle y_0, R_n \rangle \rangle\). Define \(G : \mathcal{D} \to \mathbb{K}\) by

\[G(t) = \langle \langle y_0, f(x_1) - f(\varphi(1-t)) - \sum_{i=1}^{n} \frac{t_0^i}{i!} f^{(i)}(\varphi(1-t)) (x_1-x_0) \cdots (x_1-x_0) \rangle \rangle - \| R_n \| t_0^{n+1}\]

\(\forall t \in \mathcal{D}\). 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 \(t\), \(\forall t \in \mathcal{D}\). Clearly, \(G(0) = G(1) = 0\).

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\mathbb{K} = \mathbb{R}\); Case 2: \(\mathbb{K} = \mathbb{C}\).

Case 1: \(\mathbb{K} = \mathbb{R}\). By Mean Value Theorem 9.20, \(\exists t_0 \in I^0\) such that \(0 = G(1) - G(0) = D G(t_0)\). Then, we have

\[0 = \langle \langle y_0, f^{(1)}(\varphi(1-t_0))(x_1-x_0) - \sum_{i=1}^{n} \frac{t_0^i}{(i-1)!} f^{(i)}(\varphi(1-t_0)) (x_1-x_0) \cdots (x_1-x_0) \rangle \rangle + \sum_{i=1}^{n} \frac{t_0^i}{i!} f^{(i)}(\varphi(1-t_0)) (x_1-x_0) \cdots (x_1-x_0) \rangle \rangle - (n+1) \| R_n \| t_0^n\]

\[= \langle \langle y_0, \frac{t_0^n}{n!} f^{(n+1)}(\varphi(1-t_0)) (x_1-x_0) \cdots (x_1-x_0) \rangle \rangle - (n+1) \| R_n \| t_0^n\]

Then, we have

\[\| R_n \| = \frac{1}{(n+1)!} \langle \langle y_0, f^{(n+1)}(\varphi(1-t_0)) (x_1-x_0) \cdots (x_1-x_0) \rangle \rangle\]
where the last two inequalities follows from Proposition 7. 64 and \( \bar{t}_0 \). Hence,

\[
\text{Case 2: } K = \mathbb{C}. \text{ By Lemma 9.22, } \exists t_0 \in I^o \text{ such that } \Re (G(1) - G(0)) = \Re (DG(t_0)). \text{ Then, we have}
\]

\[
0 = \Re \left( \sum_{i=1}^{n} \frac{t_0^{i-1}}{(i-1)!} f^{(i)}(1-t_0) \langle x_1 - x_0 \rangle \right) \]

\[
- \Re \left( \sum_{i=1}^{n} \frac{t_0^{i}}{i!} f^{(i)}(1-t_0) \langle x_1 - x_0 \rangle \right) \]

\[
+ \Re \left( \sum_{i=1}^{n} \frac{t_0^{i+1}}{(i+1)!} f^{(i+1)}(1-t_0) \langle x_1 - x_0 \rangle \right) \]

\[
-(n+1) \| R_n \| t_0^n
\]

Hence,

\[
\| R_n \| = \frac{1}{(n+1)!} \Re \left( \sum_{i=1}^{n} \frac{t_0^{i+1}}{(i+1)!} f^{(i+1)}(1-t_0) \langle x_1 - x_0 \rangle \right) \]

\[
\leq \frac{1}{(n+1)!} \| y_* \| \| f^{(n+1)}(1-t_0) \langle x_1 - x_0 \rangle \| \]

\[
\leq \frac{1}{(n+1)!} \| f^{(n+1)}(\bar{t}_0) \| \| x_1 - x_0 \|^{n+1}
\]

where the last two inequalities follows from Proposition 7.64 and \( \bar{t}_0 = 1 - t_0 \in I^o \).

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

### 9.5 Mapping Theorems

**Definition 9.49** Let \( X := (X, \rho) \) 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 \( \exists \alpha \in [0, 1) \subset \mathbb{R} \text{ such that } \forall x_1, x_2 \in S, \text{ we have } \rho(T(x_1), T(x_2)) \leq \alpha \rho(x_1, x_2) \).

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, $\lim_{n \to \infty} x_n = x_0$.

(iii) $\rho(x_n, x_0) \leq \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)$, $\rho(x_n, x_0) \leq \frac{\alpha}{1-\alpha} \rho(x_{n-1}, x_n)$, and $\rho(x_n, x_0) \leq \alpha \rho(x_{n-1}, x_0)$, $\forall n \in \{2, 3, \ldots\}$.

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, $\rho(x_0, \bar{x}) \leq \alpha \rho(x_0, \bar{x})$ and $\rho(x_0, \bar{x}) = 0$. Hence, $\bar{x} = x_0$. Hence the statements (i) and (ii) are true.

Note that, by Propositions 3.66 and 4.30, $\forall n \in \{2, 3, \ldots\}$,

$\rho(x_n, x_0) = \lim_{m \to \infty} \rho(x_n, x_m) \leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=n+1}^{\infty} \alpha^{i-2} \rho(x_2, x_1)$

$\rho(x_n, x_0) \leq \frac{\alpha^{n-1}}{1-\alpha} \rho(x_2, x_1)$

$\rho(x_n, x_0) \leq \sum_{i=n}^{\infty} \rho(x_{i+1}, x_i) \leq \sum_{i=1}^{\infty} \alpha^i \rho(x_n, x_{n-1}) = \frac{\alpha}{1-\alpha} \rho(x_n, x_{n-1})$

$\rho(x_n, x_0) = \rho(T(x_{n-1}), T(x_0)) \leq \alpha \rho(x_{n-1}, x_0)$

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 $D \cap B_X(x_0, \delta_0) := D$ 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 $\forall x \in D \cap B_X(x_0, \delta) =: D$, we have $\|f^{(1)}(x) - f^{(1)}(x_0)\| < \epsilon$. Define $\gamma : D \to Y$ 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 $\gamma^{(1)}(x) = f^{(1)}(x) - f^{(1)}(x_0)$, $\forall x \in D$. Clearly, $D$ is convex. Then, $\forall x_1, x_2 \in D$, by Mean Value Theorem 9.23 and Proposition 7.64, we have $\|f(x_1) - f(x_2) - f^{(1)}(x_0)(x_1 - x_2)\| = \|\gamma(x_1) - \gamma(x_2)\| \leq \|\gamma^{(1)}(t_0x_1 + (1-t_0)x_2)(x_1 - x_2)\| \leq \epsilon \|x_1 - x_2\|$, for some $t_0 \in (0, 1) \subset \mathbb{R}$. This completes the proof of the lemma. □
Theorem 9.52 (Injective Mapping Theorem) Let $X$ and $Y$ be Banach spaces over $K$, $D \subseteq X$, $F : D \to Y$ be $C_1$ at $x_0 \in D$. Assume that $F(1)(x_0) \in B(X, Y)$ is injective and $M := R(F(1)(x_0)) \subseteq Y$ is closed. Then, $\exists \delta > 0$ 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, $M \subseteq Y$ is a Banach space. Then, $F(1)(x_0) : X \to M$ is bijective. By Open Mapping Theorem 7.103, the inverse $A$ of $F(1)(x_0) : X \to M$ belongs to $B(M, X)$. Fix $h \in X$, we have $\|h\| = \|AF(1)(x_0)h\| \leq \|A\|\|F(1)(x_0)h\|$. Then, $\exists r > 0$ such that $r\|A\| \leq 1$ and $r\|h\| \leq \|F(1)(x_0)h\|$, $\forall h \in X$.

By Lemma 9.51, $\exists \delta > 0$ such that $\forall x_1, x_2 \in U$, we have $\|F(1)(x_0)(x_1 - x_2)\| \leq r\|x_1 - x_2\|$. Hence, $F(1)(x_0) : D \to M$ is surjective. Then, $F(1)(x_0)|_U : U \to F(U)$ is injective and surjective. Hence, $F(1)|_U$ is uniformly continuous. This completes the proof of the theorem.

Theorem 9.53 (Surjective Mapping Theorem) Let $X$ and $Y$ be Banach spaces over $K$, $D \subseteq X$, $F : D \to Y$, $x_0 \in D$, and $y_0 := F(x_0) \in Y$. Assume that $F$ is $C_1$ at $x_0$, and $F(1)(x_0) \in B(X, Y)$ is surjective. Then, $\exists r > 0$ such that $\forall x \in B_X(x_0, r)$ 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 $\bar{y} = F(\bar{x})$, $\forall y \in B_Y(y_0, \delta/2)$, $\exists x \in B_X(x_0, r)$ with $\|x - \bar{x}\| \leq c_1 \|y - \bar{y}\|$, we have $y = F(x)$.

Proof Let $M := N(F(1)(x_0))$, which is a closed subspace by Proposition 7.45, the quotient space $X/M$ is a Banach space. By Proposition 7.70, $F(1)(x_0) = A \circ \phi$, where $\phi : X \to X/M$ is the natural homomorphism, and $A \in B(X/M, Y)$ is injective. Since $F(1)(x_0)$ is surjective, $A$ is injective and, by Open Mapping Theorem 7.103, $A^{-1} \in B(Y, X/M)$. Let $c_1 := 4\|A^{-1}\| \in (0, \infty) \subseteq R$.

Define $\gamma : D \to X/M$ 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 $\gamma(1)(x) = A^{-1}(F(1)(x) - F(1)(x_0))$, $\forall x \in B_X(x_0, r_0)$ with $r_0 \in (0, \infty) \subseteq R$. Clearly, $\gamma(1)(x_0) = 0$. Then, by Lemma 9.51, $\exists r > 0 \in R$ such that $\forall x_1, x_2 \in B_X(x_0, r)$ with $r \in (0, \infty)$, we have $\|\gamma(x_1) - \gamma(x_2)\| \leq c_1 \|x_1 - x_2\|/4$.

Fix $\delta > 0$ such that $c_1 \delta \leq r$. Fix any $\bar{y} \in B_Y(y_0, \delta/2)$ and any $\bar{x} \in B_X(x_0, r/2)$ with $\bar{y} = F(\bar{x})$. Fix any $y \in B_Y(y_0, \delta/2)$. Recursively define $x_1 := \bar{x}$, $x_{k+1} = x_k + A^{-1}(y - F(x_k))$ and select $x_{k+1} \in [x_k, x_k + A^{-1}(y - F(x_k))]$ such that $\|x_{k+1} - x_k\| \leq 2\|\bar{x} - x_k\|$, $\forall k \in \mathbb{N}$, 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
9.5. MAPPING THEOREMS

\[ \|x_2 - x_1\| \leq 2 \|x_2 - x_1\| = 2 \|A^{-1}(y - \tilde{y})\| \leq 2 \|A^{-1}\| \|y - \tilde{y}\| = c_1 \|y - \tilde{y}\| / 2 < r/4, \quad \text{where the second inequality follows from Proposition 7.64.} \]

Then, \(x_2 \in \mathcal{B}_\chi(x_0, r)\). Assume that \(x_1, \ldots, x_k \in \mathcal{B}_\chi(x_0, r)\) for some \(k \in \{2, 3, \ldots\}\). Note that \(\forall i \in \{2, \ldots, k\}, \|x_{i+1} - x_i\| \leq 2 \|(x_{i+1} - x_i)\| / 2 < r/2\). This implies that \(x_{k+1} \in \mathcal{B}_\chi(x_0, r)\). Inductively, we have \((x_k)_{k=0}^\infty \subseteq \mathcal{B}_\chi(x_0, r)\) and \(\|x_{k+1} - x_k\| \leq \|x_{k+1} - x_k\| / 2^k < r/2^k, \forall k \in \mathbb{N}\).

Hence, \(\|x_{k+1} - \tilde{x}\| \leq \sum_{i=1}^k \|x_{i+1} - x_i\| \leq \sum_{i=1}^k \|x_{i+1} - x_i\| / 2^k = (2 - 1/2^k) \|x_{k} - x_1\| < r/2\). This completes the proof of the theorem. \(\square\)

**Theorem 9.54 (Open Mapping Theorem)** Let \(\mathcal{X}\) and \(\mathcal{Y}\) be Banach spaces over \(\mathbb{K}\), \(D \subseteq \mathcal{X}\) be open, and \(F : D \to \mathcal{Y}\) be \(C_1\). Assume that \(F^{(1)}(x) \in B(\mathcal{X}, \mathcal{Y})\) 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 \(\mathcal{X}\). We will show that \(F(U)\) is an open set in \(\mathcal{Y}\). Fix any \(y_0 \in F(U)\), there exists \(x_0 \in U\) such that \(y_0 = F(x_0)\). Then, \(\exists \tilde{r} \in (0, \infty) \subseteq \mathbb{R}\) such that \(\mathcal{B}_\chi(x_0, r) \subseteq U\). It is easy to check that all assumptions of Surjective Mapping Theorem are satisfied at \(x_0\). Then, there exist an open set \(V \subseteq \mathcal{Y}\) with \(y_0 \in V \subseteq V\) and \(c_1 \in (0, \infty) \subseteq \mathbb{R}\) such that \(\forall y \in V, \exists \tilde{x} \in D\) with \(\|\tilde{x} - x_0\| \leq c_1 \|y - y_0\|\), we have \(\bar{y} = F(\tilde{x})\). Take \(\delta \in (0, \infty) \subseteq \mathbb{R}\) such that \(c_1 \delta \leq r\) and \(\mathcal{B}_y(y_0, \delta) \subseteq \mathcal{Y}\). Then, \(\forall y \in \mathcal{B}_y(y_0, \delta), \|\bar{y} - x_0\| \leq c_1 \|\bar{y} - y_0\| < r\). Then, \(\tilde{x} \in \mathcal{B}_\chi(x_0, r) \subseteq U\) and \(\bar{y} \in F(U)\). Hence, \(\mathcal{B}_y(y_0, \delta) \subseteq F(U)\). Therefore, \(y_0 \in (F(U)^o)\). By the arbitrariness of \(y_0\), we have \(F(U)\) is open in \(\mathcal{Y}\). By the arbitrariness of \(U, F\) is an open mapping. This completes the proof of the theorem. \(\square\)

**Proposition 9.55** Let \(\mathcal{X}\) and \(\mathcal{Y}\) be Banach spaces over \(\mathbb{K}\) and \(A \in B(\mathcal{X}, \mathcal{Y})\) be bijective. Then, \(VT \in B(\mathcal{X}, \mathcal{Y})\) with \(\|T - A\| \|A^{-1}\| < 1\), we have \(T\) is bijective and \(\|T^{-1} - A^{-1}\| \leq \|A^{-1}\|^2 \|T - A\| < 1\).

**Proof** By Open Mapping Theorem 7.103, \(A^{-1} \in B(\mathcal{Y}, \mathcal{X})\). We will first prove the result for the special case \(\mathcal{Y} = \mathcal{X}\) and \(A = id_\mathcal{X}\). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \(\mathcal{X}\) is a singleton set; Case 2: \(\exists \tilde{x} \in \mathcal{X}\) such that \(\tilde{x} \neq 0_\mathcal{X}\).

Case 1: \(\mathcal{X}\) is a singleton set. Then, \(\|A^{-1}\| = \|A\| = 0\). \(\forall T \in B(\mathcal{X}, \mathcal{X})\), we have \(T = id_\mathcal{X}\). Then, \(T\) is bijective and \(\|T^{-1} - id_\mathcal{X}\| = 0\). The result holds for this case.
Case 2: \( \exists \bar{x} \in X \) such that \( \bar{x} \neq \varnothing \). Then, \( \|A\| = \|A^{-1}\| = 1 \). \( \forall T \in B(\bar{X}, X) \) 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 \( \phi : X \to X \) 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, \( T^{-1} \in B(\bar{X}, X) \).

\[ \forall y \in Y, \text{let } x = T^{-1}y. \text{ Then, } y = Tx = x + \Delta x \text{ and } x = y - \Delta x. \] By Proposition 7.64, we have \( \|x\| \leq \|y\| + \|\Delta\| \|x\| \) and \( \|x\| \leq \frac{\|x\|}{1 - \|\Delta\|} \). 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. \( \forall T \in B(\bar{X}, Y) \) with \( \|T - A\| \|A^{-1}\| < 1 \), we have \( \bar{T} := A^{-1}T \in B(\bar{X}, \bar{X}) \). Note that \( \|\bar{T} - \text{id}_{\bar{X}}\| = \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\| < 1 \). Then, \( \|\bar{T} - \text{id}_{\bar{X}}\| \|\text{id}_X\| < 1 \). By the special case, we have \( \bar{T} \) is bijective and \( \|\bar{T}^{-1} - \text{id}_X\| \leq \frac{\|\text{id}_X\|^2 \|\bar{T} - \text{id}_{\bar{X}}\|}{1 - \|\text{id}_X\| \|\bar{T} - \text{id}_{\bar{X}}\|} \). Then, \( T \) is bijective and, by Proposition 7.64,

\[
\|\bar{T} - \text{id}_X\| = \|A^{-1}(T - A)\| \leq \|A^{-1}\| \|T - A\|
\|A^{-1}\| \|\text{id}_X\| = \|A^{-1}\|
\|T^{-1} - A^{-1}\| = \|(\bar{T}^{-1} - \text{id}_{\bar{X}})A^{-1}\| \leq \|A^{-1}\| \frac{\|\text{id}_X\|^2 \|\bar{T} - \text{id}_{\bar{X}}\|}{1 - \|\text{id}_X\| \|\bar{T} - \text{id}_{\bar{X}}\|}
\leq \frac{\|A^{-1}\|^2 \|T - A\|}{1 - \|A^{-1}\| \|T - A\|}
\]

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

**Proposition 9.56** Let \( X \) and \( Y \) be Banach spaces over \( K \), \( D := \{ L \in B(\bar{X}, Y) \mid L \text{ is bijective} \} \), and \( f : D \to B(Y, \bar{X}) \) be given by \( f(A) = A^{-1}, \forall A \in D \). Then, \( D \) is open in \( B(\bar{X}, Y) \), \( f \) is \( C_\infty \), and \( f^{(1)}(A)(\Delta) = -A^{-1} \Delta A^{-1}, \forall A \in D, \forall \Delta \in B(\bar{X}, Y) \).

**Proof** By Proposition 9.55 and Open Mapping Theorem 7.103, \( D \) is open and \( f \) is continuous. \( \forall A \in D, \text{span}(A_D(A)) = B(\bar{X}, Y) \) since \( D \) is open and \( A \in D^e \). Define \( L : B(\bar{X}, Y) \to B(Y, \bar{X}) \) by \( L(\Delta) = -A^{-1} \Delta A^{-1}, \forall \Delta \in B(\bar{X}, Y) \). Clearly, \( L \) is a linear operator. Note that

\[
\|L\| = \sup_{\Delta \in B(\bar{X}, Y), \|\Delta\| \leq 1} \|L(\Delta)\| \leq \|A^{-1}\|^2
\]
9.5. MAPPING THEOREMS

where the inequality follows from Proposition 7.64. Hence, \( L \) is a bounded linear operator.

\[
\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \text{ by the continuity of } f, \ \exists \delta \in (0, \infty) \subset \mathbb{R} \text{ such that } \\
\| f(\tilde{A}) - f(A) \| < \epsilon, \ \forall \tilde{A} \in B_{B(X,Y)}(A, \delta). \text{ Then, } \forall \tilde{A} \in B_{B(X,Y)}(A, \delta), \text{ we have} \\
\| f(\tilde{A}) - f(A) - L(\tilde{A} - A) \| = \| \tilde{A}^{-1} - A^{-1} + A^{-1}(\tilde{A} - A)A^{-1} \| \\
= \| -A^{-1}(\tilde{A} - A)\tilde{A}^{-1} + A^{-1}(\tilde{A} - A)A^{-1} \| \\
\leq \| A^{-1} \| \| \tilde{A} - A \| \| f(\tilde{A}) - f(A) \| \leq \epsilon \| A^{-1} \| \| \tilde{A} - A \|
\]

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, 9.7, 3.12, and 3.32, \( f^{(1)} \) 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 \( C_{k+1} \). Then, \( f \) is \( C_{\infty} \). Note that \( f^{(1)}(A) = -f(A)ro(f(A)), \forall A \in D \). By Propositions 9.45, 9.42, and 9.44, \( f^{(1)} \) is \( C_k \). Then, \( f \) is \( C_{k+1} \). This completes the proof of the proposition.

\[\square\]

**Theorem 9.57 (Inverse Function Theorem)** Let \( X \) and \( Y \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq X \), \( F : D \rightarrow Y \) be \( C_1 \) at \( x_0 \in D \). Assume that \( F^{(1)}(x_0) \in B(Y,X) \). Then \( F \) is differentiable, and \( F^{(1)} \) is continuous at \( y_0 \) such that

(i) \( F|_U : U \rightarrow V \) is bijective;

(ii) the inverse mapping \( F_i : V \rightarrow U \) of \( F|_U \) is differentiable, \( F^{(1)}_i : V \rightarrow B(Y,X) \) is given by \( F^{(1)}_i(y) = (F^{(1)}(F_i(y)))^{-1}, \forall y \in V \), and \( F^{(1)}_i \) 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, \( (F^{(1)}(x_0))^{-1} \in B(Y,X) \).

By \( F \) being \( C_1 \) at \( x_0 \in D \), \( \exists \tilde{r} \in (0, \infty) \subset \mathbb{R} \) such that \( F^{(1)}(x) \) exists, \( \forall x \in B_X(x_0, \tilde{r}) \subseteq D \). Define \( T : B_X(\partial_X, \tilde{r}) \rightarrow X \) by \( T(x) = (F^{(1)}(x_0))^{-1}(F(x) + x_0 - y_0) \), \( \forall x \in B_X(\partial_X, \tilde{r}) \subseteq D - x_0 \). Clearly, \( T(\partial_X) = \partial_X \). By Propositions 9.45, 9.38, 9.34, and 9.44, \( T \) is \( C_1 \) at \( \partial_X \). \( T \) is differentiable, and \( T^{(1)}(x) = (F^{(1)}(x_0))^{-1}F^{(1)}(x) + x_0 \), \( \forall x \in B_X(\partial_X, \tilde{r}) \). Clearly, \( T^{(1)}(\partial_X) = id_X \).

Define \( \psi : B_X(\partial_X, \tilde{r}) \rightarrow X \) by \( \psi(x) = T(x) - x \). Then, by Propositions 9.38, 9.45, and 9.44, \( \psi \) is differentiable, \( \psi \) is \( C_1 \) at \( \partial_X \), and \( \psi^{(1)}(x) = T^{(1)}(x) - id_X, \forall x \in B_X(\partial_X, \tilde{r}) \). Clearly, \( \psi(\partial_X) = \partial_X \) and \( \psi^{(1)}(\partial_X) = \partial_{B(X,X)} \). Fix any \( \alpha \in (0,1) \subset \mathbb{R} \). Then, \( \exists r_1 \in (0, \tilde{r}) \subset \mathbb{R} \).
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, $\|\psi(x_1) - \psi(x_2)\| \leq \sup_{t \in (0, 1)} \|\psi^{(1)}(t_0 x_1 + (1 - t_0)x_2)(x_1 - x_2)\| \leq \alpha \|x_1 - x_2\|$, 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, $\phi : \overline{B}_X(\vartheta_X, r_1) \to \overline{B}_X(\vartheta_X, r_1) \subseteq \overline{B}_X(\vartheta_X, r_1)$.  

It is easy to see that $\phi$ is a contraction mapping with contraction index $\alpha$.  

By Contraction Mapping Theorem, $\exists \bar{x} \in \overline{B}_X(\vartheta_X, r_1)$ such that $\bar{x} = \phi(\bar{x}) \in \overline{B}_X(\vartheta_X, r_1)$, which is equivalent to $\bar{x} = T(\bar{x})$.  

Let $\bar{U} := T_{\overline{B}_X}(\vartheta_X, (1 - \alpha)r_1) \cap \overline{B}_X(\vartheta_X, r_1)$ and $\bar{V} := \overline{B}_X(\vartheta_X, (1 - \alpha)r_1)$.  

Note that $\bar{U}$ and $\bar{V}$ are open sets in $X$ since $T$ is continuous by Proposition 9.7.  

Then, $T|_{\bar{U}} : \bar{U} \to \bar{V}$ is bijective.  

Since $T(\vartheta_X) = \vartheta_X$, then $\vartheta_X \in \bar{U}$.

Hence, there exists an inverse mapping $T_i : \bar{V} \to \bar{U}$ of $T|_{\bar{U}}$.  

Let $U := \bar{U} + x_0$ and $V := F^{(1)}(x_0)(\bar{V}) + y_0$.  

Clearly, $U$ and $V$ are open sets in $X$ and $Y$, respectively.  

Note that $F|_U(x) = F^{(1)}(x_0)T|_{\bar{U}}(x - x_0) + y_0$, $\forall x \in U$.  

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 $F_i(y) = F_i((F^{(1)}(x_0))^{-1}(y - y_0)) + x_0$, $\forall y \in V$.  

Hence, the statement (i) holds.

Next, we will show that $F_i : V \to U$ is differentiable.  

Note that $\forall x \in \bar{V}$, $x = T(T_i(x)) = T_i(x) + \psi(T_i(x))$ and $T_i(x) = x - \psi(T_i(x))$.  

$\forall x_1, x_2 \in \bar{V}$, we have $T_i(x_1), T_i(x_2) \in U \subseteq \overline{B}_X(\vartheta_X, r_1)$ 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** \(\forall x \in U, \text{let } y = F(x). \text{ Then, } F_i \text{ is differentiable at } F(x) \text{ and } F_i^{(1)}(y) = (F^{(1)}(x))^{-1}\).

**Proof of claim:** Fix any $x \in U$ and let $y = F(x)$.  

Then, $x - x_0 \in \bar{U} \subseteq \overline{B}_X(\vartheta_X, r_1)$ and $\|\psi^{(1)}(x - x_0)\| \leq \alpha$.  

Note that $T = \text{id}_X + \psi$ and $T^{(1)}(x - x_0) = \text{id}_X + \psi^{(1)}(x - x_0)$.  

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 $c_1 := \|F^{(1)}(x)\|^{-1} < \infty$.

$\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$ with $\epsilon c_1 < 1$, by the differentiability of $F$ at $x$, $\exists \delta_1 \in (0, \infty) \subseteq \mathbb{R}$ 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 $F_i$, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $\forall u \in Y$ with $\|u\| < \delta$, we have $y + u \in V$ and $\|F_i(y + u) - F_i(y)\| < \delta_1$.  


\( \forall u \in Y \) with \( \|u\| < \delta \), let \( \beta := \|F_i(y + u) - F_i(y) - (F_i^{(1)}(x))^{-1}u\| \geq 0 \). Let \( h := F_i(y + u) - x \in X \). Then, \( \|h\| = \|F_i(y + u) - F_i(y)\| < \delta_1 \). Note that

\[
\beta = \|x + h - (F_i^{(1)}(x))^{-1}u\| = \|(F_i^{(1)}(x))^{-1}(u - (F_i^{(1)}(x))h)\| \\
\leq \| (F_i^{(1)}(x))^{-1} \| \| u - (F_i^{(1)}(x))h \| \\
= c_1 \| F(x + h) - F(x) - (F_i^{(1)}(x))h \| \\
\leq c_1 \| F(y + u) - F(y) \| \\
\leq c_1 \| F(y + u) - F(y) - (F_i^{(1)}(x))^{-1}u\| + \|(F_i^{(1)}(x))^{-1}u\| \\
\leq c_1 (\beta + c_1 \|u\|)
\]

Then, \( \beta \leq \frac{c_1^2 \|u\|}{\|F_i^{(1)}(x)\|} \). Hence, \( F_i \) is differentiable at \( y \) and \( F_i^{(1)}(y) = (F_i^{(1)}(x))^{-1} \). This completes the proof of the claim. \( \square \)

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 \( k \)-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 \( \mathcal{C}_{\bar{k}} \) at \( x_0 \). By the inductive assumption, \( F_i \) is \( \mathcal{C}_{\bar{k}} \) at \( y_0 \). By (ii) and Propositions 9.45 and 9.56, \( F_i^{(1)} \) is \( \mathcal{C}_k \) at \( y_0 \). Hence, \( F_i \) is \( \mathcal{C}_{\bar{k} + 1} \) at \( y_0 \). This completes the induction process.

When \( k = 0 \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( F \) is \( \mathcal{C}_0 \) at \( x \), \( \forall x \in \mathcal{B}_Y(x_0, \delta) \subseteq D \), \( \forall i \in \mathbb{N} \). By taking \( r_1 < \delta \) in the proof of (i), we have \( U \subseteq \mathcal{B}_X(x_0, \delta) \). Then, by the induction conclusion, \( F_i \) is \( \mathcal{C}_1 \) at \( y \), \( \forall y \in F(U) = V \), \( \forall i \in \mathbb{N} \). Hence, \( F_i \) is \( \mathcal{C}_\infty \) at \( y_0 \).

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

**Theorem 9.58 (Implicit Function Theorem)** Let \( \mathcal{X} := (X, O) \) be a topological space, \( Y \) and \( Z \) be Banach spaces over \( \mathbb{K} \), \( D \subseteq \mathcal{X} \times Y \), \( F : D \to Z \) be continuous. Assume that \( F \) is partial differentiable with respect to \( y \) and \( \frac{\partial F}{\partial y} \) is continuous at \( (x_0, y_0) \in D \), \( F(x_0, y_0) = \partial_Z \), and \( \frac{\partial F}{\partial y}(x_0, y_0) \in \mathcal{B}(Y, Z) \) is bijective. Then, the following statements hold.

(i) There exist an open set \( U_0 \in O \) with \( x_0 \in U_0 \) and \( r_1 \in (0, \infty) \subset \mathbb{R} \) such that \( U_0 \times \mathcal{B}_y(y_0, r_1) \subseteq D \) and \( \forall x \in U_0 \), \( \exists \, y \in \mathcal{B}_y(y_0, r_1) \)
satisfying $F(x, y) = \partial \Xi$. This defines a function $\phi : U_0 \to \mathcal{B}_y (y_0, r_1)$ by $\phi(x) = y, \forall x \in U_0$. Then, $\phi$ is continuous.

(ii) $\forall (x, y) \in U_0 \times \mathcal{B}_y (y_0, r_1), \frac{\partial F}{\partial y}(x, y)$ is bijective and $\left\| \left( \frac{\partial F}{\partial y}(x, y) \right)^{-1} \right\| < +\infty$.

**Proof** By Open Mapping Theorem 7.103, $\left( \frac{\partial F}{\partial y}(x, y_0) \right)^{-1} \in \mathcal{B}_z(y)$. Define a mapping $\psi : D \to \mathcal{Y}$ by

$$
\psi(x, y) = y - \left( \frac{\partial F}{\partial y}(x, y_0) \right)^{-1} F(x, y); \quad \forall (x, y) \in D
$$

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

$$
\frac{\partial \psi}{\partial y}(x, y) = \text{id}_y - \left( \frac{\partial F}{\partial y}(x, y_0) \right)^{-1} \frac{\partial F}{\partial y}(x, y); \quad \forall (x, y) \in D
$$

By the continuity of $\frac{\partial F}{\partial y}$ at $(x_0, y_0)$, then $\frac{\partial \psi}{\partial y}$ is continuous at $(x_0, y_0) \in D^o$ and $\frac{\partial \psi}{\partial y}(x_0, y_0) = \partial \mathcal{B}(y, y)$. 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 $U_1 \times \mathcal{B}_y (y_0, r_1) \subseteq D$ and $\left\| \frac{\partial \psi}{\partial y}(x, y) \right\| \leq \alpha, \forall (x, y) \in U_1 \times \mathcal{B}_y (y_0, r_1)$. By the continuity of $\psi$, $\exists U_0 \in \mathcal{O}$ with $x_0 \in U_0 \subseteq U_1$ such that $\left\| \psi(x, y_0) - y_0 \right\| = \left\| \psi(x, y_0) - \psi(x_0, y_0) \right\| < (1 - \alpha)r_1, \forall x \in U_0$.

Fix any $x \in U_0$, define mapping $\gamma_x : \mathcal{B}_y (y_0, r_1) \to \mathcal{Y}$ by $\gamma_x(y) = \psi(x, y), \forall y \in \mathcal{B}_y (y_0, r_1)$. We will show that $\gamma_x$ is a contraction mapping with contraction index $\alpha$. $\forall y \in \mathcal{B}_y (y_0, r_1)$, we have

$$
\left\| \gamma_x(y) - y_0 \right\| = \left\| \psi(x, y) - \psi(x_0, y_0) \right\|
\leq \left\| \psi(x, y) - \psi(x, y_0) \right\| + \left\| \psi(x_0, y_0) - \psi(x_0, y_0) \right\|
< \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0)(y - y_0) \right\| + (1 - \alpha)r_1
\leq \sup_{t \in (0, 1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x, ty + (1 - t)y_0) \right\| \left\| y - y_0 \right\| + (1 - \alpha)r_1
\leq \alpha r_1 + (1 - \alpha)r_1 = r_1
$$

where the second inequality follows from Mean Value Theorem and the third inequality follows from Proposition 7.64. Then, $\gamma_x : \mathcal{B}_y (y_0, r_1) \to \mathcal{B}_y (y_0, r_1) \subseteq \mathcal{B}_y (y_0, r_1)$. $\forall y_1, y_2 \in \mathcal{B}_y (y_0, r_1)$, we have

$$
\left\| \gamma_x(y_1) - \gamma_x(y_2) \right\| = \left\| \psi(x, y_1) - \psi(x, y_2) \right\|
$$
9.5. MAPPING THEOREMS

\[
\begin{align*}
&\leq \sup_{t \in (0,1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x,ty_1 + (1-t)y_2)(y_1 - y_2) \right\| \\
&\leq \sup_{t \in (0,1) \subset \mathbb{R}} \left\| \frac{\partial \psi}{\partial y}(x,ty_1 + (1-t)y_2) \right\| \|y_1 - y_2\| \\
&\leq \alpha \|y_1 - y_2\|
\end{align*}
\]

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 $y = \gamma_x(y) \in \mathcal{B}_y(y_0, r_1)$, $y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, 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 $\|\gamma_{x,n}(y_0) - y\| \leq \frac{\alpha^n}{1-\alpha} \|\gamma_{x}(y_0) - y\| < r_1\alpha^{-1}$, \( \forall n \in \{2, 3, \ldots \} \). By $\gamma_x(y) = y$, we conclude that $F(x, y) = \vartheta_x$. Hence, $\forall x \in U_0$, \( \exists y \in \mathcal{B}_y(y_0, r_1) \) such that $F(x, y) = \vartheta_x$, since $F(x, y) = \vartheta_x \Leftrightarrow y = \gamma_x(y)$. Then, we may define $\phi : U_0 \to \mathcal{B}_y(y_0, r_1)$ by $\phi(x) = y = \lim_{n \in \mathbb{N}} \gamma_{x,n}(y_0)$, \( \forall x \in U_0 \). Hence, $F(x, \phi(x)) = \vartheta_x$, \( \forall x \in U_0 \).

Next, we show that $\phi$ is continuous. Fix any $\bar{x} \in U_0$, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \eta_0 \in \mathbb{N} \) with $n_0 > 1$ such that $\alpha^{n_0-1} < \epsilon/3$. 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,

\[
\|\phi(x_1) - \phi(\bar{x})\| \leq \|\phi(x_1) - \gamma_{x_1,n_0}(y_0)\| + \|\gamma_{x_1,n_0}(y_0) - \gamma_{x,n_0}(y_0)\| \\
+ \|\gamma_{x,n_0}(y_0) - \phi(\bar{x})\| \leq r_1\alpha^{-1} + \epsilon/3 + r_1\alpha^{-1} < \epsilon
\]

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 $\left\| \frac{\partial F}{\partial y}(x, y) \right\| \leq \alpha$.

By Proposition 9.55, $\text{id}_Y - \frac{\partial \psi}{\partial y}(x, y)$ is bijective with continuous inverse.

Note that $\text{id}_Y - \frac{\partial \psi}{\partial y}(x, y) = \left( \frac{\partial F}{\partial y}(x_0, y_0) \right)^{-1} \frac{\partial F}{\partial y}(x, y)$. Therefore, $\frac{\partial F}{\partial y}(x, y)$ is bijective with continuous inverse.

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

**Theorem 9.59 (Implicit Function Theorem)** Let $X$ be normed linear space over $\mathbb{K}$, $Y$ and $Z$ be Banach spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $F : D \to Z$ 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 $\mathcal{B}_X(x_0, r_0) \times \mathcal{B}_Y(y_0, r_1) \subseteq D$ and $\forall x \in \mathcal{B}_X(x_0, r_0)$, \( \exists ! y \in \mathcal{B}_Y(y_0, r_1) \) satisfying $F(x, y) = \vartheta_z$.

This defines a function $\phi : \mathcal{B}_X(x_0, r_0) \to \mathcal{B}_Y(y_0, r_1)$ by $\phi(x) = y$, $\forall x \in \mathcal{B}_X(x_0, r_0)$. Then, $\phi$ is continuous.
Further, by the Implicit Function Theorem 9.58, the statement (i) holds. Proof

(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 $\bar{x} \in B_{(x_0,0)}$. 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, $\forall (x,y) \in B_{\mathbb{X}}(x_0,0) \times B_{\mathbb{Y}}(y_0,1)$, $\frac{\partial F}{\partial y}(x,y)$ is bijective and

$$
\left\| \left( \frac{\partial F}{\partial y}(x,y) \right)^{-1} \right\| < +\infty.
$$

(ii) Fix some $x \in B_{\mathbb{X}}(x_0,0)$ such that $F$ is differentiable at $(x,\phi(x))$. Let $y := \phi(x) \in B_{\mathbb{Y}}(y_0,1)$. 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 $L := \left[ \frac{\partial F}{\partial x}(x,y) \right]^{-1} \frac{\partial F}{\partial x}(x,\phi(x)) \in B(\mathbb{X}, \mathbb{Y})$ 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 $(x,y)$, $\exists h,y \in (0,\infty) \subset \mathbb{R}$ such that $\forall (h,k) \in B_{\mathbb{X} \times \mathbb{Y}}((x,\phi(x)), \delta_1)$, we have $(x+h, y+k) \in D$ and

$$
\left\| F(x+h, y+k) - F(x,y) - \frac{\partial F}{\partial x}(x,y)h - \frac{\partial F}{\partial y}(x,y)k \right\| \leq \epsilon \| (h,k) \|.
$$

By the continuity of $\phi$ at $x$, $\exists h,y \in (0,\infty)$ such that $\forall (h,k) \in B_{\mathbb{X} \times \mathbb{Y}}((x,\phi(x)), \delta_1)$, we have

$$
\| \phi(x+h) - \phi(x) \| \leq \| (h,k) \| < 1.
$$

Then, by Proposition 7.64,

$$
\beta = \left\| \left( \frac{\partial F}{\partial y}(x,y) \right)^{-1} \left( \frac{\partial F}{\partial y}(x,y)(\phi(x+h) - \phi(x)) + \frac{\partial F}{\partial x}(x,y)h \right) \right\|
$$

$$
\leq c_1 \| F(x+h, \phi(x+h)) - F(x,y) - \frac{\partial F}{\partial x}(x,y)h
$$

$$
- \frac{\partial F}{\partial y}(x,y)(\phi(x+h) - \phi(x)) \| \leq c_1 \epsilon \| (h,k) \| + \| \phi(x+h) - \phi(x) \|
$$

$$
\leq c_1 \epsilon (1+\|L\|) \| (h,k) \|
$$

Then, we have $\beta \leq c_1 \epsilon (1+\|L\|) \| h \|$. 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 $\mathcal{X} := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ and $\mathcal{Z}$ be
normed linear spaces over $\mathbb{K}$, $D \subseteq X \times Y$, $f : D \to Z$ be partial differentiable with respect to $y$ and $\partial_2 f(x,y) = \partial_{B(y,z)}$, $\forall (x,y) \in D$. 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 $\phi : D \to Z$ such that $f(x,y) = \phi(x)$, $\forall (x,y) \in D$. Furthermore, the following statements hold.

(i) If $f$ is continuous at $(x_0, y_0) \in D^\circ$, then $\phi$ is continuous at $x_0$.

(ii) If $X$ is a normed linear space $X$ over $\mathbb{K}$ and $f$ is $C_k$ at $(x_0, y_0) \in D^\circ$, where $k \in \mathbb{N} \cup \{ \infty \}$, then $\phi$ is $C_k$ at $x_0$.

**Proof**

$\forall x \in D = \pi_X(D)$, $D_x \neq \emptyset$. By Axiom of Choice, $\exists g : D \to Y$ such that $g(x) \in D_x$, $\forall x \in D$. Define $\phi : D \to \mathbb{X}$ 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, $\exists \epsilon \in (0, 1) \subset \mathbb{R}$ such that $\|f(x,y) - \phi(x)\| = \|f(x,y) - f(x,g(x))\| \leq \|\frac{\partial f}{\partial y}(x,t_0y + (1-t_0)g(x))(y - g(x))\| = 0$. Hence, $f(x,y) = \phi(x)$.

(i) Let $f$ be continuous at $(x_0, y_0) \in D^\circ$. Then, $\exists U \subseteq \mathbb{X} \cap U$ 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 $X$ over $\mathbb{K}$ and $f$ be $C_k$ at $(x_0, y_0) \in D^\circ$, 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.

\[ \square \]

### 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 $\sigma : \mathcal{Z} \to \mathcal{X}$. 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 $\forall x \in \mathcal{X}$, $\exists U \subseteq \mathcal{X}$ with $x \in U$ and $\exists V \subseteq \mathcal{Y}$ with $F(x) \in V$ such that $F\lvert_U : U \to V$ 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.
9.6. GLOBAL INVERSE FUNCTION THEOREM

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 $S := \{ s \in [a, b] \subset \mathbb{R} \mid \theta_1(t) = \theta_2(t) \forall t \in [a, s] \subset \mathbb{R} \}$ 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 $a \leq t < \xi$. There exists $(t_n)_{n=1}^\infty \subseteq S$ such that $\lim_{n \to \infty} t_n = \xi$.

By Proposition 3.66 and the continuity of $\theta_1$ and $\theta_2$, we have $\theta_1(\xi) = \lim_{n \to \infty} \theta_1(t_n) = \lim_{n \to \infty} \theta_2(t_n) = \theta_2(\xi)$, 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, $\exists U \in \mathcal{O}_x$ 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. By the continuity of $\theta_1$ and $\theta_2$, $\exists \xi \in (\xi, b)$ such that $\theta_1(t), \theta_2(t) \in U$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Then, $\sigma(t) = F(\theta_1(t)) = F(\theta_2(t)) \in V$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Since $F|_U : U \to V$ is a homeomorphism, then $\theta_1(t) = \theta_2(t)$, $\forall t \in [\xi, \xi] \subset \mathbb{R}$. Then, $\xi \in S$ and $\xi < \xi \leq \sup S = \xi$. 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 $S := \{ s \in [a, b] \subset \mathbb{R} \mid$ there exists a continuous $\theta : [a, s] \to \mathcal{X}$ that inverts $F$ along $\sigma|_{[a, s]}$ 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 $\xi \notin S$. Then, $a < \xi \leq b$ and $\exists (t_n)_{n=1}^\infty \subseteq S$, which is nondecreasing, such that $\lim_{n \to \infty} t_n = \xi$. $\forall n \in \mathbb{N}$, there exists a continuous $\theta_n : [a, t_n] \to \mathcal{X}$ that inverts $F$ along $\sigma|_{[a, t_n]}$ with $\theta_n(a) = x_0$. By the uniqueness property that we have shown, we have $\theta_n = \theta_{n+1}|_{[a, t_n]}$. Hence, we may define $\theta : [a, \xi] \to \mathcal{X}$ such that $\theta(t) = \theta_n(t)$, $\forall t \in [a, t_n] \subset \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, $\theta$ is continuous and inverts $F$ along $\sigma|_{[a, \xi]}$ with $\theta(a) = x_0$. Note that $\sigma(t_n) = F(\theta(t_n))$, $\forall n \in \mathbb{N}$. By continuity of $\sigma$ and Proposition 3.66, we have $\lim_{n \to \infty} \sigma(t_n) = \sigma(\xi) \in \mathcal{Y}$, 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, $(F(\theta(t_n)))_{n=1}^\infty = (\sigma(t_n))_{n=1}^\infty$ admits a cluster point $F(x)$. Since $\mathcal{Y}$ is Hausdorff, then $F(x) = \sigma(\xi) = : y$.

By the assumption of the lemma, $\exists U \in \mathcal{O}_x$ 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 $\sigma_{\text{inv}}(V)$ is open in $[a, b] \subset \mathbb{R}$. Since $\xi \in \sigma_{\text{inv}}(V)$, then $\exists \xi \in [a, \xi] \subset \mathbb{R}$ such that $[\xi, \xi] \subseteq \sigma_{\text{inv}}(V)$. By $\lim_{n \to \infty} t_n = \xi$, $\exists N \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n \geq N$, $t_n \in [\xi, \xi] \subset \mathbb{R}$. 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 $\theta(t_{n_0}) \in U$. Clearly, $\sigma([t_{n_0}, \xi]) \subseteq V$. Define $\theta_1 := (F|_U)_{\text{inv}} \circ \sigma|_{[t_{n_0}, \xi]} : [t_{n_0}, \xi] \to U$. Clearly, $\theta_1$ is continuous and inverts $F$ along $\sigma|_{[t_{n_0}, \xi]}$ with $\theta_1(t_{n_0}) = \theta(t_{n_0})$. Define
We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( a = b \); Case 2: \( a < b \). The result holds by Lemma 9.62.

Case 2: \( a < b \). “Uniqueness” Let \( \theta_1 : [a, b] \times [c, d] \rightarrow X \) and \( \theta_2 : [a, b] \times [c, d] \rightarrow X \) 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 \( \bar{\sigma} : [0, 1] \rightarrow Y' \), \( \bar{\theta}_1 : [0, 1] \rightarrow X \), and \( \bar{\theta}_2 : [0, 1] \rightarrow X \) be defined by, \( \forall \lambda \in [0, 1] \subset \mathbb{R} \),

\[
\bar{\sigma}(\lambda) = \sigma(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c)
\]

\[
\bar{\theta}_1(\lambda) = \theta_1(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c)
\]

\[
\bar{\theta}_2(\lambda) = \theta_2(\lambda t + (1 - \lambda)a, \lambda r + (1 - \lambda)c)
\]

This completes the proof of the lemma.
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, $\sigma, \theta_1, \theta_2$ 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 $\theta_2(0) = x_0$. 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 $\sigma_{c,c} : [a,b] \to \mathcal{Y}$ by $\sigma_{c,c}(t) = \sigma(t,c)$, $\forall t \in [a,b] \subset \mathbb{R}$.
By Proposition 3.12, $\sigma_{c,c}$ is continuous with $\sigma_{c,c}(a) = y_0$. By Lemma 9.62, there exists a unique continuous mapping $\theta_{c,c} : [a,b] \to \mathcal{X}$ that inverts $F$
along $\sigma_{c,c}$ with $\theta_{c,c}(a) = x_0$. 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 $\sigma_t(c) = \sigma(t,c) = \sigma_{c,c}(t) = F(\theta_{c,c}(t))$. By Lemma 9.62, there exists a
unique continuous mapping $\theta_t : [c,d] \to \mathcal{X}$ that inverts $F$ along $\sigma_t$ with
$\theta_t(c) = \theta_{c,c}(t)$.

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, $\forall (t,r) \in [a,b] \times [c,d], \sigma(t,r) = \sigma_t(r) = F(\theta_t(t)) = F(\theta(t,r))$.
Hence, $\theta$ inverts $F$ along $\sigma$. $\theta_t(c) = \theta_{c,c}(t) = \theta_{c,c}(a) = x_0$. All we need to
show is that $\theta$ is continuous to complete the proof of the lemma. Define $S := \{r \in [c,d] \subset \mathbb{R} : \theta_{[a,b] \times [c,r]}$ is continuous $\} \subset \mathbb{R}$ and $\xi = \sup S$.
Clearly, $c \in S$ as $c \leq \xi \leq d$.

We will show that $\xi \in S$ by an argument of contradiction. Suppose $\xi \notin S$. Then, $c < \xi \leq d$ and $\exists (r_n)_{n=1}^{\infty} \subseteq S$, which is nondecreasing, such
that $\lim_{n \to \infty} r_n = \xi, \forall (t,r) \in [a,b] \times [c,\xi] \subset \mathbb{R}^2$, 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, $\exists U \subseteq \mathcal{O}_x$ with $x \in U$ and $\exists V \subseteq \mathcal{O}_y$ with $y \in V$ such that $F^{-1}_U : U \to V$ 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 $c_t < \xi < c_t < d_t$ 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 $\tilde{\theta}(\bar{t}, \bar{r}) = (F|_U)_{inv}(\sigma(\bar{t}, \bar{r})) \forall (\bar{t}, \bar{r}) \in D_t$. By Proposition 3.12, $\tilde{\theta}$ is continuous.

Claim 9.63.1 $\theta(\bar{t}, \bar{r}) = \tilde{\theta}(\bar{t}, \bar{r}), \forall (\bar{t}, \bar{r}) \in D_t$.

Proof of claim: Fix any $\bar{t}, \bar{r} \in D_t$. Note that $\tilde{\theta}$ inverts $F$ along $\sigma_{D_t}$.
Define $\tilde{\theta}_t : D_{t,1} \to \mathcal{X}$ by $\tilde{\theta}_t(\bar{r}) = \tilde{\theta}(\bar{t}, \bar{r}), \forall \bar{r} \in D_{t,1} := [c_t, d_t] \cap [c, d] \subset \mathbb{R}$. Then, $\tilde{\theta}_t$ is continuous and inverts $F$ along $\sigma_{D_{t,1}}$ with $\tilde{\theta}_t(\xi) = x$. Note that $\tilde{\theta}_{D_{t,1}}$ is also continuous and inverts $F$ along $\sigma_{D_{t,1}}$ with $\tilde{\theta}_t(\xi) = x$.
By Lemma 9.62, we have $\tilde{\theta}_t = \theta_{D_{t,2}}$ and, in particular, $\tilde{\theta}_t(c_t) = \theta_t(c_t) = \theta_{c_t}(t) = \theta(t,c_t)$.
Define $\tilde{\theta}_{c_t} : D_{t,1} \to \mathcal{X}$ by $\tilde{\theta}_{c_t}(\bar{t}) = \tilde{\theta}(\bar{t}, c_t)$, $\forall \bar{t} \in D_{t,1} := (a_t, b_t) \cap [a, b] \subset \mathbb{R}$. 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 $\tilde{\theta}_{c_t}(\bar{t}) = \tilde{\theta}(\bar{t}, c_t), \forall \bar{t} \in [a, b]$. Then, $\tilde{\theta}_{c_t}$ is continuous and inverts $F$ along $\sigma_{c_t}|_{D_{t,1}}$ with $\tilde{\theta}_{c_t}(t) = \theta(t,c_t)$. Since $c_t \in [c, \xi] \subset \mathbb{R}$, then $\tilde{\theta}_{c_t}|_{D_{t,1}}$ is continuous and inverts $F$ along $\sigma_{c_t}|_{D_{t,1}}$.
with \( \theta_{c_1}(t) = \theta(t, c_1) \). By Lemma 9.62, we have \( \theta_{c_1}|_{D_{t,1}} = \tilde{\theta}_{c_1} \) and, in particular, \( \theta(t, c_1) = \theta_{c_1}(t) = \tilde{\theta}_{c_1}(t) = \tilde{\theta}(t, c_1) \). Define \( \tilde{\theta}_t : D_{t,2} \to \mathcal{X} \) by \( \tilde{\theta}_t(\bar{r}) = \theta(t, \bar{r}) \), \( \forall \bar{r} \in D_{t,2} \). Then, \( \tilde{\theta}_t \) is continuous and inverts \( F \) along \( \sigma_{t}|_{D_{t,2}} \) with \( \tilde{\theta}_t(c_1) = \theta(t, c_1) \). Note that \( \tilde{\theta}_t|_{D_{t,2}} \) is also continuous and inverts \( F \) along \( \sigma_{t}|_{D_{t,2}} \) with \( \tilde{\theta}_t(c_1) = \theta(t, c_1) \). By Lemma 9.62, we have \( \tilde{\theta}_t = \theta(t)|_{D_{t,2}} \) and, in particular, \( \tilde{\theta}(t, \bar{r}) = \tilde{\theta}_t(\bar{r}) = \theta(t, \bar{r}) \). This completes the proof of the claim.

Then, \( \theta|_{D_t} \) is continuous. Then, \( \theta \) is continuous at \( (t, \xi) \). Then, \( \theta|_{[a,b] \times [c,d]} \) is continuous and \( \xi \in S \). This contradicts with the hypothesis \( \xi \not\in S \). 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, \( \exists \mathcal{U} \in \mathcal{O}_X \) with \( x \in \mathcal{U} \) and \( \exists \mathcal{V} \in \mathcal{O}_Y \) with \( y \in \mathcal{V} \) such that \( F|_{\mathcal{U}} : \mathcal{U} \to \mathcal{V} \) 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 \( c_t < \xi < d_t \leq d \) such that \( \sigma(D_t) \subseteq \mathcal{V} \), where \( D_t := ([a_t, b_t] \times [c_t, d_t]) \cap ([a, b] \times [c, d]) \subseteq \mathbb{R}^2 \). Define \( \tilde{\theta} : D_t \to \mathcal{U} \) by \( \tilde{\theta}(\bar{r}) = (F|_{\mathcal{U}})(\sigma(\bar{r})) \), \( \forall \bar{r} \in D_t \). By Proposition 3.12, \( \tilde{\theta} \) is continuous.

**Claim 9.63.2** \( \theta(t, \bar{r}) = \tilde{\theta}(t, \bar{r}), \forall \bar{r} \in D_t \).

**Proof of claim:** Fix any \( (t, \bar{r}) \in D_t \). Note that \( \tilde{\theta} \) inverts \( F \) along \( \sigma|_{D_t} \).

Define \( \tilde{\theta}_t : D_{t,1} \to \mathcal{X} \) by \( \tilde{\theta}_t(t) = \tilde{\theta}(t, \xi), \forall t \in D_{t,1} := (a_t, b_t) \cap [a, b] \subseteq \mathbb{R} \).

Define \( \sigma_{\xi} : [a, b] \to \mathcal{Y} \) by \( \sigma_{\xi}(t) = \sigma(t, \xi), \forall t \in [a, b] \). Define \( \theta_{\xi} : [a, b] \to \mathcal{X} \) by \( \theta_{\xi}(t) = \theta(t, \xi), \forall t \in [a, b] \). Then, \( \theta_{\xi} \) is continuous inverts \( F \) along \( \sigma_{\xi}|_{D_{t,1}} \) with \( \theta_{\xi}(t) = t \). Since \( \xi \in S \), then \( \theta_{\xi}|_{D_{t,1}} \) is continuous and inverts \( F \) along \( \sigma_{\xi}|_{D_{t,1}} \) with \( \theta_{\xi}(t) = \theta(t, \xi) = t \). By Lemma 9.62, we have \( \theta_{\xi}|_{D_{t,2}} = \tilde{\theta}_t \) and, in particular, \( \tilde{\theta}_t(t) = \theta(t, \xi) = \theta_{\xi}(t) \).

Define \( \tilde{\theta}_t : D_{t,2} \to \mathcal{X} \) by \( \tilde{\theta}_t(\bar{r}) = \theta(t, \bar{r}), \forall \bar{r} \in D_{t,2} := (c_t, d_t) \cap [c, d] \subseteq \mathbb{R} \).

Then, \( \tilde{\theta}_t \) is continuous inverts \( F \) along \( \sigma_{\xi}|_{D_{t,2}} \) with \( \tilde{\theta}_t(\xi) = \theta(t, \xi) = \theta_{\xi}(\xi) = \theta(t, \xi) \). Note that \( \theta_{\xi}|_{D_{t,2}} \) is also continuous and inverts \( F \) along \( \sigma_{\xi}|_{D_{t,2}} \) with \( \theta_{\xi}(\xi) = \theta(t, \xi) \). By Lemma 9.62, we have \( \tilde{\theta}_t = \theta_{\xi}|_{D_{t,2}} \) and, in particular, \( \tilde{\theta}(t, \bar{r}) = \tilde{\theta}_t(\bar{r}) = \theta_{\xi}(\bar{r}) = \theta(t, \bar{r}) \). This completes the proof of the claim.

Then, \( \theta|_{D_t} \) is continuous. Note that \( [a, b] \subseteq \bigcup_{t \in [a, b] \cap \mathbb{R}} (a_t, b_t) \).

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 \( \theta|_{[a, b] \times [c, \xi]} \) are continuous. Then, by Theorem 3.11, \( \theta|_{D} \) is continuous, where \( D := ([a, b] \times [c, \xi]) \cap \bigcup_{t \in T_N} (a_t, b_t) \). Then, \( \theta|_{[a, b] \times [c, d]} \) is continuous and \( \bar{d} \in S \). This leads to the contradiction \( \xi < \bar{d} \leq \sup S = \xi \). Therefore, \( \xi = d \). Then, \( \theta \) is continuous. This completes the proof of Case 2.

This completes the proof of the lemma.

\( \square \)
9.6. GLOBAL INVERSE FUNCTION THEOREM

Theorem 9.64 (Global Inverse Function Theorem) Let $X$ and $Y$ be Hausdorff topological spaces, $X \neq \emptyset$, $F : X \to Y$ be continuous and countably proper. Assume that $\forall x \in X$, $\exists U \in \mathcal{O}_X$ with $x \in U$ and $\exists V \in \mathcal{O}_Y$ with $F(x) \in V$ such that $F|_U : U \to V$ is a homeomorphism, $X$ is arcwise connected, and $Y$ is simply connected. Then, $F : X \to Y$ is a homeomorphism.

Proof Fix $x_0 \in X \neq \emptyset$, let $y_0 = F(x_0) \in Y$. $\forall y \in Y$, since $Y$ is simply connected, then $Y$ is arcwise connected. Then, there exists a curve $\sigma : I \to Y$, 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 $\theta : I \to X$ that inverts $F$ along $\sigma$ with $\theta(0) = x_0$. Then, $y = \sigma(1) = F(\theta(1))$. Hence, $F$ is surjective.

Fix $x_1, x_2 \in X$ such that $F(x_1) = F(x_2) = y$. Since $X$ is arcwise connected, then there exists a curve $\delta : I \to X$ 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 $Y$ is simply connected, then $\eta$ is homotopic to a single point $\bar{y} \in Y$. Then, there exists a continuous mapping $\gamma : I \times I \to Y$ 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 $\zeta : I \times I \to X$ that inverts $F$ along $\gamma$ with $\zeta(0, 0) = x_1$.

Define $\gamma_0 : I \to Y$ by $\gamma_0(t) = \gamma(0, t)$, $\gamma_1 : I \to Y$ by $\gamma_1(t) = \gamma(1, t)$, $\gamma_0 : I \to Y$ by $\gamma_0(t) = \gamma(t, 0)$, and $\gamma_1 : I \to Y$ by $\gamma_1(t) = \gamma(t, 1)$, $\forall t \in I$. Then, $\gamma_0 = \gamma_1$. Define $\zeta_0 : I \to X$ by $\zeta_0(t) = \zeta(0, t)$, $\zeta_1 : I \to X$ by $\zeta_1(t) = \zeta(1, t)$, $\zeta_0 : I \to X$ by $\zeta_0(t) = \zeta(t, 0)$, and $\zeta_1 : I \to X$ by $\zeta_1(t) = \zeta(t, 1)$, $\forall t \in I$. Then, $\zeta_0$ is continuous and inverts $F$ along $\gamma_0$. Set $\bar{x} = \zeta_0(1) = \zeta(0, 1) \in X$, 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 $\lambda : I \to X$ 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 $\zeta_1(0) = \zeta(1, 0) = \zeta_0(0) = \zeta(0, 0) = x_1$. Note that $\zeta_0$ is continuous and inverts $F$ along $\gamma_0$ 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 $F_i : Y \to X$. $\forall y \in Y$, let $x = F_i(y) \in X$. Then, $\exists U \in \mathcal{O}_X$ with $x \in U$ and $\exists V \in \mathcal{O}_Y$ with $y \in V$ such that $F_i|_U : U \to V$ 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.

$\square$

Theorem 9.65 (Global Inverse Function Theorem) Let $X$ and $Y$ be Hausdorff topological spaces, $F : X \to Y$ be continuous and countably proper.
Thus, \( G := \{ x \in X \mid \exists U \in O_X \text{ with } x \in U \ni F|_U : U \to F(U) \in O_Y \text{ is a homeomorphism} \} \subseteq X, \Sigma := X \setminus H, \Sigma_0 := \text{inv}(F(\Sigma)) \), \( X_0 := X \setminus \Sigma_0 \), and \( Y_0 := Y \setminus F(\Sigma) \). 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 \( O_{X_0} \) and \( O_{Y_0} \) be the subset topology on \( X_0 \) and \( Y_0 \), respectively.

**Claim 9.65.1** \( \Sigma \subseteq X \) is closed.

**Proof of claim:** \( \forall x \in H, \exists U_x \in O_X \text{ with } x \in U_x \text{ such that } F|_{U_x} : U_x \to F(U_x) \in O_Y \text{ is a homeomorphism.} \forall \bar{x} \in U_x, \bar{x} \in H. \) Then, \( U_x \subseteq H \) and \( H = \bigcup_{x \in H} U_x \). Hence, \( H \in O_X \). Then, \( \Sigma := X \setminus H \) is closed. \( \square \)

**Claim 9.65.2** \( F_{\text{inv}}(Y_0) = X_0. \ G : X_0 \to Y_0 \text{ is continuous and countably proper.} \)

**Proof of claim:** By Proposition 2.5, \( F_{\text{inv}}(Y_0) = F_{\text{inv}}(Y \setminus F(\Sigma)) \subseteq F_{\text{inv}}(Y) \setminus F_{\text{inv}}(F(\Sigma)) = X \setminus \Sigma_0 = X_0 \) and \( F(X_0) = F(F_{\text{inv}}(Y_0)) \subseteq Y_0 \). Hence, \( G \) is a function of \( X_0 \) to \( Y_0 \).

Fix any \( K \subseteq Y_0 \) such that \( K \) is compact in \( O_{Y_0} \). Then \( K \) is a compact set in \( O_Y \). Then, \( G_{\text{inv}}(K) = F_{\text{inv}}(K) \subseteq X_0 \). By the countable properness of \( F \), we have \( F_{\text{inv}}(K) \subseteq X \) is countably compact in \( O_X \). Then, it is easy to show that \( G_{\text{inv}}(K) \) is countably compact in \( O_{X_0} \). Hence, \( G \) is countably proper.

\( \forall V_0 \in O_{Y_0}, V_0 = Y_0 \cap V, \text{ where } V \in O_Y. \ G_{\text{inv}}(V_0) = F_{\text{inv}}(V_0) = F_{\text{inv}}(V) \cap F_{\text{inv}}(V) = X_0 \cap F_{\text{inv}}(V). \) Since \( F \) is continuous, \( F_{\text{inv}}(V) \in O_X \). Thus, \( G_{\text{inv}}(V_0) \in O_{X_0} \). Hence, \( G \) is continuous. \( \square \)

By Proposition 2.5, \( \Sigma_0 = F_{\text{inv}}(F(\Sigma)) \supseteq \Sigma \cap \text{dom}(F) = \Sigma \). Then, \( X_0 = X \setminus \Sigma_0 \subseteq X \setminus \Sigma = \Sigma. \forall x \in X_0 \subseteq H, \exists U \in O_X \text{ with } x \in U \text{ such that } F|_U : U \to F(U) \in O_Y \text{ is a homeomorphism.} \) Let \( U_0 := X_0 \cap U \in O_{X_0} \), \( V := F(U) \in O_Y \), and \( V_0 := F(U) \cap Y_0 \in O_{Y_0} \). 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. \( \forall \bar{y} \in V_0, \bar{y} \in V \text{ then } \exists \bar{x} \in U \text{ such that } \bar{y} = F(\bar{x}). \) Note that \( \bar{y} \in Y_0 \) and \( F_{\text{inv}}(Y_0) = X_0 \), 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 \( X := (X, O) \) 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 $O_{\alpha} = O \cap A_{\alpha} = \emptyset$. Let $\mathcal{M} := \{ A \subseteq O \mid x_0 \in A \}$ and $A_0 := \bigcup_{A \in \mathcal{M}} A$. Then, $x_0 \in A_0 \subseteq O$ and, by Proposition 9.66, $A_0$ is arcwise connected. Then, by Proposition 9.66, $A_0 \cup B_X(x, \delta)$ is arcwise connected. This completes the proof of the proposition.

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 $F_n : D \to Y$, $\forall n \in \mathbb{N}$. Assume that

(i) $\exists \delta_0 \in (0, \infty) \subset \mathbb{R}$ such that $D := D \cap B_X(x_0, \delta_0) - x_0$ is a conic segment;
(ii) $F_n$ is differentiable, $\forall n \in \mathbb{N}$;

(iii) $\left( F_n^{(1)} \right)_{n=1}^\infty$ converges uniformly to $G : D \to \mathbb{B}(X, Y)$;

(iv) $\forall x \in D, \lim_{n \to \infty} F_n(x) = F(x)$, where $F : D \to Y$.

Then, $F$ is differentiable at $x_0$ and $F^{(1)}(x_0) = G(x_0) = \lim_{n \to \infty} F_n^{(1)}(x_0)$.

**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 $n \geq n_0$, $\| F_n^{(1)}(x) - G(x) \| < \epsilon/2$, $\forall x \in D$. $\forall n, m \in \mathbb{N}$ with $n \geq n_0$ and $m \geq n_0$, by Proposition 9.15, $g : D \to Y$, defined by $g(x) = F_n(x) - F_m(x)$, $\forall x \in D$, is differentiable and $g^{(1)}(x) = F_n^{(1)}(x) - F_m^{(1)}(x)$, $\forall x \in D$. 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 $m \geq n_0$, $\| F_n(x) - F_m(x) - F_n(x_0) + F_m(x_0) \| = \| g(x) - g(x_0) \| \leq \| g^{(1)}(x)(x - x_0) \| \leq \| F_n^{(1)}(x) - F_m^{(1)}(x) \| \| x - x_0 \| \leq \epsilon \| x - x_0 \| /2$, where $x = t_0 x + (1 - t_0) x_0 \in D + x_0$ and $t_0 \in (0, 1) \subset \mathbb{R}$. By $F_n^{(1)}(x_0) \in \mathbb{B}(X, Y)$, $\exists \delta_0 \in (0, \delta] \subset \mathbb{R}$ such that, $\forall x \in D := D \cap B_X(x_0, \delta) \subseteq D + x_0$, we have $\| F_n(x) - F_n(x_0) - F_n^{(1)}(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \| /4$.

Fix any $x \in D$. $\forall m \in \mathbb{N}$ with $m \geq n_0$, we have $\| F_m(x) - F_m(x_0) - G(x_0)(x - x_0) \| \leq \| F_m(x) - F_m(x_0) - F_n(x_0) + F_n(x_0) \| + \| F_n(x_0) - F_n^{(1)}(x_0)(x - x_0) \| + \| F_n^{(1)}(x_0)(x - x_0) - G(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \| /4 + \epsilon \| x - x_0 \| /4 + \epsilon \| x - x_0 \| /4 = \epsilon \| x - x_0 \|$. Take limit as $m \to \infty$, we have $\| F(x) - F(x_0) - G(x_0)(x - x_0) \| \leq \lim_{m \to \infty} \| F_m(x) - F_m(x_0) - G(x_0)(x - x_0) \| \leq \epsilon \| x - x_0 \|$. By the arbitrariness of $x$, $F$ is differentiable at $x_0$ and $F^{(1)}(x_0) = G(x_0) = \lim_{n \to \infty} DF_n(x_0)$. 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 $\text{span} (A_\Omega(x)) = X$, $\forall x \in \Omega$, $k \in \mathbb{N}$, and $\mathcal{W} := \mathcal{C}(\Omega, Y)$ be the normed linear space defined in Example 7.31. Define $Z := \{ f \in \mathcal{C}(\Omega, Y) \mid f \in \mathcal{C}_k \}$. By Proposition 9.40, $Z$ is a subspace of $\mathcal{C}(\Omega, Y)$. Then, $Z := (Z, \oplus, \ominus, \odot)$ is vector space over the field $\mathbb{K}$.

Now, define a norm on $Z$ by $\| f \|_{\mathcal{C}_k} := \left( \sum_{i=0}^k \| f^{(i)} \|_c^2 \right)^{1/2}$, $\forall f \in Z$, $\forall i \in \{0, \ldots, k\}$, $f^{(i)} \in \mathcal{C}(\Omega, B_1(X, Y))$ with norm $\| f^{(i)} \|_c \subset \mathbb{R}$.

By Proposition 9.40, $\| f \|_{\mathcal{C}_k} = 0$, then $f = \vartheta_\Omega$. $\forall f, g \in Z$, by Proposition 9.40, $\| f + g \|_{\mathcal{C}_k} = \left( \sum_{i=0}^k \| f^{(i)} + g^{(i)} \|_c^2 \right)^{1/2} \leq \left( \sum_{i=0}^k \| f^{(i)} \|_c^2 \right)^{1/2} + \left( \sum_{i=0}^k \| g^{(i)} \|_c^2 \right)^{1/2} = \| f \|_{\mathcal{C}_k} + \| g \|_{\mathcal{C}_k}$, where the inequality follows from Minkowski’s Inequality, Theorem 7.9. $\forall \alpha \in \mathbb{K}$, $\forall f \in Z$, by Proposition 9.40, $\| \alpha f \|_{\mathcal{C}_k} = \| f \|_{\mathcal{C}_k}$.
9.7. INTERCHANGE DIFFERENTIATION AND LIMIT

\[ \left( \sum_{i=0}^{k} \| \alpha f(i) \|^2 \right)^{1/2} = \left( \sum_{i=0}^{k} |\alpha|^2 \| f(i) \|^2 \right)^{1/2} = |\alpha| \| f \|_{C_k} \]. This shows that \( Z, K, \| \cdot \| \) is a normed linear space, which will be denoted \( C_k(\Omega, \gamma) \).

Example 9.70 Let \( X \) be a normed linear space over \( K \) and \( \gamma \) be a Banach space over \( K \), \( \Omega \subseteq X \) be a compact set, and \( k \in \mathbb{N} \). Assume that, \( \forall x \in \Omega \), \( \exists d_1 \in (0, \infty) \subset \mathbb{R} \) such that \( \Omega \cap B_X(x, d_1) = X \). Let \( C_k(\Omega, \gamma) \) be the normed linear space defined in Example 9.69. We will show that \( C_k(\Omega, \gamma) \) is also a Banach space over \( K \).

Fix any Cauchy sequence \( (f_n)_{n=1}^{\infty} \subseteq C_k(\Omega, \gamma) \). By the definition of the norm \( \| \cdot \|_{C_k} \), \( (f_n)_{n=1}^{\infty} \subseteq C(\Omega, B_1(X, \gamma)) =: W_i \) is a Cauchy sequence, \( \forall i \in \{0, \ldots, k\} \). \( \forall i \in \{0, \ldots, k\} \), by Example 7.50 and Proposition 7.66, \( C(\Omega, B_1(X, \gamma)) \) is a Banach space. Then, \( \exists g_i \in C(\Omega, B_1(X, \gamma)) \) such that \( \lim_{n \in \mathbb{N}} f_n(i) = g_i \) in \( C(\Omega, B_1(X, \gamma)) \). Then, \( (f_n)_{n=1}^{\infty} \) converges uniformly to \( g_i \) and \( g_i \) is continuous. By Proposition 9.68, we have \( g_i(0) = g_i, \forall i \in \{0, \ldots, k\} \). Then, \( \forall i \in \{0, \ldots, k\} \), \( g_0 \in C_k(\Omega, \gamma) \), and \( g_0 \in C_k(\Omega, \gamma) \). Then, \( \| f_n - g_0 \|_{C_k} = \left( \sum_{i=0}^{k} \left\| f_n(i) - g_0(i) \right\|^2 \right)^{1/2} \rightarrow 0 \) as \( n \rightarrow \infty \), where the first equality follows from Proposition 9.40. Then, \( \lim_{n \in \mathbb{N}} f_n = g_0 \) in \( C_k(\Omega, \gamma) \). Hence, \( C_k(\Omega, \gamma) \) 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 \( X := (\mathcal{X}, \mathcal{O}) \) be a topological space, \( \gamma \) be a normed linear space over the field \( K \), and \( C_0(\mathcal{X}, \gamma) \) be the vector space of all continuous functions of \( X \) to \( \gamma \) as defined in Example 7.50 with null vector \( \vartheta \). Define a function \( \| \cdot \| : C_0(\mathcal{X}, \gamma) \rightarrow \mathbb{R} \) by \( \| f \| = \max \{ \sup_{x \in \mathcal{X}} \| f(x) \|_{\gamma}, 0 \}, \forall f \in C_0(\mathcal{X}, \gamma) \). Consider the set \( \mathcal{M} := \{ f \in C_0(\mathcal{X}, \gamma) \mid \| f \| < +\infty \} \). Clearly, \( \vartheta \in \mathcal{M} \). \( \forall f_1, f_2 \in \mathcal{M}, \forall \alpha, \beta \in K, \| \alpha f_1 + \beta f_2 \| = \max \{ \sup_{x \in \mathcal{X}} \| \alpha f_1(x) + \beta f_2(x) \|_{\gamma}, 0 \} \leq \max \{ \sup_{x \in \mathcal{X}} (|\alpha| \| f_1(x) \|_{\gamma} + |\beta| \| f_2(x) \|_{\gamma}), 0 \} < +\infty \). Then, \( \alpha f_1 + \beta f_2 \in \mathcal{M} \). Hence, \( \mathcal{M} \) is a subspace of \( C_0(\mathcal{X}, \gamma) \).

Clearly, \( \forall f \in \mathcal{M}, \| f \| \in (0, \infty) \subset \mathbb{R} \) and \( \| f \| = 0 \leftrightarrow f = \vartheta \). \( \forall f_1, f_2 \in \mathcal{M}, \forall \alpha \in K, \| f_1 + f_2 \| = \max \{ \sup_{x \in \mathcal{X}} \| f_1(x) + f_2(x) \|_{\gamma}, 0 \} \leq \max \{ \sup_{x \in \mathcal{X}} \| f_1(x) \|_{\gamma} + \sup_{x \in \mathcal{X}} \| f_2(x) \|_{\gamma}, 0 \} = \| f_1 \| + \| f_2 \| \), where the first inequality follows from Proposition 3.81.

\[ \alpha f_1 \| = \max \{ \sup_{x \in \mathcal{X}} |\alpha f_1(x)|, 0 \} = \max \{ \sup_{x \in \mathcal{X}} |\alpha| \| f_1(x) \|_{\gamma}, 0 \} = \begin{cases} \max \{ |\alpha| \sup_{x \in \mathcal{X}} \| f_1(x) \|_{\gamma}, 0 \} & \alpha \neq 0 \\ 0 & \alpha = 0 \end{cases} = |\alpha| \| f_1 \|, \] where the third equality follows from Proposition 3.81. Hence, \( (\mathcal{M}, K, \| \cdot \|) \) is a normed linear space, which will be denoted by \( C_0(\mathcal{X}, \gamma) \).
Example 9.72  Let \( \mathcal{X} := (\mathcal{X}, \mathcal{O}) \) be a topological space and \( \mathcal{Y} \) be a Banach space over the field \( \mathbb{K} \) (with norm \( \| \cdot \|_{\mathcal{Y}} \)). Consider the normed linear space \( C_{b}(\mathcal{X}, \mathcal{Y}) \) (with norm \( \| \cdot \|_{\mathcal{Y}} \)) 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 \( (f_{n})_{n=1}^{\infty} \subseteq C_{b}(\mathcal{X}, \mathcal{Y}) \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists N \in \mathbb{N} \) such that \( \forall n, m \geq N, 0 \leq \|f_{n}(x) - f_{m}(x)\|_{\mathcal{Y}} \leq \|f_{n} - f_{m}\| < \epsilon, \forall x \in \mathcal{X} \). 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. This defines a function \( f: \mathcal{X} \to \mathcal{Y} \). 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. \( \forall x \in \mathcal{X}, \|f(x)\|_{\mathcal{Y}} \leq \|f_{n}(x) - f(x)\|_{\mathcal{Y}} + \|f_{n}(x)\|_{\mathcal{Y}} = \lim_{n \in \mathbb{N}} \|f_{n}(x) - f_{m}(x)\|_{\mathcal{Y}} + \|f_{n}(x)\|_{\mathcal{Y}} \leq \epsilon + \|f_{n}\|_{\mathcal{Y}}. \) Hence, \( \|f\|_{\mathcal{Y}} \leq \|f_{n}\|_{\mathcal{Y}} + \epsilon. \) 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 \( \mathbb{K}, \mathcal{X}, \mathcal{Y} \subseteq \mathcal{X} \) be endowed with the subset topology, which satisfies \( \text{span}(\mathcal{A}_{\mathcal{Y}}(x)) = \mathcal{X} \), \( \forall x \in \mathcal{X}, k \in \mathbb{N} \), and \( \mathcal{W} := C_{b}(\mathcal{X}, \mathcal{Y}) \) be the normed linear space defined in Example 9.71 with null vector \( \mathcal{O}_{\mathcal{Y}} \). Define \( \mathcal{Z} := \{ f \in C_{b}(\mathcal{X}, \mathcal{Y}) \mid f \text{ is } C_{k} \text{ and } f^{(i)} \in C_{b}(\mathcal{X}, \mathcal{Y}), i = 1, \ldots, k \} \). By Proposition 9.40, \( \mathcal{Z} \) is a subspace of \( \mathcal{W} \). Then, \( \mathcal{Z} := (Z, \oplus_{\mathcal{W}}, \otimes_{\mathcal{W}}, \mathcal{O}_{\mathcal{W}}) \) is vector space over the field \( \mathbb{K} \). \( \forall f \in \mathcal{Z}, \forall i \in \{0, \ldots, k\}, f^{(i)} \in C_{b}(\mathcal{X}, \mathcal{Y}) \) with norm \( \|f^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})} \). Now, define a norm on \( \mathcal{Z} \) by \( \|f\|_{C_{b}(k)} := \left( \sum_{i=0}^{k} \|f^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} \subseteq [0, \infty) \subset \mathbb{R}, \forall f \in \mathcal{Z} \). If \( f = \mathcal{O}_{\mathcal{W}} \), by Proposition 9.33, \( f \) is \( C_{\infty} \) and \( f^{(i)}(x) = \mathcal{O}_{B_{\mathcal{Y}}(x, \mathcal{Y})}, \forall x \in \mathcal{X}, \forall i \in \mathbb{N} \), then, \( \|f\|_{C_{b}(\mathcal{X}, \mathcal{Y})} = 0 \). On the other hand, if \( \|f\|_{C_{b}(\mathcal{X}, \mathcal{Y})} = 0 \), then \( f = \mathcal{O}_{\mathcal{W}} \). \( \forall f_{1}, f_{2} \in \mathcal{Z} \), by Proposition 9.40, \( \|f_{1} + f_{2}\|_{C_{b}(k)} = \left( \sum_{i=0}^{k} \|f_{1}^{(i)} + f_{2}^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} \leq \left( \sum_{i=0}^{k} \|f_{1}^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} + \left( \sum_{i=0}^{k} \|f_{2}^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} = \|f_{1}\|_{C_{b}(\mathcal{X}, \mathcal{Y})} + \|f_{2}\|_{C_{b}(\mathcal{X}, \mathcal{Y})} \), where the inequality follows from Minkowski's Inequality, Theorem 7.9. \( \forall \alpha \in \mathbb{K} \), \( \forall f \in \mathcal{Z} \), by Proposition 9.40, \( \|\alpha f\|_{C_{b}(k)} = \left( \sum_{i=0}^{k} \|\alpha f^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} = \left( \sum_{i=0}^{k} \|f^{(i)}\|_{C_{b}(\mathcal{X}, \mathcal{Y})}^{2} \right)^{1/2} = \|f\|_{C_{b}(\mathcal{X}, \mathcal{Y})} \). This shows that \( (\mathcal{Z}, \mathbb{K}, \|\cdot\|_{C_{b}(k)}) \) is a normed linear space, which will be denoted \( C_{b,k}(\mathcal{X}, \mathcal{Y}) \).

Example 9.74  Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \) and \( \mathcal{Y} \) be a Banach space over \( \mathbb{K}, \mathcal{X}, \mathcal{Y} \subseteq \mathcal{X} \) be endowed with the subset topology, and \( k \in \mathbb{N} \). Assume that, \( \forall x \in \mathcal{X}, \exists B_{x} \in (0, \infty) \subset \mathbb{R} \) such that \( \mathcal{X} \cap B_{x}(x, \mathcal{Y}) = \emptyset \) is a conic segment and \( \text{span}(\mathcal{A}_{\mathcal{Y}}(x)) = \mathcal{X} \). Let \( C_{b,k}(\mathcal{X}, \mathcal{Y}) \) be the normed
linear space defined in Example 9.73. We will show that \( C_{b,k}(\Omega, y) \) is also a Banach space over \( K \).

Fix any Cauchy sequence \( (f_n)_{n=1}^\infty \subseteq C_{b,k}(\Omega, y) \). By the definition of the norm \( \| \cdot \|_{C_{b,k}} \), \( (f_n)_{n=1}^\infty \subseteq C_{b}(\Omega, B_i(\mathcal{X}, y)) =: W_i \) is a Cauchy sequence, \( \forall i \in \{0, \ldots, k\} \). \( \forall i \in \{0, \ldots, k\} \), by Example 9.72 and Proposition 7.66, \( C_{b}(\Omega, B_i(\mathcal{X}, y)) \) is a Banach space. Then, \( \exists g_i \in C_{b}(\Omega, B_i(\mathcal{X}, y)) \) such that \( \lim_{n \in \mathbb{N}} f_n(i) = g_i \) in \( C_{b}(\Omega, B_i(\mathcal{X}, y)) \). Then, \( (f_n(i))_{n=1}^\infty \) converges uniformly to \( g_i \) and \( g_i \) is continuous. By Proposition 9.68, we have \( g_i(i) = g_{i+1}, \forall i \in \{0, \ldots, k-1\} \). Then, we have \( g_0(i) = g_i, \forall i \in \{0, \ldots, k\} \), and \( g_0 \in C_{b,k}(\Omega, y) \). Furthermore, \( \| f_n - g_0 \|_{C_{b,k}} = \left( \sum_{i=0}^{k} \left\| f_n(i) - g_{i} \right\|_{C_{b}}^{2} \right)^{1/2} = \left( \sum_{i=0}^{k} \left\| f_n(i) - g_{i} \right\|_{C_{b}}^{2} \right)^{1/2} \rightarrow 0 \) as \( n \rightarrow \infty \), where the first equality follows from Proposition 9.40. Then, \( \lim_{n \in \mathbb{N}} f_n = g_0 \) in \( C_{b,k}(\Omega, y) \). Hence, \( C_{b,k}(\Omega, y) \) is complete and therefore a Banach space.

9.8 Tensor Algebra

**Definition 9.75** Let \( m \in \mathbb{Z}_+, \mathcal{X}_i \) be a normed linear space over \( K \), \( i = 1, \ldots, m \), \( \mathcal{Z} \) be a normed linear space over \( K \). A bounded linear operator \( A \in B(\mathcal{X}_m, B(\mathcal{X}_{m-1}, \ldots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)) \) is said to be an \( m \)th order \( \mathcal{Z} \)-valued tensor. Let \( B \in B(\mathcal{Y}_n, B(\mathcal{Y}_{n-1}, \ldots, B(\mathcal{Y}_1, \mathcal{X}_m) \cdots)) \) be another \( n \)th order \( \mathcal{X}_m \)-valued tensor. We define \( AB := A \cdot B \in B(\mathcal{Y}_n, B(\mathcal{Y}_{n-1}, \ldots, B(\mathcal{Y}_1, \mathcal{X}_m) \cdots, B(\mathcal{X}_{m-1}, \ldots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)) \) to be an \( (n+m-1) \)st order \( \mathcal{Z} \)-valued tensor such that 

\[
(AB)(y_n) \cdots (y_1)(x_{m-1}) \cdots (x_1) = A(B(y_n)) \cdots (y_1)(x_{m-1}) \cdots (x_1) \in \mathcal{Z}, \forall y_i \in \mathcal{Y}_j, i = 1, \ldots, n, \forall x_j \in \mathcal{X}_j, j = 1, \ldots, m-1. \]

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 \( A^{T_{n_1,\ldots,n_m}} \in B(\mathcal{X}_{n_m}, B(\mathcal{X}_{n_{m-1}}, \ldots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)) \) such that

\[
A^{T_{n_1,\ldots,n_m}}(x_{n_m}) \cdots (x_1) = A(x_{n_m}) \cdots (x_1) \quad \forall x_i \in \mathcal{X}_i, \ i = 1, \ldots, m
\]

**Proposition 9.76** Let \( m, n \in \mathbb{Z}_+, \mathcal{X}_i, i = 1, \ldots, m, \mathcal{Y}_j, j = 1, \ldots, n, \mathcal{Z} \) be normed linear spaces over \( K \), \( A, A_k \in W_1 := B(\mathcal{X}_m, B(\mathcal{X}_1, \mathcal{Z}) \cdots), k = 1, 2, \) be \( m \)th order \( \mathcal{Z} \)-valued tensors, \( B, B_i \in W_2 := B(\mathcal{Y}_n, B(\mathcal{Y}_{n-1}, \ldots, B(\mathcal{Y}_1, \mathcal{X}_m) \cdots), l = 1, 2, \) and \( W_3 := B(\mathcal{Y}_n, B(\mathcal{Y}_{n-1}, \ldots, B(\mathcal{Y}_1, B(\mathcal{X}_{m-1}, \ldots, B(\mathcal{X}_1, \mathcal{Z}) \cdots)) \cdots) \). Then, the following statements hold.

(i) \( \| AB \| \leq \| A \| \| B \| \).

(ii) \( \forall a_k, \beta_j \in K, \ k = 1, 2, \ l = 1, 2, \) 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 \( A \in B(\mathcal{X}_2, B(\mathcal{X}_1, \mathbb{K})) \) be a second order \( \mathbb{K} \)-valued tensor. Then, \( A^{T_{2,1}} = A'^{\phi_{X_1}} \), where \( \phi_{X_1} : \mathcal{X}_1 \to \mathcal{X}_1^{**} \) is the natural mapping as defined in Remark 7.88.

(iv) Let \( f : \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{W}_3 \) be defined by \( f(A, B) = AB \in \mathcal{W}_3, \forall A \in \mathcal{W}_1, \forall B \in \mathcal{W}_2 \). Then, \( f \) is \( C_0 \), \( f^{(1)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = A_0 \Delta_{2,1} + \Delta_{1,1} B_0, f^{(2)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1})(\Delta_{1,2}, \Delta_{2,2}) = \Delta_{1,2} \Delta_{2,1} + \Delta_{1,1} \Delta_{2,2} \), and \( f^{(i+2)}(A_0, B_0) = \vartheta_{B_{i+2}}(\mathcal{W}_1 \times \mathcal{W}_2, \mathcal{W}_3), \forall (A_0, B_0) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall (\Delta_{1,1}, \Delta_{2,1}) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall (\Delta_{1,2}, \Delta_{2,2}) \in \mathcal{W}_1 \times \mathcal{W}_2, \forall i \in \mathbb{N} \).

(v) Let \( \mathcal{X} \) be a normed linear space over \( \mathbb{K} \), \( x_0 \in D \subseteq \mathcal{X} \), \( A : D \to \mathcal{W}_1 \) and \( B : D \to \mathcal{W}_2 \) be tensor-valued functions that are Fréchet differentiable at \( x_0 \), and \( C : D \to \mathcal{W}_3 \) be defined by \( C(x) = A(x)B(x), \forall x \in D \). Then,

\[
C^{(1)}(x_0) = \left( (A^{(1)}(x_0))^{T_{1,\ldots,m-1,m+1,m}} B(x_0) \right)^{T_{1,\ldots,m-1,m+1,n+1,m}} + A(x_0)B^{(1)}(x_0)
\]

Proof

(i) and (ii) These are straightforward, and are therefore omitted.

(iii) Note that \( A \in B(\mathcal{X}_2, \mathcal{X}_2^*) \). Then, \( A' \in B(\mathcal{X}_2^*, \mathcal{X}_2^*) \) and \( A^{T_{2,1}} \in B(\mathcal{X}_1, B(\mathcal{X}_2, \mathbb{K})) = B(\mathcal{X}_1, \mathcal{X}_2^*) \). Then, \( A^{T_{2,1}}(x_1)(x_2) = A(x_2)(x_1) = \langle \langle A(x_2), x_1 \rangle, A(x_2) \rangle = \langle \langle A'(\phi_{X_1}(x_1)), x_2 \rangle, A'(\phi_{X_1}(x_1))(x_2) \rangle, \forall x_1 \in \mathcal{X}_i, i = 1, 2 \). Then, we have \( A^{T_{2,1}} = A'^{\phi_{X_1}} \).

(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.

\[\square\]

Definition 9.77 Let \( \mathcal{X}_i, i = 1, \ldots, m, \mathcal{Z} \) be normed linear spaces over \( \mathbb{K} \), \( A \in B(\mathcal{X}_m, \mathbb{K}) \) be an \( n \)-th order \( \mathbb{Z} \)-valued tensor, and \( B \in B(\mathcal{Y}_n, \mathbb{K}) \) be an \( n \)-th order \( \mathbb{K} \)-valued tensor. Define the tensor product of \( A \) and \( B \) to be an \( (n+m) \)-th order \( \mathbb{Z} \)-valued tensor \( C := A \otimes B \in B(\mathcal{Y}_n, B(\mathcal{X}_m, \mathbb{K})) \) such that \( C(y_n) \cdots (y_1)(x_m) \cdots (x_1) = B(y_n) \cdots (y_1)A(x_m) \cdots (x_1) \in \mathcal{Z} \), \( \forall x_i \in \mathcal{X}_i, i = 1, \ldots, m, \forall y_j \in \mathcal{Y}_j, j = 1, \ldots, n \). Similarly, we may define the tensor product of \( B \) and \( A \) to be an \( (n+m) \)-th order \( \mathbb{Z} \)-valued tensor \( C := B \otimes A \in B(\mathcal{X}_m, B(\mathcal{Y}_n, \mathbb{K})) \) such that...
Theorem 9.80 (Taylor Series)  
\[ \forall \delta \in (0, \infty), x_0 \in D, \]  
\[ C(x_m) \cdots (x_1)(y_n) \cdots (y_1) = B(y_n) \cdots (y_1)A(x_m) \cdots (x_1) \in \mathbb{Z}, \forall x_i \in X_i, \]
\[ i = 1, \ldots, m, \forall y_j \in y_j, j = 1, \ldots, n. \]

Proposition 9.78 Let \( m, n \in \mathbb{Z}_+, \) \( X, i = 1, \ldots, m, \) \( y_j, j = 1, \ldots, n, \) \( Z \), \( \tau = 1, 2, \) be normed linear spaces over \( \mathbb{K} \) with \( Z_1 = \mathbb{K} \) or \( Z_2 = \mathbb{K}, \) \( Z = Z_1 \)
\( \) if \( Z_2 = \mathbb{K}, \) \( Z = Z_2 \) if \( Z_1 = \mathbb{K}, \) \( A, A_k \in W_1 := B(X_m, \ldots, B(X_1, Z_1) \cdots), \)
\( k = 1, 2, \) be \( n \)th order \( Z_1 \)-valued tensors, \( B, B_i \in W_2 := \) \( B(y_n, \ldots, B(y_1, Z_2) \cdots) \) be \( n \)th order \( Z_2 \)-valued tensors, \( l = 1, 2, \) and \( W_3 := B(y_n, \ldots, B(y_1, B(X_m, \ldots, B(X_1, Z) \cdots)) \cdots). \) Then, the following statements hold.

(i) \( \| A \otimes B \| \leq \| A \| \| B \|. \)

(ii) \( \forall \alpha_k, \beta_i \in \mathbb{K}, k = 1, 2, l = 1, 2, \) we have \( \langle \alpha_1 A_1 + \alpha_2 A_2 \rangle \otimes \langle \beta_1 B_1 + \beta_2 B_2 \rangle = \alpha_1 \beta_1 A_1 \otimes B_1 + \alpha_1 \beta_2 A_1 \otimes B_2 + \alpha_2 \beta_1 A_2 \otimes B_1 + \alpha_2 \beta_2 A_2 \otimes B_2. \)

(iii) Let \( f : W_1 \times W_2 \to W_3 \) be defined by \( f(A, B) = A \otimes B \in W_3, \forall A \in W_1, \forall B \in W_2. \) Then, \( f \) is \( C_\infty, f^{(1)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = A_0 \otimes \Delta_{2,1} + \Delta_{1,1} \otimes B_0, f^{(2)}(A_0, B_0)(\Delta_{1,1}, \Delta_{2,1}) = \Delta_{1,2} \otimes \Delta_{2,1} + \Delta_{1,1} \otimes \Delta_{2,2}, \) and \( f^{(l+2)}(A_0, B_0) = \varphi_{\beta_1, \beta_2}(\Delta_{1,1}, \Delta_{2,1}), \forall (A_0, B_0) \in W_1 \times W_2, \forall (\Delta_{1,1}, \Delta_{2,1}) \in W_1 \times W_2, \forall \varphi \in \mathbb{N}. \)

(iv) Let \( X \) be a normed linear space over \( \mathbb{K}, \) \( x_0 \in D \subseteq X, A : D \to W_1 \) 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 \( C(x) = A(x) \otimes B(x), \forall x \in D. \) Then,

\[ C^{(1)}(x_0) = (A^{(1)}(x_0) \otimes B(x_0))^{T_1 \cdots m_1 \cdots \cdots n + m + 1} \]
\[ + A(x_0) \otimes B^{(1)}(x_0). \]

Proof  These are straightforward, and are therefore omitted. \( \square \)

9.9 Analytic Functions

Definition 9.79 Let \( \mathfrak{X} \) and \( \mathfrak{Y} \) be normed linear spaces over \( \mathbb{K}, D \subseteq \mathfrak{X}, \)
\( f : D \to \mathfrak{Y}, \) and \( x_0 \in D^0. \) \( f \) is said to be analytic at \( x_0 \) if \( f \) is \( C_\infty \) at \( x_0, \) and \( \mathfrak{J} \in (0, \infty) \subseteq \mathbb{R}, \mathfrak{J} \in [0, \infty) \subseteq \mathbb{R}, \) and \( \exists M \in (0, \infty) \subseteq \mathbb{R}, \) such that \( \| f^{(n)}(x) \| \leq cnM^n, \forall n \in \mathbb{Z}_+, \forall x \in B_X(x_0, \mathfrak{J}). \) In this case, \( \forall \mathfrak{J} \in [0, \mathfrak{J}] \subseteq \mathbb{R} \) with \( M\mathfrak{J} < 1 \) is called an analytic radius of \( f \) at \( x_0. \) If \( D \supseteq D_1 = \mathcal{O}_x \) and \( f \) is analytic at \( x, \forall x \in D_1, \) then we say that \( f \) is analytic on \( D_1. \) When \( f \) is analytic on \( D, \) then we say that \( f \) is analytic or an analytic function.

Theorem 9.80 (Taylor Series) Let \( \mathfrak{X} \) and \( \mathfrak{Y} \) be normed linear spaces over \( \mathbb{K}, D \subseteq \mathfrak{X}, f : D \to \mathfrak{Y}, \) and \( x_0 \in D^0. \) Assume that \( f \) is analytic
at $x_0$ with an analytic radius $\delta \in (0, \infty) \subset \mathbb{R}$. Then, $\forall x \in \mathcal{B}_X(x_0, \delta)$,
\[
f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0) (x - x_0)^{n-\text{times}}
\]

**Proof** Since $f$ is analytic at $x_0 \in D^c$ with an analytic radius $\delta \in (0, \infty) \subset \mathbb{R}$, then $\exists c \in [0, \infty) \subset \mathbb{R}$ and $\exists M \in (0, \infty) \subset \mathbb{R}$ such that $\|f^{(n)}(x)\| \leq cn!M^n$, $\forall n \in \mathbb{Z}_+$, $\forall x \in \mathcal{B}_X(x_0, \delta) \subset D$, and $M\delta < 1$. Then, $\forall x \in \mathcal{B}_X(x_0, \delta)$, $\forall n \in \mathbb{Z}_+$, let
\[
R_n(x, x_0; f) := f(x) - \sum_{i=0}^{n} \frac{1}{i!} f^{(i)}(x_0) (x - x_0)^{i-\text{times}}
\]
By Taylor Theorem 9.48, $\exists \delta_0 \in (0, 1) \subset \mathbb{R}$ such that $\|R_n(x, x_0; f)\| \leq \frac{1}{(n+1)!} \|f^{(n+1)}(t_0x + (1 - t_0)x_0)\| \|x - x_0\|^{n+1} < c(M\delta)^{n+1}$. 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. $\Box$

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 $K$, $D_1 \subset X$, $f : D_1 \to B(Y, Z)$, $g : D_1 \to B(Z, W)$, $n \in \mathbb{Z}_+$, and $x_0 \in D_1^c$. Assume that $f$ and $g$ are $C_n$ at $x_0$. Define $h : D_1 \to B(Y, W)$ 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
\[
h^{(k)}(x)(\xi_{x_0})\cdots(\xi_{x_1})(\xi_y) = \sum_{i=0}^{k} \sum_{A_1 = \{j_1, \ldots, j_k\} \text{ is any } i \text{-element subset of } A = \{1, \ldots, k\}} \sum_{A_2 = \{j_{i+1}, \ldots, j_k\} = A \setminus A_1} g^{(i)}(x)(\xi_{x_{j_1}})\cdots(\xi_{x_{j_k}})
\]
$\forall x \in \mathcal{B}_X(x_0, \delta)$, $\forall \xi_{x_1} \in X$, $l = 1, \ldots, k$, $\forall \xi_y \in Y$, $\forall k \in \mathbb{Z}_+$ with $k \leq n$.

**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 $f^{(n)}(x) \in B_{\mathfrak{S}_n}(X, B(Y, Z))$ and $g^{(n)}(x) \in B_{\mathfrak{S}_n}(X, B(Z, W))$, $\forall x \in \mathcal{B}_X(x_0, \delta)$. We will use mathematical induction on $k$ to prove the result.

1° $k = 0$. We have $\forall x \in \mathcal{B}_X(x_0, \delta)$, $\forall \xi_y \in Y$,
\[
LHS = h^{(0)}(x)(\xi_y) = h(x)(\xi_y) = g(x)f(x)(\xi_y) = g(x)(f(x)(\xi_y)) = RHS
\]
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 $l < n$.  

9.9. ANALYTIC FUNCTIONS

3° Consider the case when \( k = l + 1 \). By the inductive assumption, \( \forall x \in B_X(x_0, \delta), \forall \xi_{x_s} \in X, s = 1, \ldots, l, \forall \xi_y \in Y, \)

\[
h^{(l)}(x)(\xi_{x_1}) \cdots (\xi_{x_l})(\xi_y) = \\
\sum_{i=0}^{l} \Lambda_1 := \{j_1, \ldots, j_i\} \text{ is any } i\text{-element subset of } \Lambda = \{1, \ldots, l\}, A := \{j_{i+1}, \ldots, j_l\} = \Lambda \setminus \Lambda_1 \}
\sum \left( g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\
(\xi_{x_{j_{i+1}}})(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y) \right) \\
(\xi_{x_{j_{i+1}}})(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y) \\
\right) \\
(\xi_{x_{j_{l+1}}})(\xi_{x_{j_{l+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y) \\
\right)
\end{align*}

Then, by Proposition 9.27, we have \( \forall x \in B_X(x_0, \delta), \forall \xi_{x_s} \in X, s = 1, \ldots, l+1, \forall \xi_y \in Y, \)

\[
h^{(l+1)}(x)(\xi_{x_{l+1}}) \cdots (\xi_{x_l})(\xi_y) = \tilde{h}^{(1)}(x)(\xi_{x_{l+1}}) \\
(\xi_{x_{l+1}}) \cdots (\xi_{x_l})(\xi_y) \\
(\xi_{x_{l+1}}) \cdots (\xi_{x_l})(\xi_y) \\
\right)
\end{align*}

where \( \tilde{h} : B_X(x_0, \delta) \rightarrow B(\mathbb{Y}, \mathbb{W}) \) is defined by \( \tilde{h}(x) = h^{(l)}(x)(\xi_{x_1}) \cdots (\xi_{x_l}), \forall x \in B_X(x_0, \delta), \). This implies that, by the Chain Rule (Theorem 9.18) and Proposition 9.42,

\[
h^{(l+1)}(x)(\xi_{x_{l+1}}) \cdots (\xi_{x_l})(\xi_y) \\
\sum_{i=0}^{l+1} \Lambda_1 := \{j_1, \ldots, j_i\} \text{ is any } i\text{-element subset of } \Lambda = \{1, \ldots, l\}, A := \{j_{i+1}, \ldots, j_{l+1}\} = \Lambda \setminus \Lambda_1 \}
\sum \left( g^{(i)}(x)(\xi_{x_{j_1}}) \cdots (\xi_{x_{j_i}}) \\
(\xi_{x_{j_{i+1}}})(\xi_{x_{j_{i+1}}}) \cdots (\xi_{x_{j_{l+1}}})(\xi_y) \\
(\xi_{x_{j_{l+1}}})(\xi_{x_{j_{l+1}}}) \cdots (\xi_{x_{j_l}})(\xi_y) \\
\right)
\end{align*}

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. \hfill \Box

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 $\exists \delta_2 \in (0, \infty) \subset \mathbb{R}$, $\exists c_2 \in [0, \infty) \subset \mathbb{R}$, and $\exists M_g \in (0, \infty) \subset \mathbb{R}$ such that $\|g^{(n)}(y)\| \leq c_2 n! M_g^n$, $\forall n \in \mathbb{Z}^+$, $\forall y \in B_y(y_0, \delta_2) \subseteq D_2$. Since $f$ is analytic at $x_0$, then $\exists \delta_1 \in (0, \infty) \subset \mathbb{R}$, $\exists c_1 \in [0, \infty) \subset \mathbb{R}$, and $\exists M_f \in (0, \infty) \subset \mathbb{R}$ such that $\|f^{(n)}(x)\| \leq c_1 n! M_f^n$ and $f(x) \in B_y(y_0, \delta_1)$, $\forall n \in \mathbb{Z}^+$, $\forall x \in B_x(x_0, \delta_1)$. By Proposition 9.45, $h$ is $C_\infty$ at $\forall x \in B_x(x_0, \delta_1)$.

Define nonnegative constants $K_{n,i} := \sup_{x \in B_x(x_0, \delta_1)} \|D^n (g^{(i)} \circ f)(x)\|$, $\forall i, n \in \mathbb{Z}^+$. Clearly, $\forall i \in \mathbb{Z}^+$, $\forall i \in \mathbb{Z}^+$, by Chain Rule (Theorem 9.18),

$$K_{n+1,i} = \sup_{x \in B_x(x_0, \delta_1)} \|D^{n+1} (g^{(i)} \circ f)(x)\| = \sup_{x \in B_x(x_0, \delta_1)} \|D^n ((g^{(i+1)} \circ f)(f^{(1)}))(x)\|$$

Define $l_{i+1} : B_x(x_0, \delta_1) \to B_{B_{H+1}}(y, \mathbb{Z})$ by $l_{i+1}(x) := g^{(i+1)}(f(x))$, $\forall x \in B_x(x_0, \delta_1)$, $\forall i \in \mathbb{Z}^+$. Then, $\forall i, n \in \mathbb{Z}^+$,

$$K_{n+1,i} = \sup_{x \in B_x(x_0, \delta_1)} \|D^n ((g^{(i+1)} \circ f)(f^{(1)}))(x)\|$$

$$= \sup_{x \in B_x(x_0, \delta_1)} \|D^n (l_{i+1} f^{(1)})(x)\|$$

$$= \sup_{x \in B_x(x_0, \delta_1)} \sup_{\xi \in \mathbb{X}, \xi \in \mathbb{X}, \|\xi\| \leq 1} \|D^n (l_{i+1} f^{(1)})(x)(\xi_{x_0}) \cdots (\xi_{x_1})(\xi_{x})\|$$

$$= \sum_{k=0}^{n} \sum_{\Lambda} \left( \sum_{J_{k+1}} \left( \sum_{J_{k+1}} (f^{(i+1)}(x))(\xi_{x_{j_1} j_1}) \cdots (\xi_{x_{j_k} j_k}) \right) \right)$$

where $\Lambda = \{1, \ldots, n\}$, $\Lambda_{k} = \{j_{k+1}, \ldots, j_{n}\} \subseteq \Lambda \setminus \Lambda_{k-1}$.
9.9. ANALYTIC FUNCTIONS

\[ \left\| f^{(n-k+1)}(x) \right\| \leq \sup_{x \in B_k(x_0, \delta)} \sum_{k=0}^{n} \left( \sum_{\Lambda_2 = (j_{k+1}, \ldots, j_n) = \Lambda \setminus \Lambda_1} \left( \sum_{\lambda_i = (i_1, \ldots, i_k) \text{ is any } k \text{-element subset of } \Lambda = (1, \ldots, n)} \right) \right) \left\| i^{(k)}_{i+1}(x) \right\| . \]

where the fourth equality follows from Proposition 9.81. Hence, \( 0 \leq K_{n+1, i} \leq \sum_{k=0}^{n} (n)_{k} K_{k, i+1} c_f(n-k+1)! M_f^{n-k+1}, \forall i, n \in \mathbb{Z}_+ \).

We will use mathematical induction on \( n \in \mathbb{N} \) to show that \( \forall n \in \mathbb{N}, \forall i \in \mathbb{Z}_+, \)

\[ 0 \leq K_{n, i} \leq (n+1)! c_g M_g^{i+1} M_f^{n-i} \sum_{j=0}^{n-1} \binom{n}{j+1} \binom{n-1}{j} \binom{n+i}{n-j-1}^{-1} c_f \frac{1}{M_f^{j+1}} \]

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,

\[ K_{l+1, i} \leq \sum_{k=0}^{l} \binom{l}{k} K_{k, i+1} c_f(l-k+1)! M_f^{l-k+1} \]

\[ \leq c_f (l+1)! M_f^{l+1} c_g(i+1)! M_g^{i+1} + \sum_{k=1}^{l} \binom{l}{k} c_f(l-k+1)! M_f^{l-k+1} \]

\[ - (k+i+1)! c_g M_g^{i+2} M_f^{k-1} \sum_{j=0}^{k-1} \binom{k}{j+1} \binom{k-1}{j} \binom{k+i+1}{k-j-1}^{-1} c_f \frac{1}{M_f^{j+1}} M_g^j \]

\[ = c_f c_g M_g^{i+1} M_f^{i+1} (l+i+1)! (l+1) \left( \binom{l+i+1}{l}^{-1} \right) \]

\[ + \sum_{k=1}^{l} \sum_{j=0}^{l-k-1} \frac{(l-k)!}{k!(l-k)!} (j+1)! (k-j-1)! c_f \frac{1}{M_f^{j+1}} M_g^{i+2} c_g M_f^{i+1} M_g^{j+i+2} \]

\[ = c_f c_g M_g^{i+1} M_f^{i+1} (l+i+1)! (l+1) \left( \binom{l+i+1}{l}^{-1} \right) \]

\[ + \sum_{j=0}^{l-1} \sum_{k=j+1}^{l} \frac{l!(l-k+1)!}{(l-k)!} \frac{(k-j-1)!}{(j+1)!} c_f \frac{1}{M_f^{j+1}} M_g^{i+2} c_g M_f^{i+1} M_g^{j+i+2} \]
This completes the induction process.

Hence, \( \forall n \in \mathbb{N} \),

\[
\sup_{x \in B_{\epsilon} (x_0, \delta)} \| h^{(n)} (x) \| = \sup_{x \in B_{\epsilon} (x_0, \delta)} \| D^n (g \circ f)(x) \| = K_{n,0} \\
\leq n! c_g M_g^n \sum_{j=0}^{n-1} {n \choose j+1} {n-1 \choose j} {n \choose n-j-1} c_f^{j+1} M_f^j = n! c_f c_g M_g^n M_f^n (1 + c_f M_g)^{n-1}
\]
By Proposition 9.56, \( f \) 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 \( \mathbb{K} \), \( D := \{ L \in B(X,Y) \mid L \text{ is bijective} \} \), 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 \( C_\infty \), 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 \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) 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 \( \| f(T) - f(A) \| \leq \| f(A) \| \| [T-A] \| \leq \| f(A) \|^2 \delta (1 - \| f(A) \| \| \delta \|^{-1})^{-1} \). This implies that \( \| f(T) \| \leq \| f(A) \| + \| f(A) \|^2 \delta (1 - \| f(A) \| \| \delta \|^{-1})^{-1} = \| f(A) \| (1 - \| \delta \|^{-1})^{-1} = \alpha \in [0, \infty) \subset \mathbb{R} \). This leads to

\[
\| f^{(1)}(T) \| = \sup_{\Delta \in B(X,Y), \| \Delta \| \leq 1} \| f^{(1)}(T)(\Delta) \| \leq \| f(T) \|^2 \leq \alpha^2
\]

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 \( n = k + 1 \). \( \forall T \in U \), \( \forall \Delta_1 \in B(X,Y) \),

\[
f^{(k)}(T)(\Delta_1) = -f(T)\Delta_1 f(T) = g_{\Delta_1}(T), \text{ where } g_{\Delta_1} : U \to B(Y,X) \text{. By Proposition 9.27, we have, } \forall \Delta_i \in B(X,Y), i = 2, \ldots, k + 1,
\]

\[
g^{(k)}_{\Delta_i}(T)(\Delta_{k+1}) \cdots (\Delta_2) = f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_2)(\Delta_1)
\]

This implies that

\[
\| f^{(k+1)}(T) \| = \sup_{\Delta_i \in B(X,Y), \| \Delta_i \| \leq 1, i=1,\ldots,k+1} \| f^{(k+1)}(T)(\Delta_{k+1}) \cdots (\Delta_1) \|
\]
\[ \exists \delta \in B(x, y), \| \delta \| \leq 1, \ldots, k+1 \]

By Proposition 9.81, we have the following equality,

\[
g^{(k)}_{\Delta_1}(T)(\Delta_{k+1}) \cdots (\Delta_2) = \\
- \sum_{i=0}^{k} \sum_{\Lambda_1 = \{j_1, \ldots, j_k\} \text{ is any } i\text{-element subset of } \Lambda = (2, \ldots, k+1): \Lambda_2 = (j_{k+1}, \ldots, j_k) = \Lambda \setminus \Lambda_1} f^{(i)}(T)(\Delta_{j_1}) \cdots (\Delta_{j_k})
\]

(\Delta_1 f^{(k-1)}(T)(\Delta_{j_1+1}) \cdots (\Delta_{j_k}))

Then,

\[
\| f^{(k+1)}(T) \| \leq \sum_{i=0}^{k} \sum_{\Lambda_1 = \{j_1, \ldots, j_k\} \text{ is any } i\text{-element subset of } \Lambda = (2, \ldots, k+1): \Lambda_2 = (j_{k+1}, \ldots, j_k) = \Lambda \setminus \Lambda_1} \| f^{(i)}(T) \| \| f^{(k-i)}(T) \|
\]

\[
= \sum_{i=0}^{k} \binom{k}{i} \| f^{(i)}(T) \| \| f^{(k-i)}(T) \|
\]

\[
\leq \sum_{i=0}^{k} \binom{k}{i} i! \alpha^{i+1} (k-i)! \alpha^{k+1-i} = \sum_{i=0}^{k} k! \alpha^{k+2} = (k+1)! \alpha^{k+2}
\]

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 \( \mathbb{K} \), \( D \subseteq X \), \( F : D \to Y \) be analytic at \( x_0 \in D^n \), and \( y_0 = F(x_0) \in Y \). Assume that \( F^{(1)}(x_0) \in \mathcal{B}(X, Y) \) is bijective. Then, \( \exists \delta \in (0, \infty) \subset \mathbb{R}, \exists U \subseteq D \) with \( x_0 \in U \in \mathcal{O}_X \) such that

(i) \( F|_U : U \to \mathcal{B}_Y (y_0, \delta) =: V \subseteq Y \) is bijective;

(ii) the inverse mapping \( F_i : V \to U \) is \( C_\infty \) and \( F_i^{(1)} : V \to \mathcal{B}(Y, X) \) is given by \( F^{(1)}(F_i(y))^{-1} \), \( \forall y \in V \);

(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 \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R}, \exists \delta \in [0, \infty) \subset \mathbb{R} \), and \( \exists M \in (0, \infty) \subset \mathbb{R} \) 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, \( \exists \delta_2 \in (0, \delta] \subset \mathbb{R} \) such that \( F_i(\mathcal{B}_Y(y_0, \delta_2)) \subseteq \bar{U} \).
9.9. ANALYTIC FUNCTIONS

Define \( h : \bar{U} \to B(y, X) \) 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, \( \exists \delta_1 \in (0, \delta] \subset \mathbb{R}, \exists c_h \in [0, \infty) \subset \mathbb{R}, \text{and } \exists M_h \in (0, \infty) \subset \mathbb{R} \text{ such that } \|h^{(n)}(x)\| \leq c_h n! M_h^n, \forall n \in \mathbb{Z}_+, \forall x \in B_X(x_0, \delta_1). \) By choosing \( \delta_1 \) sufficiently small, without loss of generality, we may take \( \delta_1 = \delta_1 \).

Note that \( F^{(1)}(x) = h(x), \forall x \in \bar{U} \). Then, \( h^{(1)}(x) = D(F^{(1)}(x)) = (F^{(2)}(x) - F(x))F^{(1)}(x) \) and \( F^{(2)}(x) = (h^{(1)}(x))h(x), \forall x \in \bar{U} \). Define \( h_k : \bar{U} \to B_{S_k}(y, X) \) by \( h_k(x) = F^{(k)}(x), \forall x \in \bar{U}, \forall 2 \leq k \in \mathbb{N} \). This leads to the recursive relationship

\[
(Dh_k(x))h(x) = F^{(k+1)}(x) \circ F(x) = h_{k+1}(x), \forall x \in \bar{U}, \forall 2 \leq k \in \mathbb{N}
\]

with the initial condition \( h_2(x) = h^{(1)}(x)h(x), \forall x \in \bar{U} \). Define the nonnegative constants \( K_{k,j} := \sup_{x \in \bar{U}} \|D^j h_2(x)\|, \forall j \in \mathbb{Z}_+, \forall 2 \leq k \in \mathbb{N} \). We will use mathematical induction on \( 2 \leq k \in \mathbb{N} \) to show that

\[
0 \leq K_{k,j} \leq \frac{(j + 2k - 2)!}{(k-1)!(k-2)!} c_h^k M_h^{j+k-1}, \forall j \in \mathbb{Z}_+, \forall 2 \leq k \in \mathbb{N}
\]

1° \( k = 2, \forall j \in \mathbb{Z}_+, \) by the initial condition of the recursion,

\[
K_{2,j} = \sup_{x \in \bar{U}} \|D^j h_2(x)\| = \sup_{x \in \bar{U}} \|D^j (h^{(1)}h)(x)\|
\]

\[
= \sup_{x \in \bar{U}} \sup_{\xi \in x \|\xi\| \leq 1} \|D^j (h^{(1)}h)(\xi_1, \ldots, \xi_j)(\xi_1, \xi_2)\|
\]

\[
= \sup_{x \in \bar{U}} \sup_{\xi \in x \|\xi\| \leq 1} \sum_{\Lambda_1 = \{x_1, \ldots, x_{j-1}\} \text{ is any } k-1\text{-element subset of } \Lambda = \{1, \ldots, j\}} \sum_{\Lambda_2 = \{x_{j+1}, \ldots, x_{j+\ell}\} = \Lambda \setminus \Lambda_1} \|h^{(j+\ell)}(x)(\xi_1, \ldots, \xi_{j+\ell})(\xi_1, \ldots, \xi_{j+\ell})\|
\]

\[
= \sup_{x \in \bar{U}} \sum_{\ell=0}^j \sum_{\Lambda_2 = \{x_{j+1}, \ldots, x_{j+\ell}\}} \|h^{(j+\ell)}(x)\| \|h^{(j-\ell)}(x)\|
\]

\[
= \sum_{\ell=0}^j \sum_{\Lambda_2 = \{x_{j+1}, \ldots, x_{j+\ell}\}} \frac{j!}{\ell!(j-\ell)!} c_h (l+1)! M_h^{l+1} c_h (j-l)! M_h^{j-l}
\]
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,

\[
K_{n+1,j} = \sup_{x \in U} \| D^j h_{n+1}(x) \| = \sup_{x \in U} \| D^j (h^{(1)}_n h)(x) \|
\]

\[
= \sup_{x \in U} \sup_{\xi_x \in X, \xi_y \in \mathbb{Y}, \| \xi_y \| \leq 1} \left\| D^j (h^{(1)}_n h)(x)(\xi_{x_j}) \cdots (\xi_{x_1})(\xi_y) \right\|
\]

\[
\leq \sup_{x \in U} \left\| \sum_{l=0}^j h^{(l+1)}_n(x)(\xi_{x_j}) \cdots (\xi_{x_1})(\xi_y) \right\|
\]

\[
\leq \sup_{x \in U} \sum_{l=0}^j \left( \frac{j!}{l!(j-l)!} \right) h^{(l+1)}_n(x) \| h^{(j-l)}(x) \|
\]

\[
= \sup_{x \in U} \sum_{l=0}^j \left( \frac{j!}{l!(j-l)!} \right) K_{n,l+1} \| h^{(j-l)}(x) \|
\]

\[
\leq \sum_{l=0}^j \left( \frac{j!}{l!(j-l)!} \right) \frac{c_h M_h^{l+n} c_h (j-l)! M_h^{j-l}}{(n-1)! 2^{n-1}}
\]

\[
= c_h^{n+1} M_h^{j+n} \frac{j!}{(n-1)! 2^{n-1}} \sum_{l=0}^j \prod_{s=1}^{2n-1} (l+s) = \frac{(j+2n)!}{n! 2^n} c_h^{n+1} M_h^{j+n} = \text{RHS}
\]
Based on this inequality, we have the following bounds for the high order derivatives of $F_i$ around $y_0$, $\forall y \in B_\delta(y_0, \delta_2)$, $\forall 2 \leq n \in \mathbb{N}$,

$$\|F_i^{(n)}(y)\| \leq \sup_{x \in U} \|F_i^{(n)} \circ F(x)\| = \sup_{x \in U} \|h_n(x)\| \leq K_{n,0}$$

$$\leq \frac{(2n - 2)!}{(n - 1)!2^{n-1}c_h^nM_h^{n-1}} \leq (n - 1)!2^{n-1}c_h^nM_h^{n-1} = \frac{1}{2M_hn}n!(2c_hM_h)$$

Hence, $F_i$ is analytic at $y_0$. This completes the proof of the theorem. □

**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 $r_0, r_1 \in (0, \infty) \subset \mathbb{R}$ 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 $\exists c \in [0, \infty) \subset \mathbb{R}$ and $\exists M \in (0, \infty) \subset \mathbb{R}$, $\|F^{(n)}(x, y)\| \leq cnM^n$, $\forall n \in \mathbb{Z}_+$, $\forall (x, y) \in U \times V$.

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,

$$\tilde{F}^{(1)}(x_0, y_0) = \begin{bmatrix} \text{id}_X & \vartheta_B(y, x) \\ \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \end{bmatrix}$$

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 $(x_0, y_0) \in U \in \mathcal{O}_{X \times Y}$. $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\tilde{F}\big|_U : \tilde{U} \to B_X \times Z((x_0, \vartheta_Z), \delta)$ is bijective, whose inverse function $\bar{F}_i : B_{X \times Z}((x_0, \vartheta_Z), \delta) \to \tilde{U}$ is analytic at $(x_0, \vartheta_Z)$. By taking $r_0$ and $r_1$ sufficiently small, we may assume that $\bar{U} \subseteq \tilde{U}$. Then, $\exists \delta_1 \in (0, \min\{\delta, r_0\}) \subset \mathbb{R}$, $\exists W \subseteq U \times V$ such that $\bar{F}_i |_W : B_{X \times Z}((x_0, \vartheta_Z), \delta_1) \to \tilde{U}$ is bijective. By the uniqueness of $y \in V$ that solves $F(x, y) = \vartheta_Z$, we have $\phi(x) = \pi_y \circ \bar{F}_i(x, \vartheta_Z)$, $\forall x \in B_X(x_0, \delta_1)$. By Theorem 9.82 and Proposition 9.83, $\phi$ is analytic at $x_0$. This completes the proof of the theorem. □
**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 $\exp(A) := \sum_{n=0}^{\infty} \frac{1}{n!} A^n$, 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 $\|A\|^n_{B(X,X)} < \infty$. 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 $A \in B(X,X)$. For all $t, \tau \in K$, we have

\( (i) \ h : K \to B(X,X) \text{ defined by } h(t) = \exp(At), \forall t \in K, \text{ is an analytic function.} \forall t \in K, h \text{ admits arbitrarily large analytic radius at } t. \)

\( (ii) \ \exp(A(t+\tau)) = \exp(At) \exp(A\tau) \text{ and } \exp(\partial_{B(X,X)}) = \text{id}_X; \)

\( (iii) \ \frac{d}{dt} \exp(At) = A \exp(At) = \exp(At)A; \)

\( (iv) \ \text{if } B \in B(X,X) \text{ and } AB = BA, \text{ then } \exp(At)B = B \exp(At), \text{ and } \exp(A)\exp(B) = \exp(B)\exp(A) = \exp(A+B). \)

\( (v) \ (\exp(A))^{-1} = \exp(-A) \text{ and } (\exp(A))^\prime = \exp(A'). \)

**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 $\frac{d}{dt} \exp(At) = \frac{d}{dt} \sum_{n=0}^{\infty} \frac{1}{n!} A^n t^n = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} A^n t^{n-1} = A \exp(At) = \exp(At)A$.

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, $\exp(At)A = \exp(At) \sum_{n=0}^{\infty} \frac{1}{n!} A^n B^n = \sum_{n=0}^{\infty} \frac{1}{n!} B^n \exp(A) = \exp(At)A$, where the third equality follows from the first. $B = B \exp(At)$. $\exp(A+B) = \sum_{n=0}^{\infty} \frac{1}{n!} (A+B)^n = \sum_{n=0}^{\infty} \frac{1}{n!} B^n \exp(A)\exp(B) = \exp(At)A\exp(B)$, where the second equality follows from the first and the third 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). $\exp(A)^\prime = \left( \sum_{n=0}^{\infty} \frac{1}{n!} (A')^n \right)^\prime = \sum_{n=0}^{\infty} \frac{1}{n!} (A')^n = \exp(A')$, 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 \( X := (X, \mathcal{O}) \) be a topological space, \( f: X \to \mathbb{R} \), and \( x_0 \in X \). \( x_0 \) is said to be a point of minimum for \( f \) if \( f(x_0) \leq f(x) \), \( \forall x \in X \). It is said to be the point of strict minimum for \( f \) if \( f(x_0) < f(x) \), \( \forall x \in X \setminus \{x_0\} \). 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 \( X \) be a real normed linear space, \( D \subseteq X, x_0 \in D, f: D \to \mathbb{R} \), \( u \in A_D(x_0) \). 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 \( X \) be a real normed linear space and \( A \in B_{S^2}(X, \mathbb{R}) \). \( A \) is said to be positive definite if \( \exists m \in (0, \infty) \subseteq \mathbb{R} \) such that \( A(x)(x) = \langle \langle Ax, x \rangle \rangle \geq m \|x\|^2 \), \( \forall x \in X \). 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 \( \exists m \in (0, \infty) \subset \mathbb{R} \) 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 \( S_{\text{nsd}} X \). 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 \( S_{\text{nsd}} X = -S_{\text{psd}} X \);

(ii) \( S_+ X \) and \( S_- X \) are open sets in \( B_{S_2}(X, \mathbb{R}) = S_X \);

(iii) \( S_{\text{psd}} X \) and \( S_{\text{nsd}} X \) are closed convex cones in \( S_X \);

(iv) \( S_+ X \subseteq S_{\text{psd}}^0 X \) and \( S_- X \subseteq S_{\text{nsd}}^0 X \).

**Proof**

(i) This is clear. (ii) Fix \( A \in S_+ X \). Then, \( \exists m \in (0, \infty) \subset \mathbb{R} \) 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 \( S_X \). Therefore, \( S_- X = -S_+ X \) is open in \( S_X \).

(iii) Clearly, \( \partial B_{S_2}(X, \mathbb{R}) \in S_{\text{psd}} X \). \( \forall A \in S_{\text{psd}} X, \forall \alpha \in [0, \infty) \subset \mathbb{R}, \alpha A \in S_{\text{psd}} X \). 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 \( m \in (0, \infty) \subset \mathbb{R} \) 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 \( B(x_0)(x_0) = A(x_0)(x_0) + (B - A)(x_0)(x_0) < -m \| x_0 \|^2 + \| B - A \| \| x_0 \|^2 \leq -m \| x_0 \|^2/2 < 0 \), where the first inequality follows from Proposition 7.64. This implies that \( B \in M \) and \( A \in M^c \). Then, \( M = M^c \). Hence, \( S_{\text{psd}} X \) is open. \( S_{\text{nsd}} X = -S_{\text{psd}} X \) is clearly also a closed convex cone.

(iv) Clearly, \( S_+ X \subseteq S_{\text{psd}} X \). Then, \( S_+ X \subseteq S_{\text{psd}}^0 X \). This further implies that \( S_- X \subseteq S_{\text{nsd}}^0 X \). This completes the proof of the proposition. \( \square \)

**Definition 10.5** Let \( (\mathcal{X}, \mathbb{K}) \) 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 \( \forall x_1, x_2 \in K, \forall \alpha \in (0, 1) \subset \mathbb{R}, \alpha x_1 + (1 - \alpha)x_2 \in M \) 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 \( x_0 \neq \partial X \), 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.

\[ \square \]

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 \( X_{weak} \). Let \( \mathcal{M} = \{ E \subseteq M \mid E \) is a weakly compact extreme subset of \( K \} \). Clearly, \( \mathcal{M} \) is empty. \( \exists \) defines an antisymmetric partial ordering on \( \mathcal{M} \), where smaller sets are further down the stream. Next, we will use Zorn’s Lemma to show that \( \mathcal{M} \) admits a maximal element.

Let \( \mathcal{E} \subseteq \mathcal{M} \) be a nonempty totally ordered (by \( \supseteq \) ) subcollection. Let \( E_0 := \bigcap_{E \in \mathcal{E}} E \). \( \forall E \in \mathcal{E}, E \subseteq M \) 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 \( \alpha x_1 + (1 - \alpha)x_2 \in E_0 \). \( \forall E \in \mathcal{E}, \alpha x_1 + (1 - \alpha)x_2 \in E \). 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 \( E_0 \neq \emptyset \). \( \forall E \in \mathcal{E}, E \) is nonempty. Since \( \mathcal{E} \) is totally ordered by \( \supseteq \), then the intersection of finite number of sets in \( \mathcal{E} \) is again in \( \mathcal{E} \), and hence nonempty. By Proposition 5.12, \( E_0 \neq \emptyset \). Then, \( E_0 \) is a weakly compact extreme subset of \( K \). Clearly, \( E_0 \) is an upper bound of \( \mathcal{E} \) (in terms of \( \supseteq \)). Then, by Zorn’s Lemma, \( \mathcal{M} \) 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 \( f(\alpha (x_2 - x_1)) = \alpha, \forall \alpha \in \mathbb{R} \). Clearly, \( f \) is a linear functional on \( N \), and \( \| f \|_{\mathcal{N}} = 1/\| x_2 - x_1 \| < \infty \). By Hahn-Banach Theorem 7.83, there exists \( x_0 \in X^* \) with \( \| x_0 \| = 1/\| x_2 - x_1 \| \) such that \( \langle \langle x_0, \alpha (x_2 - x_1) \rangle \rangle = \alpha, \forall \alpha \in \mathbb{R} \). 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 \( X_{weak} \) 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 \( x_0 \in E_M \). 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, \( E_m \supseteq \emptyset \) is nonempty. Note that \( H \) is weakly closed. Then, by Proposition 5.5, \( E_m \subseteq E_M \) is weakly compact. \( \forall \bar{x}_1, \bar{x}_2 \in K, \forall \tilde{\alpha} \in (0, 1) \subseteq \mathbb{R} \), let \( \tilde{\alpha}\bar{x}_1 + (1 - \tilde{\alpha})\bar{x}_2 \in E_M \subseteq E_M \). Since \( E_M \) is an extreme subset of \( K \), then we have \( \bar{x}_1, \bar{x}_2 \in E_M \).
This further implies that \( \langle x_0, \bar{x}_1 \rangle \geq c \) and \( \langle x_0, \bar{x}_2 \rangle \geq c \). Note that
\[
\langle x_0, \alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2 \rangle = c.
\]
Then, we must have \( \langle x_0, \bar{x}_1 \rangle = c = \langle x_0, \bar{x}_2 \rangle \) and \( \bar{x}_1, \bar{x}_2 \in E_m \). This shows that \( E_m \) is a extreme subset of \( K \). Then \( E_m \in M \). 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. \( \square \)

**Proposition 10.8** Let \( X \) be a real reflexive Banach space, \( K \subseteq X \) be a nonempty bounded closed convex set, and \( E \) be the set of extreme points of \( K \). Then, \( K = \overline{\text{co}}(E) \).

**Proof** Let \( C := \overline{\text{co}}(E) \). By Proposition 7.15, \( C \) is closed and convex and \( C \subseteq K \). We will prove the result by an argument of contradiction. Suppose \( K \supsetneq C \). Then, \( \exists x_0 \in K \setminus C \). By Proposition 8.10, there exists \( x_0 \in X^* \) such that \( \langle x_0, x_0 \rangle < \inf_{x \in C} \langle x_0, x \rangle \). 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, \( c_0 := \langle x_0, x_1 \rangle = \inf_{x \in K} \langle x_0, x \rangle \in \mathbb{R} \) for some \( x_1 \in K \). Clearly, \( c_0 \leq \langle x_0, x_0 \rangle < \inf_{x \in C} \langle x_0, x \rangle \). 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 \( H := \{ x \in X \mid \langle x_0, x \rangle = c_0 \} \). \( H \) 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 \( x_m \in C_m \). 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, \( K = \overline{\text{co}}(E) \).

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

**Proposition 10.9** Let \( X \) be a real vector space, \( \Omega \subseteq X \) 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) \( \forall n \in \mathbb{N}, \forall x_1, \ldots, x_n \in \Omega, \forall \alpha_1, \ldots, \alpha_n \in [0, 1] \subseteq \mathbb{R} \text{ with } \sum_{i=1}^n \alpha_i = 1 \), we have \( f_1 \left( \sum_{i=1}^n \alpha_i x_i \right) \leq \sum_{i=1}^n \alpha_i f_1(x_i) \). 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 \( f_1 \left( \sum_{i=1}^n \alpha_i x_i \right) < \sum_{i=1}^n \alpha_i f_1(x_i) \).

(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 \( Y \) be a real vector space, \( A : Y \to X \) be an affine operator, \( D \subseteq Y \) be convex, and \( A(D) \subseteq \Omega \). Then, \( f_1 \circ A|_D : D \to \mathbb{R} \) is convex.
10.1. BASIC NOTION

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

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

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

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

(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

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

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

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_1)}{x_3 - x_1}$$

(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

$$f(y) - f(x) \geq \lim_{\alpha \to 0^+} (f(x + \alpha(y - x)) - f(x))/\alpha = Df(x; y - x)$$

$$= f^{(1)}(x)(y - x)$$
By Proposition 10.9, \( f \) is convex.

"Sufficiency" Let \( f(y) \geq f(x) + f^{(1)}(x)(y - x) \), \( \forall x, y \in \Omega \). \( \forall x_1, x_2 \in \Omega, \forall \alpha \in [0, 1] \subset \mathbb{R} \), 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, \( f(x_1) = f(x + (1 - \alpha)(x_1 - x_2)) \geq f(x) + (1 - \alpha)f^{(1)}(x)(x_1 - x_2) \). Then, \( \alpha f(x_1) + (1 - \alpha)f(x_2) \geq f(x) \). 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 \( y - x \in A \Omega(x) \). \( \forall \alpha \in (0, 1) \subset \mathbb{R} \), 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 a affine operator and \( A(I) \subseteq \Omega \), where \( I = [0, 1] \subset \mathbb{R} \).

By Proposition 10.9, \( g := f \circ A|_I \) 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 \( x_1 \neq x_2 \), \( \forall \alpha \in (0, 1) \subset \mathbb{R} \), 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, \( \forall x \in \Omega \). \( \frac{\partial}{\partial x} f \).

(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, \( \exists t_0 \in (0, 1) \subset \mathbb{R} \) we have

\[
f(y) = f(x) + f^{(1)}(x)(y - x) + \frac{1}{2} f^{(2)}(t_0 y + (1 - t_0)x)(y - x)(y - x)
\]

By the assumption, \( f^{(2)}(t_0 y + (1 - t_0)x)(y - x)(y - x) \geq 0 \). Then, \( f(y) \geq f(x) + f^{(1)}(x)(y - x) \). 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, \( \exists h \in \mathbb{X} \) such that \( f^{(2)}(x_0)(h)(h) < 0 \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_0 \in \Omega^\circ \); Case 2: \( x_0 \notin \Omega^\circ \). Case 1: \( x_0 \in \Omega^\circ \). By the continuity of \( f^{(2)} \) and \( x_0 \in \Omega^\circ \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) 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, \( \exists \delta_0 \in (0, 1) \subset \mathbb{R} \) we have

\[
f(x_1) = f(x_0) + f^{(1)}(x_0)(x_1 - x_0) + \frac{\delta^2}{2} f^{(2)}(x_0 + \delta_0 h)(h)(h)
\]

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: \( x_0 \notin \Omega^\circ \). Then, \( x_0 \in (\Omega^\circ \cap \Omega^\circ \setminus \Omega^\circ) \). Then, by Proposition 4.13, there exists \( (x_n)_{n=1}^\infty \subseteq \Omega^\circ \) 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 \( x_{n_0} \in \Omega^\circ \). 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

\[
\mu_0 := \inf_{x \in \Omega} f(x) \tag{10.1}
\]

where \( \mathbb{X} \) is a real normed linear space, \( \Omega \subseteq \mathbb{X} \) is a set, and \( f : \Omega \to \mathbb{R} \) is a functional.

**Proposition 10.13** Let \( \mathbb{X} \) be a real normed linear space, \( \Omega \subseteq \mathbb{X} \) 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, \( \exists \varepsilon \in (0, \infty) \subseteq \mathbb{R} \) such that \( f(x) \geq f(x_0) \), \( \forall x \in \Omega \cap \mathcal{B}_{x_0}(\varepsilon) \). \( \forall y \in \Omega \), \( \exists \alpha \in (0, 1) \subseteq \mathbb{R} \) such that \( x_0 + \alpha (y - x_0) \in \mathcal{B}_{x_0}(\varepsilon) \). 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, \( \exists \varepsilon \in (0, \infty) \subseteq \mathbb{R} \) such that \( f(x) > f(x_0) \), \( \forall x \in (\Omega \cap \mathcal{B}_{x_0}(\varepsilon)) \setminus \{ x_0 \} \). \( \forall y \in \Omega \) with \( y \neq x_0 \), \( \exists \alpha \in (0, 1) \subseteq \mathbb{R} \) such that \( x_0 + \alpha (y - x_0) \in \mathcal{B}_{x_0}(\varepsilon) \). Then, \( f(x_0) < f(x_0 + \alpha (y - x_0)) = f(\alpha y + (1 - \alpha)x_0) \leq \alpha f(y) + (1 - \alpha)f(x_0) \). This implies that \( f(x_0) < f(y) \). 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 \( I = (-\infty, f(x_1)) \subseteq \mathbb{R} \). 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 \( f^{(1)}(x_0) = \partial{\mathcal{X}}^\circ \).

**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 \( f^{(1)}(x_0) = \partial{\mathcal{X}}^\circ \). This completes the proof of the proposition.
10.2. UNCONSTRAINED OPTIMIZATION

Proposition 10.16 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^0$, and $f : \Omega \rightarrow \mathbb{R}$ be $C_2$ at $x_0$. Assume that $x_0$ is a point of relative minimum for $f$, then $f'(x_0) = \nabla \chi^*$. and $f''(x_0)$ is positive semi-definite.

Proof By Proposition 10.15, $f'(x_0) = \nabla \chi^*$. We will show that $f''(x_0)$ is positive semi-definite by an argument of contradiction. Suppose $f''(x_0)$ is not positive semi-definite. Then, $\exists h \in \mathcal{X}$ such that $f''(x_0)(h)(h) < 0$. By the continuity of $f''(x_0)$ at $x_0$, there exists $\delta \in (0, \infty) \subseteq \mathbb{R}$ with $D := B_{\chi}(x_0, \delta) \subseteq \Omega$ such that $f''(x_0)(h)(h) < 0, \forall x \in D$. 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

\[
 f(x_1) = f(x_0) + \frac{1}{2} f''(x_0 + t_0 \alpha h)(x_1 - x_0)(x_1 - x_0) \\
 = f(x_0) + \alpha^2 f''(x_0 + t_0 \alpha h)(h)(h) < f(x_0)
\]

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. \hfill \Box

Proposition 10.17 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^0$, and $f : \Omega \rightarrow \mathbb{R}$ be twice differentiable at $x$, $\forall x \in D := B_{\chi}(x_0, \delta_0) \subseteq \Omega$, where $\delta_0 \in (0, \infty) \subseteq \mathbb{R}$ is some constant. Assume that $f'(x_0) = \nabla \chi^*$ and $f''(x) = \alpha h$ 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

\[
 f(x) = f(x_0) + \frac{1}{2} f''(x_0 + t_0 (x - x_0))(x - x_0)(x - x_0) \geq f(x_0)
\]

Hence, $x_0$ is a point of relative minimum for $f$. This completes the proof of the proposition. \hfill \Box

Proposition 10.18 Let $\mathcal{X}$ be a real normed linear space, $\Omega \subseteq \mathcal{X}$, $x_0 \in \Omega^0$, and $f : \Omega \rightarrow \mathbb{R}$ be $C_2$ at $x_0$. Assume that $f'(x_0) = \nabla \chi^*$ and $f''(x_0)$ is positive definite. Then, $x_0$ is a point of relative strict minimum for $f$.

Proof By Proposition 10.4 and the continuity of $f''(x)$ at $x_0$, $\exists \delta \in (0, \infty) \subseteq \mathbb{R}$ such that $f''(x)$ is positive definite, $\forall x \in D := B_{\chi}(x_0, \delta) \subseteq \Omega$. $\forall x \in D \setminus \{x_0\}$, by Taylor’s Theorem, there exists $t_0 \in (0, 1) \subseteq \mathbb{R}$ such that

\[
 f(x) = f(x_0) + \frac{1}{2} f''(x_0 + t_0 (x - x_0))(x - x_0)(x - x_0) > f(x_0)
\]

Hence, $x_0$ is a point of relative strict minimum for $f$. This completes the proof of the proposition. \hfill \Box
10.3 Optimization with Equality Constraints

The basic problem to be considered in this section is

\[ \mu_0 := \inf_{x \in \Omega} f(x) \quad \text{subject to } H(x) = \vartheta_y \]

(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, \( \Omega \subseteq \mathcal{X} \), \( H : \Omega \to \mathcal{Y} \) be \( C_1 \) at \( x_0 \in \Omega^0 \). \( x_0 \) is said to be a regular point of \( H \) if \( H^{(1)}(x_0) \in B(\mathcal{X}, \mathcal{Y}) \) is surjective.

**Lemma 10.20** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real Banach spaces, \( \Omega \subseteq \mathcal{X} \), \( f : \Omega \to \mathbb{R} \) 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, \( \forall u \in \mathcal{X} \) with \( H^{(1)}(x_0)u = \vartheta_y \), we have \( f^{(1)}(x_0)u = 0 \).

**Proof** Define \( T : \Omega \to \mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \) by \( T(x) = (f(x), H(x), \vartheta_y) \), \( \forall x \in \Omega \). By Proposition 9.44, \( T \) is \( C_1 \) at \( x_0 \) and \( T^{(1)}(x_0) = \begin{bmatrix} f^{(1)}(x_0) \\ H^{(1)}(x_0) \end{bmatrix} \).

We will prove the result using an argument of contradiction. Suppose the result is not true. Then, \( \exists u_0 \in \mathcal{X} \) with \( H^{(1)}(x_0)u_0 = \vartheta_y \), we have \( f^{(1)}(x_0)u_0 \neq 0 \). We will show that \( T^{(1)}(x_0) \) is surjective. \( \forall (r, y) \in \mathcal{X} \times \mathcal{Y} \), by \( x_0 \) being a regular point of \( H \), \( \exists u_1 \in \mathcal{X} \) such that \( H^{(1)}(x_0)u_1 = y \).

Let \( u = \frac{f^{(1)}(x_0)u_1}{f^{(1)}(x_0)u_0}u_0 + u_1 \). Then, \( T^{(1)}(x_0)u = (r, y) \). Hence, \( T^{(1)}(x_0) \) is surjective.

Note that \( T(x_0) = (f(x_0), \vartheta_y) \) and \( \mathcal{X} \times \mathcal{Y} \) is a real Banach space, by Propositions 7.22 and 4.31. By Surjective Mapping Theorem 9.53, \( \exists \delta_r \in (0, \infty) \subset \mathcal{R} \), \( \exists \delta \in (0, \infty) \subset \mathcal{R} \), and \( c_1 \in (0, \infty) \subset \mathcal{R} \) with \( c_1 \delta \leq \delta_r \) such that \( \forall (r, y) \in B_{\mathcal{X} \times \mathcal{Y}} (T(x_0), \delta_r) \), \( \exists x \in B_{\mathcal{X}} (x_0, \delta_r) \subseteq \Omega \) with \( \| x - x_0 \| \leq c_1 \| (r, y) - T(x_0) \| \), we have \( T(x) = (r, y) \). Then, \( \forall \delta_r \in (0, \delta_r) \subset \mathcal{R} \), let \( r_1 = f(x_0) - \delta_r \in \mathcal{R} \), \( (r_1, \vartheta_y) \in B_{\mathcal{X} \times \mathcal{Y}} (T(x_0), \delta_r) \). Then, \( \exists x_1 \in \Omega \) with \( \| x_1 - x_0 \| \leq c_1 \| x_1 \| - T(x_0) \| = c_1 \| r_1 - f(x_0) \| = c_1 \delta_r < (1 + c_1 \delta_r) \), such that \( T(x_1) = (r_1, \vartheta_y) \). Then, \( H(x_1) = \vartheta_y \) and \( x_1 \in \Omega_c \cap B_{\mathcal{X}} (x_0, (1 + c_1 \delta_r)) \).

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.

**Proposition 10.21 (Lagrange Multiplier)** Let \( \mathcal{X} \) and \( \mathcal{Y} \) be real Banach spaces, \( \Omega \subseteq \mathcal{X} \), \( f : \Omega \to \mathbb{R} \) 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 \( \Omega_c := \{ x \in \Omega \mid H(x) = \partial_y \} \); and \( x_0 \) is a regular point of \( H \). Then, there exists a Lagrange multiplier \( y_0 \in \mathbb{Y}^* \) such that the Lagrangian \( L : \Omega \times \mathbb{Y}^* \rightarrow \mathbb{R} \) defined by \( L(x,y) = f(x) + \langle (y_*,H(x)) \rangle \), \( \forall (x,y) \in \Omega \times \mathbb{Y}^* \), is stationary at \( (x_0,y_0) \), that is \( L^{(1)}(x_0,y_0) = \partial_{B(\mathbb{X} \times \mathbb{Y}^*,\mathbb{R})} \).

**Proof** By Lemma 10.20, \( \forall u \in \mathbb{X} \) with \( H^{(1)}(x_0)u = \partial_y \), we have \( f^{(1)}(x_0) = 0 \). Then, \( f^{(1)}(x_0) \in \left( \mathcal{N} \left( H^{(1)}(x_0) \right) \right)^\perp \). Since \( x_0 \) is a regular point of \( H \), then \( \mathcal{R} \left( H^{(1)}(x_0) \right) = \mathbb{Y} \) is closed. By Proposition 7.114, we have \( \left( \mathcal{N} \left( H^{(1)}(x_0) \right) \right)^\perp = \mathcal{R} \left( \left( H^{(1)}(x_0) \right)' \right) \). Then, there exists \( y_0 \in \mathbb{Y}^* \) such that \( f^{(1)}(x_0) = \left( H^{(1)}(x_0) \right)' y_0 \). By Propositions 9.34, 9.41, 9.38, 9.37, 9.44, and 9.45, the Lagrangian \( L \) is \( C_1 \) at \( (x_0,y_0) \) and, \( \forall (u,v) \in \mathbb{X} \times \mathbb{Y}^* \),

\[
L^{(1)}(x_0,y_0)(u,v) = f^{(1)}(x_0)u + \langle (v_*,H(x_0)) \rangle + \langle (y_0,H^{(1)}(x_0)u) \rangle = 0
\]

where the second equality follows from that fact that \( x_0 \in \Omega_c \). Hence, \( L^{(1)}(x_0,y_0) = \partial_{B(\mathbb{X} \times \mathbb{Y}^*,\mathbb{R})} \). This completes the proof of the proposition. \( \square \)

**Proposition 10.22 (Generalized Lagrange Multiplier)** Let \( \mathbb{X} \) and \( \mathbb{Y} \) be real Banach spaces, \( \Omega \subseteq \mathbb{X} \), \( f : \Omega \rightarrow \mathbb{R} \) be \( C_1 \) at \( x_0 \in \Omega^* \), and \( H : \Omega \rightarrow \mathbb{Y} \) 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 \( \Omega_c := \{ x \in \Omega \mid H(x) = \partial_y \} \); and \( \mathcal{R} \left( H^{(1)}(x_0) \right) \subseteq \mathbb{Y} \) is closed. Then, there exists a Lagrange multiplier \( (r_0,y_0) \in \mathbb{R} \times \mathbb{Y}^* \) with \( (r_0,y_0) \neq (0,\partial_y) \) such that the Lagrangian \( L : \Omega \times \mathbb{Y}^* \rightarrow \mathbb{R} \) defined by \( L(x,y) = r_0 f(x) + \langle (y_*,H(x)) \rangle \), \( \forall (x,y) \in \Omega \times \mathbb{Y}^* \), is stationary at \( (x_0,y_0) \), that is \( L^{(1)}(x_0,y_0) = \partial_{B(\mathbb{X} \times \mathbb{Y}^*,\mathbb{R})} \).

**Proof** We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mathcal{R} \left( H^{(1)}(x_0) \right) = \mathbb{Y} \). Take \( r_0 = 1 \). Clearly, \( x_0 \) is a regular point of \( H \). The result follows from Proposition 10.21.

Case 2: \( \mathcal{R} \left( H^{(1)}(x_0) \right) \subset \mathbb{Y} \). Take \( r_0 = 0 \). Clearly, \( x_0 \) is not a regular point of \( H \). Then, \( \exists y_0 \in \mathbb{Y} \setminus \mathcal{R} \left( H^{(1)}(x_0) \right) \). Let \( M = \mathcal{R} \left( H^{(1)}(x_0) \right) \), 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 \( \delta = \max_{y_* \in M^*} \| y_* \| \leq 1 \| (y_*,y_0) \| \), where the maximum is achieved at \( y_0 \in M^* \). Then, \( y_0 \neq \partial_y \) and, by Proposition 7.112, \( y_0 \in \mathcal{N} \left( \left( H^{(1)}(x_0) \right)' \right) \).

The Lagrangian \( L \) is given by \( L(x,y) = \langle (y_*,H(x)) \rangle \), \( \forall (x,y) \in \Omega \times \mathbb{Y}^* \). By Propositions 9.34, 9.41, 9.37, 9.44, and 9.45, \( L \) is \( C_1 \) at \( (x_0,y_0) \) and, \( \forall (u,v) \in \mathbb{X} \times \mathbb{Y}^* \),

\[
L^{(1)}(x_0,y_0)(u,v) = \langle (v_*,H(x_0)) \rangle + \langle (y_0,H^{(1)}(x_0)u) \rangle = \langle \left( \left( H^{(1)}(x_0) \right)' y_0, u \right) \rangle = \langle (y_*^*,u) \rangle = 0
\]
Then, Claim 10.23.1 follows from that fact that \( x_0 \in \Omega_c \). Hence, \( L^{(1)}(x_0, y_{*0}) = \partial_{B(\chi \times y_{*} \cdot \mathbb{R})} \). This case is proved.

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

**Proposition 10.23** Let \( \mathcal{X} \) and \( y \) be real Banach spaces, \( \Omega \subseteq \mathcal{X} \), \( f : \Omega \to \mathbb{R} \) be \( C_2 \) at \( x_0 \in \Omega^o \), \( H : \Omega \to \mathbb{R} \) be \( C_2 \) at \( x_0 \). Consider the optimization problem (10.2). Assume that

(i) \( H(x_0) = \vartheta_y \) and \( \mathcal{R}(H^{(1)}(x_0)) \subseteq \mathcal{Y} \) is closed;

(ii) the Lagrangian \( L : \Omega \times \mathcal{Y}^* \to \mathbb{R} \) defined by \( L(x, y_{*}) = f(x) + \langle (y_{*}, H(x)) \rangle \), \( \forall (x, y_{*}) \in \Omega \times \mathcal{Y}^* \), is stationary at \( (x_0, y_{*0}) \), where \( y_{*0} \in \mathcal{Y}^* \) is a Lagrange multiplier;

(iii) \( \partial_{y_{*0}}^2 L(x_0, y_{*0}) \) is positive definite on the subspace \( \mathcal{M} := \mathcal{N}(H^{(1)}(x_0)) \), that is, \( \exists m \in (0, \infty) \in \mathbb{R} \) such that \( \partial_{y_{*0}}^2 L(x_0, y_{*0})(h) \geq m \| h \|^2 \), \( \forall h \in \mathcal{N}(H^{(1)}(x_0)) \).

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 \( (x_0, y_{*0}) \). By Proposition 9.9 and (ii),

\[
L^{(1)}(x_0, y_{*0}) = \left[ \frac{\partial L}{\partial y_{*}}(x_0, y_{*0}) \right] = \partial_{B(\chi \times y_{*} \cdot \mathbb{R})}
\]

Then, \( \exists \delta_0 \in (0, \infty) \in \mathbb{R} \) such that \( f^{(2)}(x), H^{(2)}(x), \) and \( \partial_{y_{*0}}^2 L(x, y_{*}) \) exists, \( \forall x \in \mathcal{B}_\chi(x_0, \delta_0) \subseteq \Omega \) and \( \forall y_{*} \in \mathcal{B}_y(\delta_0, y_{*0}) \), and, by Proposition 9.46, \( \partial_{y_{*0}}^2 L \) is continuous at \( (x_0, y_{*0}) \). Then, \( \exists \delta_1 \in (0, \delta_0) \in \mathbb{R} \) and \( \exists \delta_1 \in [0, \infty) \in \mathbb{R} \) such that \( \| \partial_{y_{*0}}^2 L(x, y_{*0}) - \partial_{y_{*0}}^2 L(x_0, y_{*0}) \| < m/5 \) and \( \| H^{(2)}(x) \| \leq c_1 \), \( \forall x \in D_1 := \mathcal{B}_\chi(x_0, \delta_1) \). By Propositions 7.114 and 7.98 and (i), \( \mathcal{M}^1 = \mathcal{R}\left(\left(H^{(1)}(x_0)\right)^{\dagger}\right) \) is closed. Then, by Proposition 7.113, \( \exists \epsilon_2 \in [0, \infty) \in \mathbb{R} \) such that, \( \forall x_{*} \in \mathcal{R}\left(\left(H^{(1)}(x_0)\right)^{\dagger}\right) \), there exists \( y_{*} \in \mathcal{Y}^* \) such that \( x_{*} = \left(H^{(1)}(x_0)\right)^{\dagger} y_{*} \) and \( \| y_{*} \| \leq c_2 \| x_{*} \| \).

**Claim 10.23.1** \( \forall x \in \Omega_c \cap D_1 \), \( \exists \delta_0 \in M \) such that \( \| x - x_0 - h_0 \| \leq c_1 \| x - x_0 \|^2, \| x - x_0 \| - c_1 \| x - x_0 \| \leq \| h_0 \| \leq \| x - x_0 \| + c_1 \| x - x_0 \| \).

**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, \( \exists \delta_0 \in (0, 1) \in \mathbb{R} \) such that

\[
\| H^{(1)}(x_0)(x - x_0) \| = \| H(x) - H(x_0) - H^{(1)}(x_0)(x - x_0) \| \leq \frac{1}{2} \| H^{(2)}(t(x_0 + (1 - t_0)x_0))(x - x_0) \| \leq c_1 \| x - x_0 \|^2 / 2
\]

By Proposition 7.97, we have \( \inf_{x \in M} \| x - x_0 - h \| = \max_{x \in M, \| x \| \leq 1} \| (x_{*}, x - x_0) \| \), where the maximum is achieved at \( x_{*1} \in M^* = \mathcal{R}\left(\left(H^{(1)}(x_0)\right)^{\dagger}\right) \) with \( \| x_{*1} \| \leq 1 \). Then, \( \exists y_{*1} \in \mathcal{Y}^* \) such that
10.3. Optimization with equality constraints

\[ x_{x_1} = (H^{(1)}(x_0))^\prime_{y_{x_1}} \text{ and } \|y_{x_1}\| \leq c_2. \] 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 (y_{x_1}, H^{(1)}(x_0)(x - x_0)) \rangle = 2 \langle (y_{x_1}, H^{(1)}(x_0)(x - x_0)) \rangle \leq 2 \|y_{x_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. \( \square \)

Let \( c_3 = \left\| \frac{\partial^2 L}{\partial x \partial y}(x_0, y_0) \right\| \). Let \( \delta \in (0, \delta_1] \subset \mathbb{R} \) such that \( c_1 c_2 \delta^2 \leq 1/4, \) \( c_1 c_2^2 (c_3 + m/5) \delta^2 \leq m/5, \) and \( 5 c_1 c_2^2 (c_3 + m/5) \delta^2 \leq m/5. \) \( \forall x \in \Omega_c \cap B_{\mathbb{R}^2} (x_0, \delta), \) 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, \( \exists t_1 \in (0, 1) \subset \mathbb{R} \) such that

\[
\begin{align*}
f(x) - f(x_0) &= L(x, y_0) - L(x_0, y_0) - \frac{\partial L}{\partial x}(x_0, y_0) (x - x_0) \\
&= \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(x, y_0) (h_0)(h_0) + \frac{\partial^2 L}{\partial x^2}(x, y_0) (x - x_0 - h_0)(h_0) \\
&+ \frac{\partial^2 L}{\partial x^2}(x, y_0) (x - x_0 - h_0)(x - x_0 - h_0) \\
&\geq \frac{1}{2} \left( \frac{\partial^2 L}{\partial x^2}(x, y_0) (h_0)(h_0) + \frac{\partial^2 L}{\partial x^2}(x, y_0) (x - x_0 - h_0)(h_0) \\
&\quad - \left\| \frac{\partial^2 L}{\partial x^2}(x, y_0) \right\| (2 \|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2) \right) \\
&\geq \frac{1}{2} \left( m \|h_0\|^2 - m \|h_0\|/5 - \left\| \frac{\partial^2 L}{\partial x^2}(x, y_0) \right\| \right) (2 \|x - x_0 - h_0\| \|h_0\| + \|x - x_0 - h_0\|^2) \\
&\geq \frac{1}{2} \left( m \|h_0\|^2 - (c_3 + m/5) (\frac{5}{2} c_1 c_2 \|x - x_0\|^3 + c_1 c_2^2 \|x - x_0\|^4) \right) \\
&\geq \frac{1}{2} \left( m \|x_0\|^2 - \frac{m}{16} \frac{c_3 + m/5}{2} \|x_0\|^3 + c_1 c_2^2 \|x_0\|^4 \right) \\
&\geq \frac{1}{2} \left( m \|x_0\|^2 - \frac{m}{16} (c_3 + m/5) \|x_0\|^3 + c_1 c_2^2 \|x_0\|^4 \right)
\end{align*}
\]

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 10.24** Let \( X \) and \( Y \) be real Banach spaces, \( \Omega \subset X, \ f : \Omega \to \mathbb{R} \) be \( C^2 \) at \( x_0 \in \Omega, \ H : \Omega \to Y \) 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 \( \Omega_c := \{ x \in \Omega \mid H(x) = \vartheta \} \). Then, there exists a Lagrange multiplier \( y_{x_0} \in Y^* \) such that
(i) the Lagrangian $L : \Omega \times \mathcal{Y}^* \to \mathbb{R}$ defined by $L(x, y_\ast) = f(x) + \langle (y_\ast, H(x)) \rangle$, $\forall (x, y_\ast) \in \Omega \times \mathcal{Y}^*$, is stationary at $(x_0, y_{\ast 0})$;

(ii) $\frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0})$ is positive semi-definite on the subspace $\mathcal{N}(H^{(1)}(x_0)) =: M$, that is, $\frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0})(h)(h) \geq 0$, $\forall h \in \mathcal{N}(H^{(1)}(x_0))$.

Proof
Under the assumption of the proposition, by Proposition 10.21, there exists a Lagrange multiplier $y_{\ast 0} \in \mathcal{Y}^*$ 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 $\frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0})(h_0)(h_0) < -m < 0$ 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 $\exists r_0 \in [0, \infty) \subset \mathbb{R}$ with $c_1\delta_1 \leq r_1$ such that $\forall \bar{y} \in \mathcal{B}_y(\delta y, \delta_1/2)$, $\forall \bar{x} \in \mathcal{B}_x(x_0, r_1/2)$ with $\bar{y} = H(\bar{x})$, $\forall y \in \mathcal{B}_y(\bar{y}, \delta_1/2)$, $\exists \bar{x} \in \mathcal{B}_x(x_0, r_1)$ such that $\|x - \bar{x}\| \leq c_1 \|y - \bar{y}\|$, 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 $(x_0, y_{\ast 0})$. By Proposition 9.9 and (1),

$$L^{(1)}(x_0, y_{\ast 0}) = \left[ \frac{\partial L}{\partial x}(x_0, y_{\ast 0}) \frac{\partial L}{\partial y_{\ast}}(x_0, y_{\ast 0}) \right] = \vartheta_{\mathcal{B}(\mathcal{X} \times \mathcal{Y}^*)}$$

By Proposition 9.46, $\exists r_2 \in (0, r_1] \subset \mathbb{R}$ such that $\frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0})$ exists and $\left| \frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0}) - \frac{\partial^2 L}{\partial x^2}(x_0, y_{0 \ast}) \right| < m/5$, $\forall x \in \mathcal{B}_x(x_0, r_2)$. Since $H$ is $C_2$ at $x_0$, then $\exists r_3 \in (0, r_2] \subset \mathbb{R}$ such that $\|H^{(2)}(x)\| \leq c_2$, $\forall x \in \mathcal{B}_x(x_0, r_3)$. Let $c_3 := \left| \frac{\partial^2 L}{\partial x^2}(x_0, y_{\ast 0}) \right|$.

$\forall \delta \in (0, r_3/2] \subset \mathbb{R}$ such that $c_2 \delta^2 < \delta_1$, $c_1c_2\delta^2/2 \leq 1/4$, $c_2\delta^2/4 \leq m/5$, and $c_1c_2(c_3 + m/5)\delta \leq m/5$. By Taylor’s Theorem 9.48, $\exists t_0 \in (0, 1) \subset \mathbb{R}$ such that

$$\|H(x_0 + \delta h_0)\| = \|H(x_0 + \delta h_0) - H(x_0) - \delta H^{(1)}(x_0)h_0\| \\ \leq \frac{1}{2} \left\| \frac{\partial^2 L}{\partial x^2}(x_0 + t_0\delta h_0) \right\| \delta^2 \leq c_2 \delta^2/2 < \delta_1/2$$

Let $\bar{y}_\delta := H(x_0 + \delta h_0) \in \mathcal{B}_y(\delta y, \delta_1/2)$ and $\bar{x}_\delta := x_0 + \delta h_0 \in \mathcal{B}_x(x_0, r_1/2)$. Note that $\bar{y}_\delta = H(\bar{x}_\delta)$ and $\bar{y} \in \mathcal{B}_y(\bar{y}_\delta, \delta_1/2)$. Then, $\exists x_\delta \in \mathcal{B}_x(x_0, r_1)$ with $\|x_\delta - \bar{x}_\delta\| \leq c_1 \|\bar{y}_\delta\| \leq c_1c_2\delta^2/2$ such that $H(x_\delta) = \bar{y}_\delta$. Then, we have $\|x_\delta - x_0\| \geq \|\delta h_0\| - \|x_\delta - \bar{x}_\delta\| \geq \delta - c_1c_2\delta^2/2 \geq 3\delta/4$; and $\|x_\delta - x_0\| \leq \|\delta h_0\| + \|x_\delta - \bar{x}_\delta\| \leq \delta + c_1c_2\delta^2/2 \leq 5\delta/4 < r_3$. Hence, $x_\delta \in \Omega_c \cap \mathcal{B}_x(x_0, r_3)$. By Taylor’s Theorem 9.48, $\exists t_1 \in (0, 1) \subset \mathbb{R}$ such that

$$f(x_\delta) - f(x_0) = L(x_\delta, y_{\ast 0}) - L(x_0, y_{\ast 0}) - \frac{\partial L}{\partial x}(x_0, y_{\ast 0})(x_\delta - x_0)$$

$$= \frac{1}{2} \frac{\partial^2 L}{\partial x^2}(t_1x_\delta + (1 - t_1)x_0, y_{\ast 0})(x_\delta - x_0)(x_\delta - x_0)$$
10.4 INEQUALITY CONSTRAINTS

\[
\frac{1}{2} \left( \frac{\partial^2 L}{\partial x^2}(\bar{x}, y) - \frac{\partial^2 L}{\partial x^2}(\bar{x}, y) \right)
+ \frac{\partial^2 L}{\partial x^2}(\bar{x}, y) \cdot (x - \bar{x}) \cdot (x - \bar{x})
+ \frac{\partial^2 L}{\partial x^2}(\bar{x}, y) \cdot (x - \bar{x})
\]
\[
\leq \frac{1}{2} \left( \delta + \frac{\partial^2 L}{\partial x^2}(x, y) \right)
+ \frac{\partial^2 L}{\partial x^2}(x, y) \cdot (x - \bar{x}) \cdot (x - \bar{x})
+ \frac{\partial^2 L}{\partial x^2}(x, y) \cdot (x - \bar{x})
\]
\[
< \frac{1}{2} \left( -m\delta^2 + m\delta^2/5 + \left( \frac{\partial^2 L}{\partial x^2}(\bar{x}, y) \right) \cdot \left( c_1 + c_2 + c_3 + 1/2 \delta^2 \right) \right)
\]
\[
< \frac{1}{2} \left( -m\delta^2 + m\delta^2/5 + m\delta^2/5 \right) = -m\delta^2 < 0
\]

Thus, \(\lim_{\delta \to 0^+} x = x_0, x \in \Omega, \) and \(f(x_0) < f(x_0),\) for sufficiently small \(\delta.\) Then, \(x_0\) is not a point of relative minimum for \(f\) on the set \(\Omega.\) This contradicts the assumption. Hence, (ii) must hold. This completes the proof of the proposition.

\[\square\]

10.4 Inequality Constraints

The basic problem to be considered in this section is

\[
\mu_0 := \inf_{x \in \Omega} f(x) \quad \text{subject to} \quad G(x) \triangleq \vartheta_{\mathbb{Z}}
\]

where \(\mathcal{X}\) is a real normed linear space, \(\Omega \subseteq \mathcal{X}\) is a set, \(f : \Omega \to \mathbb{R}\) is a functional, \(\mathbb{Z}\) is a real normed linear space with a positive cone \(P \subseteq \mathbb{Z},\) and \(G : \Omega \to \mathbb{Z}\) is a function.

**Theorem 10.25 (Generalized Kuhn-Tucker Theorem)** Let \(\mathcal{X}\) be a real Banach space, \(\Omega \subseteq \mathcal{X}, x_0 \in \Omega^0, \mathbb{Z}\) be a real Banach space with a positive cone \(P \subseteq \mathbb{Z},\) and \(f : \Omega \to \mathbb{R}\) and \(G : \Omega \to \mathbb{Z}\) 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 \triangleq \left\{ x \in \Omega : G(x) \triangleq \vartheta_{\mathbb{Z}} \right\}.

Then, there exists a Lagrange multiplier \(z_{*} \in \mathbb{Z}^*\) with \(z_{*} \triangleq \vartheta_{\mathbb{Z}}\) such that
Consider the small variation equivalent problem of \((10.3)\) at \(x_0\):

\[
\hat{\mu}_0 := \inf_{u \in X} f^{(1)}(x_0)u \tag{10.4}
\]

subject to

\[
G(x_0) + G^{(1)}(x_0)u \leq \vartheta_z.
\]

This problem is convex, and belongs to the class of problems studied in Section 8.8. The primal function for \((10.4)\) is \(\omega : \Gamma \to \mathbb{R}_+\), where

\[
\Gamma := \left\{ \bar{z} \in \mathbb{Z} \mid \exists u \in X \ni \exists t \ni G(x_0) + G^{(1)}(x_0)u \leq \bar{z} \right\} \tag{10.5a}
\]

\[
\omega(\bar{z}) = \inf_{u \in X, G(x_0) + G^{(1)}(x_0)u \bar{z}} f^{(1)}(x_0)u; \forall \bar{z} \in \Gamma. \tag{10.5b}
\]

Since \(G^{(1)}(x_0) \in \mathbb{B}(X, \mathbb{Z})\), then \(\bar{z} = \mathbb{Z}\). Note that \(\hat{\mu}_0 = \omega(\bar{z})\).

Since \(x_0 \in \Omega_c\), then \(z_0 := G(x_0) \leq \bar{z}\). Hence, \(\hat{\mu}_0 \leq 0\).

**Claim 10.25.1** \(\hat{\mu}_0 = 0\).

**Proof of claim:** We will prove this using an argument of contradiction. Suppose \(\hat{\mu}_0 < 0\). Then, \(\exists u_0 \in X\) with \(\|u_0\| > 0\) such that \(f^{(1)}(x_0)u_0 = -m \|u_0\| < 0\) and \(G(x_0) + G^{(1)}(x_0)u_0 \leq \bar{z}\). By Surjective Mapping Theorem 9.53, \(\exists t_1 \in (0, \infty) \subset \mathbb{R}\), \(\exists \delta_1 \in (0, \infty) \subset \mathbb{R}\), and \(\exists \epsilon_1 \in [0, \infty) \subset \mathbb{R}\) with \(c_2 \delta_1 \leq \epsilon_1 \) such that \(\forall z \in \mathbb{B}_z(z_0, \delta_1/2), \forall \bar{z} \in \mathbb{B}_X(x_0, r_1/2)\) with \(z = G(\bar{x}), \bar{z} \in \mathbb{B}_z(z, \delta_1/2)\), \(x \in \mathbb{B}_X(x_0, r_1) \subset \Omega\) with \(\|x - \bar{x}\| \leq c_3\|z - \bar{z}\|\), 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 \(x_0\), \(\exists \epsilon_2 \in (0, r_1) \subset \mathbb{R}\) such that \(\forall \epsilon \in \mathbb{B}_X(x_0, r_2), \|f^{(1)}(x) - f^{(1)}(x_0)\| < m/3\) and \(\|G^{(1)}(x) - G^{(1)}(x_0)\| < 3(1 + c_1 + c_2(\epsilon_1 + m/3))\).
10.4. INEQUALITY CONSTRAINTS

\[ f^{(1)}(t_1 x_\delta + (1 - t_1) \bar{x}_\delta)(x_\delta - \bar{x}_\delta) + \delta f^{(1)}(x_0 + t_2 \delta u_0) u_0 \]
\[ \leq (\| f^{(1)}(x_0) \| + \| f^{(1)}(t_1 x_\delta + (1 - t_1) \bar{x}_\delta) - f^{(1)}(x_0) \|) \| x_\delta - \bar{x}_\delta \| \\
+ \delta f^{(1)}(x_0) u_0 + \delta \| f^{(1)}(x_0 + t_2 \delta u_0) - f^{(1)}(x_0) \| \| u_0 \| \\
\leq -m\delta \| u_0 \| + (c_3 + m/3) \frac{c_1 m}{3(1 + c_1 + c_3 + m/3)} \delta \| u_0 \| \\
+ m\delta \| u_0 \| /3 < -m\delta \| u_0 \| /3 < 0 \]

Note that \( \| x_\delta - x_0 \| \leq \| x_\delta - \bar{x}_\delta \| + \delta \| u_0 \| < (1 + m/3)\delta \| u_0 \|. \) Then, 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 with the assumption that \( x_0 \) is a point of relative minimum for \( f \) on \( \Omega_c. \) Hence, \( \bar{u}_0 = 0. \)

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

Then, \( \omega(\vartheta_z) = 0. \) By Fact 8.52, \( \omega : \mathcal{Z} \to \mathbb{R} \) is real-valued and convex.

Next, we will show that \( \omega \) is continuous at \( \vartheta_z \) by Proposition 8.22. By Proposition 7.113, \( \exists c_4 \in [0, \infty) \subset \mathbb{R}, \forall \bar{z} \in \mathcal{Z}, \exists u \in \bar{X} \) such that \( \bar{z} = G^{(1)}(x_0) u + \| u \| \leq c_4 \| \bar{z} \|. \) Then, \( G(x_0) + G^{(1)}(x_0) u = G(x_0) + \bar{z} \leq \bar{z} \) and \( f^{(1)}(x_\delta) u \leq \| f^{(1)}(x_0) \| \| u \| \leq c_4 \| f^{(1)}(x_0) \| \| z \|. \) This implies that \( \forall \bar{z} \in B_z(\vartheta_z, 1/2), \) \( \omega(\bar{z}) \leq c_4 \| f^{(1)}(x_0) \| /2 = : r_0 - 1/2. \) Then, \( B_{\mathbb{Z}}(\bar{x}_0, 1/2) \subseteq [\omega, \bar{Z}]. \) Hence, \( (r_0, \vartheta_z) \in \bar{X}_z \). By Proposition 8.22, \( \omega \) is continuous at \( \vartheta_z. \) By Proposition 8.23, \( \omega \) is continuous.

By Proposition 8.34, we have
\[ 0 = \bar{u}_0 = \omega(\vartheta_z) = \max_{z_* \in \Omega_{\omega \text{conj}}} (-\omega_{\omega \text{conj}}(z_*)) \]
where \( \Omega_{\omega \text{conj}} \) and \( \omega_{\omega \text{conj}} : \Omega_{\omega \text{conj}} \to \mathbb{R} \) are the conjugate set and conjugate functional of \( \omega. \) By Fact 8.53, \( \Omega_{\omega \text{conj}} \subseteq P^\omega. \) Let the maximum be achieved at \( -z_* \in \Omega_{\omega \text{conj}} \), where \( z_* \in \Omega_{\omega \text{conj}} \). Define \( \bar{\omega} : P^\omega \to \mathbb{R} \) by \( \bar{\omega}(z_*) = \sup_{x \in \bar{X}} (\langle z_*, G(x) \rangle - \omega(z)), \forall z_* \in P^\omega. \) By Fact 8.54, we have \( \max_{z_* \in \Omega_{\omega \text{conj}}} (-\omega_{\omega \text{conj}}(z_*)) = \max_{z_* \in P^\omega} (-\bar{\omega}(z_*)), \) where both maximums are achieved at \( -z_* \), and \( \forall z_* \in P^\omega, \)
\[ -\bar{\omega}(z_*) = \inf_{u \in \bar{X}} (f^{(1)}(x_0) u + \langle z_*, G(x_0) + G^{(1)}(x_0) u \rangle) \]
Then, we have
\[ 0 = \bar{u}_0 = \max_{z_* \in \Omega_{\omega \text{conj}}} \inf_{u \in \bar{X}} (f^{(1)}(x_0) u + \langle z_*, G(x_0) + G^{(1)}(x_0) u \rangle) \]
where the maximum is achieved at \( z_* \). This implies that \( \langle (z_0, G(x_0)) \rangle + \inf_{u \in \bar{X}} \left\langle (f^{(1)}(x_0) + (G^{(1)}(x_0)) z_0, u) \right\rangle = 0 \) by Proposition 8.37. For the infimum to be finite, we must have \( f^{(1)}(x_0) + (G^{(1)}(x_0)) z_0 = \vartheta_x. \) Hence, the Lagrangian \( L \) is stationary at \( x_0. \) Then, \( 0 = \langle (z_0, G(x_0)) \rangle. \)

This completes the proof of the proposition. \( \square \)
Proposition 10.26 Let \( X \) be a real Banach space, \( Y \) be a real Banach space with a positive cone \( P \subseteq Y \), and \( A \in B(X,Y) \) be surjective. Then,

\[
(A_{\text{inv}}(P))^{\mathfrak{B}} = A'(P^{\mathfrak{B}})
\]

Proof. \( \forall x_* \in A'(P^{\mathfrak{B}}), \exists y_* \in P^{\mathfrak{B}} \) such that \( x_* = A'y_* \). \( \forall x \in A_{\text{inv}}(P) \), we have \( Ax \in P \). Then, \( \langle \langle x_*,x \rangle \rangle = \langle \langle A'y_*,x \rangle \rangle = \langle \langle y_*,Ax \rangle \rangle \geq 0 \). By the arbitrariness of \( x \), \( x_* \in (A_{\text{inv}}(P))^{\mathfrak{B}} \). By the arbitrariness of \( x_* \), we have \( A'(P^{\mathfrak{B}}) \subseteq (A_{\text{inv}}(P))^{\mathfrak{B}} \).

On the other hand, fix any \( x_* \in (A_{\text{inv}}(P))^{\mathfrak{B}} \). \( \forall x \in \mathcal{N}(A), x \in A_{\text{inv}}(P) \), since \( \vartheta y \in P \). Then, \( \langle \langle x_*,x \rangle \rangle \geq 0 \). Since \( -x \in \mathcal{N}(A) \) as well, then \( \langle \langle x_*,x \rangle \rangle = 0 \). Hence, \( x_* \in (\mathcal{N}(A))^\perp \). By Proposition 7.114, \( (\mathcal{N}(A))^\perp = \mathcal{R}(A') \). Then, \( \exists y_* \in Y^* \) such that \( x_* = A'y_* \). \( \forall y \in P \), since \( A \) is surjective, then \( \exists x \in A_{\text{inv}}(P) \) such that \( y = Ax \). Then, we have \( \langle \langle y_*,y \rangle \rangle = \langle \langle y_*,Ax \rangle \rangle = \langle \langle A'y_*,x \rangle \rangle = \langle \langle x_*,x \rangle \rangle \geq 0 \). By the arbitrariness of \( y \), we have \( y_* \in P^{\mathfrak{B}} \). Then, \( x_* \in A'(P^{\mathfrak{B}}) \). By the arbitrariness of \( x_* \), we have \( (A_{\text{inv}}(P))^{\mathfrak{B}} \subseteq A'(P^{\mathfrak{B}}) \).

Therefore, we have \( A'(P^{\mathfrak{B}}) = (A_{\text{inv}}(P))^{\mathfrak{B}} \). This completes the proof of the proposition. \( \square \)

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 \( X \) be a real Banach space, \( \Omega \subseteq X \), \( x_0 \in \Omega^o \), \( Z \) 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

(i) \( G(x_0) \leq \vartheta Z \) and \( x_0 \) is regular point of \( G \);

(ii) \( z_{x_0} \in P^{\mathfrak{B}} \) is the Lagrange multiplier, the Lagrangian \( L : \Omega \to \mathbb{R} \) defined by \( L(x) = f(x) + \langle \langle z_{x_0},G(x) \rangle \rangle \), \( \forall x \in \Omega \), is stationary at \( x_0 \);

(iii) \( \langle \langle z_{x_0},G(x_0) \rangle \rangle = 0 \);

(iv) \( \exists m \in (0,\infty) \subset \mathbb{R} \) such that \( L^{(2)}(x_0)(u)(u) \geq m \| u \|^2, \forall u \in M_c := \{ u \in X \mid G(x_0) + G^{(1)}(x_0)u \leq \vartheta Z \} \), 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 \( \Omega_c := \{ x \in X \mid G(x) \leq \vartheta Z \} \).

Proof. By Propositions 9.34, 9.37, 9.40, 9.41, 9.44, and 9.45, \( L \) is \( C_2 \) at \( x_0 \). By (ii), \( L^{(1)}(x_0) = \vartheta_{B(X,\mathbb{R})} = \vartheta_X \). Then, \( \exists \delta_0 \in (0,\infty) \subset \mathbb{R} \) such that \( f^{(2)}(x_0), G^{(2)}(x_0) \) and \( L^{(2)}(x) \) exist, \( \forall x \in B_X(x_0,\delta_0) \subseteq \Omega \) and \( L^{(2)} \) is continuous at \( x_0 \). Then, \( \exists \delta_1 \in (0,\delta_0] \subset \mathbb{R} \) and \( \exists c_1 \in [0,\infty) \subset \mathbb{R} \)
such that \( \| L^{(2)}(x) - L^{(2)}(x_0) \| < m/5 \) and \( \| G^{(2)}(x) \| \leq c_1 \), \( \forall x \in D_1 := \mathcal{B}_x(x_0, \delta_1) \). By Propositions 7.114 and 7.98 and (i), \( (\mathcal{N}(G^{(1)}(x_0)))^\perp = \mathcal{R}(\langle G^{(1)}(x_0) \rangle^t) \) is closed. Then, by Proposition 7.113, \( \exists \delta_2 \in [0, \infty) \subset \mathbb{R} \) such that \( \forall x_* \in \mathcal{R}(\langle G^{(1)}(x_0) \rangle^t) \), there exists \( z_* \in \mathcal{Z}^* \) such that \( x_* = (G^{(1)}(x_0))^t z_* \) and \( \| z_* \| \leq \delta_2 \| x_* \| \).

Claim 10.27.1 \( \forall x \in \Omega \cap D_1 \), \( \exists h_0 \in M_c \) 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 \( x \in \Omega \cap D_1 \). Then, \( G(x) \leq \mathcal{Y} \). It is easy to see that \( M_c \supseteq \mathcal{Y} \) is a nonempty closed convex set. Then, by Proposition 8.15, we have \( 0 \leq \delta := \inf_{h \in M_c} \| x - x_0 - h \| = \max_{x_* \in M_{c_{supp}}, \| x_* \| \leq 1} \langle \langle x_* - x_0 \rangle, g(x_*) \rangle \geq 0 \), where \( g : M_{c_{supp}} \rightarrow \mathbb{R} \) is the support functional of \( M_c \). We will show that \( M_{c_{supp}} \subseteq \langle (G^{(1)}(x_0))_{\text{inv}}(P) \rangle^\oplus \). Fix any \( x_* \in M_{c_{supp}} \), \( \forall x \in \langle (G^{(1)}(x_0))_{\text{inv}}(P) \rangle^\oplus \), we have \( (G^{(1)}(x_0))_{\text{inv}}(P) \subseteq \mathcal{Z} \). It is easy to show that \( -\alpha \in M_c \). Suppose \( \langle \langle x_* \rangle, u \rangle \rangle < 0 \). Then, \( \sup_{u \in M_c} \langle \langle x_* \rangle, u \rangle \rangle \geq \sup_{u \in [0, \infty)} \mathcal{R}(\langle \langle x_* \rangle, -\alpha u \rangle \rangle = +\infty \). This implies that \( x_* \notin M_{c_{supp}} \), which is a contradiction. Hence, we must have \( \langle \langle x_* \rangle, u \rangle \rangle \geq 0 \). By the arbitrariness of \( u \), we have \( x_* \in \langle (G^{(1)}(x_0))_{\text{inv}}(P) \rangle^\oplus \). Then, \( M_{c_{supp}} \subseteq \langle (G^{(1)}(x_0))_{\text{inv}}(P) \rangle^\oplus \).

Then, \( \delta = \inf_{x \in \Omega} \| x - x_0 \| - g(x_0) \) for some \( x_0 \in M_{c_{supp}} \subseteq \langle (G^{(1)}(x_0))_{\text{inv}}(P) \rangle^\oplus \) with \( \| x_0 \| \leq 1 \). By (i) and Proposition 10.26, we have \( x_0 \in (G^{(1)}(x_0))(P^\oplus) \subseteq \mathcal{R}(\langle (G^{(1)}(x_0))^t \rangle) \). Then, \( \exists x_1 \in \mathcal{Z}^* \) such that \( x_0 = (G^{(1)}(x_0))^t z_1 \) and \( \| z_1 \| \leq c_1 \| x_0 \| \leq 2 \). This further implies that

\[
\delta = \inf_{z_1} \langle \langle z_1, G^{(1)}(x_0)(x - x_0) \rangle \rangle - \sup_{u \in M_c} \langle \langle z_1, G^{(1)}(x_0)u \rangle \rangle
\]

where the second equality follows from Proposition 8.37. By (i), we have \( \mathcal{R}(G^{(1)}(x_0)) = \mathcal{Z} \) and \( \exists u_1 \in \mathcal{X} \) 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 \( x \in \Omega \) and \( G(x) \leq \mathcal{Y} \). Then, \( \delta \leq \langle \langle z_1, -G(x) + G(x_0) + G^{(1)}(x_0)(x - x_0) \rangle \rangle \).

By Taylor’s Theorem 9.48, \( \exists \delta_2 \in (0, 1) \subset \mathbb{R} \) such that \( \| G(x) - G(x_0) - G^{(1)}(x_0)(x - x_0) \| \leq \delta_2 \| G^{(2)}(t_0 x + (1-t_0) x_0) \| \| x - x_0 \| + \delta_2 = \frac{\delta}{2} \leq c_1 \| x - x_0 \|^2 / 2. \)

Then, we have \( \delta \leq \| z_1 \| \| G(x) - G(x_0) - G^{(1)}(x_0)(x - x_0) \| \leq c_1 \| x - x_0 \|^2 / 2. \) 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 \( \delta \in (0, \delta_1) \subseteq \mathbb{R} \) 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 \( 5c_1 c_2 (c_3 + m/5) \delta/2 \leq m/5 \). \( \forall x \in \Omega_c \cap \mathcal{B}_x (x_0, \delta) \), by Claim 10.27.1, \( \exists h_0 \in M_c \) 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, \( \exists t_1 \in (0, 1) \subseteq \mathbb{R} \) such that

\[
f(x) - f(x_0) \geq L(x) - L(x_0) - L^{(1)}(x_0)(x - x_0)
\]

\[
= \frac{1}{2} L^{(2)}(\xi)(x - x_0)(x - x_0)
\]

\[
= \frac{1}{2} \left( L^{(2)}(\hat{x})(h_0)(h_0) + L^{(2)}(\hat{x})(x - x_0 - h_0)(h_0) + L^{(2)}(\hat{x})(x - x_0 - h_0)(x - x_0 - h_0) \right)
\]

\[
\geq \frac{1}{2} \left( L^{(2)}(x_0)(h_0)(h_0) + L^{(2)}(\hat{x})(x - x_0)(h_0) \right)
\]

\[
\geq \frac{1}{2} \left( m \| h_0 \|^2 - m \| h_0 \|^2 / 5 \right)
\]

\[
\geq \frac{1}{2} \left( \frac{4m}{5} \| h_0 \|^2 - \left( c_3 + m/5 \right) \left( \frac{5}{2} c_1 c_2 \| x - x_0 \|^3 + \frac{c_1^2 c_2^2}{2} \| x - x_0 \|^4 \right) \right)
\]

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, \( \Omega \subseteq X \), \( x_0 \in \Omega^o \), \( Z \) be a real Banach space with a positive cone \( P \subseteq Z^* \), and \( f : \Omega \rightarrow \mathbb{R} \) and \( G : \Omega \rightarrow Z \) 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 \( \Omega_c := \{ x \in X \mid G(x) \subseteq \partial Z \} \). Then, there exists a Lagrange multiplier \( z_{*0} \in P^\oplus \) such that

(i) 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 \( \langle (z_{*0}, G(x_0)) \rangle = 0 \);

(ii) \( L^{(2)}(x_0) \) is positive semi-definite on the subspace \( \mathcal{N} (G^{(1)}(x_0)) =: M \), that is, \( G^{(2)}(x_0)(h)(h) \geq 0 \), \( \forall h \in \mathcal{N} (G^{(1)}(x_0)) \).

**Proof** By the Generalized Kuhn-Tucker Theorem 10.25, there exists a Lagrange multiplier \( z_{*0} \in Z^* \) with \( z_{*0} \trianglerighteq \partial Z \) 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 \( L(x_0(h_0)(h_0) < -m < 0 \) for some \( m \in (0, \infty) \subset \mathbb{R} \). Let \( z_0 := G(x_0) \in \mathbb{Z} \). By Surjective Mapping Theorem 9.53, \( \exists_1 \in (0, \infty) \subset \mathbb{R}, \, \exists_2 \in (0, \infty) \subset \mathbb{R}, \) and \( \exists_1 \in (0, \infty) \subset \mathbb{R} \) with \( c_1 \delta \leq r_1 \) such that \( \forall \exists \in B_2 (z_0, \delta_1/2), \forall \exists \in B_2 (x_0, r_1/2) \) with \( \exists = G(\bar{x}), \forall \exists \in B_2 (\bar{x}, \delta_1/2), \exists \in B_2 (x_0, r_1) \subset \Omega \) with \( \| x - \bar{x} \| \leq c_1 \| z - \exists \| \), 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), \( L^1(x_0) = \delta B(x_0) = \delta B^\gamma \). Then, \( \exists r_2 \in (0, r_1] \subset \mathbb{R} \) such that \( L^2(x) \| L^2(x) - L^2(x_0)\| < m/5, \forall \exists \in B_2 (x_0, r_2) \). Since \( G \) is \( C_2 \) at \( x_0 \), then \( \exists_2 \in (0, \infty) \subset \mathbb{R} \) and \( \exists_3 \in (0, r_2] \subset \mathbb{R} \) such that \( \| G^2(x) \| = c_2, \forall \exists \in B_2 (x_0, r_3) \). Let \( c_3 := \| L^2(x_0) \| \). 

\( \forall \delta \in (0, r_3/2) \subset \mathbb{R} \) such that \( c_2 \delta^2 < \delta_1, \, c_1 c_2 \delta^2/2 \leq 1/4, \, c_1^2 \delta^2(x_3 + m/5) \delta^2/4 \leq m/5, \) 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 

\[
\| G(x_0 + \delta h_0) - z_0 \| = \| G(x_0 + \delta h_0) - G(x_0) - \delta G(x_0) h_0 \| 
\leq \frac{1}{2} \| L^2(x_0 + t_0 \delta h_0) \| \delta^2 \leq c_2 \delta^2/2 < \delta_1/2
\]

Let \( \bar{x} := G(x_0 + \delta h_0) \in B_2 (x_0, \delta_1/2) \) and \( \bar{x}_3 := x_0 + \delta h_0 \in B_2 (x_0, r_1/2) \). Note that \( \bar{x}_3 = G(\bar{x}_3) \) and \( \exists h_0 \in \mathbb{Z}_2 (x_0, \delta_1/2) \). Then, \( \exists \bar{x}_3 \in B_2 (x_0, r_1) \) with \( \| x - x \| \leq c_1 \| z - \exists \| \leq c_1 c_2 \delta^2/2 \) such that \( G(x_3) = z_0 \). Then, we have \( \| x - x_0 \| \geq \| \delta h_0 \| - \| x - \bar{x}_3 \| \geq \delta - c_1 c_2 \delta^2/2 \geq 3\delta/4 \) \( \| x - x_0 \| \geq \| \delta h_0 \| + \| x - \bar{x}_3 \| \leq \delta + c_1 c_2 \delta^2/2 \leq 5\delta/4 < r_3 \). Hence, \( x_0 \in \Omega_0 \subset B_2 (x_0, r_3) \). By Taylor’s Theorem 9.48, \( \exists_1 \in (0, 1) \subset \mathbb{R} \) such that 

\[
f(x_0) - f(x_0) = L(x_0) - L(x_0) - L_1(x_0)(x_0 - x_0) - \langle z_0, G(x_0) \rangle 
+ \langle z_0, G(x_0) \rangle = L(x_0) - L_1(x_0)(x_0 - x_0)
= \frac{1}{2} L^2(t_0 \bar{x}_3 + (1 - t_1) x_0)(x_0 - x_0)(x_0 - x_0)
= \frac{1}{2} (L^2(\bar{x}_3)(\bar{x}_3 - x_0)(x_0 - x_0) + L^2(\bar{x}_3)(\bar{x}_3 - x_0)(x_0 - x_0)
+ L^2(\bar{x}_3)(\bar{x}_3 - x_0)(x_0 - x_0) + L^2(\bar{x}_3)(\bar{x}_3 - x_0)(x_0 - x_0))
\leq \frac{1}{2} \| L^2(\bar{x}_3)(h_0)(h_0) + \delta^2(L^2(\bar{x}_3) - L^2(\bar{x}_3)(x_0)(h_0) + \| L^2(\bar{x}_3)(\| x_0 - \bar{x}_3 \| = \| x_0 - \bar{x}_3 \|)
+ \frac{1}{2} \left( -m \delta^2 + m \delta^2/5 + \frac{1}{2} \left( L^2(\bar{x}_3) - L^2(\bar{x}_3)(x_0) + \| L^2(\bar{x}_3)(x_0) \| \right)
\cdot (c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4)
\right)
\leq \frac{1}{2} \left( -\frac{4m}{5} \delta^2 + (c_3 + m/5)(c_1 c_2 \delta^3 + c_1^2 c_2^2 \delta^4/4) \right)
\]
Thus, \( \lim_{\delta \to 0^+} x_\delta = x_0, \) 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 \( \Omega. \) 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 hold.

(i) \( L' = L. \)

(ii) If \( L \in S_{+ X}, \) that is, \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( L(x)(x) \geq \delta \| x \|^2_X, \) \( \forall x \in X, \) then, \( L \in B(X, X^*) \) is bijective, \( L^{-1} \in S_{+ X^*}, \) that is, \( L^{-1}(x)(x) \geq \frac{\delta}{c + \| L^{-1} \|_{B(X, X^*)}} \| x \|^2_{X^*}, \) \( \forall x \in X^*, \) \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}. \)

**Proof**

(i) \( \forall x, y \in X, \) we have

\[
L(x)(y) = L(y)(x) = \langle \langle L(x), L(y) \rangle \rangle = \langle \langle L'(\phi_X(x), L(x)) \rangle \rangle = \langle \langle L'(y), x \rangle \rangle
\]

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 \( \phi_X = id_X. \) This implies that

\[
L(x)(y) = \langle \langle L'(y), x \rangle \rangle = L(y)(x) = \langle \langle L(y), x \rangle \rangle
\]

By the arbitrariness of \( x \) and \( y, \) we have \( L = L'. \)

(ii) By the assumption, \( L \in B_{S_2}(X, \mathbb{R}) \subset B(X, X^*). \) \( \forall x \in X \) with \( x \neq \vartheta_X, L(x) \neq \vartheta_X, \) since \( L(x)(x) > 0. \) Hence, \( \mathcal{N}(L) = \{ \vartheta_X \} \) and \( L \) is injective. By (i), \( L = L'. \) Then, \( \mathcal{N}'(L') = \{ \vartheta_X \}. \) By Proposition 7.112, \( \mathcal{R}'(L) = \mathcal{N}'(L'). \) This shows that \( \mathcal{R}'(L) = X^*. \) By Proposition 7.98, \( \mathcal{R}'(L) = X^*. \) Hence, \( \mathcal{R}'(L) \) is dense in \( X^*. \) Clearly, \( \vartheta_X \in \mathcal{R}'(L). \) \( \forall x \in X^* \) with \( x \neq \vartheta_X, \) there exists \( (x_i)_{i=1}^\infty \subset \mathcal{R}'(L) \) such that \( x_i = \lim_{i \to \infty} x_{\epsilon_i}. \) Without loss of generality, we may assume that \( x_{\epsilon_i} \neq \vartheta_X, \) \( \forall i \in \mathbb{N}. \) Let \( x_{\epsilon_i} = L(x_i), \) \( \forall i \in \mathbb{N}, \) where \( x_i \in X. \) Clearly, \( x_i \neq \vartheta_X \) since \( x_{\epsilon_i} \neq \vartheta_X, \) \( \forall i \in \mathbb{N}. \) Then, we have

\[
\| x_{\epsilon_i} \|_{X^*} \geq \| x_i \|_X \geq \| x_{\epsilon_i} \|_{X^*} \geq \| x_i \|_X \geq \delta \| x_i \|^2_X; \quad \forall i \in \mathbb{N}
\]

This implies that \( \| x_{\epsilon_i} \|_{X^*} \geq \delta \| x_i \|_X \) and \( \| x_i \|_X \leq \| x_{\epsilon_i} \|_{X^*} / \delta, \forall i \in \mathbb{N}. \) Since \( (x_{\epsilon_i})_{i=1}^\infty \) is convergent, then there exists \( c \in (0, \infty) \subset \mathbb{R} \) such that \( \| x_{\epsilon_i} \|_{X^*} \leq c, \forall i \in \mathbb{N}. \) Then, \( \| x_i \|_X \leq c / \delta =: c_1 \in (0, \infty) \subset \mathbb{R}, \forall i \in \mathbb{N}. \) This shows that \( (x_i)_{i=1}^\infty \subset \overline{B}(x, c_1) \) \( \cong S_1 \subset X = X^{**}. \) By Alaoglu Theorem 7.122, \( S_1 \subset X^{**} \) is weak* compact. By Propositions 5.22 and 5.26, \( (x_i)_{i=1}^\infty \) has a cluster point \( x \in X^{**} = X \) in weak*
10.4. INEQUALITY CONSTRAINTS

...
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, \(\mu\left(\bigcup_{i=1}^{n} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \mu(E_i)\).

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, \(\mu(A) \leq \mu(B)\).
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, $\lim_{n \to \infty} \mu(E_n) = \mu(\bigcap_{n=1}^{\infty} E_n)$.

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 $E_1 = E \cup \left( \bigcup_{n=1}^{\infty} \left( E_n \cap \widetilde{E}_{n+1} \right) \right)$. Clearly, the right-hand-side of the above equality is a disjoint union of measurable sets. Therefore,

$$\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} \mu\left( E_n \cap \widetilde{E}_{n+1} \right) \quad (11.1)$$

On the other hand, note that $E_n = E_{n+1} \cup \left( E_n \cap \widetilde{E}_{n+1} \right)$, $n = 1, 2, \ldots$, which implies that $\mu(E_n) = \mu(E_{n+1}) + \mu\left( E_n \cap \widetilde{E}_{n+1} \right)$, 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 $\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \mu(E_1) - \lim_{n \to \infty} \mu(E_n)$. Since $\mu(E_1) < +\infty$, then $\mu(E) < +\infty$ by Proposition 11.4 and $E \subseteq E_1$. Therefore, we have $\lim_{n \to \infty} \mu(E_n) = \mu(E)$.

This completes the proof of the proposition. \qed

Proposition 11.6 Let $(X, \mathcal{B}, \mu)$ be a measure space. A sequence $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$. Then, $\mu\left( \bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \mu(E_n)$.

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, $\mu\left( \bigcup_{i=1}^{\infty} G_i \right) = \sum_{i=1}^{\infty} \mu(G_i)$. 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,

$$\mu\left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mu(G_i) \quad (11.2)$$

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 $\mu\left( \bigcup_{i=1}^{\infty} E_i \right) \leq \sum_{i=1}^{\infty} \mu(E_i)$.

This completes the proof of the proposition. \qed

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, $\mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{i \to \infty} \mu(A_i) \in \mathbb{R}_e$. 

Proof The set $B \cap \tilde{A} \in \mathcal{B}$ since $\mathcal{B}$ is a $\sigma$-algebra. Note that $A \cap (B \cap \tilde{A}) = \emptyset$, then, $\mu(B) = \mu\left( A \cup (B \cap \tilde{A}) \right) = \mu(A) + \mu(B \cap \tilde{A}) \geq \mu(A)$, by Fact 11.3. This completes the proof of the proposition. \qed
11.1. MEASURE SPACES

Proof Let $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$. 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. Then, by countable additivity of measure, $\mu(A) = \mu(A_1) + \sum_{i=1}^{\infty} \mu(A_{i+1} \setminus A_i) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{i=1}^{n} \mu(A_{i+1} \setminus A_i)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{n \in \mathbb{N}} \mu(A_i)$, 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, $\forall E, E_2 \in \mathcal{B}$ with $\mu(E_1 \Delta E_2) = 0$, we have $\mu(E_1) = \mu(E_2)$.

Proof $0 = \mu(E_1 \Delta E_2) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1)$. Then, $\mu(E_1 \setminus E_2) = \mu(E_2 \setminus E_1) = 0$. Hence, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_2 \setminus E_1) = \mu(E_2)$. This completes the proof of the proposition. $\square$

Lemma 11.9 (Borel-Cantelli) 11 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 $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists N \in \mathbb{N}$ 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 equals 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 $\bigcup_{n=1}^{\infty} X_n = X$ 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 $\tilde{E}_1 = \tilde{B}_1 \cap \tilde{A}_1 = \tilde{B}_1 \cap \left( \tilde{A}_1 \setminus \tilde{C}_1 \right) \cup \tilde{C}_1$, since $\tilde{C}_1 \subseteq \tilde{A}_1$.

This implies that $\tilde{E}_1 = \left( \tilde{B}_1 \cap \left( \tilde{A}_1 \setminus \tilde{C}_1 \right) \right) \cup \left( \tilde{B}_1 \cap \tilde{C}_1 \right)$. Since $\tilde{B}_1 \cap \tilde{C}_1 \in \mathcal{B}$ and $\tilde{B}_1 \cap \left( \tilde{A}_1 \setminus \tilde{C}_1 \right) = \tilde{B}_1 \cap \tilde{A}_1 \setminus \tilde{C}_1 \subseteq \tilde{A}_1$ and $\mu(C_1) = 0$, then, $\tilde{E}_1 \in \mathcal{B}_0$.

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, $\bigcup_{n=1}^{\infty} E_n = \left( \bigcup_{n=1}^{\infty} A_n \right) \cup \left( \bigcup_{n=1}^{\infty} B_n \right)$. 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 $0 \leq \mu\left( \bigcup_{n=1}^{\infty} C_n \right) \leq \sum_{n=1}^{\infty} \mu(C_n) = 0$. Then, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}_0$.

This proves that $\mathcal{B}_0$ is a $\sigma$-algebra.

Define a mapping $\mu_0 : \mathcal{B}_0 \rightarrow [0, \infty] \subseteq \mathbb{R}$ 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,

$$E = A_1 \cup B_1 = A_2 \cup B_2$$
Hence, \( X \), and (ii). It is shown in the following that \( B, C \) \( \in B \), \( X, \emptyset \) Since \( \mu \) we have \( \mu \left( \bigcup_{n=1}^{\infty} 0 \right) \). This shows that \( \mu \left( \widetilde{B} \cap B_1 \right) = 0 \). Therefore, \( \mu \left( B_1 \right) \leq \mu \left( B_2 \right) \).

By an argument similar to the above, it can be concluded that \( \mu \left( B_2 \right) \leq \mu \left( B_1 \right) \). Combining the two inequalities, yields \( \mu \left( B_2 \right) = \mu \left( B_1 \right) \).

This shows that \( \mu_0(E) \) 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 \( B \).

In the following, it will be shown that the nonnegative, extended real-valued function \( \mu_0 \) defines a measure on the measurable space \((X, B_0)\).

Since \( \emptyset \in B \), then \( \mu_0(\emptyset) = \mu(\emptyset) = 0 \).

Consider any sequence of subsets \( \left( E_n \right)_{n=1}^{\infty} \subseteq B_0 \), which are pairwise disjoint. There exists \( A_n, B_n, C_n \subseteq X \), such that \( E_n = A_n \cup B_n, B_n, C_n \in B \), \( A_n \subseteq C_n \), and \( \mu(C_n) = 0 \), \( n = 1, 2, \ldots \). Since \( B_n \subseteq E_n \), \( n = 1, 2, \ldots \), then, \( \left( B_n \right)_{n=1}^{\infty} \) are pairwise disjoint. Hence, \( \mu \left( \bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} \mu \left( B_n \right) \).

Since \( \bigcup_{n=1}^{\infty} E_n = \left( \bigcup_{n=1}^{\infty} A_n \right) \cup \left( \bigcup_{n=1}^{\infty} B_n \right), \cup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n, 0 \leq \mu \left( \bigcup_{n=1}^{\infty} C_n \right) \leq \sum_{n=1}^{\infty} \mu(C_n) = 0 \), and \( \sum_{n=1}^{\infty} B_n, \bigcup_{n=1}^{\infty} C_n \in B \) then, \( \mu_0 \left( \bigcup_{n=1}^{\infty} E_n \right) = \mu \left( \bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} \mu \left( B_n \right) = \sum_{n=1}^{\infty} \mu_0 \left( E_n \right) \). This proves that \( \mu_0 \) is countably additive for disjoint \( B_0 \)-measurable sets.

Hence, \( \mu_0 \) is a measure on \((X, B_0)\).

In the following, it is proved that \((X, B_0, \mu_0)\) is a complete measure space.

\( \forall E_0 \subseteq X \), such that there exists \( E \in B_0 \) with \( E_0 \subseteq E \) and \( \mu_0 \left( E \right) = 0 \). Then, there exists \( A, B, C \subseteq X \) such that \( E = A \cup B, B, C \in B, A \subseteq C, \) and \( \mu(C) = 0 \), by the definition of \( B_0 \) and \( E \in B_0 \). By the definition of \( \mu_0 \), we have \( \mu_0 \left( E \right) = \mu \left( B \right) = 0 \). Hence, \( E_0 = E_0 \cup \emptyset \) and \( E_0 \subseteq A \cup B \subseteq C \cup B \).

Since \( \emptyset, C \cup B \in B \), and \( 0 \leq \mu \left( C \cup B \right) \leq \mu(C) + \mu(B) = 0 \), then \( E_0 \in B_0 \). This proves that \((X, B_0, \mu_0)\) is complete.

Let \((X, B_1, \mu_1)\) be any complete measure space that satisfies the (i) and (ii). It is shown in the following that \( B_0 \subseteq B_1 \) by an argument of contradiction. Suppose there exist a set \( E \in B_0 \setminus B_1 \), then, \( E \notin B \). Since \( E \in B_0 \), then, there exists \( A, B, C \subseteq X \), such that \( E = A \cup B, A \subseteq C, B, C \in B, \) and \( \mu(C) = 0 \). This implies \( B, C \in B_1 \) and \( \mu_1 \left( C \right) = \mu(C) = 0 \).

Hence, \( A \notin B_1 \), otherwise, we have \( E \in B_1 \). But, \( A \subseteq C, \) with \( \mu_1 \left( C \right) = 0 \), which contradicts the fact that \((X, B_1, \mu_1)\) being complete.

Therefore, we have \( B_0 \subseteq B_1 \).
∀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(A) + 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 measure 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 := ∪ᵢ=1→∞ Eᵢ ∈ B and µ|B(E) = ∫∞₁₀ µ(Eᵢ) = ∫∞₁₀ µ|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. ∀Eᵢ ∈ B, ∀i ∈ N, A ⊆ Eᵢ ∈ B. Then, A ⊆ ∪ᵢ=1→∞ Eᵢ ∈ B and ∪ᵢ=1→∞ Eᵢ ∈ 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 ∈ Bₐ 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) = µₐ(E) = µₐ(E). ∀E ⊆ A, assume that E = Eₐ ∪ Eₐ, with Eₐ, Eₐ ∈ Bₐ, Eₐ ⊆ Eₐ, and µ|Bₐ(Eₐ) = 0. Then, µₐ(Eₐ) = 0 and Eₐ ∈ 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 ∪ Eₐ with Eₐ, Eₐ ∈ Bₐ, Eₐ ⊆ Eₐ, and µ(Eₐ) = 0. Then, Eₐ, Eₐ ∩ A ∈ Bₐ, Eₐ ⊆ Eₐ ∩ A, and 0 ≤ µₐ(Eₐ ∩ A) = µ(Eₐ ∩ A) ≤ µ(Eₐ) = 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 $\mu_o : \mathcal{X}_2 \to [0, \infty] \subset \mathbb{R}_e$ satisfying

(i) $\mu_o(\emptyset) = 0$;

(ii) if $A \subseteq B \subseteq X$, then $\mu_o(A) \leq \mu_o(B)$; (monotonicity)

(iii) if $E \subseteq \bigcup_{i=1}^\infty E_i \subseteq X$, then $\mu_o(E) \leq \sum_{i=1}^\infty \mu_o(E_i)$. (countable subadditivity)

Note that (i) and (iii) above implies finite subadditivity.

Definition 11.15 Let $X$ be a set and $\mu_o : \mathcal{X}_2 \to [0, \infty] \subset \mathbb{R}_e$ be an outer measure. $E \subseteq X$ is said to be measurable with respect to $\mu_o$ if $\forall A \subseteq X$, we have $\mu_o(A) = \mu_o(A \cap E) + \mu_o(A \cap \overline{E})$.

Lemma 11.16 Let $X$ be a set and $\mu_o : \mathcal{X}_2 \to [0, \infty] \subset \mathbb{R}_e$ 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 $\mu_o(A \cap \overline{E}) \leq \mu_o(A)$. Then, we have $\mu_o(A) \leq \mu_o(A \cap E) + \mu_o(A \cap \overline{E}) \leq \mu_o(A)$. Hence, $E$ is measurable with respect to $\mu_o$. This completes the proof of the lemma.

Theorem 11.17 Let $X$ be a set, $\mu_o : \mathcal{X}_2 \to [0, \infty] \subset \mathbb{R}_e$ be an outer measure, $\mathcal{B} := \{ E \subseteq X \mid E \text{ is measurable with respect to } \mu_o \}$, and $\bar{\mu} := \mu_o|_{\mathcal{B}} : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_e$. Then, $\mathcal{B}$ is a $\sigma$-algebra on $X$ and $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space.

Proof Clearly $\emptyset, X \subseteq \mathcal{B}$. $\forall E \in \mathcal{B}$, we have $\overline{E} \in \mathcal{B}$. $\forall E_1, E_2 \in \mathcal{B}$, $\forall A \subseteq X$, by the measurability of $E_1$ and $E_2$, we have $\mu_o(A) = \mu_o(A \cap E_1) + \mu_o(A \cap E_2) + \mu_o(A \cap E_1 \cap E_2) = \mu_o(A \cap E_1) + \mu_o(A \cap E_2) + \mu_o(A \cap E_1 \cap E_2)$.

Note that $A \cap (E_1 \cup E_2) = A \cap (E_1 \cup (E_1 \cap E_2)) = (A \cap E_1) \cup (A \cap E_2) \cap (A \cap E_1 \cap E_2)$. Then, we have $\mu_o(A \cap (E_1 \cup E_2)) \leq \mu_o(A \cap E_1) + \mu_o(A \cap E_2)$. Then, we have $\mu_o(A) \geq \mu_o(A \cap (E_1 \cup E_2)) + \mu_o(A \cap E_1 \cap E_2) = \mu_o(A \cap (E_1 \cup E_2)) + \mu_o(A \cap E_1 \cap E_2) \geq \mu_o(A)$. Hence, $E_1 \cup E_2 \in \mathcal{B}$. Thus, $\mathcal{B}$ 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 $\overline{E}_1 := G_1 \in \mathcal{B}$, $\overline{E}_{n+1} := G_{n+1} \setminus \bigcup_{i=1}^n E_i$, $\forall n \in \mathbb{N}$. Then, $E_1, E_2, \ldots$ are pairwise disjoint. Let $E := \bigcup_{i=1}^\infty E_i = \bigcup_{i=1}^\infty \overline{E}_i$, $\forall A \subseteq X$, we have $\mu_o(A) = \mu_o(A \cap G_n) + \mu_o(A \cap \overline{G}_n) \geq \mu_o(A \cap G_n) + \mu_o(A \cap \overline{E}_n)$ and $\mu_o(A \cap G_{n+1}) = \mu_o(A \cap G_{n+1} \cap \overline{E}_{n+1}) + \mu_o(A \cap G_{n+1} \cap \overline{E}_{n+1}) = \mu_o(A \cap \overline{E}_{n+1}) + \mu_o(A \cap G_{n+1})$, $\forall n \in \mathbb{N}$. Hence, we have $\mu_o(A \cap G_n) = \sum_{i=1}^n \mu_o(A \cap \overline{E}_i)$ and $\mu_o(A) \geq \sum_{i=1}^\infty \mu_o(A \cap \overline{E}_i) + \mu_o(A \cap \overline{E})$, $\forall A \subseteq X$. Therefore, $(X, \mathcal{B}, \bar{\mu})$ is a complete measure space.
\( \forall n \in \mathbb{N} \). Therefore, \( \mu_o(A) \geq \sum_{i=1}^{\infty} \mu(A \cap E_i) + \mu_o(A \cap \bar{E}) \). Note that \( A \cap E = \bigcup_{i=1}^{\infty} (A \cap E_i) \), which implies that \( \mu_o(A \cap \bar{E}) \leq \sum_{i=1}^{\infty} \mu(A \cap E_i) \), by countable subadditivity of outer measures. Then, we have \( \mu_o(A) \geq \sum_{i=1}^{\infty} \mu(A \cap E_i) + \mu_o(A \cap \bar{E}) \geq \mu_o(A) \). 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 \( \mathcal{A} \subseteq 2^X \) be an algebra on \( X \). A measure on algebra \( \mathcal{A} \) is a function \( \mu : \mathcal{A} \to [0, \infty] \subseteq \mathbb{R}_e \) that satisfies:

(i) \( \mu(\emptyset) = 0 \);

(ii) \( \forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) with \( A := \bigcup_{i=1}^{\infty} A_i \subseteq \mathcal{A} \) 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, \( \mathcal{A} \subseteq 2^X \) be an algebra on \( X \), and \( \mu : \mathcal{A} \to [0, \infty] \subseteq \mathbb{R}_e \) be a measure on algebra \( \mathcal{A} \). Define \( \mu_o : 2^X \to [0, \infty] \subseteq \mathbb{R}_e \) by

\[
\mu_o(A) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}, A \subseteq \bigcup_{i=1}^{\infty} A_i} \sum_{i=1}^{\infty} \mu(A_i), \quad \forall A \subseteq X
\]

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 \( \bar{\mu} := \mu_o|_\mathcal{B} : \mathcal{B} \to [0, \infty] \subseteq \mathbb{R}_e \). 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 \( \mu \).
11.2. OUTER MEASURE AND THE EXTENSION THEOREM

(v) If \( \mu \) is \( \sigma \)-finite, then the measure space \( (X, \mathcal{B}, \bar{\mu}) \) is the completion of the measure space \( (X, \mathcal{B}, \mu|_{\mathcal{B}}) \).

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 \( G_{n+1} := A \cap A_{n+1} \cap (\bigcap_{i=1}^{n} \bar{A}_i) \in \mathcal{A}, \ \forall n \in \mathbb{N} \). 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 \( \mu_0(A) = \mu(A) \).

**Claim 11.19.2** \( \mu_0 \) is an outer measure.

**Proof of claim:** Clearly, \( \mu_0(\emptyset) = \mu(\emptyset) = 0 \) since \( \emptyset \in \mathcal{A} \). \( \forall A_1 \subseteq A_2 \subseteq X, \) by the definition of \( \mu_0, \) we have \( \mu_0(A_1) \leq \mu_0(A_2) \). We need only to show the countable subadditivity for \( \mu_0 \). Fix any \( E \subseteq \bigcup_{i=1}^\infty E_i \subseteq X \).

We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( \mu_0(E_i) < \infty, \ \forall i \in \mathbb{N}; \) Case 2: \( \exists i_0 \in \mathbb{N} \) such that \( \mu_0(E_{i_0}) = \infty \).

Case 1: \( \mu_0(E_i) < \infty, \ \forall i \in \mathbb{N}. \) \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \ \forall i \in \mathbb{N}, \ \exists (A_{i,j})_{j=1}^\infty \subseteq \mathcal{A} \) such that \( E_i \subseteq \bigcup_{j=1}^\infty A_{i,j} \) and \( \mu_0(E_i) \leq \sum_{j=1}^\infty \mu(A_{i,j}) < \mu_0(E_i) + \epsilon/2^i \).

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 \( \mu_0, \) we have \( \mu_0(E) \leq \sum_{i=1}^\infty \sum_{j=1}^\infty \mu(A_{i,j}) \leq \sum_{i=1}^\infty \mu_0(E_i) + \epsilon. \) By the arbitrariness of \( \epsilon, \) we have \( \mu_0(E) \leq \sum_{i=1}^\infty \mu_0(E_i). \)

Case 2: \( \exists i_0 \in \mathbb{N} \) such that \( \mu_0(E_{i_0}) = \infty. \) Then, \( \mu_0(E) \leq \infty = \sum_{i=1}^\infty \mu_0(E_i). \)

In both cases, we have shown that \( \mu_0(E) \leq \sum_{i=1}^\infty \mu_0(E_i). \) Hence, \( \mu_0 \) is an outer measure. This completes the proof of the claim.

Hence, the statement (i) holds.

(ii) By Theorem 11.17, \((X, \mathcal{B}, \bar{\mu})\) is a complete measure space.

**Claim 11.19.3** \( \mathcal{A} \subseteq \mathcal{B}. \)

**Proof of claim:** Fix any \( A \in \mathcal{A}, \ \forall \bar{A} \subseteq X, \) we will show that \( \mu_0(\bar{A}) = \mu_0(\bar{A} \cap A) + \mu_0(\bar{A} \cap \overline{A}) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \mu_0(\bar{A}) = \infty; \) Case 2: \( \mu_0(\bar{A}) < \infty. \) Case 1: \( \mu_0(\bar{A}) = \infty. \) Then, \( \mu_0(\bar{A}) \geq \mu_0(\bar{A} \cap A) + \mu_0(\bar{A} \cap \overline{A}) \geq \mu_0(\bar{A}), \) where the second inequality follows from the countable subadditivity of an outer measure. This case is proved.

Case 2: \( \mu_0(\bar{A}) < \infty. \) \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \ \exists (A_i)_{i=1}^\infty \subseteq \mathcal{A} \) such that \( \bar{A} \subseteq \bigcup_{i=1}^\infty A_i \) and \( \mu_0(\bar{A}) \leq \sum_{i=1}^\infty \mu(A_i) < \mu_0(\bar{A}) + \epsilon. \) By the countable additivity of \( \mu, \) we have \( \mu(A_i) = \mu(A_i \cap A) + \mu(A_i \cap \overline{A}), \ \forall i \in \mathbb{N}. \) Then,
\[ \mu_\circ (\bar{A}) + \epsilon > \sum_{i=1}^{\infty} (\mu(A_i \cap A) + \mu(A_i \cap \bar{A})) \geq \mu_\circ (\bar{A} \cap A) + \mu_\circ (\bar{A} \cap \bar{A}) \geq \mu_\circ (\bar{A}), \]

where the second inequality follows from the definition of \( \mu_\circ \); and the third inequality follows from the countable subadditivity of \( \mu_\circ \). By the arbitrariness of \( \epsilon \), we have \( \mu_\circ (\bar{A}) = \mu_\circ (\bar{A} \cap A) + \mu_\circ (\bar{A} \cap \bar{A}) \). This case is proved.

Hence, in both cases, we have \( \mu_\circ (\bar{A}) = \mu_\circ (A \cap B) + \mu_\circ (A \cap \bar{A}) \). By the arbitrariness of \( \bar{A} \), we have \( A \) is measurable with respect to \( \mu_\circ \). 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 \( \bar{\mu}(A) = \mu_\circ (A) = \mu(A) \). 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 \( \bar{\mu} : \mathcal{B}_a \rightarrow [0, \infty) \subseteq \mathbb{R} \) 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, \( \bar{\mu}_{\mid \mathcal{B}_a} \) is a measure on \( (X, \mathcal{B}_a) \) and it is an extension of \( \mu \).

Let \( \mathcal{A}_\circ := \{ E \subseteq X \mid \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni E = \bigcup_{i=1}^{\infty} A_i \} \). Since \( \mathcal{B} \supseteq \mathcal{A} \) and is a \( \sigma \)-algebra on \( X \), then \( \mathcal{A}_\circ \subseteq \mathcal{B}_a \subseteq \mathcal{B} \).

**Claim 11.19.4** \( \forall E \in \mathcal{A}_\circ \), \( \bar{\mu}(E) = \bar{\mu}(E) \).

**Proof of claim:** Fix any \( E \in \mathcal{A}_\circ \). Then, \( \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \ni \quad E = \bigcup_{i=1}^{\infty} A_i \). 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, \( \mu_\circ (B) = \mu_\circ (B) < \infty \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), by the definition of \( \mu_\circ \), \( \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) such that \( B \subseteq \bigcup_{i=1}^{\infty} A_i \Rightarrow E \in \mathcal{A}_\circ \subseteq \mathcal{B}_a \) and \( \bar{\mu}(B) = \mu_\circ (B) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \bar{\mu}(A_i) = \bar{\mu}(E) \). Then, \( \bar{\mu}(E) \leq \bar{\mu}(E) = \mu_\circ (E) \leq \sum_{i=1}^{\infty} \mu(A_i) < \bar{\mu}(B) + \epsilon \), 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, \( \bar{\mu}(E) = \bar{\mu}(E) + \mu_\circ (E \setminus B) < \bar{\mu}(B) + \epsilon \). Since \( \bar{\mu}(B) < \infty \), then \( \bar{\mu}(E \setminus B) < \epsilon \). Note that \( E \setminus B \in \mathcal{B}_a \). This implies that \( \bar{\mu}(E) \leq \bar{\mu}(E) = \bar{\mu}(B) + \mu_\circ (E \setminus B) \leq \bar{\mu}(B) + \mu_\circ (E \setminus B) < \bar{\mu}(B) + \epsilon \), 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 $\exists (X_i)_{i=1}^{\infty} \subseteq A$ 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 $\hat{\mu}(B \cap X_i) \leq \hat{\mu}(X_i) < \infty$ with $B \cap X_i \in \mathcal{B}_a$, $\forall i \in \mathbb{N}$. This implies that $\hat{\mu}(B) = \sum_{i=1}^{\infty} \hat{\mu}(B \cap X_i) = \sum_{i=1}^{\infty} \hat{\mu}(B \cap X_i) = \hat{\mu}(B)$, where the first and last equality follows from the countable additivity of measures; and the second equality follows from Claim 11.19.5. Hence, $\hat{\mu}|_{\mathcal{B}_a} = \hat{\mu}$. Therefore, $\hat{\mu}|_{\mathcal{B}_a}$ 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 $(X, \mathcal{B}_a, \hat{\mu}|_{\mathcal{B}_a})$. By Proposition 11.12, this measure space admits the completion $(X, \mathcal{B}, \hat{\mu})$.

Since $(X, \mathcal{B}, \hat{\mu})$ is a complete measure space that agrees with $\hat{\mu}|_{\mathcal{B}_a}$ on $\mathcal{B}_a$, by Proposition 11.12, we have $\mathcal{B} \subseteq \mathcal{B}$ and $\hat{\mu} = \hat{\mu}|_{\mathcal{B}}$.

**Claim 11.19.6** $\forall B \in \mathcal{B}$, $\exists U \in \mathcal{B}_a$ with $B \subseteq U$ such that $\hat{\mu}(U \setminus B) = 0$.

**Proof of claim:** $\forall B \in \mathcal{B}$, $\hat{\mu}(B) = \inf_{(A_i)_{i=1}^{\infty} \subseteq A, B \subseteq \bigcup_{i=1}^{\infty} A_i} \sum_{i=1}^{\infty} \mu(A_i)$.

We first prove the special case $\hat{\mu}(B) < +\infty$. $\forall i \in \mathbb{N}$, $\exists (U_{i,j})_{j=1}^{\infty} \subseteq A$ with $B \subseteq \bigcup_{j=1}^{\infty} U_{i,j} =: U_i \in \mathcal{B}_a$ such that $\hat{\mu}(B) \leq \hat{\mu}(U_i) = \hat{\mu}(U_i) \leq \sum_{j=1}^{\infty} \mu(U_{i,j}) < \hat{\mu}(B) + 2^{-i} < +\infty$. Then, $\hat{\mu}(U_i) = \hat{\mu}(B) + \hat{\mu}(U_i \setminus B) < \hat{\mu}(B) + 2^{-i} < +\infty$. This implies that $\hat{\mu}(U_i \setminus B) < 2^{-i}$. Let $U := \bigcap_{i=1}^{\infty} U_i \in \mathcal{B}_a$. Clearly, $B \subseteq U$ and $\hat{\mu}(U \setminus B) \leq \hat{\mu}(U_i \setminus B) < 2^{-i}$, $\forall i \in \mathbb{N}$. Hence, $\hat{\mu}(U \setminus B) = 0$.

Now, we prove the general case $\hat{\mu}(B) \in [0, \infty] \subset \mathbb{R}_e$. $B = \bigcup_{i=1}^{\infty} (X_i \cap B) =: \bigcup_{i=1}^{\infty} B_i$. $\forall i \in \mathbb{N}$, $\hat{\mu}(B_i) \leq \hat{\mu}(X_i) = \mu(X_i) < +\infty$. By the special case, $\exists U_i \in \mathcal{B}_a$ with $B_i \subseteq U_i$ such that $\hat{\mu}(U_i \setminus B_i) = 0$. Then, $B \subseteq \bigcup_{i=1}^{\infty} U_i =: \tilde{U} \in \mathcal{B}_a$ and $0 \leq \hat{\mu}(U \setminus B) \leq \sum_{i=1}^{\infty} \hat{\mu}(U_i \setminus B_i) = 0$, where the second inequality follows from the countable subadditivity of measure. This completes the proof of the claim.

$\forall B \in \mathcal{B}$, by Claim 11.19.6, $\exists U, \tilde{U} \in \mathcal{B}_a$ with $B \subseteq U$ and $\tilde{B} \subseteq \tilde{U}$ such that $\hat{\mu}(U \setminus B) = \hat{\mu}(\tilde{U} \setminus \tilde{B}) = 0$. Let $F := \tilde{U} \in \mathcal{B}_a$. Then, $F \subseteq B$ and $\hat{\mu}(B \setminus F) = \hat{\mu}(\tilde{U} \setminus \tilde{B}) = 0$. Then, $B = F \cup (B \setminus F)$, $B \setminus F \subseteq U \setminus F$, $F \setminus U \subseteq \mathcal{B}_a$ and $\hat{\mu}(U \setminus F) = \hat{\mu}|_{\mathcal{B}_a}(U \setminus F) = \hat{\mu}(U \setminus F) = \hat{\mu}(U \setminus B) + \hat{\mu}(B \setminus F) = 0$. This shows that $B \in \mathcal{B}$, by the completeness of $(X, \mathcal{B}, \hat{\mu})$. By the arbitrariness of $B$, we have $\mathcal{B} \subseteq \mathcal{B}$. Hence, $\mathcal{B} = \mathcal{B}$ and $\hat{\mu} = \hat{\mu}$. Hence $(X, \mathcal{B}, \hat{\mu})$ is the completion of $(X, \mathcal{B}_a, \hat{\mu}|_{\mathcal{B}_a})$.

This completes the proof of the theorem.

**Definition 11.20** An interval $I$ in $\mathbb{R}$ is any of the following sets: (i) $\emptyset$; (ii) $(a, b) \subseteq \mathbb{R}$ with $a, b \in \mathbb{R}_e$ and $a < b$; (iii) $[a, b) \subseteq \mathbb{R}$ with $a \in \mathbb{R}, b \in \mathbb{R}_e$, and $a < b$; (iv) $(a, b] \subseteq \mathbb{R}$ with $a \in \mathbb{R}_e, b \in \mathbb{R}$, and $a < b$; and (v) $[a, b] \subseteq \mathbb{R}$ with $a, b \in \mathbb{R}$ and $a \leq b$.

**Example 11.21** Let $X = \mathbb{R}$,

$$\mathcal{A} := \{E \subseteq \mathbb{R} \mid E \text{ is the union of finitely many intervals in } \mathbb{R}\}$$
It is clear that $\mathcal{A}$ is an algebra on $\mathbb{R}$. $\forall E \in \mathcal{A}$, $E$ 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

$$\mu(I) = \begin{cases} 0 & \text{if } I = \emptyset \\ b - a & \text{if } I \text{ is any other type of interval} \end{cases}$$

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, $\mu : \mathcal{A} \to [0, \infty] \subseteq \mathbb{R}$, is well-defined and $\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$, $\forall E_1, E_2 \in \mathcal{A}$ with $E_1 \cap E_2 = \emptyset$. Hence, $\mu$ is finitely additive.

Fix any interval $I = \bigcup_{i=1}^{\infty} E_i$, where $(E_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ 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 $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. Case 1: $\sum_{i=1}^{\infty} \mu(E_i) = \infty$. Then, $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i) = \infty$. Hence, we have $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$. Case 3: $I \neq \emptyset$ and $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. 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 $I \neq \emptyset$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\forall i \in \mathbb{N}$, $\forall j \in \{1, \ldots, n\}$, we may enlarge $I_{i,j}$ to an open interval $I_{i,j} \supseteq I_{i,j}$ with $\mu(I_{i,j}) \leq \mu(I_{i,j}) + \frac{2 \epsilon}{n_i}$. Then, $\bar{I} \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} I_{i,j}$. Clearly, we have $b - a = \mu(I) \leq \sum_{i=1}^{N} \sum_{j=1}^{n_i} \mu(I_{i,j}) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j}) < \infty$. By the arbitrariness of $\epsilon, a, b$, we have $\mu(I) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} \mu(I_{i,j})$. Then, $\mu(I) \leq \sum_{i=1}^{\infty} \mu(E_i)$. Hence, in all three cases, we have $\mu(I) = \sum_{i=1}^{\infty} \mu(E_i)$.

$\forall E \in \mathcal{A}, \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with $E = \bigcup_{i=1}^{n} E_i$ and $E_1, E_2, \ldots$ being pairwise disjoint, we have $E = \bigcup_{i=1}^{\infty} I_i$, where $n \in \mathbb{N}$ and $I_1, I_2, \ldots$ are pairwise disjoint intervals. Then, $I_j = \bigcup_{i=1}^{\infty} (I_j \cap E_i)$, $\forall j \in \{1, \ldots, n\}$. By the above argument, we have $\mu(I_j) = \sum_{i=1}^{\infty} \mu(I_j \cap E_i)$. Then, $\mu(E) = \sum_{j=1}^{n} \mu(I_j) = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \mu(I_j \cap E_i) = \sum_{j=1}^{n} \mu(I_j \cap E_i) = \sum_{i=1}^{\infty} \mu(E_i)$, where the last equality follows from the finite additivity of $\mu$.

Hence, $\mu$ is a measure on algebra $\mathcal{A}$ and is $\sigma$-finite. By Carathéodory Extension Theorem 11.19, $\mu$ extends uniquely to a $\sigma$-finite complete measure space $(X, \mathcal{B}, \mu)$. This measure space is called the Lebesgue measure space. Since $\mathcal{B}$ is a $\sigma$-algebra on $\mathbb{R}$ and the topology $\mathcal{O}_\mathbb{R}$ on $\mathbb{R}$ is second countable, then $\mathcal{B} \supseteq \mathcal{O}_\mathbb{R}$. We will denote the Lebesgue measure space by $(\mathbb{R}, \mathcal{B}_L, \mu_L)$, where $\mathcal{B}_L$ 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 $\mu_L$ be the Lebesgue outer measure induced by $\mu$ as defined in Carathéodory Extension Theorem 11.19. We claim that

$$\mu_L(E) = \inf_{(I_i)_{i=1}^\infty \text{ are open intervals}} E \subseteq \bigcup_{i=1}^\infty I_i \sum_{i=1}^\infty \mu(I_i); \quad \forall E \subseteq \mathbb{R} \quad (11.3)$$

To prove this, we note that $\mu_L(E) = \inf_{(E_i)_{i=1}^\infty \subseteq \mathcal{A}} E \subseteq \bigcup_{i=1}^\infty E_i \sum_{i=1}^\infty \mu(E_i)$. Then, clearly, we have

$$\mu_L(E) \leq \inf_{(I_i)_{i=1}^\infty \text{ are open intervals}} E \subseteq \bigcup_{i=1}^\infty I_i \sum_{i=1}^\infty \mu(I_i)$$

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 $\mu(\bar{I}_{i,j}) \leq \mu(I_{i,j}) + \frac{2^{-i} \epsilon}{n_i}$. 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 $\sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(I_{i,j}) \geq \sum_{i=1}^\infty \sum_{j=1}^{n_i} \left( \mu(\bar{I}_{i,j}) - \frac{2^{-i} \epsilon}{n_i} \right) = \sum_{i=1}^\infty \sum_{j=1}^{n_i} \mu(\bar{I}_{i,j}) - \epsilon \geq \mu \left( \bigcup_{i=1}^\infty \bigcup_{j=1}^{n_i} \bar{I}_{i,j} \right) - \epsilon \geq \mu_L(E) - \epsilon$. Thus, that $\mu_L(E) \geq \mu_L(E) - \epsilon$. Hence, by the arbitrariness of $\epsilon$, we have

$$\mu_L(E) \geq \inf_{(I_i)_{i=1}^\infty \text{ are open intervals}} E \subseteq \bigcup_{i=1}^\infty I_i \sum_{i=1}^\infty \mu(I_i)$$

Therefore, (11.3) is true.

By Carathéodory Extension Theorem 11.19, $\mu_L = \mu_L|_{\mathcal{B}_L}$ and $(\mathbb{R}, \mathcal{B}_L, \mu_L)$ is the completion of $(\mathbb{R}, \mathcal{B}_o, \mu|_{\mathbb{R}})$, where $\mathcal{B}_o$ is the $\sigma$-algebra on $\mathbb{R}$ generated by $\mathcal{A}$. Clearly, $\mathcal{O}_L \subseteq \mathcal{B}_o$ and $\mathcal{B}_o$ is the $\sigma$-algebra on $\mathbb{R}$ generated by $\mathcal{O}_L$.

**Definition 11.22** Let $X := (X, \mathcal{O})$ be a topological space. The collection of Borel sets on $X$ is the smallest $\sigma$-algebra on $X$ that contains $\mathcal{O}$, which will be denoted by $\mathcal{B}_B(X)$. The smallest algebra on $X$ that contains $\mathcal{O}$ will be denoted by $\mathcal{A}(X)$.

Clearly, we have $\mathcal{B}_B(\mathbb{R}) \subseteq \mathcal{B}_L$.

**Proposition 11.23** Let $E \subseteq \mathbb{R}$. Then, the following statements are equivalent.

(i) $E \in \mathcal{B}_L$.

(ii) $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \mathcal{O} \in \mathcal{O}_L$ with $E \subseteq \mathcal{O}$ such that $\mu_L(O \setminus E) < \epsilon$. 

(iii) ∀ε ∈ (0, ∞) ⊂ ℝ, ∃F ⊆ ℝ with ̃F ∈ O_ℝ and F ⊆ E such that μ_{Ło}(E \ F) < ε.

If μ_{Ło}(E) < +∞. Then, (i)–(iii) are also equivalent to

(iv) ∀ε ∈ (0, ∞) ⊂ ℝ, ∃U ∈ O_ℝ which is finite union of open intervals, such that μ_{Ło}(U △ E) < ε.

Hence, the Lebesgue measure space (ℝ, ℬ_Ł, μ_L) is the completion of the Borel measure space (ℝ, ℬ_B(ℝ), μ_B := μ_L|ℬ_Ł(ℝ)).

Proof Fix E ⊆ ℝ.

(i) ⇒ (ii). First, we show the special case that μ_{Ło}(E) < +∞. Let E ∈ ℬ_Ł. By Example 11.21,

\[
\mu_{Ło}(E) = \mu_L(E) = \inf_{(I_i)_{i=1}^{∞} \text{ are open intervals, } E \subseteq \bigcup_{i=1}^{∞} I_i} \sum_{i=1}^{∞} \mu_L(I_i)
\]

Then, ∀ε ∈ (0, ∞) ⊂ ℝ, ∃ a sequence of open intervals (I_i)_{i=1}^{∞} ⊆ O_ℝ with E ⊆ \bigcup_{i=1}^{∞} I_i := O ∈ O_ℝ such that \mu_L(O) ≤ μ_L(O) ≤ \sum_{i=1}^{∞} \mu_L(I_i) < \mu_L(E) + ε < +∞. Then, ε > μ_L(O) − μ_L(E) = μ_L(O \ E) = μ_{Ło}(O \ E),

where the first equality follows from the countable additivity of μ_L. Hence, (ii) holds.

Next, we show the general case μ_{Ło}(E) ∈ [0, ∞] ⊂ ℝ_ε. Let E ∈ ℬ_Ł and I_1 := (−1, 1] ⊂ ℝ, I_n := (−n, −n + 1] ∪ (n − 1, n] ⊂ ℝ, ∀n = 2, 3, …

Clearly, E = \bigcup_{i=1}^{∞} I_i ⊆ E_i. Fix any ε ∈ (0, ∞) ⊂ ℝ. ∀i ∈ ℕ, I_i ∈ ℬ_Ł and E_i ∈ ℬ_Ł with μ_L(E_i) ≤ μ_L(I_i) = 2. By the special case we have just shown, ∃O_i ∈ O_ℝ ⊆ ℬ_Ł with E_i ⊆ O_i such that μ_{Ło}(O_i \ E_i) = μ_L(O_i \ E_i) < 2^{–i}ε. Let O := \bigcup_{i=1}^{∞} O_i ∈ O_ℝ. Then, μ_{Ło}(O \ E) = μ_L(O \ E) = μ_L((\bigcup_{i=1}^{∞} O_i) \ (∪_{i=1}^{∞} E_i)) = μ_L(∪_{i=1}^{∞} (O_i \ (\bigcap_{j=1}^{∞} E_j))) ≤ \sum_{i=1}^{∞} μ_L(O_i \ (\bigcap_{j=1}^{∞} E_j)) ≤ \sum_{i=1}^{∞} μ_L(O_i \ E_i) = \lim_{n \to ∞} μ_L(O \ E_i) < ε. Hence (ii) holds.

(ii) ⇒ (i). Fix any A ⊆ ℝ. ∀ε ∈ (0, ∞) ⊂ ℝ, ∃O ∈ O_ℝ with E ⊆ O such that μ_{Ło}(O \ E) < ε. Note that A ∩ E = (A ∩ E ∩ O) ∪ (A ∩ ̃E ∩ O) ⊆ (O \ E) ∪ (A ∩ ̃E) and A ∩ E ⊆ A ∩ O. By Definition 11.14, we have μ_{Ło}(A) ≤ μ_{Ło}(A ∩ E) + μ_{Ło}(A ∩ ̃E) ≤ μ_{Ło}(A ∩ O) + μ_{Ło}(O \ E) + μ_{Ło}(A ∩ ̃E) ≤ μ_{Ło}(A) + ε, where the third inequality follows from the fact that O ∈ O_ℝ ⊆ ℬ_Ł and Definition 11.15. By the arbitrariness of ε, we have μ_{Ło}(A) = μ_{Ło}(A ∩ E) + μ_{Ło}(A ∩ ̃E). By the arbitrariness of A, we have E ∈ ℬ_Ł.

(i) ⇔ (iii). E ∈ ℬ_Ł ⇔ ̃E ∈ ℬ_Ł ⇔ ∀ε ∈ (0, ∞) ⊂ ℝ, ∃O ∈ O_ℝ with ̃E ⊆ O such that μ_{Ło}(O \ ̃E) < ε, where the last ⇔ follows from (i) ⇔ (ii).

Let F := ̃Ω, we have F ⊆ E and μ_{Ło}(E \ F) = μ_{Ło}(O \ ̃E) < ε.

Next, we show (i) − (iii) are equivalent to (iv) under the additional assumption that μ_{Ło}(E) < +∞.
(i) ⇒ (iv). Let \( E \in \mathcal{B}_L \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by Example 11.21, \( \exists \) a sequence of open intervals \( (I_i)_{i=1}^\infty \) with \( E \subseteq \bigcup_{i=1}^\infty I_i \) such that \( \mu_{\mathcal{L}}(E) = \mu_\mathcal{L}(E) \leq \sum_{i=1}^\infty \mu_\mathcal{L}(I_i) \leq \mu_\mathcal{L}(E) + \epsilon/2 < +\infty \). Then, \( \exists N \in \mathbb{N} \) such that \( \sum_{i=N+1}^\infty \mu_\mathcal{L}(I_i) < \epsilon/2 \). Let \( U := \bigcup_{i=1}^N I_i \in \mathcal{O}_\mathbb{R} \). We have \( \mu_{\mathcal{L}}(U \triangle E) = \mu_\mathcal{L}(U \setminus E) + \mu_\mathcal{L}(E \setminus U) \leq \mu_\mathcal{L}(O \setminus E) + \mu_\mathcal{L}(E \cap \tilde{U}) \). Note that \( \mu_\mathcal{L}(O) = \mu_\mathcal{L}(E) + \mu_\mathcal{L}(O \setminus E) \leq \sum_{i=1}^\infty \mu_\mathcal{L}(I_i) \leq \mu_\mathcal{L}(E) + \epsilon/2 < +\infty \) and \( E \cap \tilde{U} = (E \cap \tilde{U} \cap O) \cup (E \cap \tilde{U} \cap \tilde{O}) = E \cap \tilde{U} \cap O \subseteq O \setminus U \subseteq \bigcup_{i=N+1}^\infty I_i \). Then, \( \mu_{\mathcal{L}}(U \triangle E) \leq \epsilon/2 + \mu_\mathcal{L}(\bigcup_{i=N+1}^\infty I_i) \leq \epsilon/2 + \sum_{i=N+1}^\infty \mu_\mathcal{L}(I_i) < \epsilon \). Hence, (iv) holds.

(iv) ⇒ (i). Fix any set \( A \subseteq \mathbb{R} \). By (iv), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists U \in \mathcal{O}_\mathbb{R} \) which is finite union of open intervals, such that \( \mu_{\mathcal{L}}(U \triangle E) < \epsilon/2 \). Note that \( A \cap E = (A \cap U \cap E) \cup (A \cap E \cap \tilde{U}) \subseteq (A \cap U) \cup (E \setminus U) \) and \( A \cap \tilde{E} = (A \cap \tilde{E} \cap U) \cup (A \cap \tilde{E} \cap \tilde{U}) \subseteq (U \setminus E) \cup (A \cap \tilde{U}) \). By Definition 11.14, \( \mu_{\mathcal{L}}(A) \leq \mu_{\mathcal{L}}(A \cap U) + \mu_{\mathcal{L}}(A \cap \tilde{E}) \leq \mu_{\mathcal{L}}((A \cap U) \cup (E \setminus U)) + \mu_{\mathcal{L}}((\tilde{E} \cap U) \cup (A \cap \tilde{U}) \cup (\tilde{E} \cap \tilde{U}) \cup \mu_{\mathcal{L}}(E \setminus U) + \mu_{\mathcal{L}}(U \setminus E) + \mu_{\mathcal{L}}(A \cap \tilde{U}) = \mu_{\mathcal{L}}(A) + \mu_{\mathcal{L}}(E \setminus U) + \mu_{\mathcal{L}}(U \setminus E) \leq \mu_{\mathcal{L}}(A) + 2\mu_{\mathcal{L}}(E \triangle U) \leq \mu_{\mathcal{L}}(A) + \epsilon \), where the equality follows from the fact that \( U \in \mathcal{O}_\mathbb{R} \subseteq \mathcal{B}_L \) and Definition 11.15. By the arbitrariness of \( \epsilon \), we have \( \mu_{\mathcal{L}}(A) = \mu_{\mathcal{L}}(A \cap U) + \mu_{\mathcal{L}}(A \cap \tilde{E}) \). 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 \((\mathbb{R}, \mathcal{B}_B(\mathbb{R}), \mu_{\mathcal{L}}|_{\mathcal{B}_B(\mathbb{R})})\). This completes the proof of the proposition.

\( \square \)

**Proposition 11.24** Let \( X_i := (X_i, \mathcal{O}_i) \) be a second countable topological space, \( i = 1, 2 \), \( X := X_1 \times X_2 =: (X_1 \times X_2, \mathcal{O}) \) be the product topological space, \( \mathcal{E} := \{ B_1 \times B_2 \subseteq X_1 \times X_2 \mid B_i \in \mathcal{B}_B(X_i), i = 1, 2 \} \), and \( \mathcal{B} \) be the \( \sigma \)-algebra on \( X_1 \times X_2 \) generated by \( \mathcal{E} \). Then, \( \mathcal{B} = \mathcal{B}_B(X) \).

**Proof** Let \( \mathcal{O}_{B_i} \subseteq \mathcal{O}_i \) 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 \( \mathcal{O}_B := \{ O_{B_1} \times O_{B_2} \subseteq X_1 \times X_2 \mid O_{B_1} \in \mathcal{O}_{B_1}, i = 1, 2 \} \subseteq \mathcal{O} \). Clearly, \( \mathcal{O}_{B_1} \subseteq \mathcal{O}_i \subseteq \mathcal{B}_B(X_i), i = 1, 2 \). Then, \( \mathcal{O}_B \subseteq \mathcal{E} \). \( \forall O \in \mathcal{O}, O = \bigcup_{i=1}^\infty O_i \), where \( O_i \in \mathcal{O}_B, \forall i \in \mathbb{N} \). Hence, \( O \in \mathcal{B} \). This implies that \( \mathcal{O} \subseteq \mathcal{B} \). Since \( \mathcal{B} \) is a \( \sigma \)-algebra, then \( \mathcal{O} \subseteq \mathcal{B}_B(X) \subseteq \mathcal{B} \).

On the other hand, note that \( \mathcal{B}_B(X) \) is the \( \sigma \)-algebra generated by \( \mathcal{O} \). Then, \( \forall O_1 \in \mathcal{O}_1, \forall O_2 \in \mathcal{O}_2, \) we have \( O_1 \times O_2 \in \mathcal{O} \subseteq \mathcal{B}_B(X) \).

**Claim 11.24.1** \( \forall O_1 \in \mathcal{O}_1, \forall B_2 \in \mathcal{B}_B(X_2), \) we have \( O_1 \times B_2 \in \mathcal{B}_B(X) \).

**Proof of claim:** Fix any \( O_1 \in \mathcal{O}_1 \). Define \( \mathcal{F} := \{ E \subseteq X_2 \mid O_1 \times E \in \mathcal{B}_B(X) \} \). Clearly, \( \mathcal{O}_2 \subseteq \mathcal{F} \). Then, \( \emptyset, X_2 \in \mathcal{O}_2 \subseteq \mathcal{F}, \forall E \in \mathcal{F}, \) we have \( O_1 \times E \in \mathcal{B}_B(X) \). Note that \( O_1 \times X_2 \in \mathcal{B}_B(X) \). This leads to \( O_1 \times (X_2 \setminus E) = (O_1 \times X_2) \setminus (O_1 \times E) \in \mathcal{B}_B(X) \). Thus, \( X_2 \setminus E \in \mathcal{F} \). \( \forall (E_i)_{i=1}^\infty \subseteq \mathcal{F}, \) we have \( O_1 \times E_i \in \mathcal{B}_B(X), \forall i \in \mathbb{N} \). This implies that
Then, \( \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} (O_1 \times E_i) \in B_B(\mathcal{X}) \) and \( \bigcup_{i=1}^{\infty} E_i \in \mathcal{F} \). The above shows that \( \mathcal{F} \) is a \( \sigma \)-algebra on \( X_2 \). Then, \( O_2 \subseteq B_B(X_2) \subseteq \mathcal{F} \). Hence, \( \forall B_2 \in B_B(X_2) \), 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.

\[ \Box \]

**Claim 11.24.2** \( \forall B_1 \in B_B(X_1), \forall B_2 \in B_B(X_2) \), we have \( B_1 \times B_2 \in B_B(\mathcal{X}) \).

**Proof of claim:** Fix any \( B_2 \in B_B(X_2) \). Define \( \mathcal{F} := \{ E \subseteq X_1 \mid E \times B_2 \in B_B(\mathcal{X}) \} \). 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 \( E_i \times B_2 \in B_B(\mathcal{X}) \), \( \forall i \in \mathbb{N} \). 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 \( X_1 \). Then, \( O_1 \subseteq B_B(X_1) \subseteq \mathcal{F} \). Hence, \( \forall B_1 \in B_B(X_1) \), 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, \( \mathcal{E} \subseteq B_B(\mathcal{X}) \). 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.

\[ \Box \]

**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 \( E \supseteq \bigcup_{i=1}^{\infty} A_i \in B_B(\mathcal{X}) \), \( \forall i \in \mathbb{N} \). Then, \( E \supseteq \bigcup_{i=1}^{\infty} A_i \in B_B(\mathcal{X}) \) 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 \( C_2 := C \cap (X \setminus E) \in B_B(\mathcal{E}) \).

**Proof of claim:** Define \( \mathcal{F} := \{ A_1 \cup A_2 \subseteq X \mid A_1 \in B_B(\mathcal{E}), A_2 \in B_B(\mathcal{E}) \} \). \( \forall O \in \mathcal{O}, O \cap E \in \mathcal{O}_E \subseteq B_B(\mathcal{E}) \) and \( O \cap (X \setminus E) \in \mathcal{O}_{X \setminus E} \subseteq B_B(\mathcal{E}) \). Then, \( O = (O \cap E) \cup (O \cap (X \setminus E)) \in \mathcal{F} \) and \( \mathcal{O} \subseteq \mathcal{F} \).

Clearly, \( \emptyset, E \in B_B(\mathcal{E}) \) and \( \emptyset, X \setminus E \in B_B(\mathcal{E}) \). 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 \( \exists A_2 \in B_B(\mathcal{E}) \) 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

X \setminus A = (E \cup (X \setminus E)) \setminus A = (E \setminus A) \cup ((X \setminus E) \setminus A) = (E \setminus A_1) \cup ((X \setminus E) \setminus A_2),
E \setminus A_1 \in B_B(\mathcal{E}), \text{ and } (X \setminus E) \setminus A_2 \in B_B\left(\hat{\mathcal{E}}\right). \text{ This implies that } X \setminus A \in \mathcal{F},

\forall (A_i)_{i=1}^\infty \subseteq \mathcal{F}, \text{ let } A_{i,1} := E \cap A_i \in B_B(\mathcal{E}) \text{ and } A_{i,2} := (X \setminus E) \cap A_i \in B_B\left(\hat{\mathcal{E}}\right), \forall i \in \mathbb{N}. \text{ Note that } \bigcup_{i=1}^\infty A_i = \bigcup_{i=1}^\infty (A_{i,1} \cup A_{i,2}) = \bigcup_{i=1}^\infty A_{i,1} \cup \bigcup_{i=1}^\infty A_{i,2} \in B_B\left(\hat{\mathcal{E}}\right). \text{ This implies that } \bigcup_{i=1}^\infty A_i \in \mathcal{F}. \text{ The above shows that } \mathcal{F} \text{ is a } \sigma\text{-algebra on } X. \text{ Then, } \mathcal{O} \subseteq B_B(\mathcal{X}) \subseteq \mathcal{F}. \text{ This completes the proof of the claim.} \quad \Box

\forall C \in B_E, \text{ we have } E \supseteq C \in B_B(\mathcal{X}). \text{ By Claim 11.25.1, } C = C \cap E \in B_B(\mathcal{E}). \text{ Therefore, } B_E \subseteq B_B(\mathcal{E}). \text{ This completes the proof of the proposition.} \quad \Box

**Definition 11.26** Let \( \mathcal{X} := (X, \mathcal{O}) \) 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 \( \mathcal{B} = B_B(\mathcal{X}) \) and \( \forall E \in B_B(\mathcal{X}), \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists U \in \mathcal{O} \text{ with } U \subseteq E \text{ such that } \mu(U \setminus E) < \epsilon. \) 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 \( \forall \text{ compact } K \subseteq \mathcal{X}, \mu(K) < \infty. \)

Let \( \mathcal{X} := (X, \rho) \) be a metric space, \( \mathcal{O} \) be the natural topology on \( \mathcal{X} \) 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 \( \mathcal{K} \), \( \mathcal{O} \) 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. Depending on whether \( \mathcal{K} = \mathbb{R} \) or \( \mathcal{K} = \mathbb{C} \), 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 \( \mathcal{B} = B_B(\mathcal{X}) \). The triple \( \mathcal{X} := (X, \mathcal{B}, \mu) \) is topological measure space if, and only if, \( \forall E \in B_B(\mathcal{X}), \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists X \setminus F \in \mathcal{O} \text{ with } F \subseteq E \text{ such that } \mu(E \setminus F) < \epsilon. \)

**Proof** “Necessity” \( \forall E \in \mathcal{B}, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, X \setminus E \in \mathcal{B}. \) By \( \mathcal{X} \) being a topological measure space, \( \exists U \in \mathcal{O} \text{ with } X \setminus U \subseteq U \text{ such that } \mu(U \setminus (X \setminus E)) < \epsilon. \) Then, \( X \setminus U \subseteq E \) and \( \mu(E \setminus (X \setminus U)) = \mu(E) - \mu(U \setminus (X \setminus U)) < \epsilon. \) Hence, \( F := X \setminus U \) is the set we seek.
Let $\mathbb{R}$ be the real Banach space with norm defined to be the absolute value. Then, by Proposition 11.23 and Example 11.21, $\mathbb{R} := ((\mathbb{R}, \mathbb{R}, \cdot)), B_\mathcal{B}(\mathbb{R}), \mu_\mathcal{B})$ is a $\sigma$-finite real Banach measure space.

**Proposition 11.29** Let $X := (X, \mathcal{B}, \mu)$ be a topological measure space, where $X := (X, \mathcal{O})$ is a topological space, $E \in \mathcal{B}, \mathcal{E} := (E, \mathcal{O}_E)$ be the topological space with the subset topology of $X$, and $(E, \mathcal{B}_E, \mu_E)$ be the measure subspace of $X$ as defined in Proposition 11.13. Then, $(E, \mathcal{B}_E, \mu_E)$ is a topological measure space and is said to be the topological measure subspace of $X$.

**Proof** By Definition 11.26, $\mathcal{B} = \mathcal{B}_\mathcal{B}(X)$. 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 $\mathcal{B}_E = \mathcal{B}_\mathcal{B}(E)$.

A measure space, $\forall \in (0, \infty) \in \mathbb{R}, E \supseteq A \in \mathcal{B}$. By $X$ being a topological measure space, $\exists U \in \mathcal{O}$ with $A \subseteq U$ such that $\mu(U \setminus A) < \epsilon$. Let $U := U \cap E \in \mathcal{O}_E \subseteq \mathcal{B}_E$. Clearly, $A \subseteq U$ and $0 \leq \mu_E(U \setminus A) = \mu((U \cap E) \setminus A) < \epsilon$. Hence, $E$ is a topological measure space. This completes the proof of the proposition.

**Definition 11.30** Let $X$ be a set and $\mathcal{C} \subseteq 2^X$ 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 $\mathcal{C} \subseteq 2^X$ 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 $\emptyset \in A$. 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, $X = C_1 \cup C_1 = \bigcup_{i=1}^n C_i$, where the sets in the union are pairwise disjoint. Hence, $X \in \mathcal{A}$. $\forall A_1, A_2 \in \mathcal{A}$, then $A_i = \bigcup_{j=1}^m C_{i,j}$, 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 $\tilde{1} = \bigcap_{i=1}^{n_1} \tilde{C_{i,1}}$. $\forall j \in \{1, \ldots, n_1\}$, $\tilde{C_{i,1}} = \bigcup_{i=1}^m C_{i,j,1}$, where $m_j \in \mathbb{Z}_+$ and $(C_{i,j,1})_{j=1}^{m_j} \subseteq \mathcal{C}$ is pairwise disjoint. Then,

$$\tilde{A_1} = \bigcap_{j=1}^{n_1} \bigcup_{l=1}^{m_j} C_{l,j,1} = \bigcap_{l=1}^{m_1} \bigcup_{j=2}^{m_2} \bigcup_{l=1}^{m_j} (C_{l,1,1} \cap (\bigcup_{j=2}^{m_2} \bigcup_{l=1}^{m_j} C_{l,j,1}))$$

$$= \bigcup_{l=1}^{m_1} \bigcap_{j=1}^{n_1} \bigcup_{l=1}^{m_j} \bigcup_{j=2}^{m_2} (C_{l,1,1} \cap (\bigcup_{j=3}^{m_3} \bigcup_{l=1}^{m_j} C_{l,j,1}))$$
By (ii), we have
\[
\sum_{i=1}^{n_1} \bigcup_{j=1}^{m_1} (C_{i_1,1} \cap C_{i_2,2,1} \cap \left( \bigcap_{j=3}^{m_1} C_{i,j,1} \right))
\]
\[
= \bigcup_{i_1=1}^{m_1} \bigcup_{i_2=1}^{m_2} (C_{i_1,1} \cap C_{i_2,2,1} \cap \left( \bigcap_{j=3}^{m_1} C_{i,j,1} \right))
\]

When \( n_1 = 0 \), clearly \( \widetilde{A}_1 = X \in \mathcal{A} \). When \( n_1 \geq 1 \), \( \bigcap_{j=1}^{n_1} C_{i,j,1} \in \mathcal{C} \) 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 \( \tilde{l}_1, \ldots, \tilde{l}_{n_1} \), \( \bigcap_{j=1}^{n_1} C_{i,j,1} \cap \left( \bigcap_{j=1}^{n_1} C_{\tilde{i},j,1} \right) = \emptyset \). Hence, \( \widetilde{A}_1 \in \mathcal{A} \). Note that \( A_1 \cap A_2 = \bigcup_{j=1}^{m_1} \bigcup_{j=1}^{m_2} (C_{i_1,1} \cap C_{i_2,2,1}) \). \( C_{i,1,1} \cap C_{i,2,2} \in \mathcal{C} \) for any admissible choice of indices \( j_1, j_2 \). For different admissible choices of indices \( j_1, j_2 \) and \( j_1, j_2 \), \( (C_{i_1,1} \cap C_{i_2,2,1}) \cap (C_{i_1,1} \cap C_{i_2,2,1}) = \emptyset \). 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. \( \square \)

**Proposition 11.32** Let \( X \) be a set, \( \mathcal{C} \) be a semialgebra on \( X \), and \( \mu : \mathcal{C} \to [0, +\infty] \subseteq \mathbb{R} \). Assume that

(i) \( \forall C \in \mathcal{C}, \forall n \in \mathbb{Z}_+, \forall \text{ pairwise disjoint } (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) with \( C = \bigcup_{i=1}^{n} C_i \), then \( \mu(C) = \sum_{i=1}^{n} \mu(C_i) \); 

(ii) \( \forall C \in \mathcal{C}, \forall \text{ pairwise disjoint } (C_i)_{i=1}^{\infty} \subseteq \mathcal{C} \) 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, \( \mathcal{A} = \{ A \subseteq X \mid A \text{ is the finite disjoint union of sets in } \mathcal{C} \} \). Define \( \bar{\mu} : \mathcal{A} \to [0, \infty] \subseteq \mathbb{R} \) 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 \( A_1 \cap A_2 = \emptyset \), \( A_1 = \bigcup_{i=1}^{n_1} C_{i,l}, \quad A_2 = \bigcup_{i=1}^{n_2} C_{i,l} \), where \( n_1, n_2 \in \mathbb{Z}_+ \) and \( (C_{i,l})_{i=1}^{n_1}, (C_{i,l})_{i=1}^{n_2} \subseteq \mathcal{C} \) are pairwise disjoint, \( \forall l = 1, 2 \). \( A_1 \cup A_2 = \bigcup_{i=1}^{n_1} \bigcup_{i=1}^{n_2} C_{i,l} \) where \( C_{i,l}, l = 1, 2, \quad i = 1, \ldots, n \) are pairwise disjoint. Then, \( \bar{\mu}(A_1 \cup A_2) = \sum_{i=1}^{n_1} \sum_{l=1}^{n_1} \mu(C_{i,l}) = \bar{\mu}(A_1) + \bar{\mu}(A_2) \). Hence, \( \bar{\mu} \) is finitely additive.

\[ \forall (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \text{ with } A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A} \text{ and } A_1, A_2, \ldots \text{ being pairwise disjoint. By finite additivity } \bar{\mu}(A) = \bar{\mu}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \bar{\mu}(A_i), \quad \forall n \in \mathbb{N} \text{. Then, } \bar{\mu}(A) \geq \sum_{i=1}^{\infty} \bar{\mu}(A_i). \text{ On the other hand, since } A \in \mathcal{A} \text{, then } A = \bigcup_{i=1}^{n} C_i, \text{ where } n \in \mathbb{Z}_+ \text{ and } (C_i)_{i=1}^{n} \subseteq \mathcal{C} \text{ is pairwise disjoint. } \forall i \in \mathbb{N}, A_i = \bigcup_{i=1}^{n} C_{i,j}, \text{ where } n_i \in \mathbb{Z}_+ \text{ and } (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C} \text{ is pairwise disjoint. Then, } \forall i \in \{1, \ldots, n\}, C_i = C_i \cap (\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} C_{i,j} \cap C_i. \text{ By (ii), we have } \mu(C_i) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu(C_{i,j} \cap C_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \mu(C_{i,j} \cap C_i) \].
Hence, $\mu$ is a measure on algebra $A$. Clearly, it is an extension of $\mu$ to $A$ and it is the unique extension. This completes the proof of the proposition.

\section{11.3 Measurable Functions}

\begin{definition}
Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be a topological space, $D \in \mathcal{B}$, and $f : D \to \mathcal{Y}$. We will say that $f$ is $\mathcal{B}$-measurable, if $\forall O \in \mathcal{O}_Y$, $f_{\text{inv}}(O) \in \mathcal{B}$. When it is clear from the context, we will simply say that $f$ is measurable.
\end{definition}

\begin{proposition}
Let $(X, \mathcal{B})$ be a measurable space, $\mathcal{Y} := (Y, \mathcal{O}_Y)$ be a topological space, $D \in \mathcal{B}$, and $f : D \to \mathcal{Y}$. Then,

(i) $f$ is $\mathcal{B}$-measurable if, and only if, $\forall E \in \mathcal{B}_Y (\mathcal{Y})$, $f_{\text{inv}}(E) \in \mathcal{B}$;

(ii) if $\mathcal{Y}$ is second countable with a countable basis $\mathcal{B}$, then $f$ is $\mathcal{B}$-measurable if, and only if, $\forall B_Y \in \mathcal{B}$, $f_{\text{inv}}(B_Y) \in \mathcal{B}$.
\end{proposition}

\begin{proof}
(i) “Sufficiency” is obvious. “Necessity” Let $f$ be $\mathcal{B}$-measurable. Define $B_Y := \{ E \subseteq Y \mid f_{\text{inv}}(E) \in \mathcal{B} \}$. Clearly, $\mathcal{O}_Y \subseteq B_Y$. We will show that $B_Y$ is a $\sigma$-algebra on $Y$. Then, $\mathcal{B}_Y (\mathcal{Y}) \subseteq B_Y$ and the result is established.

Clearly, $f_{\text{inv}}(\emptyset) = \emptyset \in \mathcal{B}$ and $f_{\text{inv}}(Y) = D \in \mathcal{B}$, then $\emptyset, Y \in B_Y$. $\forall E \in B_Y$, by Proposition 2.5, $f_{\text{inv}}(Y \setminus E) = f_{\text{inv}}(Y) \setminus f_{\text{inv}}(E) = D \setminus f_{\text{inv}}(E) \in \mathcal{B}$. Hence, $Y \setminus E \in B_Y$. $\forall \{ E_i \}_{i=1}^\infty \subseteq B_Y$, we have $f_{\text{inv}}(E_i) \in \mathcal{B}$, $\forall i \in \mathbb{N}$. Then, by Proposition 2.5, $f_{\text{inv}}( \bigcup_{i=1}^\infty E_i ) = \bigcup_{i=1}^\infty f_{\text{inv}}(E_i) \in \mathcal{B}$. Then, $\bigcup_{i=1}^\infty E_i \in B_Y$. This shows $B_Y$ is a $\sigma$-algebra on $Y$.

(ii) “Necessity” is obvious. “Sufficiency” $\forall O \in \mathcal{O}_Y$, since $\mathcal{Y}$ is second countable, then $\exists (B_i)_{i=1}^\infty \subseteq \mathcal{B}$ such that $O = \bigcup_{i=1}^\infty B_i$. This implies that $f_{\text{inv}}(O) = \bigcup_{i=1}^\infty f_{\text{inv}}(B_i) \in \mathcal{B}$, where the equality follows from Proposition 2.5. Hence, $f$ is $\mathcal{B}$-measurable.

This completes the proof of the proposition.
\end{proof}

\begin{proposition}
Let $(X, \mathcal{B})$ be a measurable space, $D \in \mathcal{B}$, $f : D \to \mathbb{R}$. Then, the following statements are equivalent.

(i) $f$ is $\mathcal{B}$-measurable.

(ii) $\forall \alpha \in \mathbb{R}$, the set $\{ x \in D \mid f(x) < \alpha \} \in \mathcal{B}$.

(iii) $\forall \alpha \in \mathbb{R}$, the set $\{ x \in D \mid f(x) \geq \alpha \} \in \mathcal{B}$.

(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 $\{ x \in D \mid f(x) > \alpha \} \in \mathcal{B}$.
\end{proposition}
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. \(\square\)

**Proposition 11.36** Let \(X := (X, \mathcal{B})\) be a measurable space, \(D \in \mathcal{B}, f_i : D \to [0, \infty] \subseteq \mathbb{R}_e\) be \(\mathcal{B}\)-measurable, \(i = 1, 2,\) and \(c \in (0, \infty) \subseteq \mathbb{R}.\) Then, \(c f_1\) 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. \(\square\)

**Proposition 11.37** Let \(X := (X, \mathcal{B}, \mu)\) be a topological measure space, \(\mathcal{Y}\) be a topological space, \(D \in \mathcal{B},\) and \(f : D \to \mathcal{Y}\) be continuous. Then, \(f\) is \(\mathcal{B}\)-measurable.

**Proof** The proof is straightforward, and is therefore omitted. \(\square\)

**Proposition 11.38** Let \((X, \mathcal{B})\) be a measurable space, \(\mathcal{Y}\) and \(\mathcal{Z}\) be topological spaces, \(D \in \mathcal{B}, f : D \to \mathcal{Y}\) be \(\mathcal{B}\)-measurable, and \(g : \mathcal{Y} \to \mathcal{Z}\) be continuous. Then, \(h := g \circ f : D \to \mathcal{Z}\) is \(\mathcal{B}\)-measurable.

**Proof** \(\forall O \in \mathcal{O}_Z, g_{\text{inv}}(O) \in \mathcal{O}_\mathcal{Y}\) since \(g\) is continuous. \(h_{\text{inv}}(O) = f_{\text{inv}}(g_{\text{inv}}(O)) \in \mathcal{B}.\) Hence, \(h\) is \(\mathcal{B}\)-measurable. \(\square\)

**Proposition 11.39** Let \((X, \mathcal{B})\) be a measurable space, \(\mathcal{Y}_i := (Y_i, \mathcal{O}_i)\) be a second countable topological space, \(D \in \mathcal{B}, f_i : D \to \mathcal{Y}_i, \forall i \in N \subseteq \mathbb{N}, N\) is a countable index set, \(\mathcal{Y} := (Y, \mathcal{O}) := \prod_{i \in N} \mathcal{Y}_i\) be the product topological space, and \(f : D \to \mathcal{Y}\) 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 \(B_{Y_i}^i\) be a countable basis for \(\mathcal{Y}_i.\) Without loss of generality, assume \(\emptyset, Y_i \in B_{Y_i}^i.\) Then, \(\mathcal{Y}\) is second countable with a countable basis \(B_Y\) as defined in Proposition 3.28. \(\forall O \in \mathcal{O}, O = \bigcup_{i \in N} O_i\) where \(O_i \in B_{Y_i}, \forall i \in \mathbb{N}, O_{B_i} = \prod_{i \in N} B_{Y_i},\) where \(B_{Y_i} \subseteq \mathcal{B}_{Y_i}, \forall i \in \mathbb{N}, B_{Y_i} = Y_i, \forall i \in \mathbb{N} \setminus N_i,\) and \(N_i \subseteq \mathbb{N}\) is a finite set. Then, \(f_{\text{inv}}(O_{B_i}) = \bigcap_{i \in N} f_{\text{inv}}(B_{Y_i}) \in \mathcal{B}\) 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 \(O := \bigcap_{i \in N} O_i \in \mathcal{O},\) where \(O_i = Y_i, \forall i \in \mathbb{N}\) 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. \(\square\)
Lemma 11.43 said to hold

By Proposition 4.4, where it fails to hold or does not make sense belongs to

and 4.30, the function

Proof

\( f \) is \( \bar{\mathcal{B}} \)-measurable.

By arguments that are similar to the above, \( \bar{\mathcal{B}} \)-measurable.

By Definition 3.82, \( \bar{f} \) and \( \bar{f} \) are \( \mathcal{B} \)-measurable. This completes the proof of the proposition. \( \square \)

Proposition 11.41 Let \((X,B)\) be a measurable space, \((E_i)_{i=1}^{\infty} \subseteq B\) be such that \(\bigcup_{i=1}^{\infty} E_i = X\), \(Y := (Y, \mathcal{O})\) be a topological space, \(D \in \overline{B}, B_{E_i} := \{B \in B \mid B \subseteq E_i \}, \forall i \in \mathbb{N}, \) and \(f : D \to Y\). Then, \(f\) is \(B_{E_i}\)-measurable, \(\forall i \in \mathbb{N}, \) if, and only if, \(f\) is \(B\)-measurable.

Proof

(i) \(\Rightarrow\) (ii). Assume that \(f\) is \(B_{E_i}\)-measurable, \(\forall i \in \mathbb{N}, \forall O \in \mathcal{O}, \) by the assumption, we have \((f|_{E_i})_{\text{inv}}(O) \in B_{E_i} \subseteq B, \forall i \in \mathbb{N}. \)

Then, \(f|_{E_i}\) is \(B\)-measurable.

(ii) \(\Rightarrow\) (iii). Assume that \(f|_{E_i}\) is \(B\)-measurable, \(\forall i \in \mathbb{N}, \forall O \in \mathcal{O}, \) by the assumption, we have \((f|_{E_i})_{\text{inv}}(O) \in B, \forall i \in \mathbb{N}. \)

Then, \(f_{\text{inv}}(O) = \bigcup_{i=1}^{\infty} (f|_{E_i})_{\text{inv}}(O) \in B. \)

Hence, \(f\) is \(B\)-measurable.

(iii) \(\Rightarrow\) (i). Assume that \(f\) is \(B\)-measurable, \(\forall O \in \mathcal{O}, \) by the assumption, \(f_{\text{inv}}(O) \in B. \)

Then, \((f|_{E_i})_{\text{inv}}(O) = f_{\text{inv}}(O) \cap E_i \in B_{E_i}, \forall i \in \mathbb{N}. \)

Hence, \(f|_{E_i}\) is \(B_{E_i}\)-measurable, \(\forall i \in \mathbb{N}. \)

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

Definition 11.42 Let \((X, \mathcal{B}, \mu)\) be a measure space. A property \(P\) is said to hold almost everywhere in \(X\) (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 \(X\) or \(P(x)\) \(a.e.\) \(x \in X.\)

Lemma 11.43 Let \((X, B)\) be a measurable space, \((Y, \rho)\) be a separable metric space, \(D_1, D_2 \subseteq B, and f : D_1 \to Y \) and \(g : D_2 \to Y \) be \(B\)-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 \(B\)-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 \(B\)-measurable. By Propositions 11.38 and 4.30, the function \(h\) is \(B\)-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 $E_1 := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x)\}$, $E_2 := \widetilde{D}_2 \cup \widetilde{D}_3 \cup \{x \in D_2 \cap D_3 \mid f_2(x) \neq f_3(x)\}$, and $E_3 := \widetilde{D}_1 \cup \widetilde{D}_3 \cup \{x \in D_1 \cap D_3 \mid f_3(x) \neq f_1(x)\}$. 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 $E_3 = \widetilde{D}_1 \cup \widetilde{D}_3 \cup \{x \in D_1 \cap D_3 \mid \rho(f_1(x), f_3(x)) > 0\} \in \mathcal{B}$ 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, $D_1, D_2 \in \mathcal{B}$, $f_i : D_1 \to \mathcal{Y}_i$ and $g_i : D_2 \to \mathcal{Y}_i$ 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 $E_i := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid f_i(x) \neq g_i(x)\}$, $\forall i \in N$, and $E := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid f(x) \neq g(x)\}$.

“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 $E_i = \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid \rho_i(f_i(x), g_i(x)) > 0\} \in \mathcal{B}$. 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, $E := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid f_1(x) \neq f_2(x)\} \in \mathcal{B}$ and $\mu(E) = 0$. By Proposition 11.38, we have $g \circ f_i$ is $\mathcal{B}$-measurable, $i = 1, 2$. By Lemma 11.43, $E_2 \supseteq \widetilde{E} := \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid g(f_1(x)) \neq g(f_2(x))\} = \widetilde{D}_1 \cup \widetilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid \rho(g(f_1(x)), g(f_2(x))) > 0\} \in \mathcal{B}$. Then, $0 \leq \mu(E_2) \leq \mu(E) = 0$. 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 $X := (X, \mathcal{B}, \mu)$ 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 $X$. 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 $\mu(E) = 0$. \(\forall O \in \mathcal{O}_Y, \ g_{\text{inv}}(O) = g_{\text{inv}}(O) \cap (E \cup (X \setminus E)) = (g_{\text{inv}}(O) \cap E) \cup (g_{\text{inv}}(O) \cap (X \setminus E)) = (g_{\text{inv}}(O) \cap E) \cup (f_{\text{inv}}(O) \cap (X \setminus E))\). Note that $X \setminus E, f_{\text{inv}}(O) \in \mathcal{B}$. By the completeness of $X$, $E \supseteq g_{\text{inv}}(O) \cap E \in \mathcal{B}$. Then, $g_{\text{inv}}(O) \in \mathcal{B}$. Hence, $g$ is $\mathcal{B}$-measurable.

Proposition 11.48  Let $X := (X, \mathcal{B}, \mu)$ 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, $O_n = O \cap \left( \bigcap_{y \in \mathcal{Y} \setminus O} (B_Y(y, 1/n)) \right)$. Note that $\bigcap_{y \in \mathcal{Y} \setminus O} (B_Y(y, 1/n))$ 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, $\forall l \geq L + 1$, $\forall \tilde{y} \in B_Y\left(y, \frac{1}{L(L+1)}\right)$, $\exists \tilde{y} \in B_Y(\tilde{y}, 1/l)$, we have $\rho(y, \tilde{y}) \leq \rho(y, \tilde{y}) + \rho(\tilde{y}, \tilde{y}) < \frac{1}{L(L+1)} + \frac{1}{l} \leq \frac{1}{L}$. This implies that $\tilde{y} \in B_Y(y, 1/L) \subseteq O$. Then, $B_Y(\tilde{y}, 1/l) \subseteq O$. Thus, $\tilde{y} \in O_l$ and $B_Y\left(y, \frac{1}{L(L+1)}\right) \subseteq O_l$. Since $\lim_{n \to \infty} f_n(x) = f(x) = y$, then $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $\rho(f_n(x), y) < \frac{1}{L(L+1)}$. Then, $f_n(x) \in B_Y\left(y, \frac{1}{L(L+1)}\right) \subseteq O_l$ and $x \in f_{n,\text{inv}}(O_l)$. By arbitrariness of $n$, $x \in \bigcap_{j=n}^{\infty} f_{j,\text{inv}}(O_l) \subseteq U$. By arbitrariness of $l$, we have $x \in \bigcap_{k=L+1}^{\infty} \bigcap_{n=1}^{\infty} \bigcap_{j=n}^{\infty} f_{j,\text{inv}}(O_k) \subseteq U$. 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 $\rho(f_{\tilde{N}}(x), f(x)) < 1/L$. 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 the arbitrariness of $x$, $U \subseteq f_{\text{inv}}(O)$.

Hence, $f_{\text{inv}}(O) = U$. This completes the proof of the claim. \(\square\)

\(\forall k \in \mathbb{N}, \forall j \in \mathbb{N}\), by Proposition 11.34, $f_{j,\text{inv}}(O_k) \in \mathcal{B}$. By Claim 11.48.1, $f_{\text{inv}}(O) \in \mathcal{B}$. Hence, $f$ is $\mathcal{B}$-measurable. This completes the proof of the proposition. \(\square\)

Proposition 11.49  Let $X := (X, \mathcal{B}, \mu)$ 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,
∀n ∈ N, and f : D_2 → Y. Assume that \( \lim_{n \to \infty} f_n = f \) a.e. in \( X \). Then, f is \( B \)-measurable.

**Proof**  Let \( E := \{ x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^{\infty} \) does not converge to \( f(x) \} \cup \widetilde{D}_1 \cup \widetilde{D}_2 \). By the assumption, \( E \in B \) and \( \mu(E) = 0 \). Let \( g := f|_{X \setminus E} \) and \( g_n := f_n|_{X \setminus E}, \forall n \in N \). Then, \( \lim_{n \to \infty} g_n(x) = g(x), \forall x \in X \setminus E \in B \). By Proposition 11.48, \( g \) is \( B \)-measurable. \( \forall \) open set \( O \subseteq Y \), \( f_{\text{inv}}(O) = g_{\text{inv}}(O) \cup (f_{\text{inv}}(O) \cap E) \). By the completeness of \( X \), we have \( E \supseteq f_{\text{inv}}(O) \cap E \in B \). Then, \( f_{\text{inv}}(O) \in B \). Hence, \( f \) is \( B \)-measurable. This completes the proof of the proposition.

\( \square \)

**Proposition 11.50**  Let \( (X, B) \) be a measurable space, \( Y := (Y, \rho) \) be a separable metric space, \( D_1, D_2 \in B; f_n : D_1 \to Y \) be \( B \)-measurable, \( \forall n \in N \), and \( f : D_2 \to Y \) be \( B \)-measurable. Then, the set
\[
E := \{ x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^{\infty} \text{ converges to } f(x) \} \in B
\]

**Proof**  We will show that
\[
E = \bigcap_{m=1}^{\infty} \bigcup_{i=1}^{\infty} E_{k,m} =: \tilde{E}
\]
where \( E_{k,m} := \{ x \in D_1 \cap D_2 \mid \rho(f_k(x), f(x)) < 1/m \} \in B, \forall m \in N, \forall k \in N \), by Lemma 11.43. This further implies that \( \tilde{E} \in B \).

\( \forall x \in E, \forall m \in N, \exists k_0 \in N, \forall k \in N \) with \( k \geq k_0 \), \( x \in D_1 \cap D_2 \) and \( \rho(f_k(x), f(x)) < 1/m \). Then, \( x \in E_{k,m} \) and \( x \in \bigcap_{i=1}^{\infty} E_{k,m} \subseteq \bigcup_{i=1}^{\infty} E_{k,m} \). By the arbitrariness of \( m \), we have \( x \in \tilde{E} \). This implies that \( E \subseteq \tilde{E} \). On the other hand, \( \forall x \in \tilde{E}, \forall m \in N \), \( x \in \bigcup_{i=1}^{\infty} E_{k,m} \). Then, \( \exists k_0 \in N \), such that \( x \in \bigcap_{i=1}^{\infty} E_{k,m} \). \( \forall k \in N \) with \( k \geq k_0 \), \( x \in E_{k,m} \). Then, \( x \in D_1 \cap D_2 \) and \( \rho(f_k(x), f(x)) < 1/m \). This leads to \( \lim_{n \to \infty} f_n(x) = f(x) \) and \( x \in E \). Then, we have \( E \subseteq \tilde{E} \). Hence, \( E = \tilde{E} \in B \). This completes the proof of the proposition.

\( \square \)

**Proposition 11.51**  Let \( X := (X, B, \mu) \) be a measure space, \( Y := (Y, \rho) \) be a separable complete metric space, \( D \in B; f_i : D \to Y \) be \( B \)-measurable, \( \forall i \in N \), \( E := \{ x \in D \mid (f_i(x))_{i=1}^{\infty} \text{ converges in } Y \} \) and \( f : E \to Y \) defined by \( f(x) = \lim_{i \to \infty} f_i(x), \forall x \in E \). Then, \( E = \bigcap_{i=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{l=1}^{m} E_{j,l,n} =: \tilde{E} \in B \), where \( E_{j,l,n} := \{ x \in D \mid \rho(f_j(x), f_n(x)) < 1/j \} \), \( \forall j, l, n \in N \), and \( f \) is \( B \)-measurable.

**Proof**  \( \forall j, l, n \in N \), by Lemma 11.43 and the separability of \( Y \), \( E_{j,l,n} \in B \). Then, \( \tilde{E} \in B \).

\( \forall x \in \tilde{E}, \forall j \in N \), \( \exists m_0 \in N \), \( \forall l, n \in N \) with \( l \geq m_0 \) and \( n \geq m_0 \), we have \( x \in D \) and \( \rho(f_j(x), f_n(x)) < 1/j \). Then, \( (f_i(x))_{i=1}^{\infty} \subseteq Y \) is a Cauchy sequence. By the completeness of \( Y \), \( (f_i(x))_{i=1}^{\infty} \) converges in \( Y \). Hence, \( x \in E \). By the arbitrariness of \( x \), we have \( \tilde{E} \subseteq E \).
On the other hand, \( \forall x \in E \), since \((f_i(x))_{i=1}^\infty\) converges in \( \mathcal{Y} \), then it is a Cauchy sequence. \( \forall j \in \mathbb{N}, \exists m_0 \in \mathbb{N}, \forall l,n \in \mathbb{N} \) with \( l \geq m_0 \) and \( n \geq m_0 \), we have \( \rho(f_i(x), f_n(x)) < 1/j \). Then, \( x \in E_{j,l,n} \). Then, \( x \in \bar{E} \). By the arbitrariness of \( x \), we have \( E \subseteq \bar{E} \). Therefore, \( E = \bar{E} \in \mathcal{B} \).

Note that \( f(x) = \lim_{i \in \mathbb{N}} f_i|_E(x), \forall x \in E \). 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.

**Proposition 11.52** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} := (Y, \mathcal{O}) \) be a Hausdorff topological space, \( \mathcal{Z} := (Z, \rho) \) be a separable metric space, \( D_1, D_2 \in \mathcal{B}, f_n : D_1 \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, f : D_2 \to \mathcal{Z} \) be \( \mathcal{B} \)-measurable, and \( g : \mathcal{Y} \to \mathcal{Z} \) be continuous. Assume that \( \lim_{n \in \mathbb{N}} f_n = f \) a.e. in \( \mathcal{X} \). Then, \( \lim_{n \in \mathbb{N}} g \circ f_n = g \circ f \) a.e. in \( \mathcal{X} \).

**Proof** Let \( E := \{ x \in D_1 \cap D_2 \mid \lim_{n \in \mathbb{N}} f_n(x) = f(x) \} \) and \( \bar{E} := \{ x \in D_1 \cap D_2 \mid \lim_{n \in \mathbb{N}} g(f_n(x)) = g(f(x)) \} \). 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, \( \lim_{n \in \mathbb{N}} g \circ f_n = g \circ f \) a.e. in \( \mathcal{X} \). This completes the proof of the proposition.

**Proposition 11.53** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( D_1, D_2 \in \mathcal{B}, \mathcal{Y}_i := (Y_i, \mathcal{O}_i) \) be a second countable metrizable topological space, \( f_i : D_1 \to \mathcal{Y}_i \) be \( \mathcal{B} \)-measurable, \( g_{n,i} : D_2 \to \mathcal{Y}_i \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N}, \forall i \in \mathbb{N}, N \subseteq \mathbb{N} \), \( N \) is a countable index set, \( \mathcal{Y} := (Y, \mathcal{O}) := \prod_{i \in \mathbb{N}} \mathcal{Y}_i \) be the product topological space, \( f : D_1 \to \mathcal{Y} \) be given by \( \pi_i(f(x)) = f_i(x), \forall x \in D_1, \forall i \in \mathbb{N}, \) and \( g_n : D_2 \to \mathcal{Y} \) be given by \( \pi_i(g_n(x)) = g_{n,i}(x), \forall x \in D_2, \forall i \in \mathbb{N}, \forall n \in \mathbb{N} \). Then, \( \lim_{n \in \mathbb{N}} g_n = f \) a.e. in \( \mathcal{X} \) if, and only if, \( \lim_{n \in \mathbb{N}} g_{n,i} = f_i \) a.e. in \( \mathcal{X}, \forall i \in \mathbb{N} \).

**Proof** By Proposition 11.39, \( f \) and \( g_n \) are \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). “Sufficiency” Assume \( \lim_{n \in \mathbb{N}} g_{n,i} = f_i \) a.e. in \( \mathcal{X} \), \( \forall i \in \mathbb{N} \). \( \forall i \in \mathbb{N} \), let \( E_i := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid (g_{n,i}(x))_{n=1}^\infty \) does not converge to \( f_i(x) \} \). Then, \( E_i \in \mathcal{B} \) and \( \mu(E_i) = 0 \). Let \( E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{ x \in D_1 \cap D_2 \mid (g_n(x))_{n=1}^\infty \) does not converge to \( f(x) \} \). By Proposition 3.67, \( E = \bigcup_{i \in \mathbb{N}} E_i \in \mathcal{B} \). Then, \( 0 \leq \mu(E) \leq \sum_{i \in \mathbb{N}} \mu(E_i) = 0 \). Hence, \( \lim_{n \in \mathbb{N}} g_n = f \) a.e. in \( \mathcal{X} \). “Necessity” Assume \( \lim_{n \in \mathbb{N}} g_n = f \) a.e. in \( \mathcal{X} \). Let \( E \) and \( E_i \) be as defined above, \( \forall i \in \mathbb{N} \). Then, \( E \in \mathcal{B} \) and \( \mu(E) = 0 \). \( \forall i \in \mathbb{N} \), by Proposition 3.67, we have \( E_i \subseteq E \). Since \( \mathcal{Y}_i \) is 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, by Proposition 4.4, \( (Y_i, \rho_i) \) is a separable metric space. By Proposition 11.50, we have \( E_i \in \mathcal{B} \). Then, \( 0 \leq \mu(E_i) \leq \mu(E) = 0 \). Hence, \( \lim_{n \in \mathbb{N}} g_{n,i} = f_i \) a.e. in \( \mathcal{X} \). This completes the proof of the proposition.
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 $\lim_{n \to \infty} f_n = f$ a.e. in $\mathcal{X}$, $f_n = g_n$ a.e. in $\mathcal{X}$, $\forall i \in \mathbb{N}$, and $f = g$ a.e. in $\mathcal{X}$. Then, $\lim_{n \to \infty} g_n = g$ a.e. in $\mathcal{X}$.

Proof By the assumption, we have $E_1 := \widehat{D}_1 \cup \widehat{D}_2 \cup \{x \in D_1 \cap D_3 \mid (f_n(x))_{n=1}^\infty$ does not converge to $f(x)\} \in \mathcal{B}$; $E_n := \widehat{D}_1 \cup \widehat{D}_2 \cup \{x \in D_1 \cap D_2 \mid f_n(x) \neq g_n(x)\} \in \mathcal{B}, \forall n \in \mathbb{N}$; $E := \widehat{D}_3 \cup \widehat{D}_4 \cup \{x \in D_3 \cap D_4 \mid f(x) \neq g(x)\} \in \mathcal{B}$; and $\mu(E_1) = \mu(E_n) = \mu(E) = 0, \forall n \in \mathbb{N}$. Then, $E := E_1 \cup (\bigcup_{n=1}^\infty E_n) \cup E \in \mathcal{B}$ and $\mu(E) = 0$. Let $E_2 := \widehat{D}_2 \cup \widehat{D}_4 \cup \{x \in D_2 \cap D_4 \mid (g_n(x))_{n=1}^\infty$ does not converge to $g(x)\}$. By Proposition 11.50, $E_2 \in \mathcal{B}$. Clearly, $E \subseteq \overline{E_2}$ and $E_2 \subseteq E$. Then, $0 \leq \mu(E_2) \leq \mu(E) = 0$. This shows that $\lim_{n \to \infty} g_n = g$ 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 $\lim_{n \to \infty} f_n = f$ a.e. in $\mathcal{X}$. Then, $\forall \eta \in (0, \infty) \subset \mathbb{R}, \exists A \in \mathcal{B}$ with $\mu(A) < \eta$ and $A \supseteq \widehat{D}_1 \cup \widehat{D}_2$ such that $\left( f_n|_{X\setminus A} \right)_{n=1}^\infty$ converges uniformly to $f|_{X\setminus A}$.

Proof Let $E := \widehat{D}_1 \cup \widehat{D}_2 \cup \{x \in D_1 \cap D_2 \mid (f_n(x))_{n=1}^\infty$ does not converge to $f(x)\}$. By the assumption, $E \in \mathcal{B}$ and $\mu(E) = 0$. \(\forall \eta \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall m \in \mathbb{N}\), 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, $H_{m+1,n} \subseteq H_{m,n} \in \mathcal{B}$. $\forall x \in X \setminus E$, we have $\lim_{i \to \infty} f_i(x) = f(x)$, then $\exists m_x \in \mathbb{N}$ such that $x \not\in H_{m_x,n}$. Hence, $\bigcap_{m=1}^\infty H_{m,n} = \emptyset$. Note that $\mu(H_{n},n) \leq \mu(X) < \infty$. By Proposition 11.5, $\lim_{n \to \infty} \mu(H_{m,n}) = \mu(\bigcap_{m=1}^\infty H_{m,n}) = 0$. 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, $A \supseteq \widehat{D}_1 \cup \widehat{D}_2$. 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 $m_{n_0} \leq j$, $\forall x \in X \setminus A \subseteq (X \setminus E) \setminus A_{n_0}$, we have $\rho(f_j(x), f(x)) < 2^{-n_0} \leq \epsilon$. Hence,
\[(f_n|_{X \setminus A})_{n=1}^\infty\] converges uniformly to \(f|_{X \setminus A}\). This completes the proof of the proposition.

**Definition 11.56** Let \(X := (X, \mathcal{B}, \mu)\) 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 \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists N \in \mathbb{N}\) such that \(\forall n \in \mathbb{N}\) with \(n \geq N\), we have \(\mu(A_{n,\epsilon}) := \mu(D_n \cup \tilde{D} \cup \{x \in D_n \cap D \mid \rho(f_n(x), f(x)) \geq \epsilon\}) < \epsilon\). In this case, we will write \(\lim_{n \in \mathbb{N}} f_n = f\) in measure in \(\mathcal{X}\).

Note that in the above definition, \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N}\), \(A_{n,\epsilon} \in \mathcal{B}\) by Lemma 11.43. Hence, the definition is well-defined.

**Proposition 11.57** Let \(X := (X, \mathcal{B}, \mu)\) 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 \(\lim_{n \in \mathbb{N}} f_n = f\) in measure in \(\mathcal{X}\). Then, there exists a subsequence \((f_{n_k})_{k=1}^\infty\) of \((f_n)_{n=1}^\infty\) such that \(\lim_{k \in \mathbb{N}} f_{n_k} = f\) a.e. in \(\mathcal{X}\).

**Proof** Define \(A_{n,k} := \tilde{D}_1 \cup \tilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \leq 2^{-k}\}, \forall n, k \in \mathbb{N}\). 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 \(A := \bigcap_{i=1}^\infty \bigcup_{k=i}^\infty A_{n_k,k} \in \mathcal{B}\). Then, \(\mu(A) \leq \mu(\bigcup_{k=i}^\infty A_{n_k,k}) \leq \sum_{k=i}^\infty \mu(A_{n_k,k}) < 2^{-i+1}, \forall i \in \mathbb{N}\). Hence, \(\mu(A) = 0\). \(\forall x \in X \setminus A = \bigcup_{k=1}^\infty A_{n_k,k}, \forall i_0 \in \mathbb{N}\) 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, \(\lim_{k \in \mathbb{N}} f_{n_k}(x) = f(x)\). Then, \(E := \tilde{D}_1 \cup \tilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid (f_{n_k}(x))_{k=1}^\infty\) does not converge to \(f(x)\} \subseteq A\). By Proposition 11.50, we have \(E \in \mathcal{B}\) and \(\mu(E) = 0\). Therefore, \(\lim_{k \in \mathbb{N}} f_{n_k} = f\) a.e. in \(\mathcal{X}\). This completes the proof of the proposition.

**Proposition 11.58** Let \(X := (X, \mathcal{B}, \mu)\) 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 \(\lim_{n \in \mathbb{N}} f_n = f\) a.e. in \(\mathcal{X}\). Then, \(\lim_{n \in \mathbb{N}} f_n = f\) in measure in \(\mathcal{X}\).

**Proof** \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}\), by Egoroff’s Theorem 11.55, \(\exists A \in \mathcal{B}\) with \(\mu(A) < \epsilon\) and \(A \supseteq \tilde{D}_1 \cup \tilde{D}_2\) such that \((f_n|_{X \setminus A})_{n=1}^\infty\) 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 \(\rho(f_n(x), f(x)) < \epsilon, \forall x \in X \setminus A\). Then, \(A_{n,\epsilon} := \tilde{D}_1 \cup \tilde{D}_2 \cup \{x \in D_1 \cap D_2 \mid \rho(f_n(x), f(x)) \geq \epsilon\} \subseteq A\). By Lemma 11.43, \(A_{n,\epsilon} \in \mathcal{B}\) and \(\mu(A_{n,\epsilon}) \leq \mu(A) < \epsilon\). This shows that \(\lim_{n \in \mathbb{N}} f_n = f\) in measure in \(\mathcal{X}\). This completes the proof of the proposition.
Proposition 11.59 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ 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. Then, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$ if, and only if, every subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ admits a subsequence $(f_{n_{kj}})_{j=1}^{\infty}$ 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 $(f_{n_{kj}})_{j=1}^{\infty}$ 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, $\exists \varepsilon_0 \in (0, \infty) \subset \mathbb{R}, \forall n_0 \in \mathbb{N}, \exists n \in \mathbb{N}$ with $n \geq n_0$ such that $\mu(A_n) := \mu(D_n \cup \{ x \in D_n \cap D \mid \rho(f_n(x), f(x)) \geq \varepsilon_0 \}) \geq \varepsilon_0$, where $A_n \in \mathcal{B}$ by Lemma 11.43. Then, $\exists n_1 \in \mathbb{N}$ such that $\mu(A_{n_1}) \geq \varepsilon_0, \forall k \in \mathbb{N}, \exists n_{k+1} \in \mathbb{N}$ with $n_{k+1} > n_k$ such that $\mu(A_{n_{k+1}}) \geq \varepsilon_0$. This defines a subsequence $(f_{n_k})_{k=1}^{\infty}$, which satisfies that $\forall k \in \mathbb{N}, \mu(A_{n_k}) \geq \varepsilon_0$. 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. \qed

Proposition 11.60 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. Then, $\lim_{n \in \mathbb{N}} f_n = f$ in measure in $\mathcal{X}$ if, and only if, every subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ admits a subsequence $(f_{n_{kj}})_{j=1}^{\infty}$ that converges to $f$ a.e. in $\mathcal{X}$.

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 $(f_{n_{kj}})_{j=1}^{\infty}$ of $(f_{n_k})_{k=1}^{\infty}$ that converges to $f$ a.e. in $\mathcal{X}$.

“Sufficiency” Assume that every subsequence $(f_{n_k})_{k=1}^{\infty}$ of $(f_n)_{n=1}^{\infty}$ admits a subsequence $(f_{n_{kj}})_{j=1}^{\infty}$ that converges to $f$ a.e. in $\mathcal{X}$. By Proposition 11.58, $(f_{n_{kj}})_{j=1}^{\infty}$ converges to $f$ in measure. By Proposition 11.59, we have $(f_n)_{n=1}^{\infty}$ converges to $f$ in measure in $\mathcal{X}$.

This completes the proof of the proposition. \qed

Definition 11.61 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\mathcal{Y} := (Y, \rho)$ be a separable metric space, $D_n \in \mathcal{B}$, and $f_n : D_n \to \mathcal{Y}$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. We will say that $(f_n)_{n=1}^{\infty}$ is Cauchy in measure in $\mathcal{X}$ if $\forall \varepsilon \in$
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 $A := \bigcap_{i=1}^{\infty} \bigcup_{j=i}^{\infty} A_j \subseteq B$. 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, $\mu(A) = 0$. $\forall x \in X \setminus A = \bigcap_{i=1}^\infty \bigcup_{j=i}^{\infty} A_j \subseteq D$, $\exists i_0 \in \mathbb{N}$ 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 $\lim_{n \to \infty} f_n(x) = f(x)$. $\forall x \in A \cap D$, 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 $\lim_{n \to \infty} f_n(x) = f(x)$, $\forall x \in X \setminus A$.

$\forall i \in \mathbb{N}$, define $\tilde{f}_{n_i} : D \to Y$ by $\tilde{f}_{n_i}(x) = \begin{cases} f_{n_i}(x) & x \in D_{n_i} \\ f_n(x) & x \in D \setminus D_{n_i} \end{cases}$, $\forall x \in D$. By Proposition 11.41, $\tilde{f}_{n_i}$ is $B$-measurable. Clearly, $\lim_{n \to \infty} \tilde{f}_{n_i}(x) = f(x)$, $\forall x \in D$. By Proposition 11.48, $f$ is $B$-measurable.

$\forall i \in \mathbb{N}, \forall x \in \bigcup_{j=i}^{\infty} A_j = \bigcap_{j=i}^{\infty} \bar{A}_j \subseteq X \setminus A$, we have $\rho(f_{n_j}(x), f_{n_{j+1}}(x)) < 2^{-j}$, $\forall j \geq i$. Then, by Propositions 4.30, 3.66, and 3.67, $\rho(f_{n_j}(x), f(x)) = \lim_{n \to \infty} \rho(f_{n_j}(x), f_n(x)) = \lim_{n \to \infty} \sum_{i=j}^{l-1} \rho(f_{n_j}(x), f_{n_{i+1}}(x)) < 2^{-i+1}$. 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 $\mu(E_i) \leq \sum_{j=i}^{\infty} \mu(A_j) < 2^{-i+1}$.

$\forall \epsilon \in (0, \infty) \in \mathbb{R}$, $\exists i_0 \in \mathbb{N}$ such that $2^{-i_0+1} < \epsilon/2$. $\forall n \in \mathbb{N}$ with $n \geq n_{i_0}$, we have $\mu(D_n \cup \{ x \in D_n \mid \rho(f_n(x), f(x)) \geq \epsilon \}) \leq \mu(D_n \cup D_{n_{i_0}} \cup \{ x \in D_n \cap D_{n_{i_0}} \mid \rho(f_n(x), f_{n_{i_0}}(x)) + \rho(f_{n_{i_0}}(x), f(x)) < 2^{-i_0+1} \}) \leq \mu(D_n \cup D_{n_{i_0}} \cup \{ x \in D_n \cap D_{n_{i_0}} \mid \rho(f_n(x), f_{n_{i_0}}(x)) \geq \epsilon/2 \} + \mu(D_{n_{i_0}} \cup \{ x \in D_{n_{i_0}} \mid \rho(f_{n_{i_0}}(x), f(x)) \geq \epsilon/2 \} < 2^{-i_0} + 2^{-i_0+1} < \epsilon$. Hence, $\lim_{n \to \infty} f_n = f$ in measure in $X$. This completes the proof of the proposition. \(\Box\)
Definition 11.63 Let $X$ be a set and $A \subseteq X$. The indicator function $\chi_{A,X} : X \to \{0,1\} \subset \mathbb{R}$ 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 $\phi(x) = \sum_{i=1}^{n} y_i \chi_{A,X}(x)$, $\forall x \in X$. We will say that a simple function $\phi$ is in canonical representation if $y_1, \ldots, y_n$ are distinct and none equals to $\psi_y$, 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) $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\exists$ a simple function $\phi_\varepsilon : X \to U$ such that $\mu(\{x \in X \mid \| f(x) - \phi_\varepsilon(x) \| \geq \varepsilon \}) < \varepsilon$;

(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, $X = \emptyset$, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, choose $\phi_\varepsilon : X \to U$ to be the simple function with $n = 0$ as in Definition 11.64. Then, $\mu(\{x \in X \mid \| f(x) - \phi_\varepsilon(x) \| \geq \varepsilon \}) = \mu(\emptyset) = 0 < \varepsilon$. Hence, (i) holds. Case 2: $U \neq \emptyset$. Then, $\exists (y_i)_{i=1}^n \subseteq U$ such that $\forall y \in U$, $\forall \delta \in (0, \infty) \subset \mathbb{R}$, $\exists i \in \mathbb{N}$ $\exists \varepsilon \in \mathbb{R}$. Define $V_i := V_i \in \mathcal{B}(y)$, $V_{i+1} := V_{i+1} \setminus (\bigcup_{j=i+1}^n \bar{V}_j) \in \mathcal{B}(y)$, $\forall i \in \mathbb{N}$. 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, $X = f_{\text{inv}}(U) = \bigcup_{i=1}^\infty f_{\text{inv}}(V_i)$ and $(f_{\text{inv}}(V_i))_{i=1}^\infty \subseteq \mathcal{B}$ and are pairwise disjoint. Then, $\mu(X) = \sum_{i=1}^\infty \mu(f_{\text{inv}}(V_i)) < +\infty$. This implies that $\exists n_0 \in \mathbb{N}$ such that $\sum_{i=n_0+1}^\infty \mu(f_{\text{inv}}(V_i)) < \varepsilon$. Let $A_i := f_{\text{inv}}(V_i) \in \mathcal{B}$, $i = 1, \ldots, n_0$, and $A_{n_0+1} := X \setminus \bigcup_{i=1}^{n_0} A_i \in \mathcal{B}$. Define the simple function $\phi_\varepsilon : X \to U$ by $\phi_\varepsilon(x) = \sum_{i=1}^{n_0+1} y_i \chi_{A,X}(x)$, $\forall x \in X$. $\forall x \in X$, $\exists i_x \in \{1, \ldots, n_0+1\}$ such that $x \in A_{i_x}$. Then, $\phi_\varepsilon(x) = y_{i_x}$. $\forall x \in X \setminus A_{n_0+1}$, $i_x \in \{1, \ldots, n_0\}$ and $x \in A_{i_x} = f_{\text{inv}}(V_{i_x}) \subseteq f_{\text{inv}}(\bar{V}_{i_x})$. Then, $\phi_\varepsilon(x) = y_{i_x}$ and $f(x) \in V_{i_x} = \mathcal{B}(y_{i_x}, \varepsilon)$. Hence, $\| f(x) - \phi_\varepsilon(x) \| < \varepsilon$. Then, $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, let $\phi_n := \phi_{2^{-n}}$. Then, $(\phi_n)_{n=1}^\infty$ converges to $f$ in measure. This completes the proof of the proposition. □
Proposition 11.66 Let $X := (X, \mathcal{B}, \mu)$ 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 $\lim_{n \in \mathbb{N}} \varphi_n = f$ a.e. in $X$ 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 $X$ 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 $I_i := [-(M_\epsilon - i + 1)\epsilon, -(M_\epsilon - i)\epsilon) \subset \mathbb{R}$, $i = 1, \ldots, 2M_\epsilon$ and $A_i := f^{-1}(I_i) \in \mathcal{B}$. Note that $\mu(A_i) \leq \mu(X) < +\infty$, $i = 1, \ldots, 2M_\epsilon$ and $A_i$'s are pairwise disjoint. Define $\phi_x := \sum_{i=1}^{2M_\epsilon} -(M_\epsilon - i)\epsilon \mathcal{X}_{A_i,x} + \sum_{i=M_\epsilon+1}^{2M_\epsilon} -(M_\epsilon - i + 1)\epsilon \mathcal{X}_{A_i,x}$. Clearly, we have $\phi_x : X \to U$, $|\phi_x(x)| \leq |f(x)|$, $\forall x \in X$, and $\mu(\{x \in X \mid |f(x) - \phi_x(x)| > \epsilon\}) \leq \mu(A_\epsilon) < \epsilon$. Then, the sequence of simple functions $(\varphi_n)_{n=1}^\infty$ converges to $f$ in measure in $X$. By Proposition 11.57, there exists a subsequence $(\varphi_k)_{k=1}^{\infty} = (\varphi_{2^{n_k}})_{k=1}^{\infty}$ such that $\lim_{k \in \mathbb{N}} \varphi_k = f$ a.e. in $X$. The result holds in this special case.

Next, we consider the special case $X$ is finite. By Proposition 11.65, there exists a sequence of simple functions $(\psi_i)_{i=1}^\infty$, $\psi_i : X \to U$, $\forall i \in \mathbb{N}$, that converges to $f$ in measure. By Proposition 11.57, there exists a subsequence $(\psi_n)_{n=1}^\infty = (\psi_i)_{n=1}^\infty$, such that $\lim_{n \in \mathbb{N}} \psi_n = f$ a.e. in $X$. 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 $\mathcal{P} \circ f(x) = \|f(x)\|$, $\forall x \in X$. By the previous special case, 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 $\lim_{n \in \mathbb{N}} \phi_n = \mathcal{P} \circ f$ a.e. in $X$ and $0 \leq \phi_n(x) \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Let $E_1 := \{x \in X \mid (\psi_n(x))_{n=1}^\infty$ 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 $\psi_n \sum_{j=1}^{\hat{n}} y_j \mathcal{X}_{A_j,x}$, where $\nhat \in \mathbb{Z}_+, y_1, \ldots, y_{\hat{n}} \in U$ are distinct and none equals to $\partial y$, $A_1, \ldots, A_{\hat{n}} \in \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. Let $y_{\hat{n}+1} := \partial y \in U$ and $A_{\hat{n}+1} := X \setminus (\bigcup_{j=1}^{\hat{n}} A_j) \in \mathcal{B}$. Then, $\psi_n = \sum_{j=1}^{\hat{n}+1} y_j \mathcal{X}_{A_j,x}$. Let $\phi_n$ admit the canonical representation $\sum_{j=1}^{\hat{n}} a_j \mathcal{X}_{A_j,x}$, where $\nhat \in \mathbb{Z}_+, a_1, \ldots, a_{\hat{n}} \in [0, \infty) \subset \mathbb{R}$ are distinct and none equals to $0$, $\vec{A}_1, \ldots, \vec{A}_{\hat{n}} \in \mathcal{B}$ are pairwise disjoint, nonempty, and of finite measure. Let $a_{\hat{n}+1} := 0$ and $\vec{A}_{\hat{n}+1} := X \setminus (\bigcup_{j=1}^{\hat{n}} \vec{A}_j) \in \mathcal{B}$. Then, $\phi_n = \sum_{j=1}^{\hat{n}+1} a_j \mathcal{X}_{A_j,x}$. Define a simple function $\varphi_n : X \to U$ by $\varphi_n(x) = \begin{cases} y_j & \text{if } a_l / \|y_j\| \leq a_l, \forall x \in A_j \cap \vec{A}_l, j = 1, \ldots, \hat{n} + 1, l = 1, \ldots, \hat{n} + 1, \\ y_j \|y_j\| & \text{if } a_l / \|y_j\| > a_l, \forall x \in A_j \cap \vec{A}_l, j = 1, \ldots, \hat{n} + 1, l = 1, \ldots, \hat{n} + 1. \end{cases}$ 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

Let \( \mathcal{X} = (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space and \( f : X \to [0, \infty] \subset \mathbb{R} \) be \( \mathcal{B} \)-measurable. Then, 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 \( 0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \varphi_n(x) = f(x) \), \( \forall x \in X \). Additionally, \( \lim_{n \to \infty} \varphi_n(x) = f(x) \), \( \forall x \in X \).
pairwise disjoint and \(\mu(E_i) \leq \mu(X_n) < \infty, \forall i \in \{1, \ldots, n^2 + 1\}\). Define \(\phi_n := \sum_{i=1}^{n^2+1} (i-1)/n X_{E_i,n} X\). 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 \(0 \leq f(x) - \phi_n(x) \leq 1/n, \forall x \in f_{\text{inv}}([0,n]) \cap X_n\). 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 \(0 \leq f(x) - \varphi_n(x) \leq 1/n, \forall x \in f_{\text{inv}}([0,n]) \cap X_n\).

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, \(x = f(x) = \lim_{n \to \infty} \varphi_n(x)\). On the other hand, if \(f(x) < \infty\), then \(\exists n_0 \in \mathbb{N}\) such that \(x \in X_{n_0}\) and \(0 \leq f(x) \leq n_0\). Then, \(\forall n \in \mathbb{N}\) with \(n_0 \leq n\), \(|f(x) - \varphi_n(x)| \leq 1/n\). This implies that \(\lim_{n \to \mathbb{N}} \varphi_n(x) = f(x)\).

Hence, we have \(f(x) = \lim_{n \to \mathbb{N}} \varphi_n(x), \forall x \in X\). This completes the proof of the proposition. \(\square\)

**Lemma 11.68** Let \(X := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite measure space, \(\mathbb{Y}\) be a normed linear space over \(\mathbb{K}\), \(U \subseteq \mathbb{Y}\) be a \(\sigma\)-compact conic segment, and \(f : X \to U\) 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 \(\lim_{n \to \mathbb{N}} \varphi_n(x) = f(x), \forall x \in X\), \(\|\varphi_n(x)\| \leq \|f(x)\|, \forall x \in X, \forall n \in \mathbb{N}\).

**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 \(\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}\) 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 \(K_n \subseteq K_{n+1}\) 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 \(D_n = \{y_{n,1}, \ldots, y_{n,m_n}\}\), where \(m_n \in \mathbb{Z}_+\). Let \(E_{n,1} := B_{y_{n,1}}(y_{n,1},1/n)\), and \(E_{n,i} := B_{y_{n,i}}(y_{n,i},1/n) \setminus \left(\bigcup_{j=1}^{i-1} E_{n,j}\right), \forall i = 2, \ldots, m_n\). Then, \(K_n \subseteq \bigcup_{i=1}^{m_n} E_{n,i}\) and the sets in the union are pairwise disjoint. Let \(A_{n,i} := f_{\text{inv}}(E_{n,i}) \cap X_n \in \mathcal{B}\), \(i = 1, \ldots, m_n\), since \(E_{n,i} \in \mathcal{B}\) (\(\forall y\)) and \(f\) is \(\mathcal{B}\)-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 \(\psi_n := \sum_{i=1}^{m_n} y_{n,i} \chi_{A_{n,i},x}\). 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, \(\lim_{n \to \mathbb{N}} \psi_n(x) = f(x), \forall x \in X\).

By Propositions 7.21 and 11.38, \(P \circ f : X \to [0,\infty) \subset \mathbb{R}\) is \(\mathcal{B}\)-measurable. By Proposition 11.67, there exists a sequence of simple functions \((\phi_n)_{n=1}^\infty\) such that \(0 \leq \phi_n(x) \leq \phi_{n+1}(x) \leq P \circ f(x) = \|f(x)\|, \forall x \in X, \forall n \in \mathbb{N}\), and \(\lim_{n \to \mathbb{N}} \phi_n(x) = P \circ f(x), \forall x \in X\).

Fix any \(n \in \mathbb{N}\), let \(\phi_n\) admit the canonical representation \(\phi_n = \sum_{i=1}^{m_n} a_j \chi_{A_{n,j},x}\), where \(m_n \in \mathbb{Z}_+, a_1, \ldots, a_{m_n} \in [0,\infty) \subset \mathbb{R}\) are distinct and none equals to \(0, \bar{A}_{n,1}, \ldots, \bar{A}_{n,m_n} \in \mathcal{B}\) are pairwise disjoint, nonempty,
and of finite measure. Define $\varphi_n : X \to U$ by

$$
\varphi_n(x) = \begin{cases} 
\frac{y_{n,i}}{\|y_{n,i}\|} & \text{if } x \in A_{n,i} \cap \bar{A}_{n,j}, \|y_{n,i}\| \leq a_j, i = 1, \ldots, m_n, j = 1, \ldots, m_n \\
\frac{a_j}{\|y_{n,i}\|} y_{n,i} & \text{if } x \in A_{n,i} \cap \bar{A}_{n,j}, \|y_{n,i}\| > a_j, i = 1, \ldots, m_n, j = 1, \ldots, m_n \\
\varnothing & \text{else}
\end{cases}
$$

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 $\varphi_n(x) = \left\{ \psi_n(x) \mid \|\psi_n(x)\| \leq \phi_n(x) \right\}$ if $\|\psi_n(x)\| \leq \phi_n(x)$, $\varphi_n(x) = \left\{ \phi_n(x) \mid \|\psi_n(x)\| > \phi_n(x) \right\}$ if $\|\psi_n(x)\| > \phi_n(x)$, $\forall x \in X$. Then, we have $\|\varphi_n(x)\| \leq \phi_n(x) \leq \mathcal{P} \circ f(x)$, $\forall x \in X$. Furthermore, $\lim_{n \to \infty} \varphi_n(x) = f(x)$, $\forall x \in X$.

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

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 spaces 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, $\mathcal{B}(X)$ 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, $(U_\alpha)_{\alpha \in \Lambda} \subseteq \mathcal{B}(X)$ 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, $\mathcal{I}(X) := (\mathcal{R}(X), \preceq)$ is a directed system and $\preceq$ is an antisymmetric partial ordering on $\mathcal{R}(X)$. $\mathcal{I}(X)$ is said to be the integration system on $X$.

**Proof** Clearly, $R_0 := \{ (\emptyset, X) \} \in \mathcal{R}(X) \neq \emptyset$. 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 $R_1 \preceq R_2$, $\exists \alpha \in \Lambda$ such that $V_\beta \subseteq U_\alpha$. Then, we have $U_\alpha \subseteq V_\beta \subseteq U_{\bar{\beta}}$. By $R_1 \in \mathcal{R}(X)$, $(U_\alpha)_{\bar{\alpha} \in \Lambda}$ are nonempty and pairwise disjoint. Then, we must have $\alpha = \bar{\alpha}$ and $U_\alpha = V_\beta$. 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 $R_2 \in \mathcal{R}(X)$, $(V_\beta)_{\bar{\beta} \in \Gamma}$ are nonempty and pairwise disjoint. Then, we must have $\bar{\beta} = \beta$ 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 $\mathcal{R}(X)$.

Case 1: $W_{\alpha, \beta} = \emptyset$. Define $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta}^{(2)} = \emptyset$.

Case 2: $W_{\alpha, \beta} \neq \emptyset$. Define $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta} \in \mathcal{B}_B(X)$, $z_{\alpha, \beta}^{(1)} \in W_{\alpha, \beta}^{(1)}$ such that $\left\| z_{\alpha, \beta}^{(1)} \right\| < \inf_{x \in W_{\alpha, \beta}^{(1)}} \| x \| + 1$, and $W_{\alpha, \beta}^{(2)} = \emptyset$.

Case 3: $x_{\alpha} \in W_{\alpha, \beta}$ and $y_{\beta} \not\in W_{\alpha, \beta}$. Define $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta} \in \mathcal{B}_B(X)$, $z_{\alpha, \beta}^{(1)} = x_{\alpha} \in W_{\alpha, \beta}^{(1)}$, and $W_{\alpha, \beta}^{(2)} = \emptyset$. Then, $\left\| z_{\alpha, \beta}^{(1)} \right\| = \| x_{\alpha} \| < \inf_{x \in U_{\alpha}} \| x \| + 1 \leq \inf_{x \in W_{\alpha, \beta}^{(1)}} \| x \| + 1$.

Case 4: $x_{\alpha} \not\in W_{\alpha, \beta}$ and $y_{\beta} \in W_{\alpha, \beta}$. Define $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta} \in \mathcal{B}_B(X)$, $z_{\alpha, \beta}^{(1)} = y_{\beta} \in W_{\alpha, \beta}^{(1)}$, and $W_{\alpha, \beta}^{(2)} = \emptyset$. Then, $\left\| z_{\alpha, \beta}^{(1)} \right\| = \| y_{\beta} \| < \inf_{x \in V_{\beta}} \| x \| + 1 \leq \inf_{x \in W_{\alpha, \beta}^{(1)}} \| x \| + 1$.

Case 5: $x_{\alpha} = y_{\beta} \in W_{\alpha, \beta}$. Define $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta} \in \mathcal{B}_B(X)$, $z_{\alpha, \beta}^{(1)} = x_{\alpha} \in W_{\alpha, \beta}^{(1)}$, and $W_{\alpha, \beta}^{(2)} = \emptyset$. Then, $\left\| z_{\alpha, \beta}^{(1)} \right\| = \| x_{\alpha} \| < \inf_{x \in U_{\alpha}} \| x \| + 1 \leq \inf_{x \in W_{\alpha, \beta}^{(1)}} \| x \| + 1$.

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 $W_{\alpha, \beta}^{(1)} = W_{\alpha, \beta} \cap \mathcal{B}_X(x_{\alpha}, \delta/2) \in \mathcal{B}_B(X)$, $z_{\alpha, \beta}^{(1)} = x_{\alpha} \in W_{\alpha, \beta}^{(1)}$, $W_{\alpha, \beta}^{(2)} = W_{\alpha, \beta} \cap \left( \mathcal{B}_X(x_{\alpha}, \delta/2) \right)^{\complement} \in \mathcal{B}_B(X)$, and $z_{\alpha, \beta}^{(2)} = y_{\beta} \in W_{\alpha, \beta}^{(2)}$. Then, $\left\| z_{\alpha, \beta}^{(1)} \right\| = \| x_{\alpha} \| < \inf_{x \in U_{\alpha}} \| x \| + 1 \leq \inf_{x \in W_{\alpha, \beta}^{(1)}} \| x \| + 1$ and $\left\| z_{\alpha, \beta}^{(2)} \right\| = \| y_{\beta} \| < \inf_{x \in V_{\beta}} \| x \| + 1 \leq \inf_{x \in W_{\alpha, \beta}^{(2)}} \| x \| + 1$.

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, $R_3 \in \mathcal{R}(X)$, $R_1 \subseteq R_3$, and $R_2 \subseteq R_3$. Therefore, $\mathcal{J}(X)$ is a directed system. This completes the proof of the proposition.

\textbf{Definition 11.70} Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $\mathcal{Y}$ be a normed linear space with the integration system $\mathcal{J}(\mathcal{Y}) := (\mathcal{R}(\mathcal{Y}), \preceq)$, and $f : X \to \mathcal{Y}$ be $\mathcal{B}$-measurable. $R = \left\{ (y_{\alpha}, U_{\alpha}) \mid \alpha \in \Lambda \right\} \subseteq \mathcal{R}(\mathcal{Y})$. Define $F_R := \sum_{\alpha \in \Lambda} y_{\alpha} \mu(f^{-1}(U_{\alpha}) \cap \mathcal{B}_X)$. 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 $X$ and denoted by $\int_X f \, d\mu$. 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 $X$.

\textbf{Definition 11.71} Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space with $\mu(X) = \infty$, $\mathcal{Y}$ be a normed linear space, and $f : X \to \mathcal{Y}$ 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
Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} \) be a normed linear space, \( f : X \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, and \( \hat{\mathcal{X}} := (\hat{X}, \hat{\mathcal{B}}, \hat{\mu}) \) be a finite measure space that is an extension of \( \mathcal{X} \), i.e., \( \mathcal{B} \subseteq \hat{\mathcal{B}} \) and \( \mu = \hat{\mu}|_{\mathcal{B}} \). Then, \( \int_X f \, d\mu = \int_X f|_{\hat{X}} \, d\hat{\mu} \), whenever one of the integrals exists.

**Proof**

Let \( J(\mathcal{Y}) \) be the integration system on \( \mathcal{Y} \) as defined in Proposition 11.69. \( (F_R)_{R \in J(\mathcal{Y})} \) be the net for \( \int_X f \, d\mu \), and \( (\hat{F}_R)_{R \in J(\mathcal{Y})} \) be the net for \( \int_X f|_{\hat{X}} \, d\hat{\mu} \), as defined in Definition 11.70. \( \forall R := \{ (y_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in J(\mathcal{Y}) \), \( \forall \alpha \in \Lambda \), we have \( f_{\text{inv}}(U_\alpha) \in \mathcal{B} \subseteq \hat{\mathcal{B}} \) since \( f \) is \( \mathcal{B} \)-measurable, and \( \mu(f_{\text{inv}}(U_\alpha)) = \hat{\mu}(f_{\text{inv}}(U_\alpha)) \in \mathbb{R} \) since \( \mu = \hat{\mu}|_{\mathcal{B}} \). Then, \( F_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} y_\alpha \hat{\mu}(f_{\text{inv}}(U_\alpha)) = \hat{F}_R \in \mathcal{Y} \). Hence, \( \int_X f \, d\mu = \lim_{R \in J(\mathcal{Y})} F_R = \lim_{R \in J(\mathcal{Y})} \hat{F}_R = \int_X f \, d\hat{\mu} \), whenever one of the integrals exists. This completes the proof of the proposition. \( \square \)

**Lemma 11.73**

Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathcal{Y} \) be a normed linear space, \( f : X \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( \hat{\mathcal{X}} := (\hat{X}, \mathcal{B}, \hat{\mu}) \) be the finite measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13. Assume that \( \mu(X \setminus \hat{X}) = 0 \). Then, \( \int_X f \, d\mu = \int_X f|_{\hat{X}} \, d\hat{\mu} \), whenever one of the integrals exists.

**Proof**

Let \( J(\mathcal{Y}) \) be the integration system on \( \mathcal{Y} \) as defined in Proposition 11.69. \( (F_R)_{R \in J(\mathcal{Y})} \) be the net for \( \int_X f \, d\mu \), and \( (\hat{F}_R)_{R \in J(\mathcal{Y})} \) be the net for \( \int_X f|_{\hat{X}} \, d\hat{\mu} \), as defined in Definition 11.70. \( \forall R := \{ (y_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in J(\mathcal{Y}) \), \( \forall \alpha \in \Lambda \), we have \( f_{\text{inv}}(U_\alpha) \cap \hat{X} = f|_{\hat{X}} \, \text{inv}(U_\alpha) \in \mathcal{B} \subseteq \hat{\mathcal{B}} \) since \( f \) is \( \mathcal{B} \)-measurable, and \( 0 \leq \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) \leq \mu(X \setminus \hat{X}) = 0 \). Then, \( F_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} y_\alpha (\mu(f_{\text{inv}}(U_\alpha) \cap \hat{X}) + \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X}))) = \sum_{\alpha \in \Lambda} y_\alpha \mu(f|_{\text{inv}}(U_\alpha)) = \hat{F}_R \in \mathcal{Y} \). Hence, \( \int_X f \, d\mu = \lim_{R \in J(\mathcal{Y})} F_R = \lim_{R \in J(\mathcal{Y})} \hat{F}_R = \int_X f \, d\hat{\mu} \), whenever one of the integrals exists. This completes the proof of the proposition. \( \square \)

**Proposition 11.74**

Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space with \( \mu(X) = +\infty \), \( \mathcal{Y} \) be a normed linear space, \( f : X \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, and \( \hat{\mathcal{X}} := (X, \mathcal{B}, \hat{\mu}) \) be the completion of \( \mathcal{X} \). Then, \( \int_X f \, d\mu = \int_X f \, d\hat{\mu} \), whenever one of the integrals exists.
Proposition 3.70, Let \((F_A)_{A \in \mathcal{M}(X)}\) be the net for \(\int_X f \, d\mu\) and \((\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})}\) be the net for \(\int_X f \, d\tilde{\mu}\) as defined in Definition 11.71. Clearly, \(\mathcal{M}(X) \subseteq \mathcal{M}(\tilde{X})\), since \(\tilde{X}\) is the completion of \(X\). We will distinguish two exhaustive cases: Case 1: \(\int_X f \, d\mu\) exists; Case 2: \(\int_X f \, d\tilde{\mu}\) exists.

Case 1: \(\int_X f \, d\mu\) exists. Then, the net \((\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})}\) is well defined. 
\(\forall A \in \mathcal{M}(X), A \in \mathcal{M}(\tilde{X})\). By Proposition 11.72, \(F_A = \int_A f \, d\mu_A = \int_A f \, d\tilde{\mu}_A = F_A\), where \((A, B_A, \mu_A)\) is the finite measure subspace of \(X\) and \((A, B_A, \tilde{\mu}_A)\) is the finite complete measure subspace of \(X\). Then, \((F_A)_{A \in \mathcal{M}(X)}\) is well defined. \(\forall A \in \mathcal{M}(\tilde{X})\), we have \(A \in \tilde{B}\) with \(\tilde{\mu}(A) < +\infty\). By Proposition 11.12 and its proof, \(\exists A, B, C \subseteq X\) with \(A = A \cup B, B, C \subseteq B, A \subseteq C, \mu(C) = 0\), and \(\tilde{\mu}(A) = \mu(B)\). Then, \(A \subseteq B \cup C =: A \in \tilde{B}\) and \(\mu(A) \leq \mu(B) + \mu(C) = \mu(B) = \tilde{\mu}(A) < +\infty\). This shows that \(A \in \mathcal{M}(\tilde{X})\) and \(A \subseteq A\). Hence, \((F_A)_{A \in \mathcal{M}(X)}\) is a subnet 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: \(\int_X f \, d\tilde{\mu}\) exists. Then, the net \((\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})}\) is well defined. 
\(\forall A \in \mathcal{M}(\tilde{X}), A \in \mathcal{M}(\tilde{X})\). By Proposition 11.12 and its proof, \(\exists A, B, C \subseteq X\) with \(A = A \cup B, B, C \subseteq B, A \subseteq C, \mu(C) = 0\), and \(\tilde{\mu}(A) = \mu(B)\). Let \(\tilde{B} := A \in \tilde{B}\). Then, \(0 \leq \tilde{\mu}(\tilde{A} \setminus A) = \tilde{\mu}(\tilde{A} \setminus B) \leq \mu(C) = 0\) and \(\mu(A) = \tilde{\mu}(A) < +\infty\). This shows that \(A \in \mathcal{M}(\tilde{X})\) and \(A \subseteq \tilde{X}\). Let \(\tilde{A} := (\tilde{A}, B_A, \tilde{\mu}_A)\) be the finite complete measure subspace of \(A\), and \(\tilde{A} := (A, B_A, \mu_A)\) be the finite measure subspace of \(\mathcal{M}(\tilde{X})\), and \(\tilde{A} := (A, B_A, \tilde{\mu}_A)\) be the finite complete measure subspace of \(\tilde{A}\). By Lemma 11.73, we have \(\tilde{F}_A = \int_A f \, d\tilde{\mu}_A = \int_A f \, d\mu_A\), whenever one of them exists. Since \(\tilde{X}\) is the completion of \(X\), then \(\tilde{A}\) is an extension of \(A\). By Proposition 11.72, we have \(\int_A f \, d\mu_A = \int_A f \, d\tilde{\mu}_A = F_A\), hence, \(\tilde{F}_A = F_A\). Then, the net \((\tilde{F}_A)_{A \in \mathcal{M}(\tilde{X})}\) is well defined.
Hence, in both cases, we have \( \int_X f \, d\mu = \int_X f \, d\bar{\mu} \). This completes the proof of the proposition. \( \square \)

**Proposition 11.75** Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a normed linear space, \( \phi : X \to Y \) be a simple function in canonical representation, i.e., \( \exists n \in \mathbb{Z}_+ \), \( \exists y_1, \ldots, y_n \in Y \), which are distinct and none equals to \( \phi_y \), \( \exists A_1, \ldots, A_n \in \mathcal{B} \), 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, \( \int_X \phi \, d\mu = \sum_{i=1}^n y_i \mu(A_i) =: I \in Y \).

**Proof**

We will distinguish two exhaustive and mutually exclusive cases: Case 1a: \( Y \) is finite index set, \( \forall y \in Y \) exist exhaustive and mutually exclusive subcases: Case 1b: \( Y \) is a simple function in canonical representation, \( I \in Y \) such that \( \parallel \phi \parallel \leq \infty \). Hence, in both cases, we have \( \parallel \phi \parallel \leq \infty \), \( \forall y \in Y \).

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 \( \Phi_{\mathcal{B}} = \vartheta y \). Clearly, we must have \( n = 0 \), since \( y_1, \ldots, y_n \in Y \) are distinct and none equals to \( \vartheta y \). Then, \( I = \vartheta y \). Hence, \( \forall R \in \mathcal{I}(Y) \) with \( R_0 \leq R \), \( \parallel \Phi_R - I \parallel = \parallel \Phi_{R_0} - I \parallel = 0 \).

Let \( y \in Y \), \( U_{n+2} = \bigcap_{i=1}^{n+2} U_i \in \mathcal{B}_Y \), 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 \( \epsilon_R := \{y_i, U_i\} \). Then, \( \Phi_{R_0} = \vartheta y \). Clearly, we must have \( n = 0 \), since \( y_1, \ldots, y_n \in Y \) are distinct and none equals to \( \vartheta y \). Then, \( I = \vartheta y \). Hence, \( \forall R \in \mathcal{I}(Y) \) with \( R_0 \leq R \), \( \parallel \Phi_R - I \parallel = \parallel \Phi_{R_0} - I \parallel = 0 \).

In both subcases, we have \( \exists R_0 \in \mathcal{I}(Y), \forall R \in \mathcal{I}(Y) \) with \( R_0 \leq R \), we have \( \parallel \Phi_R - I \parallel = 0 < \epsilon \). Hence, \( \int_X \phi \, d\mu = \lim_{R \in \mathcal{I}(Y)} \Phi_R = I. \)

**11.4. INTEGRATION**

387
Case 2: \( \mu(X) = +\infty \). \( \forall A \in \mathcal{M}(\mathcal{X}), \mu(A) < +\infty \). \( \phi|_A : A \to \mathbb{Y} \) 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 \( \mathcal{X} \). Hence, the net \( (\Phi_A)_{A \in \mathcal{M}(\mathcal{X})} \) 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, \( A_0 \in \mathcal{M}(\mathcal{X}) \). \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_0 \subseteq A \), by Case 1, \( \Phi_A = \sum_{i=1}^n y_i \mu(A_i) = I \). Then, \( \int_X \phi \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} \Phi_A = I \).

This completes the proof of the proposition.

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 \( \int_X (\phi_1 + \phi_2) \, d\mu = \int_X \phi_1 \, d\mu + \int_X \phi_2 \, d\mu \) and \( \int_X (c\phi_1) \, d\mu = c \int_X \phi_1 \, d\mu \).

Lemma 11.76 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a finite measure space, \( \mathbb{Y} \) be a separable Banach space, \( \phi_n : X \to \mathbb{Y} \) be a simple function in canonical representation, \( \forall n \in \mathbb{N} \), and \( f : X \to \mathbb{Y} \) be \( \mathcal{B} \)-measurable. Assume that \( \lim_{n \in \mathbb{N}} \phi_n = f \) a.e. in \( \mathcal{X} \) 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 \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu \in \mathbb{Y} \).

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 \( (F_R)_{R \in \mathcal{J}(\mathbb{Y})} \) be the net for \( \int_X f \, d\mu \) as defined in Definition 11.70.

\( \forall \epsilon \in (0, 1) \subset \mathbb{R} \), by Egoroff’s Theorem 11.55, \( \exists E_1 \in \mathcal{B} \) with \( \mu(E_1) < \frac{\epsilon}{2M+2} \), such that \( \left( \phi_n|_{X \setminus E_1} \right)_{n=1}^\infty \) converges uniformly to \( f|_{X \setminus E_1} \). Clearly, \( E \subseteq E_1 \). Then, \( \exists n_0 \in \mathbb{N} \), \( \forall n \in \mathbb{N} \) with \( n \geq n_0 \), \( \forall x \in X \setminus E_1 \), \( \|\phi_n(x) - f(x)\| < \frac{\epsilon}{\mu(X \setminus E_1)} := \epsilon \). 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 \( n \in \mathbb{Z}_+ \), \( y_1, \ldots, y_n \in \mathbb{Y} \) are distinct and none equals to \( y \), and \( A_1, \ldots, A_n \in \mathcal{B} \) are nonempty, pairwise disjoint, and of finite measure. Let \( y_{n+1} := y \) 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. Define \( V_i := B_y(y_i, \epsilon), i = 1, \ldots, n+1 \). Define \( V_1 := V_1 \in \mathcal{B}_\mathbb{Y}(\mathbb{Y}), V_{i+1} := V_{i+1} \setminus \left( \bigcup_{j=1}^{n} V_j \right) \in \mathcal{B}_\mathbb{Y}(\mathbb{Y}), i = 1, \ldots, n \). Let \( V_{n+2} := \mathbb{Y} \setminus \left( \bigcup_{j=1}^{n+1} V_j \right) \). Clearly, we may form a presentation \( \tilde{R} := \{ (y_i, V_i) \mid i = 1, \ldots, n+2, V_i \neq \emptyset \} \in \mathcal{J}(\mathbb{Y}) \), where \( y_i \in V_i, i = 1, \ldots, n+2 \), are any vectors that satisfy the assumption of Proposition 11.69. \( \forall R \in \mathcal{J}(\mathbb{Y}) \) with \( \tilde{R} \preceq R \). Let \( R := \{ (y_\gamma, \tilde{U}_\gamma) \mid \gamma \in \Gamma \} \).

\( \forall \gamma \in \Gamma, \) \( \tilde{y}_\gamma \in \mathbb{Y} \), \( \tilde{U}_\gamma \subseteq V_i \). Define \( \tilde{\Gamma} := \{ \gamma \in \Gamma \mid i_\gamma = n+2 \} \). Let \( \tilde{A}_{i_\gamma} := A_i \cap f_{\mathbb{Y}}(\tilde{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}, i = 1, \ldots, n+1 \).

Claim 11.76.1 \( \forall \gamma \in \tilde{\Gamma}, f_{\mathbb{Y}}(\tilde{U}_\gamma) \subseteq E_1 \) and \( \|\tilde{y}_\gamma \mu(f_{\mathbb{Y}}(\tilde{U}_\gamma) \cap E_1)\| \leq (M + 1)\mu(f_{\mathbb{Y}}(\tilde{U}_\gamma)). \)
11.4. INTEGRATION

Proof of claim: Fix any $\gamma \in \Gamma$. $\hat{U}_\gamma \subseteq V_{n+2} = y \setminus (\bigcup_{j=0}^{n+1} \hat{V}_j)$. Then, $B \ni f_{\text{inv}}(\hat{U}_\gamma) \subseteq E_1$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E$; Case 2: $f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$.

Case 1: $f_{\text{inv}}(\hat{U}_\gamma) \subseteq E$. Then, $\mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) = 0$ and $\left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| = 0 = (M + 1) \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)$. Case 2: $f_{\text{inv}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset$.

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 $R \in \mathcal{F}(\mathcal{Y})$, $\| \hat{y}_\gamma \| < \inf_{y \in \hat{U}_\gamma} \| y \| + 1 \leq M + 1$. Therefore, $\left\| \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\| \leq (M + 1) \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1)$. This completes the proof of the claim.

Note that $F_R = \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma))$ and $\int_X \phi_n \, d\mu = \sum_{i=1}^n y_i \mu(A_i)$, by Proposition 11.75. Then,

$$
\left\| F_R - \int_X \phi_n \, d\mu \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^n y_i \mu(A_i) \right\|
$$

$$
= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right. \\
- \left. \sum_{i=1}^{n+1} y_i \mu(A_i \cap E_1) + \mu(A_i \cap (X \setminus E_1)) \right\|
$$

$$
= \left\| \sum_{\gamma \in \Gamma} \hat{y}_\gamma \left( \sum_{i=1}^{n+1} \mu(A_{\gamma,i}) + \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right) - \sum_{i=1}^{n+1} y_i \mu(A_i \cap E_1) \\
+ \sum_{\gamma \in \Gamma} \mu(A_{\gamma,i}) \right\|
$$

$$
= \left\| \sum_{\gamma \in \Gamma} \left( \sum_{i=1}^{n+1} \hat{y}_\gamma \mu(A_{\gamma,i}) + \sum_{i=1}^{n+1} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) - \sum_{i=1}^{n+1} y_i \mu(A_i \cap E_1) \\
- \sum_{i=1}^{n+1} y_i \mu(A_i \cap E_1) \right) \right\|
$$

$$
= \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{n+1} (\hat{y}_\gamma - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap E_1) \right\|
$$
\[
- \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1)
\]

\[
= \left\| \sum_{\gamma \in \Gamma \setminus \Gamma'} \sum_{i=1}^{\bar{n}+1} (\hat{y}_{\gamma} - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_{\gamma} - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_{\gamma} \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma \setminus \Gamma'} \sum_{i=1}^{\bar{n}+1} (\hat{y}_{\gamma} - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{y}_{\gamma} - y_i) \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \hat{y}_{\gamma} \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1) - \sum_{i=1}^{\bar{n}+1} y_i \mu(A_i \cap E_1) \right\|
\]

\[
\leq \sum_{\gamma \in \Gamma \setminus \Gamma'} \sum_{i=1}^{\bar{n}+1} \| \hat{y}_{\gamma} - y_i \| \mu(A_{\gamma,i}) + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \| \hat{y}_{\gamma} \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1) \|
\]

\[
+ \sum_{\gamma \in \Gamma} \| \hat{y}_{\gamma} \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1) \| + \sum_{i=1}^{\bar{n}+1} \| y_i \| \mu(A_i \cap E_1)
\]

\[
\leq \frac{3\mu(X)\epsilon}{\delta \mu(X) + 1} + \sum_{\gamma \in \Gamma \setminus \Gamma'} (M + 1) \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1))
\]

\[
+ \sum_{\gamma \in \Gamma} (M + 1) \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1)) + \sum_{i=1}^{\bar{n}+1} M \mu(A_i \cap E_1)
\]

\[
< \epsilon/2 + \sum_{\gamma \in \Gamma} (M + 1) \mu(f_{\text{inv}}(\hat{U}_{\gamma}) \cap E_1)) + \sum_{i=1}^{\bar{n}+1} M \mu(A_i \cap E_1)
\]

\[
= \epsilon/2 + (2M + 1) \mu(E_1) < \epsilon
\]

where the seventh equality follows from the fact that $\forall \gamma \in \Gamma$, $A_{\gamma,i} \subseteq f_{\text{inv}}(\hat{U}_{\gamma}) \cap (X \setminus E_1) = \emptyset$; and the third inequality follows from the fact that $\forall \gamma \in \Gamma \setminus \Gamma'$, $\hat{y}_{\gamma} \in \hat{U}_{\gamma} \subseteq V_{\gamma, i} \subseteq \bar{V}_{\gamma, i} = B_y(y_{\gamma, i}, \bar{\epsilon})$ and $\| \hat{y}_{\gamma} \| \leq \| y_{\gamma, i} \| + \| \hat{y}_{\gamma} - y_{\gamma, i} \| < M + \bar{\epsilon} < M + 1$.

In summary, we have shown that $\forall \epsilon \in (0, 1) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall n \in \mathbb{N}$ with $n \geq n_0$, $\exists R \in \mathcal{J}(\mathcal{Y})$ such that $\forall R \in \mathcal{J}(\mathcal{Y})$ with $R \subseteq R$, we have $\| F_R - \int_X \phi_n \, d\mu \| < \epsilon$. Then, the net $(F_R)_{R \in \mathcal{J}(\mathcal{Y})}$ is a Cauchy net. By Proposi-
11.4. INTEGRATION

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 \( \| \int_X f \, d\mu - \int_X \phi_n \, d\mu \| = \lim_{R \in \mathbb{R}} \| F_R - \int_X \phi_n \, d\mu \| \leq \varepsilon, \forall n \geq n_0 \). Hence, we have \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu \). This completes the proof of the lemma.

\[ \square \]

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 \(\lim_{n \in \mathbb{N}} f_n = f\) 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 \(\int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{Y}\).

\textbf{Proof} Let \(U := \overline{B}_y(\phi_y, M)\), 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, \(\int_X f_n \, d\mu = \lim_{n \in \mathbb{N}} \int_X \phi_{n,i} \, d\mu \in \mathcal{Y}\). By Proposition 11.58, we have \(\lim_{n \in \mathbb{N}} \phi_{n,i} = f_n\) 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, \(\lim_{n \in \mathbb{N}} f_n = f\) in measure in \(\mathcal{X}\). \(\forall \varepsilon \in (0, \infty) \subset \mathbb{R}\), \(\exists n_0 \in \mathbb{N}\) such that \(\forall n \in \mathbb{N}\), \(2^{-n} < \varepsilon/2\) and \(\mu(E_n) := \mu(\{ x \in X \mid \| f_n(x) - f(x) \| \geq \varepsilon/2 \}) < \varepsilon/2\). Note that, by Lemma 11.43, \(\mathcal{B} \ni F_n := \{ x \in X \mid \| \phi_n(x) - f(x) \| \geq \varepsilon \} \subseteq \{ x \in X \mid \| \phi_n(x) - f_n(x) \| + \| f_n(x) - f(x) \| \geq \varepsilon \} \subseteq E_n \cup \tilde{E}_n\). Then, we have \(\mu(F_n) \leq \mu(E_n) + \mu(\tilde{E}_n) < \varepsilon\). This implies that \(\lim_{n \in \mathbb{N}} \phi_n = f\) 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 \(\int_X f \, d\mu = \lim_{k \in \mathbb{N}} \int_X \phi_{n_k} \, d\mu \in \mathcal{Y}\). 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 \(\lim_{k \in \mathbb{N}} \int_X f_n \, d\mu = \lim_{k \in \mathbb{N}} \int_X \phi_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\). 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 \(\lim_{k \in \mathbb{N}} \int_X f_{n_k} \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\). This then holds for any subsequence of \((f_n)_{n=1}^{\infty}\). By Proposition 3.71, we have \(\lim_{n \in \mathbb{N}} \int_X f_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}\) and \(f\) is integrable over \(X\).

This completes the proof of the theorem.

\[ \square \]

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,

\[ \int_X f \, d\mu = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu =: S \in \mathbb{R} \]

where the supremum is over all simple functions \(\phi : X \to [0, \infty) \subset \mathbb{R}\) that satisfy \(0 \leq \phi(x) \leq f(x), \forall x \in X\).
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 $\mu(A_i) < \infty$, $i = 1, \ldots, n$. Let $y_0 := 0$ 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 $R_0 \not\subset R$, let $\overline{R} = \{ (\overline{y}_\alpha, \overline{U}_\alpha) \mid \alpha \in \Lambda \}$. $\forall \alpha \in \Lambda$, by $R_0 \not\subset R$, $\exists i_\alpha \in \{1, \ldots, n+2\}$ such that $\overline{U}_\alpha \subseteq I_{i_\alpha}$. Define $\overline{\Lambda} := \{ \alpha \in \Lambda \mid i_\alpha = n+2 \}$. Note that $\forall \alpha \in \overline{\Lambda}, \overline{U}_\alpha \subseteq I_{n+2}$ and $f_{\text{inv}}(\overline{U}_\alpha) = \emptyset$. Then,

$$F_R = \sum_{\alpha \in \Lambda} \overline{y}_\alpha \mu(f_{\text{inv}}(\overline{U}_\alpha)) = \sum_{\alpha \in \Lambda} \overline{y}_\alpha \sum_{i=1}^{n+1} \mu(A_i \cap f_{\text{inv}}(\overline{U}_\alpha))$$

$$= \sum_{\alpha \in \Lambda} \overline{y}_\alpha \sum_{i=1}^{n+1} \mu(E_{i,\alpha}) = \sum_{\alpha \in \Lambda} \overline{y}_\alpha \sum_{i=1}^{n+1} \mu(E_{i,\alpha})$$

$\forall \alpha \in \Lambda \setminus \overline{\Lambda}, \forall i \in \{1, \ldots, n+1\}$, if $E_{i,\alpha} \neq \emptyset$, then $\exists x \in E_{i,\alpha}$, $y_i = \phi(x) \leq f(x) \in \overline{U}_\alpha \subseteq I_{i_\alpha}$, which further implies that $y_i \leq y_{i_\alpha}$. Note that $\overline{y}_\alpha \in \overline{U}_\alpha \subseteq I_{i_\alpha}$ and $\overline{y}_\alpha \geq y_{i_\alpha}$. Then, $\overline{y}_\alpha \mu(E_{i,\alpha}) \geq y_{i_\alpha} \mu(E_{i,\alpha})$. If $E_{i,\alpha} = \emptyset$, clearly $0 = \overline{y}_\alpha \mu(E_{i,\alpha}) = y_{i_\alpha} \mu(E_{i,\alpha}) = 0$. Therefore, we have $\overline{y}_\alpha \mu(E_{i,\alpha}) \geq y_{i_\alpha} \mu(E_{i,\alpha})$,

$\forall \alpha \in \Lambda \setminus \overline{\Lambda}, \forall i \in \{1, \ldots, n+1\}$. This leads to

$$F_R = \sum_{\alpha \in \Lambda \setminus \overline{\Lambda}} \sum_{i=1}^{n+1} \overline{y}_\alpha \mu(E_{i,\alpha}) \geq \sum_{\alpha \in \Lambda \setminus \overline{\Lambda}} \sum_{i=1}^{n+1} y_{i_\alpha} \mu(E_{i,\alpha}) = \sum_{i=1}^{n+1} \sum_{\alpha \in \Lambda \setminus \overline{\Lambda}} y_{i_\alpha} \mu(E_{i,\alpha})$$

where the last equality follows from Proposition 11.75. In summary, we have shown that $\forall R \in \mathcal{J}(\mathbb{R})$ with $R_0 \not\subset R$, 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 $\overline{S} := \limsup_{R \in \mathcal{J}(\mathbb{R})} F_R > S$. 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 $R_0 \not\subset R$, such that $F_R > S + \epsilon_0$. Let $\epsilon_1 := \min\{1, \frac{\epsilon_0}{1 + 2\mu(X)}\} > 0$. By Proposition 11.5, $\exists E_1 \in \mathcal{B}$ with $\mu(E_1) < \epsilon_1$.
11.4. INTEGRATION

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 \( R = \{(\tilde{y}_\alpha, \tilde{U}_\alpha) \mid \alpha \in \Lambda\}. \forall \alpha \in \Lambda, \exists i_\alpha \in \{1, \ldots, n+2\} \) such that \( \tilde{U}_\alpha \subseteq U_{i_\alpha} \).

Define \( \tilde{\Lambda} := \{\alpha \in \Lambda \mid i_\alpha = n+2\}, \Lambda_1 := \{\alpha \in \Lambda \mid i_\alpha = 1, \ldots, n\}, \) and \( \Lambda_2 := \{\alpha \in \Lambda \mid i_\alpha = n+1\}. \) Clearly, \( \Lambda = \Lambda_1 \cup \Lambda_2 \cup \tilde{\Lambda} \) 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 \( \tilde{y}_\alpha \in \tilde{U}_\alpha \subseteq U_{i_\alpha} \), then \( \tilde{y}_\alpha \leq \tilde{y}_\alpha \leq y_{i_\alpha} < y_{i_\alpha} + 1 \).

Define \( \phi : X \to [0, \infty) \subset \mathbb{R} \) by \( \phi := \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \chi_{A_\alpha} \).

Clearly, \( 0 \leq \phi(x) \leq f(x), \forall x \in X \). Then, by Proposition 11.75,

\[
\int_X \phi \, d\mu = \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \mu(A_\alpha) = \sum_{\alpha \in \Lambda_1 \cup \Lambda_2} \tilde{y}_\alpha \mu(f_{\text{inv}}(\tilde{U}_\alpha)) - F_R + F_R
\]

This contradicts with the definition of \( S \). Therefore, \( \liminf_{R \in J(\mathbb{R})} F_R \leq S \).

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, \( S = \limsup_{R \in J(\mathbb{R})} F_R = \int_X f \, d\mu \).

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, \( F_A = \int_A f \, d\mu_A = \sup_{0 \leq \phi \leq f_A} \int_A \phi \, d\mu_A \in \mathbb{R}_e \), where supremum is over all simple functions \( \phi : X \to [0, \infty) \subset \mathbb{R} \) with \( 0 \leq \phi(x) \leq f(x), \forall x \in X \).

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 \( \phi = \sum_{i=1}^n y_i \chi_{A_i} \). Take \( A_0 = \)
CHAPTER 11. GENERAL MEASURE AND INTEGRATION

\[ \bigcup_{i=1}^{n} A_i \subseteq B. \] 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 follows from Proposition 11.75. Then, \( \lim \inf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq \int_X \phi \, d\mu \). By the arbitrariness of \( \phi \), we have \( \lim \inf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S. \)

On the other hand, \( \sup_{A \in \mathcal{M}(\mathcal{X})} F_A =: \bar{S} \geq \lim \inf_{A \in \mathcal{M}(\mathcal{X})} F_A \geq S. \) Suppose \( \bar{S} > S \). 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 \( \bar{\phi} : X \to [0, \infty) \subset \mathbb{R} \).

Then, clearly \( 0 \leq \bar{\phi}(x) \leq f(x), \forall x \in X \), and, by Proposition 11.75, \( \int_X \bar{\phi} \, d\mu = \int_X \phi \, d\mu_A > S \). This contradicts with definition of \( S \). Therefore, we have \( \bar{S} \leq S \).

In summary, we have \( S \leq \lim \inf_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \lim \sup_{A \in \mathcal{M}(\mathcal{X})} F_A \leq \sup_{A \in \mathcal{M}(\mathcal{X})} F_A = \bar{S} \leq S \). Then, by Proposition 3.83, we have \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_A = S \in \mathbb{R}. \)

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

**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, \( \int_X f \, d\mu \leq \lim \inf_{n \in \mathbb{N}} \int_X f_n \, d\mu. \)

**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 \( g_n := f_n \wedge \phi, \forall n \in \mathbb{N} \). 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 \(\mathcal{X}\) as defined in Proposition 11.13. By Definition 11.79, we have \(\int_E g_n|_E \, d\mu_E \leq \int_X g_n \, d\mu \leq \int_X f_n \, d\mu\). Then, \(\int_X \phi \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu\). By the arbitrariness of \(\phi\), we have \(\int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu\). This completes the proof of the theorem. \(\square\)

**Theorem 11.81 (Monotone Convergence Theorem)** 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 \(\lim_{n \to \infty} f_n = f\) a.e. in \(\mathcal{X}\) and \(f_n(x) \leq f_{n+1}(x), \forall x \in X, \forall n \in \mathbb{N}\). Then, \(\int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu\).\(^{\dagger}\)

**Proof** By the assumption, \(f_n \leq f\) a.e. in \(\mathcal{X}\), \(\forall n \in \mathbb{N}\). 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 \(\int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu \leq \limsup_{n \to \infty} \int_X f_n \, d\mu \leq \int_X f \, d\mu\). Hence, by Proposition 3.83, we have \(\int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in \mathbb{R}_e\). This completes the proof of the proposition. \(\square\)

**Proposition 11.82** Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite measure space and \(g : X \to [0, \infty] \subset \mathbb{R}_e\) be \(\mathcal{B}\)-measurable. Assume that \(\int_X g \, d\mu =: \lambda < \infty\). Then, define \(f : X \to [0, \infty] \subset \mathbb{R}\) by \(f(x) = \begin{cases} \frac{g(x)}{\lambda} & g(x) < \infty \\ 0 & g(x) = \infty \end{cases}, \forall x \in X\). Then, \(f\) is \(\mathcal{B}\)-measurable, \(f = g\) a.e. in \(\mathcal{X}\) 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 \(\mathcal{X}\) 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 \(\lim_{n \to \infty} \mu(B_n) = \mu(X \setminus E) > 0\). 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 \(\lambda = \int_X g \, d\mu \geq \int_X \tilde{g} \, d\mu \geq M \mu(B_n) = \lambda + 1 > \lambda\), 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 \(f(x) = \begin{cases} \frac{g(x)}{\lambda} & x \in E \\ 0 & x \in X \setminus E \end{cases}\). By Proposition 11.41, \(f\) is \(\mathcal{B}\)-measurable and \(f = g\) a.e. in \(\mathcal{X}\). 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, $f_i : X \to [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$-measurable, $i = 1, 2$, $g_n : X \to [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$, and $g : X \to [0, \infty] \subset \mathbb{R}_e$ be $\mathcal{B}$-measurable. Then, the following statements hold.

(i) $\forall c \in (0, \infty) \subset \mathbb{R}$, $\int_X (cg) \, d\mu = c \int_X g \, d\mu; \forall c \in [0, \infty) \subset \mathbb{R}$, 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 $\int_E g \, d\mu_E = \int_E (g\chi_E) \, d\mu =: \int_E g \chi_E \, d\mu$, where $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ is the measure subspace of $\mathcal{X}$.

(iv) $\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathbb{R}_e$.

(v) $\forall E_1, E_2 \in \mathcal{B}$ with $E_1 \cap E_2 = \emptyset$, we have $\int_{E_1 \cup E_2} g \, d\mu = \int_{E_1} g \, d\mu + \int_{E_2} g \, d\mu \in \mathbb{R}_e$.

(vi) $\forall$ pairwise disjoint $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$, we have $\int_{\bigcup_{n=1}^{\infty} E_n} g \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} g \, d\mu \in \mathbb{R}_e$.

(vii) If $g$ and $g_n$'s are integrable over $\mathcal{X}$, $\lim_{n \to \infty} g_n = g$ a.e. in $\mathcal{X}$, and $\lim_{n \to \infty} \int_X g_n \, d\mu = \int_X g \, d\mu \in \mathbb{R}$. Then, $\forall E \in \mathcal{B}$, we have $\int_E g \, d\mu = \lim_{n \to \infty} \int_E g_n \, d\mu$.

Proof

(i) and (ii) follows directly from Propositions 11.75 and 3.81 and Definition 11.79.

(iii) Note that $g|_E$ is $\mathcal{B}_E$-measurable by Proposition 11.41. By Propositions 7.23, 11.38, and 11.39, $g\chi_{E,X}$ is $\mathcal{B}$-measurable. Then, 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 $\phi : X \to [0, \infty) \subset \mathbb{R}$ satisfying the stated inequalities. Any simple function $\phi_1 : X \to [0, \infty) \subset \mathbb{R}$ and $\phi_2 : X \to [0, \infty) \subset \mathbb{R}$ 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, $\int_X \phi_1 \, d\mu + \int_X \phi_2 \, d\mu = \int_X (\phi_1 + \phi_2) \, d\mu \leq \int_X (f_1 + f_2) \, d\mu$, where the equality follows from Proposition 11.75. Hence, we have $\int_X f_1 \, d\mu + \int_X f_2 \, d\mu \leq \int_X (f_1 + f_2) \, d\mu$.

On the other hand, fix any simple function $\phi$ satisfying $0 \leq \phi \leq f_1 + f_2$. Let $f_1 := \phi \wedge f_1$ and $f_2 := \phi \wedge f_2$. By Propositions 11.40, 11.38, 11.39, and 7.23, $f_1$ and $f_2$ are $\mathcal{B}$-measurable. Then, we have $0 \leq f_1 \leq f_1$, $0 \leq f_2 \leq f_2$, $f_1(x) = f_2(x) = 0$, $\forall x \in X \setminus E$, where $E := \{x \in X \mid \phi(x) > 0\} \in \mathcal{B}$, and $0 \leq f_1(x) \leq \phi(x) \leq M$ and $0 \leq f_2(x) \leq \phi(x) \leq M$, $\forall x \in X$, for some $M \in [0, \infty) \subset \mathbb{R}$. Note that $\mu(E) < +\infty$ since $\phi$ is a simple function. Let $\mathcal{E} := (E, \mathcal{B}_E, \mu_E)$ be the finite measure subspace of $\mathcal{X}$ as
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 \(E\), 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, \(\lim_{n \to \infty} \int_E \phi_{1,n} + \phi_{2,n} = \int_E \phi_1 + \phi_2\) a.e. in \(E\). Again, by Bounded Convergence Theorem 11.77 and Proposition 11.75, we have \(\int_E \phi|_E \, d\mu_E = \lim_{n \to \infty} \int_E \phi_{1,n} + \phi_{2,n} \, d\mu_E = \lim_{n \to \infty} \left(\int_E \phi_1 \, d\mu_E + \int_E \phi_2 \, d\mu_E\right)\). By Proposition 11.75, \(\mu\) and \((\phi_{1,n} + \phi_{2,n})\), we have \(\int_X \phi \, d\mu = \int_X \phi_1 \, d\mu + \int_X \phi_2 \, d\mu\). By the arbitrariness of \(\phi\), we have \(\int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu\).

(v) By (iii) and (iv), we have \(\int_{E_1 \cup E_2} g \, d\mu = \int_X (g \chi_{E_1 \cup E_2}) \, d\mu = \int_X (g \chi_{E_1} + g \chi_{E_2}) \, d\mu = \int_X g \chi_{E_1} \, d\mu + \int_X g \chi_{E_2} \, d\mu = \int_{E_1} g \, d\mu + \int_{E_2} g \, d\mu \in \mathbb{R}^e\).

(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, \(g \chi_{E_n} \) and \(g \chi_{E_n} \) are \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\). Clearly, \(\lim_{n \to \infty} g(x) \chi_{E_n} \leq g(x) \chi_{E_n} \) for all \(x \in X\), and \(0 \leq g(x) \chi_{E_n} \leq g(x) \chi_{E_{n-1}} \) for all \(x \in X, \forall n \in \mathbb{N}\). By Monotone Convergence Theorem 11.81, \(\int_E g \, d\mu = \lim_{n \to \infty} \int_{E_n} g \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} g \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} g \, d\mu\) where the fourth equality follows from (v).

(vii) By (iii), we have \(0 \leq \int_X g \, d\mu \leq \int_X g \, d\mu < +\infty, 0 \leq \int_{X \setminus E} g \, d\mu \leq \int_{X \setminus E} g \, d\mu < +\infty, 0 \leq \int_X g \, d\mu < +\infty, 0 \leq \int_X g \, d\mu < +\infty, \forall n \in \mathbb{N}\). By Fatou’s Lemma 11.80, \(\int_E g \, d\mu \leq \liminf_{n \to \infty} \int_{E_n} g \, d\mu = \liminf_{n \to \infty} \int_{E_n} g \, d\mu\). By a similar argument, \(\int_X g \, d\mu \leq \liminf_{n \to \infty} \int_{X \setminus E_n} g \, d\mu\). By (v) and Proposition 3.83, we have \(\int_E g \, d\mu = \limsup_{n \to \infty} \int_{E_n} g \, d\mu = \limsup_{n \to \infty} \left(\int_{E_n} g \, d\mu - \int_{X \setminus E_n} g \, d\mu\right)\). By Proposition 3.83, we have \(\int_E g \, d\mu = \lim_{n \to \infty} \int_{E_n} g \, d\mu \in \mathbb{R}\).

This completes the proof of the proposition. □

**Proposition 11.84** Let \(X := (X, \mathcal{B}, \mu)\) be a measure space and \(f : X \to [0, \infty] \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that \(f\) is integrable over \(X\), that is, \(\int_X f \, d\mu \in \mathbb{R}\). Then, \(\forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta \in (0, \infty) \subseteq \mathbb{R}\) such that \(\forall A \in \mathcal{B}\) with \(\mu(A) < \delta\), we have \(\int_A f \, d\mu_A < \varepsilon\), where \(A := (A, \mathcal{B}_A, \mu_A)\) is the measure subspace of \(X\) as defined in Proposition 11.13.

**Proof** \(\forall n \in \mathbb{N}\), let \(f_n := f \wedge n : X \to [0, n] \subseteq \mathbb{R}\). Then, \(f_n(x) \leq f_{n+1}(x), \forall x \in X\) and \(\lim_{n \to \infty} f_n(x) = f(x), \forall x \in X\). By
Chapter 11. General Measure and Integration

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, 

\[
\int_X f \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu. \quad \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \text{ then, } \exists n_0 \in \mathbb{N} \text{ such that } \left| \int_X f \, d\mu - \int_X f_n \, d\mu \right| < \varepsilon/2 \text{ Choose } \delta := \frac{\varepsilon}{2n_0} \in (0, \infty) \subseteq \mathbb{R}. \quad \forall A \in \mathcal{B} \text{ with } \mu(A) < \delta, \text{ by Proposition 11.83 and Definition 11.79, we have }
\]

\[
\int_A f|_A \, d\mu_A = \int_A (f - f_n)|_A \, d\mu_A + \int_A f_n|_A \, d\mu_A \\
\leq \int_X (f - f_n) \, d\mu + n_0 \mu(A) = \int_X f \, d\mu - \int_X f_n \, d\mu + n_0 \mu(A) \\
< \varepsilon/2 + \varepsilon/2 = \varepsilon
\]

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, \( f_n : X \to [0, \infty] \subseteq \mathbb{R} \), be \( \mathcal{B} \)-measurable, \( f : X \to [0, \infty] \subseteq \mathbb{R} \) 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, \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) and \( \exists n_0 \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \text{ with } \mu(A) < \delta, \forall n \in \mathbb{N} \text{ with } n > n_0, \text{ we have } 0 \leq \int_A f|_A \, d\mu_A < \varepsilon \) and \( 0 \leq \int_A f_n|_A \, d\mu_A < \varepsilon \), where \( (A, \mathcal{B}_A, \mu_A) \) is the finite measure subspace of \( \mathcal{X} \) as defined in Proposition 11.13.

Proof. Let \( \bar{f}_n := f_n \land f, \forall n \in \mathbb{N} \). 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, \( \lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu \) 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, \( \int_X f \, d\mu = \lim_{n\to\infty} \int_X f_n \, d\mu \in \mathbb{R} \).

\( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \) by Proposition 11.84, \( \exists \delta \in (0, \infty) \subseteq \mathbb{R} \) and \( \exists n_0 \in \mathbb{N} \) such that \( \forall A \in \mathcal{B} \text{ with } \mu(A) < \delta, \forall n \in \mathbb{N} \text{ with } n > n_0, \text{ we have } 0 \leq \int_A f|_A \, d\mu_A < \varepsilon/3, \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu < \varepsilon/3, \text{ and } \int_X f \, d\mu - \int_X f_n \, d\mu < \varepsilon/3. \) Then, 

\[
0 \leq \int_A f_n|_A \, d\mu_A = \int_A (f_n - \bar{f}_n)|_A \, d\mu_A + \int_A \bar{f}_n|_A \, d\mu_A \\
\leq \int_A (f_n - \bar{f}_n)|_A \, d\mu_A + \int_A f|_A \, d\mu_A < \int_X (f_n - \bar{f}_n) \, d\mu + \epsilon/3 \\
= \int_X f_n \, d\mu - \int_X \bar{f}_n \, d\mu + \epsilon/3 \\
\leq \epsilon/3 + \int_X f \, d\mu - \int_X \bar{f}_n \, d\mu \leq \epsilon
\]
11.5. General Convergence Theorems

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 $\int_X f_n \, d\mu < +\infty$. This completes the proof of the lemma. □

Lemma 11.86 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite measure space, $Y$ be a Banach space, $\mathcal{W}$ be a separable subspace of $Y$, $\phi_n : X \to \mathcal{W}$ 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. Assume that

1. $\lim_{n \to \infty} \phi_n = f$ a.e. in $\mathcal{X}$ and $\lim_{n \to \infty} g_n = g$ a.e. in $\mathcal{X}$;
2. $g_n$'s and $g$ are integrable over $X$ and $\lim_{n \to \infty} \int_X g_n \, d\mu = \int_X g \, d\mu \in \mathbb{R}$.

Then, $f$ is integrable over $X$ and $\int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu \in \mathbb{R}$.

Proof Let $E := \{ x \in X \mid (\phi_n(x))_{n=1}^\infty$ does not converge to $f(x)$ or $(g_n(x))_{n=1}^\infty$ does not converge to $g(x) \}$. Then, by (i), $E \in \mathcal{B}$ and $\mu(E) = 0$. Let $(F_r)_{r \in (0, \infty)}$ be the net for $\int_X f \, d\mu$ as defined in Definition 11.70.

1. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, 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 $0 \leq \int_A g \, d\mu_A < \epsilon/6$ and $0 \leq \int_A g_n \, d\mu_A < \epsilon/6$, where $(A, \mathcal{B}, \mu_A)$ is the finite measure subspace of $X$. By Egoroff's Theorem 11.55, $\exists E_1 \in \mathcal{B}$ with $\mu(E_1) < \delta$ such that $\left( \phi_n \mid_{X \setminus E_1} \right)_{n=1}^\infty$ 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 $\left( \phi_n \mid_{X \setminus E_1} \right)_{n=1}^\infty$ converges uniformly to $f|_{X \setminus E_1}$. Then, $\exists n_1 \in \mathbb{N}$ with $n_1 \geq n_0$, $\forall n \in \mathbb{N}$ with $n \geq n_1$, $\forall x \in X \setminus E_1$, $\| \phi_n(x) - f(x) \| < \epsilon$. Fix any $n \geq n_1$. Let $\hat{\phi}_n$ admit the canonical representation $\hat{\phi}_n = \sum_{i=1}^n w_i \hat{\chi}_{A_i \setminus X}$, where $\hat{n} \in \mathbb{Z}_+$, $w_1, \ldots, w_n \in \mathbb{W}$ are distinct and none equals to $\hat{y}_y$, and $A_1, \ldots, A_n \in \mathcal{B}$ are nonempty, pairwise disjoint, and of finite measure. Let $w_{n+1} := \hat{y}_y$ and $A_{n+1} := X \setminus \left( \bigcup_{i=1}^n A_i \right) \in \mathcal{B}$.

Then, $X = \bigcup_{i=1}^{n+1} A_i$ and the sets in the union are pairwise disjoint and of finite measure. Define $V_i := B_{\hat{y}}(w_i, \epsilon)$, $i = 1, \ldots, n+1$. Define $V_1 := V_1 \in \mathcal{B}(Y)$, $V_{n+1} := V_{n+1} \setminus \left( \bigcup_{j=1}^n V_j \right) \in \mathcal{B}(Y)$, $i = 1, \ldots, n$. Let $V_{n+2} := \overline{Y} \setminus \left( \bigcup_{j=1}^{n+1} V_j \right) \in \mathcal{B}(Y)$. 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 $R := \{ (\tilde{y}_\gamma, \tilde{U}_\gamma) \mid \gamma \in \Gamma \}$.

$\forall \gamma \in \Gamma$, by $\tilde{R} \preceq R$, $\exists i_\gamma \in \{ 1, \ldots, n+2 \}$ such that $\tilde{U}_\gamma \subseteq V_{i_\gamma}$. Define $\Gamma := \{ \gamma \in \Gamma \mid i_\gamma = n+2 \}$. Let $A_{\gamma,i} := A_i \cap f_{\text{inv}}(\tilde{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}$, $i = 1, \ldots, n+1$. Let $A_{\gamma} := f_{\text{inv}}(\tilde{U}_\gamma) \cap (E_1 \setminus E) \in \mathcal{B}$.
Claim 11.86.1 \( \forall \gamma \in \bar{\Gamma}, f_{\text{inv}}(\bar{U}_\gamma) \subseteq E_1; \forall \gamma \in \Gamma_s := \{ \gamma \in \bar{\Gamma} \mid \bar{A}_\gamma \neq \emptyset \}, \| \hat{y}_\gamma \mu(\bar{A}_\gamma) \| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma); \) and \( \left\| \sum_{\gamma \in \bar{\Gamma}} \hat{y}_\gamma \mu(f_{\text{inv}}(\bar{U}_\gamma) \cap E_1) \right\| < \epsilon/3 \).

**Proof of claim:** Fix any \( \gamma \in \bar{\Gamma} \). \( \bar{U}_\gamma \subseteq V_{\bar{n}+2} = \bar{Y} \setminus \bigcup_{j=1}^{\bar{n}+1} \bar{V}_j \). 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 \( f(x) \in \bar{U}_\gamma, \lim_{m \in \mathbb{N}} \phi_m(x) = f(x) \), \( \| \phi_m(x) \| \leq g_m(x), \forall m \in \mathbb{N}, \) 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 \( R \in \mathcal{F}(\bar{Y}), \) \( \| \hat{y}_\gamma \| < \inf_{g \in \mathcal{U}_\gamma} \| g \| + 1 \leq \| f(x) \| + 1 \leq g(x) + 1. \) By the arbitrariness of \( x, \) we have \( \| \hat{y}_\gamma \| \leq \inf_{x \in \bar{A}_\gamma} g(x) + 1. \) Therefore, \( \| \hat{y}_\gamma \mu(\bar{A}_\gamma) \| \leq (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma). \)

Note that
\[
\left\| \sum_{\gamma \in \bar{\Gamma}} \hat{y}_\gamma \mu(f_{\text{inv}}(\bar{U}_\gamma) \cap E_1) \right\| = \left\| \sum_{\gamma \in \bar{\Gamma}} \hat{y}_\gamma (\mu(\bar{A}_\gamma) + \mu(f_{\text{inv}}(\bar{U}_\gamma) \cap E_1 \cap E)) \right\|
= \left\| \sum_{\gamma \in \bar{\Gamma}} \hat{y}_\gamma \mu(\bar{A}_\gamma) \right\| = \left\| \sum_{\gamma \in \bar{\Gamma}_s} \hat{y}_\gamma \mu(\bar{A}_\gamma) \right\| \leq \sum_{\gamma \in \bar{\Gamma}_s} (1 + \inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma) = \sum_{\gamma \in \bar{\Gamma}_s} \mu(\bar{A}_\gamma) + \sum_{\gamma \in \bar{\Gamma}_s} \mu(\bar{A}_\gamma)
\]

The simple function \( \psi : X \to [0, \infty) \subseteq \mathbb{R}, \) defined by \( \psi(x) = \sum_{\gamma \in \bar{\Gamma}_s} (\inf_{x \in \bar{A}_\gamma} g(x)) \chi_{\bar{A}_\gamma}(x), \forall x \in X, \) satisfies \( 0 \leq \psi(x) \leq g(x), \forall x \in X \) and \( \psi(x) = 0, \forall x \in \bar{X} \setminus E_1. \) Then, \( \sum_{\gamma \in \bar{\Gamma}_s} (\inf_{x \in \bar{A}_\gamma} g(x)) \mu(\bar{A}_\gamma) \neq \int_X \psi \, d\mu = \int_{E_1 \setminus \bar{V}_1} g_1 \, d\mu E_1 \leq \int_{E_1} g \, d\mu E_1 < \epsilon/6, \) where the equalities follows from Proposition 11.75; and the first inequality follows from Proposition 11.83.

Then, \( \left\| \sum_{\gamma \in \bar{\Gamma}} \hat{y}_\gamma \mu(f_{\text{inv}}(\bar{U}_\gamma) \cap E_1) \right\| < \mu(E_1) + \epsilon/6 < \delta + \epsilon/6 \leq \epsilon/3. \)

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

Claim 11.86.2 \( \left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| < \epsilon/6. \)

**Proof of claim:** Since \( \| \phi_i(x) \| \leq g_n(x), \forall x \in X, \) then \( \| w_i \| \leq g_n(x), \forall x \in A_i, i = 1, \ldots, \bar{n} + 1. \) Then, \( \left\| \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap E_1) \right\| \leq \sum_{i=1}^{\bar{n}+1} \| w_i \| \mu(A_i \cap E_1) \leq \int_{E_1} g_n \, d\mu E_1 < \epsilon/6, \) where the second inequality follows from Definition 11.79. \( \square \)

Claim 11.86.3 \( \forall \gamma \in \bar{\Gamma} \setminus \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}, \) we have \( \| \hat{y}_\gamma - w_i \| \mu(\bar{A}_{\gamma,i}) \leq 3\epsilon \mu(\bar{A}_{\gamma,i}). \)

**Proof of claim:** \( \forall \gamma \in \bar{\Gamma} \setminus \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}, \) \( w_i \in \{1, \ldots, \bar{n} + 1\} \) and \( \bar{U}_\gamma \subseteq V_{\bar{n}+1}. \) The result is trivial if \( A_{\gamma,i} = \emptyset. \) If \( A_{\gamma,i} \neq \emptyset, \) then \( \exists x \in A_{\gamma,i} = A_i \cap \)
11.5. GENERAL CONVERGENCE THEOREMS

That \( \hat{\gamma}, \hat{\mu} \)

\[
\hat{\gamma} \subset \hat{B}_y(w_i, \epsilon) \cap \hat{U}_y \subseteq \hat{B}_y(w_i, \epsilon) \cap V_i \subseteq \hat{B}_y(w_i, \epsilon) \cap \hat{V}_i = \hat{B}_y(w_i, \epsilon) \cap \hat{B}_y(w_i, \epsilon).
\]

Note that \( \hat{B}_y \subseteq \hat{B}_y(w_i, \epsilon) \).

Therefore, we have \( \| \hat{\gamma} - w_i \| \leq \| \hat{\gamma} - w_i, \| + \| w_i - f(x) \| < \epsilon \). Hence, the result holds.

Note that \( F_R = \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(f_{\text{inv}}(\hat{U}_\gamma)) \) and \( \int_X \phi_n \, d\mu = \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \), by Proposition 11.75. Then,

\[
\| F_R - \int_X \phi_n \, d\mu \| = \left\| \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \right\| \\
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1)) + \mu(f_{\text{inv}}(\hat{U}_\gamma)) \right\| \\
- \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i \cap (X \setminus E_1)) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(A_i, i) + \mu(f_{\text{inv}}(\hat{U}_\gamma)) \right\| \\
- \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i, i) \right\|
\]

\[
= \left\| \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(A_i, i) + \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(f_{\text{inv}}(\hat{U}_\gamma)) \right\| \\
- \sum_{i=1}^{\bar{n}+1} w_i \mu(A_i, i) \right\|
\]

\[
< \left\| \sum_{\gamma \in \Gamma} \hat{\gamma}_i \mu(A_i, i) \right\| + \frac{\epsilon}{3} + \frac{\epsilon}{6}
\]

\[
= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{\gamma}_i \mu(A_i, i) + \hat{\gamma}_i \mu(A_i, i)) \right\|
\]

\[
= \frac{\epsilon}{2} + \left\| \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} (\hat{\gamma}_i - w_i) \mu(A_i, i) \right\| < \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} \| \hat{\gamma}_i - w_i \| \mu(A_i, i)
\]

\[
\leq \frac{\epsilon}{2} + \sum_{\gamma \in \Gamma} \sum_{i=1}^{\bar{n}+1} 3\epsilon \mu(A_i, i) \leq \frac{\epsilon}{2} + \frac{3\mu(X)\epsilon}{6\mu(X) + 1} < \epsilon
\]
By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}$, $f$ is integrable over $X$ and $f$ is integrable with $\int_X f \, d\mu = \lim_{n \to \infty} \int_X \varphi_n \, d\mu \in \mathcal{Y}$ and $f$ is integrable over $X$. This completes the proof of the lemma.

**Lemma 11.87** Let $X := (X, \mathcal{B}, \mu)$ be a finite measure space, $Y$ be a Banach space, $W$ be a separable subspace of $Y$, $U \subseteq W$ be a conic segment, $f : X \to U \subseteq W$ be $\mathcal{B}$-measurable, and $g : X \to [0, \infty) \subseteq \mathbb{R}$ 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 $\lim_{k \to \infty} \varphi_k = f$ a.e. in $X$, $\|\varphi_k(x)\| \leq \|f(x)\|$, $\forall x \in X$, $\forall k \in \mathbb{N}$, and $\int_X \varphi_k \, d\mu \in \mathcal{Y}$.

Furthermore, if $Y$ admits a positive cone $P$ and $U \subseteq P \subseteq Y$, 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 $\lim_{k \to \infty} \varphi_k = f$ a.e. in $X$. By Lemma 11.86, $f$ is integrable over $X$ and $\int_X \varphi_k \, d\mu = \lim_{k \to \infty} \int_X \varphi_k \, d\mu \in \mathcal{Y}$.

Furthermore, if $Y$ admits a positive cone $P$ and $U \subseteq P \subseteq Y$, then, by Propositions 8.41 and 11.75, $\varphi_k : X \to U \subseteq P \subseteq Y$, $\forall k \in \mathbb{N}$, and $\int_X \varphi_k \, d\mu \in P$. Then, by Proposition 4.13, $\int_X f \, d\mu \in \mathcal{P} = P$, since $P$ is closed. This completes the proof of the lemma.

**Theorem 11.88 (Lebesgue Dominated Convergence)** Let $X := (X, \mathcal{B}, \mu)$ be a finite 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

(i) $\lim_{n \to \infty} f_n = f$ a.e. in $X$, $\lim_{n \to \infty} g_n = g$ a.e. in $X$;

(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 $\int_X g \, d\mu = \lim_{n \to \infty} \int_X g_n \, d\mu \in \mathbb{R}$.

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** $\forall n \in \mathbb{N}$, by Lemma 11.87, there exists a sequence of simple functions $(\varphi_{n,k})_{k=1}^{\infty}$, $\varphi_{n,k} : X \to W$, $\forall k \in \mathbb{N}$, such that $\lim_{k \to \infty} \varphi_{n,k} = f_n$ a.e.
11.5. GENERAL CONVERGENCE THEOREMS

Let \( X := (X, \mathcal{B}, \mu) \) be a finite measure space, \( Y \) be a Banach space over \( K \), \( W \) be a separable subspace of \( Y \), \( Z \) be a Banach space over \( K \), \( f_i : X \to W \) be \( \mathcal{B} \)-measurable, \( g_i : X \to [0, \infty) \subset \mathbb{R} \) be integrable over \( X \), \( i = 1, 2 \). 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 \( X \) and \( \int_X (f_1 + f_2) \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathbb{Y} \).

(ii) \( \forall A \in \mathcal{B}(Y,W), \ A f_1 \) is integrable over \( X \) and \( \int_X (A f_1) \, d\mu = A \int_X f_1 \, d\mu \in \mathbb{Z} \).

(iii) \( \forall c \in K, \ c f_1 \) is integrable over \( X \) and \( \int_X (c f_1) \, d\mu = c \int_X f_1 \, d\mu \in \mathbb{Y} \).

(iv) \( \forall E \in \mathcal{B}, \ f_1|_E \) is integrable over \( E \) and \( \int_X (f_1 1_E) \, d\mu = \int_E f_E \, d\mu \in \mathbb{Y} \), where \( E := (E, \mathcal{B}_E, \mu_E) \) is the finite measure subspace of \( X \) 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 \( X \) then \( \int_X f_1 \, d\mu = \int_X f_2 \, d\mu \in \mathbb{Y} \).

(vi) \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subset \mathcal{B}, \ \sum_{i=1}^\infty \int_{E_i} f_1 \, d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 \, d\mu \in \mathbb{Y} \).

(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 380.
If $\mathcal{Y}$ 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 $\int_X f_i \, d\mu = \lim_{n \to \infty} \int_X \phi_{i,n} \, d\mu \in \mathbb{Y}$.

(i) By Propositions 7.23, 11.38, and 11.39, $f_1 + f_2, g_1 + g_2$, and $\phi_{1,n} + \phi_{2,n}$ 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 $\int_X (f_1 + f_2) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{1,n} + \phi_{2,n}) \, d\mu = \lim_{n \to \infty} (\int_X \phi_{1,n} \, d\mu + \int_X \phi_{2,n} \, d\mu) = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu \in \mathbb{Y}$, where the second equality follows from Proposition 11.75, and the third equality follows from Propositions 3.66, 3.67, and 7.23.

(ii) $\forall A \in \mathcal{B}(\mathbb{Y}, \mathcal{Z})$, 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 $\mathcal{W} := \{z \in \mathcal{Z} \mid z = A w, w \in \mathcal{W}\}$ is a separable subspace of $\mathcal{Z}$. 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) $(\phi_{1,n})_{n=1}^{\infty}$ is a sequence of simple functions that converges to $f_1|_E$ a.e. in $\mathcal{E}$. Note that $\|\phi_{1,n}|_E(x)\| \leq g_1|_E(x), \forall x \in E, \forall n \in \mathbb{N}$. By Proposition 11.83, $0 \leq \int_E g_1|_E \, d\mu_E \leq \int_X g_1 \, d\mu < +\infty$. By Lebesgue Dominated Convergence Theorem 11.88, we have $\int_E f_1|_E \, d\mu_E = \lim_{n \to \infty} \int_E \phi_{1,n}|_E \, d\mu_E \in \mathbb{Y}$. Note that $(\phi_{1,n})_{n=1}^{\infty}$ is a sequence of simple functions that converges to $f_1|_E \chi_{E,X}$ a.e. in $\mathcal{X}$, and $\|\phi_{1,n}(x)\chi_{E,X}(x)\| \leq g_1(x), \forall x \in X, \forall n \in \mathbb{N}$. By Propositions 7.23, 11.38, and 11.39, $\phi_{1,n}\chi_{E,X}$ and $f_1\chi_{E,X}$ are $\mathcal{B}$-measurable. Again by Lebesgue Dominated Convergence Theorem 11.88, we have $\int_X (f_1 \chi_{E,X}) \, d\mu = \lim_{n \to \infty} \int_X (\phi_{1,n} \chi_{E,X}) \, d\mu = \lim_{n \to \infty} \int_X \phi_{1,n}|_E \, d\mu_E = \int_X f_1|_E \, d\mu_E \in \mathbb{Y}$, 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 $\int_X f_2 \, d\mu = \lim_{n \to \infty} \int_X \phi_{1,n} \, d\mu = \int_X f_1 \, d\mu$. 

404  \textit{CHAPTER 11. GENERAL MEASURE AND INTEGRATION}
11.5. General Convergence Theorems

(vi) \( \forall \) 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, \( \lim_{n \to \infty} f_1(x)\chi_{\tilde{E}_n,X}(x) = f_1(x)\chi_{X,X}(x), \forall x \in X \). Then, we have \( \sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \int_{\tilde{E}_n} f_1 \, d\mu = \int_{X} (f_1 \chi_{X,X}) \, d\mu = \int_{X} f_1 \, d\mu \in \mathbb{R} \), where the first equality follows from (iv); the second equality follows from (i); the third equality follows from Proposition 11.83; and the third inequality follows from Lebesgue Dominated Convergence Theorem 11.88; the last equality follows from (iv); and the last step follows from (iv).

(vii) Note that, by Proposition 11.75, \( \| \int_X \phi_{1,n} \, d\mu \| \leq \int_X \mathcal{P} \circ \phi_{1,n} \, d\mu \in \mathbb{R}, \forall n \in \mathbb{N} \). By Propositions 7.21 and 11.38, \( \mathcal{P} \circ \phi_{1,n} \) and \( \mathcal{P} \circ f_1 \) are \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). By Propositions 7.21 and 11.52, we have \( \lim_{n \to \infty} \mathcal{P} \circ \phi_{1,n} = \mathcal{P} \circ f_1 \) a.e. in \( X \). Note that \( 0 \leq \mathcal{P} \circ \phi_{1,n}(x) \leq g_1(x), \forall x \in X, \forall n \in \mathbb{N} \). Then, \( 0 \leq \| \int_X f_1 \, d\mu \| = \| \lim_{n \to \infty} \int_X \phi_{1,n} \, d\mu \| = \lim_{n \to \infty} \| \int_X \phi_{1,n} \, d\mu \| \leq \lim_{n \to \infty} \int_X \mathcal{P} \circ \phi_{1,n} \, d\mu = \int_X \mathcal{P} \circ f_1 \, d\mu \leq \int_X g_1 \, d\mu < +\infty \), where the second equality follows from Propositions 7.21 and 3.66; the third equality follows from Lebesgue Dominated Convergence Theorem 11.88; and the third inequality follows from Proposition 11.83.

(viii) Let \( f := f_2 - f_1 \). Then, \( f \equiv \emptyset \) a.e. in \( X \), \( f : X \to \mathcal{W} \), and \( f \) is \( \mathcal{B} \)-measurable by Propositions 11.38, 11.39, and 7.23. Note that \( P \) is a closed set, \( P \in \mathcal{B}_0(\mathcal{W}) \), and \( f_{\text{inv}}(P) \in \mathcal{B} \). Define \( \bar{f} : X \to \mathcal{W} \cap \emptyset \) by \( \bar{f}(x) = \begin{cases} f(x) & x \in f_{\text{inv}}(P) \\ \emptyset & x \in X \setminus f_{\text{inv}}(P) \end{cases} \). Then, \( f = \bar{f} \) a.e. in \( X \), \( \bar{f} : X \to \mathcal{W} \cap \emptyset \), and \( \| \bar{f}(x) \| \leq \| f(x) \| \leq \| f_1(x) \| + \| f_2(x) \| \leq g_1(x) + g_2(x), \forall x \in X \). By Proposition 11.41, \( \bar{f} \) is \( \mathcal{B} \)-measurable. By Proposition 11.83, \( g_1 + g_2 \) is integrable over \( X \). By Lemma 11.87, \( \int_X \bar{f} \, d\mu \in \mathbb{P} \). By (i) and (v), we have \( \int_X f_2 \, d\mu - \int_X f_1 \, d\mu = \int_X f \, d\mu = \int_X \bar{f} \, d\mu \equiv \emptyset \). Then, \( \int_X f_1 \, d\mu \leq \int_X f_2 \, d\mu \).

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

Definition 11.90 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} \) be a normed linear space, and \( f : X \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable. \( f \) is said to be absolutely integrable if \( \mathcal{P} \circ f \) is integrable over \( X \), that is, \( 0 \leq \int_X \mathcal{P} \circ f \, d\mu < +\infty \).

Note that \( \mathcal{P} \circ f \) is \( \mathcal{B} \)-measurable by Propositions 7.21 and 11.38. We will show that the desired properties of integration are valid if the integrand is absolutely integrable.

Theorem 11.91 (Lebesgue Dominated Convergence) Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y} \) be a Banach space, \( \mathcal{W} \) be a separable subspace of \( \mathcal{Y} \), \( f_n : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), \( f : X \to \mathcal{W} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( g_n : X \to [0, \infty) \subset \mathbb{R} \) 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) \( \lim_{n \to \infty} f_n = f \) a.e. in \( X \), \( \lim_{n \to \infty} g_n = g \) a.e. in \( X \);

(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 \( \int_X g \, d\mu = \lim_{n \to \infty} \int_X g_n \, d\mu \in \mathbb{R} \).
Then, \( f \) is integrable over \( X \) and \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in Y \). Furthermore, \( f \) is absolutely integrable and \( 0 \leq \int_X (P \circ f) \, d\mu = \lim_{n \to \infty} \int_X (P \circ f_n) \, d\mu \leq \int_X g \, d\mu < \infty \).

**Proof** We will show that \( f \) is integrable over \( X \) and \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu \in Y \) 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 \((F_A)_{A \in \mathcal{M}(X)}\) and \((F_{n,A})_{A \in \mathcal{M}(X)}\) be the nets for the functions \( f \) and \( f_n \) as defined in Definition 11.71, respectively, \( \forall n \in \mathbb{N} \). \( \forall A \in \mathcal{M}(X) \), we have \( A \in \mathcal{B} \) and \( \mu(A) < \infty \). Let \( \mathcal{A} := (A, \mathcal{B}, \mu_A) \) be the finite measure subspace of \( X \) as defined in Proposition 11.13. By Proposition 11.83, \( g|_A \) and \( g_n|_A \)'s are integrable over \( A \), \( \lim_{n \to \infty} g_n|_A = g|_A \) a.e. in \( A \), and \( \int_A g|_A \, d\mu_A = \lim_{n \to \infty} \int_A g_n|_A \, d\mu_A \). By Lebesgue Dominated Convergence Theorem 11.88, we have \( f|_A \) and \( f_n|_A \)'s are integrable over \( A \) and \( \int_A f|_A \, d\mu_A = \lim_{n \to \infty} \int_A f_n|_A \, d\mu_A \in Y \). Note that \( F_A = \int_A f|_A \, d\mu_A \in Y \) and \( F_{n,A} = \int_A f_n|_A \, d\mu_A \in Y \), \( \forall n \in \mathbb{N} \). Then, the nets \((F_A)_{A \in \mathcal{M}(X)}\) and \((F_{n,A})_{A \in \mathcal{M}(X)}\)'s are well defined.

\( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), 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 \( \int_X g \, d\mu - \epsilon/5 < \int_X \phi \, d\mu \leq \int_X g \, d\mu < +\infty \). Let \( A_0 := \{ x \in X \mid \phi(x) > 0 \} \). Then, \( A_0 \in \mathcal{B} \), \( \mu(A_0) < +\infty \), and \( A_0 \in \mathcal{M}(X) \). \( \forall A \in \mathcal{M}(X) \) with \( A_0 \subseteq A \), we have \( 0 \leq \int_{A \setminus A_0} g|_{A \setminus A_0} \, d\mu_{A \setminus A_0} = \int_A g|_A \, d\mu_A - \int_{A_0} g|_{A_0} \, d\mu_{A_0} \leq \int_X g \, d\mu - \int_{A_0} \phi|_{A_0} \, d\mu_{A_0} = \int_X g \, d\mu - \int_X \phi \, d\mu < \epsilon/5 \), where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that \( \| F_A - F_{A_0} \| \) is well defined.

\( \forall n \in \mathbb{N} \), 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 \( \int_X g_n \, d\mu - \epsilon/5 < \int_X \phi_n \, d\mu \leq \int_X g_n \, d\mu < +\infty \). Let \( A_n := \{ x \in X \mid \phi_n(x) > 0 \} \). Then, \( A_n \in \mathcal{B} \), \( \mu(A_n) < +\infty \), and \( A_n \in \mathcal{M}(X) \). \( \forall A \in \mathcal{M}(X) \) with \( A_n \subseteq A \), we have \( 0 \leq \int_{A \setminus A_n} g|_{A \setminus A_n} \, d\mu_{A \setminus A_n} = \int_A g_n|_A \, d\mu_A - \int_{A_n} g_n|_{A_n} \, d\mu_{A_n} \leq \int_X g_n \, d\mu - \int_{A_n} \phi_n|_{A_n} \, d\mu_{A_n} = \int_X g_n \, d\mu - \int_X \phi_n \, d\mu < \epsilon/5 \), where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that \( \| F_{n,A} - F_{n,A_n} \| = \int_X (P \circ f_n) \, d\mu = \lim_{n \to \infty} \int_X (P \circ f_n) \, d\mu \).
11.5. GENERAL CONVERGENCE THEOREMS

\[ \left\| \int_{A} f_{n}\, d\mu_{A} - \int_{A_{n}} f_{n}\, d\mu_{A_{n}} \right\| = \left\| \int_{A_{n}\setminus A} f_{n}\, d\mu_{A_{n}} \right\| \leq \int_{A_{n}\setminus A} g_{n}\, d\mu_{A_{n}} < \epsilon/5, \text{where the second equality and the first inequality follow from Proposition 11.89.} \]

Then, the net \((F_{n,A})_{A\in\mathcal{M}(X)}\) is a Cauchy net, which admits a limit, by Proposition 4.44, \(\int_{X} f_{n}\, d\mu = \lim_{A\in\mathcal{M}(X)} F_{n,A} \in \mathcal{Y}\).

Let \(\hat{g}_{n} := g_{n} \wedge g, \forall n \in \mathcal{N}\). Then, \(0 \leq \hat{g}_{n}(x) \leq g(x), \forall x \in X, \) and \(\hat{g}_{n}\) is \(\mathcal{B}\)-measurable by Proposition 11.40, \(\forall n \in \mathcal{N}\). Then, \(\lim_{n\in\mathcal{N}} \hat{g}_{n} = g\) a.e. in \(X\) by Propositions 11.52 and 11.53. By Fatou’s Lemma 11.80, \(\int_{X} g\, d\mu \leq \liminf_{n\in\mathcal{N}} \int_{X} \hat{g}_{n}\, d\mu\). By Proposition 11.83, \(0 \leq \int_{X} \hat{g}_{n}\, d\mu \leq \int_{X} g\, d\mu < \infty, \forall n \in \mathcal{N}\). Then, we have \(\int_{X} g\, d\mu \leq \liminf_{n\in\mathcal{N}} \int_{X} \hat{g}_{n}\, d\mu \leq \limsup_{n\in\mathcal{N}} \int_{X} \hat{g}_{n}\, d\mu \leq \int_{X} g\, d\mu\). Therefore, by Proposition 3.83, \(\int_{X} g\, d\mu = \lim_{n\in\mathcal{N}} \int_{X} \hat{g}_{n}\, d\mu \in \mathcal{R}\). This coupled with (iii) and the fact that \(F_{A_{0}} = \lim_{n\in\mathcal{N}} F_{n,A_{0}}\), implies that \(\exists \hat{g}_{0} \in \mathcal{N}, \forall n \in \mathcal{N}\) with \(n \geq n_{0}\), we have \(0 \leq \int_{X} g\, d\mu - \int_{X} \hat{g}_{0}\, d\mu \leq \epsilon/5, \forall A \in \mathcal{M}(X)\) with \(A_{0} \subseteq A\), we have \(0 \leq \int_{A\setminus A_{0}} (g_{n} - \hat{g}_{n})\, d\mu_{A_{0}} + \int_{A\setminus A_{0}} \hat{g}_{n}\, d\mu_{A_{0}} \leq \int_{X} g\, d\mu - \int_{X} \hat{g}_{0}\, d\mu + \int_{X} \hat{g}_{0}\, d\mu - \int_{X} \hat{g}_{n}\, d\mu + \epsilon/5 \leq \int_{X} g\, d\mu - \int_{X} \hat{g}_{n}\, d\mu + \epsilon/5 \leq 3\epsilon/5, \text{where the first equality, the second inequality, and the third inequality follow from Proposition 11.83. Furthermore,}\)

\[ \|F_{n,A} - F_{n,A_{0}}\| = \left\| \int_{A} f_{n}\, d\mu_{A} - \int_{A_{0}} f_{n}\, d\mu_{A_{0}} \right\| = \left\| \int_{A\setminus A_{0}} f_{n}\, d\mu_{A_{0}} \right\| \leq \int_{A\setminus A_{0}} (g_{n} - \hat{g}_{n})\, d\mu_{A_{0}} + \int_{A\setminus A_{0}} \hat{g}_{n}\, d\mu_{A_{0}} < \epsilon/5, \text{where the second equality and the first inequality follow from Proposition 11.89. This implies that}\)

\[ \left\| \int_{X} f\, d\mu - \int_{X} f_{n}\, d\mu \right\| \leq \left\| \int_{X} f\, d\mu - F_{A_{0}} \right\| + \left\| F_{A_{0}} - F_{n,A_{0}} \right\| + \left\| F_{n,A_{0}} - F_{n,A} \right\| + \left\| F_{n,A} - \int_{X} f\, d\mu \right\| < \lim_{A\in\mathcal{M}(X)} \|F_{A} - F_{A_{0}}\| + \epsilon/5 + \lim_{A\in\mathcal{M}(X)} \|F_{n,A_{0}} - F_{n,A}\| + \epsilon/5 = \epsilon, \text{where the second inequality follows from Propositions 7.21 and 3.66. Hence, } \int_{X} f\, d\mu = \lim_{n\in\mathcal{N}} \int_{X} f_{n}\, d\mu \in \mathcal{Y}.

Then, in both cases, we have shown that \(f\) is integrable over \(X\) and \(\int_{X} f\, d\mu = \lim_{n\in\mathcal{N}} \int_{X} f_{n}\, d\mu \in \mathcal{Y}\).

Note that \(0 \leq \mathcal{P} \circ f_{n}(x) \leq g_{n}(x), \forall x \in X, \forall n \in \mathcal{N}\), \(f_{n} = \mathcal{P} \circ f\) a.e. in \(X\) by Propositions 7.21 and 11.52, and \(\mathcal{P} \circ f\) and \(\mathcal{P} \circ f_{n}\)’s are \(\mathcal{B}\)-measurable by Propositions 7.21 and 11.38. Then, by what we have proved so far, \(\int_{X} \mathcal{P} \circ f\, d\mu = \lim_{n\in\mathcal{N}} \int_{X} \mathcal{P} \circ f_{n}\, d\mu \leq \lim_{n\in\mathcal{N}} \int_{X} g_{n}\, d\mu = \int_{X} g\, d\mu < \infty\). Hence, \(f\) is absolutely integrable.

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

**Proposition 11.92** Let \(X := (X, \mathcal{B}, \mu)\) 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 \(\mathbb{K}\), \(f_{i} : X \to \mathcal{W}\) be absolutely integrable over \(X\), \(i = 1, 2\). Then, the following statements hold.

(i) \(f_{i}\) is integrable over \(X\) and \(\int_{X} f_{i}\, d\mu \in \mathcal{Y}, i = 1, 2\).

(ii) \(f_{1} + f_{2}\) is absolutely integrable over \(X\) and \(\int_{X} (f_{1} + f_{2})\, d\mu = \int_{X} f_{1}\, d\mu + \int_{X} f_{2}\, d\mu \in \mathcal{Y}\).
The results follow immediately from Propositions 11.89 and 11.83 and their
\[ c f_1 \text{ is absolutely integrable over } X \] and \( f \in \mathbb{B} \)
We will distinguish two exhaustive and mutually exclusive
Propositions 7.21 and 11.38.

\[ \{X, f_1 \} \subset \mathbb{B}, \sum_{i=1}^{\infty} f_1 d\mu = \int_X f_1 d\mu < \infty. \]

\( f \text{ is absolutely integrable over } X \) as defined in Proposition 11.13. We will henceforth denote
\( f \in \mathbb{B} \), \( f \in \mathbb{B} \).

If \( f_1 = f_2 \) a.e. in \( X \) then \( f_1 = f_2 \).

\[ \forall \mu \in \mathcal{M}(X) \text{ admits a positive cone } P \text{ and } f_1 \leq f_2 \text{ a.e. in } X, \text{ then } f_1 d\mu \leq f_2 d\mu. \]

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

\[ \forall \mu \in \mathcal{M}(X) \text{ admits a positive cone } P \text{ and } f_1 \leq f_2 \text{ a.e. in } X, \text{ then } f_1 d\mu \leq f_2 d\mu. \]

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

\[ \forall \mu \in \mathcal{M}(X) \text{ admits a positive cone } P \text{ and } f_1 \leq f_2 \text{ a.e. in } X, \text{ then } f_1 d\mu \leq f_2 d\mu. \]

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

\[ \forall \mu \in \mathcal{M}(X) \text{ admits a positive cone } P \text{ and } f_1 \leq f_2 \text{ a.e. in } X, \text{ then } f_1 d\mu \leq f_2 d\mu. \]

**Proof** We will distinguish two exhaustive and mutually exclusive
cases: Case 1: \( \mu(X) < +\infty \); Case 2: \( \mu(X) = +\infty \). Case 1: \( \mu(X) < +\infty \). The results follow immediately from Propositions 11.89 and 11.83 and their

\[ \forall \mu \in \mathcal{M}(X) \text{ admits a positive cone } P \text{ and } f_1 \leq f_2 \text{ a.e. in } X, \text{ then } f_1 d\mu \leq f_2 d\mu. \]
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 $X$. Let $(\tilde{F}_E)_{E \in \mathcal{M}(X)}$ be the net for $\int_X (Af_1) \, d\mu$ as defined in Definition 11.71. $\forall E \in \mathcal{M}(X)$, $\tilde{F}_E = \int_E (Af_1) \, d\mu_E = A \int_E f_1 \, d\mu_E = AF_{1,E}$, where the second equality follows from Proposition 11.89. Then, by Propositions 3.66 and 7.62, we have $\int_X (Af_1) \, d\mu = \lim_{E \in \mathcal{M}(X)} \tilde{F}_E = \lim_{E \in \mathcal{M}(X)} AF_{1,E} = A \int_X f_1 \, d\mu \in \mathbb{Z}$. Hence, $Af_1$ is integrable over $X$.

(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, $f_1|_H$ is absolutely integrable over $\mathcal{H}$. Then, by (i), $f_1|_H$ is integrable over $\mathcal{H}$. Again, by (i), $f_1|_{\mathcal{H},X}$ is integrable over $X$. $\forall t \in (0, \infty) \subseteq \mathbb{R}$, $\exists A_0 \in \mathcal{M}(X)$ such that $\forall A \in \mathcal{M}(X)$ with $A_0 \subseteq A$, we have $\|\int_A (f_1|_{\mathcal{H},X}) \, d\mu_A - \int_X (f_1|_{\mathcal{H},X}) \, d\mu_X\| < \epsilon/2$. 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 $\|\int_E f_1 \cdot \chi_{E_0} \, d\mu_E - \int_H f_1 \cdot \chi_{E_0} \, d\mu_H\| < \epsilon/2$. Let $A_1 := A_0 \cup E_0 \in \mathcal{M}(X)$ and $E_1 := A_1 \cap H \in \mathcal{M}(\mathcal{H})$. By Proposition 11.89, $\int_{A_1} (f_1 \cdot \chi_{E_1}) \, d\mu_{A_1} = \int_{H \cap A_1} (f_1 \cdot \chi_{E_1}) \, d\mu_{H \cap A_1} = \int_{E_1} f_1 \cdot \chi_{E_1} \, d\mu_{E_1}$. Then, we have $\|\int_{E_1} (f_1 \cdot \chi_{E_1}) \, d\mu_E - \int_{H \cap E_1} f_1 \cdot \chi_{E_1} \, d\mu_H\| \leq \|\int_X (f_1 \cdot \chi_{E_1}) \, d\mu_X\| < \epsilon/2 + \epsilon/2 = \epsilon$.

By the arbitrarienss of $\epsilon$, we have $\int_X (f_1 \cdot \chi_{E_1}) \, d\mu = \int_H f_1 \, d\mu_H \in \mathbb{R}$.

Case 2b: $\mu(H) < +\infty$. Let $A_1 := A_0 \cup H \in \mathcal{M}(X)$. By Proposition 11.89, $\int_{A_1} (f_1 \cdot \chi_{E_1}) \, d\mu_{A_1} = \int_{H \cap A_1} f_1 \cdot \chi_{E_1} \, d\mu_{H \cap A_1}$. Then, we have $\|\int_X (f_1 \cdot \chi_{E_1}) \, d\mu - \int_H f_1 \, d\mu_H\| = \|\int_X (f_1 \cdot \chi_{E_1}) \, d\mu_X\| < \epsilon/2$. By the arbitrarienss of $\epsilon$, we have $\int_X (f_1 \cdot \chi_{E_1}) \, d\mu = \int_H f_1 \, d\mu_H \in \mathbb{R}$.

Hence, in both subcases, $\int_X (f_1 \cdot \chi_{E_1}) \, d\mu = \int_H f_1 \, d\mu_H \in \mathbb{R}$.

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.91.

(vii) $\forall$ pairwise disjoint $(E_i)_{i=1}^n \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^n E_i \in \mathcal{B}$ and $E_n := \bigcup_{i=1}^n E_i \in \mathcal{B}$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \to \infty} f_1(x) = f_1(x)$, $\forall x \in X$. Then, we have $\sum_{i=1}^n \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \sum_{i=1}^n \int_{E_i} f_1 \, d\mu = \int_X (f_1 \cdot \chi_{E_n}) \, d\mu_X = \int_X (f_1 \cdot \chi_{E_i}) \, d\mu = \int_X f_1 \, d\mu \in \mathbb{R}$, where the first equality follows from (v); the second equality follows from (ii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.91; the last equality follows from (v); and the last inclusion follows from (v).

(viii) By (i), $\mathcal{P} \circ f_1$ is integrable over $X$. $\forall E \in \mathcal{M}(X)$, by Proposition 11.89, $\|F_{1,E}\| \leq \int_E \mathcal{P} \circ f_1 \, d\mu \leq \int_X \mathcal{P} \circ f_1 \, d\mu < +\infty$, where the second inequality follows from Proposition 11.83. Then, $0 \leq \|\int_X f_1 \, d\mu\| = \lim_{E \in \mathcal{M}(X)} \|F_{1,E}\| \leq \int_X \mathcal{P} \circ f_1 \, d\mu < +\infty$, where the equality follows from Propositions 3.66 and 7.21.
(ix) \( \forall E \in \mathcal{M}(\mathcal{X}) \), by Proposition 11.89 and (v), we have \( \int_E (f_2 - f_1)|E| \cdot d\mu \leq \int_E f_2|E| \cdot d\mu - \int_E f_1|E| \cdot d\mu \in \mathbb{P} \). Note that, by (i) and (iv), \( \int_X f_2 \cdot d\mu \leq \int_X f_1 \cdot d\mu = \int_X (f_2 - f_1) \cdot d\mu = \lim_{E \uparrow \mathbb{P}(\mathcal{X})} \int_E (f_2 - f_1)|E| \cdot d\mu \in \mathbb{P} \), where the second to last step follows from Proposition 3.68; and the last step follows from the fact that \( \mathbb{P} \) is closed. Therefore, we have \( \int_X f_1 \cdot d\mu \leq \int_X f_2 \cdot d\mu \).

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

**Proposition 11.93** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( \mathcal{Y}_i \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W}_i \subseteq \mathcal{Y}_i \) be a separable subspace, \( f_i : X \to \mathcal{W}_i \), \( i = 1, 2 \), and \( h : X \to \mathcal{W}_1 \times \mathcal{W}_2 \) 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 \( \int_X h \cdot d\mu = (\int_X f_1 \cdot d\mu, \int_X f_2 \cdot d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \).

**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, \( \mathbb{P} \circ h \) is \( \mathcal{B} \)-measurable. \( \forall x \in X \), \( \mathbb{P} \circ h(x) = \| (f_1(x), f_2(x)) \| = (\| f_1(x) \|^2 + \| f_2(x) \|^2)^{1/2} \leq \mathbb{P} \circ f_1(x) + \mathbb{P} \circ f_2(x) \). By Proposition 11.83, \( 0 \leq \int_X (\mathbb{P} \circ h) \cdot d\mu \leq \int_X (\mathbb{P} \circ f_1 + \mathbb{P} \circ f_2) \cdot d\mu = \int_X f_1 \cdot d\mu + \int_X f_2 \cdot d\mu \leq \int_X (\mathbb{P} \circ f_1 \cdot d\mu + \mathbb{P} \circ f_2 \cdot d\mu) \leq \mathbb{P} \circ h \cdot d\mu < \infty \). Hence, \( h \) is absolutely integrable over \( \mathcal{X} \). By Propositions 7.31 and 7.22, \( \mathcal{Y}_1 \times \mathcal{Y}_2 \) is a Banach space over \( \mathbb{K} \). By Propositions 7.22, 3.28, and 4.4, \( \mathcal{W}_1 \times \mathcal{W}_2 \subseteq \mathcal{Y}_1 \times \mathcal{Y}_2 \) is a separable subspace. By Proposition 11.92, \( h \), \( f_1 \), and \( f_2 \) are integrable over \( \mathcal{X} \) and \( \int_X h \cdot d\mu \in \mathcal{Y}_1 \times \mathcal{Y}_2 \).

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, \( \mathbb{P} \circ f_i \) is \( \mathcal{B} \)-measurable. \( \forall x \in X \), \( \mathbb{P} \circ f_i(x) \leq \mathbb{P} \circ h(x) \). Then, by Proposition 11.83, \( 0 \leq \int_X (\mathbb{P} \circ f_i) \cdot d\mu \leq \int_X (\mathbb{P} \circ h) \cdot d\mu < \infty \). Hence, \( f_i \) is absolutely integrable over \( \mathcal{X} \). By Proposition 11.92, \( f_i \) is integrable over \( \mathcal{X} \) and \( \int_X f_i \cdot d\mu \in \mathcal{Y}_i \).

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 \( \int_X h \cdot d\mu = (\int_X f_1 \cdot d\mu, \int_X f_2 \cdot d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \) 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 \( (\varphi_{i,n})_{n=1}^{\infty} \) such that \( \lim_{n \to \infty} \varphi_{i,n} = f_i \) a.e. in \( X \), \( \| \varphi_{i,n}(x) \| \leq \| f_i(x) \| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), and \( \int_X f_i \cdot d\mu = \lim_{n \to \infty} \int_X \varphi_{i,n} \cdot d\mu \in \mathcal{Y}_i \). By Proposition 11.53, we have \( \lim_{n \to \infty} \psi_n = h \) a.e. in \( X \), where \( \psi_n : X \to \mathcal{W}_1 \times \mathcal{W}_2 \) 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, \( \int_X h \cdot d\mu = \lim_{n \to \infty} \int_X \psi_n \cdot d\mu = \lim_{n \to \infty} (\int_X \varphi_{1,n} \cdot d\mu, \int_X \varphi_{2,n} \cdot d\mu) = (\int_X f_1 \cdot d\mu, \int_X f_2 \cdot d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \), 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 $(F_i)_{E \in M(X)}$ be the nets for $\int_X h \, d\mu$ and $\int_X f_i \, d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2$, $\forall E \in M(\mathcal{X})$, by Case 1,

$$H_E = \int_E h \mid_E \, d\mu_E = (\int_E f_1 \mid_E \, d\mu_E, \int_E f_2 \mid_E \, d\mu_E) = (F_1, F_2, E)$$

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 $\int_X h \, d\mu = \lim_{E \in M(X)} H_E = \lim_{E \in M(X)} (F_1, F_2, E) = (\int_X f_1 \, d\mu, \int_X f_2 \, d\mu) \in \mathcal{Y}_1 \times \mathcal{Y}_2$, 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 $\mathcal{X}$, $\lim_{n \in \mathbb{N}} \int_X \varphi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \varphi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in [0, \infty) \subseteq \mathbb{R}$, 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 $\lim_{n \in \mathbb{N}} \int_X (\mathcal{P} \circ \varphi_n) \, d\mu = \int_X (\mathcal{P} \circ f) \, d\mu \in [0, \infty) \subseteq \mathbb{R}$. 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 $\|\phi_n(x)\| \leq \mathcal{P} \circ f(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, $\lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu = \int_X f \, d\mu \in \mathcal{Y}$, $\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in [0, \infty) \subseteq \mathbb{R}$, 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 \( \int_X P \circ f \, d\mu < +\infty \), then, by Definition 11.71, \( \forall n \in \mathbb{N} \), \( \exists E_n \in \mathcal{B} \) with \( \mu(E_n) < +\infty \) such that \( \int_X P \circ f \, d\mu - 2^{-n-1} < \int_{E_n} P \circ f \, d\mu \leq \int_X P \circ f \, d\mu \). By Case 1, \( \exists \) a simple function \( \tilde{\phi}_n : E_n \rightarrow U \) such that \( \| \tilde{\phi}_n(x) \| \leq P \circ f(x) \), \( \forall x \in E_n \), \[
 \left\| \int_{E_n} \tilde{\phi}_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu \right\| < 2^{-n-1}, \quad \left\| \int_{E_n} P \circ \tilde{\phi}_n \, d\mu_{E_n} - \int_{E_n} P \circ f \, d\mu \right\| < 2^{-n-1}, \quad \text{and} \quad 0 \leq \int_{E_n} P \circ (\tilde{\phi}_n - f(E_n)) \, d\mu_{E_n} < 2^{-n-1}, \quad \text{where} \ (E_n, \mathcal{B}_{E_n}, \mu_{E_n}) \text{ is the finite measure subspace of } X. \]

This simple function \( \tilde{\phi}_n \) may be extended to a simple function \( \phi_n : X \rightarrow U \) such that \( \phi_n|_{E_n} = \tilde{\phi}_n \) and \( \phi_n(x) = \vartheta_y \), \( \forall x \in X \setminus E_n \). Then, we have \( \| \phi_n(x) \| \leq P \circ f(x) \), \( \forall x \in X \), and \( \| \int_X \phi_n \, d\mu - \int_X f \, d\mu \| = \left\| \int_X (\phi_n|_{E_n}) \, d\mu - \int_{E_n} f \, d\mu - \int_{X \setminus E_n} f \, d\mu \right\| = \left\| \int_{E_n} \phi_n \, d\mu_{E_n} - \int_{E_n} f \, d\mu \right\| + \left\| \int_{X \setminus E_n} f \, d\mu \right\| < 2^{-n-1} + \int_{X \setminus E_n} P \circ f \, d\mu = 2^{-n-1} + \int_X P \circ f \, d\mu - \int_{E_n} P \circ f \, d\mu < 2^{-n}, \quad \text{where the first and the second equalities follow from Proposition 11.92; the second inequality follows from Proposition 11.92; and the third equality follows from Proposition 11.83.} \]

Next, we have \( \left\| \int_X P \circ \phi_n \, d\mu - \int_X P \circ f \, d\mu \right\| \leq \left\| \int_X ((P \circ \phi_n)|_{E_n}) \, d\mu - \int_{E_n} P \circ f \, d\mu \right\| + \left\| \int_X P \circ f \, d\mu - \int_{E_n} P \circ f \, d\mu \right\| < \int_{E_n} P \circ \phi_n \, d\mu_{E_n} - \int_{E_n} P \circ f \, d\mu + 2^{-n-1} < 2^{-n}, \quad \text{where the second inequality follows from Proposition 11.92.} \]

Finally, we have \( 0 \leq \int_X P \circ (\phi_n - f) \, d\mu = \int_{E_n} P \circ ((\phi_n - f)|_{E_n}) \, d\mu_{E_n} + \int_{X \setminus E_n} P \circ ((\phi_n - f)|_{X \setminus E_n}) \, d\mu_{X \setminus E_n} = \int_{E_n} P \circ (\phi_n - f|_{E_n}) \, d\mu_{E_n} + \int_{X \setminus E_n} P \circ (f|_{X \setminus E_n}) \, d\mu_{X \setminus E_n} < 2^{-n-1} + \int_{X \setminus E_n} P \circ f \, d\mu < 2^{-n}, \quad \text{where the first equality follows from Propositions 11.83.} \]

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 \( f_1 : X \rightarrow \mathbb{R} \) and \( f_2 : X \rightarrow \mathbb{R} \) be \( \mathcal{B} \)-measurable. Assume that \( \int_{E_i} f_1 \, d\mu \leq \int_{E_i} f_2 \, d\mu, \forall E_i \in \mathcal{B} \) with \( \int_E P \circ f_i \, d\mu < \infty, i = 1, 2 \). Then, \( f_1 \leq f_2 \) a.e. in \( X \).

**Proof** By Propositions 7.21, 7.23, 11.38, and 11.39, \( f_1 - f_2, P \circ f_1, \) and \( P \circ f_2 \) are \( \mathcal{B} \)-measurable. Let \( E := \{ x \in X \mid f_2(x) - f_1(x) < 0 \} \) \( \subseteq \mathcal{B} \). Suppose \( \mu(E) > 0 \). \( E = \bigcup_{n=1}^\infty E_n := \bigcup_{n=1}^\infty \{ x \in X \mid f_2(x) - f_1(x) < -1/n, |f_1(x)| \leq n, |f_2(x)| \leq n \} \). By Proposition 11.7, \( \mu(E) = \lim_{n \to \infty} \mu(E_n) \) 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, \( \mu(E_n) = \lim_{m \to \infty} \mu(E_n \cap X_m) \). Then, \( \exists m \in \mathbb{N} \) such that \( \mu(E) := \mu(E \cap X_m) \in (0, +\infty) \subseteq \mathbb{R} \). By Definition 11.79, \( \int_{E_i} P \circ f_i \, d\mu \leq n\mu(E) < \infty, i = 1, 2 \). Then, by Propositions 11.75 and 11.92, \( \mathbb{R} \supseteq \int_{E_i} f_1 \, d\mu - \int_{E_i} f_2 \, d\mu - \mu(E)/n = \int_{E_i} (f_1 - f_2 - 1/n) \, d\mu \geq 0 \). This contradicts with the assumption. Hence, \( \mu(E) = 0 \) and \( f_1 \leq f_2 \) a.e. in \( X \). \( \square \)
Proposition 11.97 Let $X := (X, B, \mu)$ be a $\sigma$-finite measure space, $f_i : X \rightarrow [0, \infty] \subset \mathbb{R}$ be $B$-measurable, $i = 1, 2$. Assume that $f_1 \leq f_2$ a.e. in $X$, \( \mu(\{x \in X \mid f_1(x) < f_2(x)\}) > 0 \), 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, \( f_1 \) and \( f_2 \) are absolutely integrable over \( X \). Let \( g := f_2 - f_1 \). Then, by Propositions 7.23, 11.38, and 11.39, we have \( g : X \rightarrow \mathbb{R} \) is $B$-measurable and \( g \geq 0 \) a.e. in $X$. By Proposition 11.92, \( g \) is absolutely integrable over \( X \) and \( \int_X g \, d\mu = 0 \). \( \forall E \in B \), by Proposition 11.92, we have \( \int_E g \, d\mu + \int_{X \setminus E} g \, d\mu = \int_X g \, d\mu = 0 \) and both of the summands on the left-hand-side are nonnegative. Then, \( \int_E g \, d\mu = 0 \). By Proposition 11.96, we have \( g = 0 \) a.e. in $X$. This contradicts the fact that \( \mu(\{x \in X \mid g(x) > 0\}) = \mu(\{x \in X \mid f_1(x) < f_2(x)\}) > 0 \). 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 $X := (X, B, \mu)$ 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, $f : X \rightarrow \Omega \cap \mathcal{W}$ be absolutely integrable over $X$, and $G : \Omega \rightarrow \mathcal{W}$ be a convex functional. Assume that $G \circ f$ is absolutely integrable over $X$ and the epigraph $[G, \Omega]$ is closed. Then, \( \int_X f \, d\mu \in \Omega \) and \( G(\int_X f \, d\mu) \leq \int_X (G \circ f) \, d\mu \in \mathcal{W} \).

**Proof** By Proposition 11.92, \( y_0 := \int_X f \, d\mu \in \mathcal{Y} \). We will first show that \( y_0 \in \Omega \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \vartheta \in \Omega \); Case 2: \( \vartheta \notin \Omega \).

Case 1: \( \vartheta \in \Omega \). Then, \( \Omega \cap \mathcal{W} \) is a conic segment. By Lemma 11.87, there exists a sequence of simple functions \((\varphi_n)_{n=1}^\infty \), \( \varphi_n : X \rightarrow \Omega \cap \mathcal{W} \), \( \forall n \in \mathbb{N} \), such that \( \lim_{n \to \infty} \varphi_n = f \) a.e. in $X$, \( \| \varphi_n(x) \| \leq \| f(x) \| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), 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 \( \varphi_n = \sum_{i=1}^n y_i x A_i \). Let \( y_{n+1} := \vartheta \in \Omega \cap \mathcal{W} \) and \( A_{n+1} := X \setminus (\bigcup_{i=1}^n A_i) \in \mathcal{B} \). Then, by Proposition 11.75, \( \int_X \varphi_n \, d\mu = \sum_{i=1}^{n+1} y_i \mu(A_i) \). Note that \( 1 = \mu(X) = \sum_{i=1}^{n+1} \mu(A_i) \) and the summands are nonnegative. Since \( \Omega \) is convex, then \( \int_X \varphi_n \, d\mu \in \mathcal{W} \). Since \( \Omega \) is closed, then, by Proposition 4.13, \( y_0 = \lim_{n \to \infty} \int_X \varphi_n \, d\mu \in \Omega \).

Case 2: \( \vartheta \notin \Omega \). Note that \( \mu(X) = 1 \) implies that \( X \neq \emptyset \). Then, \( \Omega \cap \mathcal{W} \neq \emptyset \). Let \( \bar{y} \in \Omega \cap \mathcal{W} \) and \( \Omega := \Omega - \bar{y} \). Then, \( \vartheta \notin \Omega \) and, by Proposition 7.16, \( \Omega \) is a closed convex set. Let \( \bar{f} := f - \bar{y} \). Then, we have \( \bar{f} : X \rightarrow \Omega \cap \mathcal{W} \). By Propositions 11.38, 11.39, and 7.23, \( \bar{f} \) is $B$-measurable. Note that \( \mathcal{P} \circ f(x) = \| f(x) - \bar{y} \| \leq \mathcal{P} \circ f(x) + \| \bar{y} \| \), \( \forall x \in X \). Then, by Propositions 11.83 and 11.75, \( 0 \leq \int_X (\mathcal{P} \circ f) \, d\mu \leq \int_X (\mathcal{P} \circ f + \| \bar{y} \|) \, d\mu = \int_X (\mathcal{P} \circ f) \, d\mu + \int_X \| \bar{y} \| \, d\mu = \int_X (\mathcal{P} \circ f) \, d\mu + \| \bar{y} \| < +\infty \). Hence, \( f \) is absolutely integrable over $X$. 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 \( G(y_0) = \sup_{y_\ast \in \Omega_{\text{conj}}} (\langle y_\ast, y_0 \rangle - G_{\text{conj}}(y_\ast)) \), where \( G_{\text{conj}} : \Omega_{\text{conj}} \to \mathbb{R} \) is the conjugate functional to \( G \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists y_\ast \in \Omega_{\text{conj}} \) such that \( G(y_0) - \epsilon \leq \langle y_\ast, y_0 \rangle - G_{\text{conj}}(y_\ast) \leq \langle y_\ast, y_0 \rangle + G(y) = G(y) - \langle y_\ast, y - y_0 \rangle, \forall y \in \Omega \). Then, \( \forall x \in X \), we have \( G(y_0) - \epsilon \leq G(f(x)) - \langle y_\ast, f(x) - y_0 \rangle \). By Proposition 11.92, we have

\[
G(y_0) - \epsilon = \int_X (G(y_0) - \epsilon) \, d\mu \leq \int_X (G \circ f - \langle y_\ast, f - y_0 \rangle) \, d\mu
\]

\[
= \int_X (G \circ f) \, d\mu - \int_X \langle y_\ast, f \rangle \, d\mu + \int_X \langle y_\ast, y_0 \rangle \, d\mu
\]

\[
= \int_X (G \circ f) \, d\mu - \langle y_\ast, \int_X f \, d\mu \rangle + \langle y_\ast, y_0 \rangle = \int_X (G \circ f) \, d\mu \in \mathbb{R}
\]

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.

\section*{11.6 Banach Space Valued Measures}

\textbf{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

\begin{enumerate}
  \item \( \mu(\emptyset) = 0 \);
  \item \( \forall \{E_i\}_{i=1}^\infty \subseteq \mathcal{B} \), which is pairwise disjoint, we have \( \sum_{i=1}^\infty \|\mu(E_i)\| < +\infty \) and \( \mu(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i) \).
\end{enumerate}

Then, the triple \( (X, \mathcal{B}, \mu) \) is said to be a \( \mathcal{Y} \)-valued pre-measure.

Define \( \mathcal{P} \circ \mu : \mathcal{B} \to [0, \infty] \subseteq \mathbb{R} \) by, \( \forall E \in \mathcal{B} \), \( \mathcal{P} \circ \mu(E) := \sup_{\epsilon, \{E_i\}_{i=1}^n \subseteq \mathcal{B}, \cup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(E_i)\| \), \( \mathcal{P} \circ \mu \) is said to be the total variation of \( \mu \).

\textbf{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 \( (X, \mathcal{B}) \).

\begin{proof}
Clearly, \( \mathcal{P} \circ \mu(\emptyset) = 0 \). \( \forall \) pairwise disjoint \( \{E_i\}_{i=1}^\infty \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^\infty E_i, \forall i \in \mathbb{N} \) 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 \( \{E_{i,j}\}_{j=1}^{n_i} \subseteq \mathcal{B} \) 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, \( 0 = t_i \leq \|\mu(E_{i,1})\| \leq \mathcal{P} \circ \mu(E_i) = 0 \).

\textbf{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

\[
\sum_{j=1}^{n} \| \mu(A_j) \| = \sum_{j=1}^{n} \| \mu \left( \bigcup_{i=1}^{\infty} (E_i \cap A_j) \right) \| = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(E_i \cap A_j) \| \\
\leq \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(E_i \cap A_j) \| = \sum_{i=1}^{\infty} \sum_{j=1}^{n} \| \mu(E_i \cap A_j) \| \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)
\]

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 \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued pre-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 a property holds almost everywhere in \( \mathcal{X} \) if it holds almost everywhere in \( \mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu) \). We will say that \( \mathcal{X} \) is complete if the measure space \( \mathcal{X} \) 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 \( E_{n+i} = 0, \forall i \in \mathbb{N} \). 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 \(\mathcal{Y}\)-valued pre-measure space, where \(\mathcal{Y}\) 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, \(\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(\bigcap_{n=1}^{\infty} E_n) \in \mathcal{Y}\).

Proof Let \(E := \bigcap_{n=1}^{\infty} E_n \in \mathcal{B}\). 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, \(\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} \mu(E_n \setminus E_{n+1}) + \sum_{n=1}^{\infty} ||\mu(E_n \setminus E_{n+1})|| < +\infty\). 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 \(\mu(E_1) = \mu(E) + \sum_{n=1}^{\infty} (\mu(E_n) - \mu(E_{n+1})) = \mu(E) + \lim_{n \in \mathbb{N}} \mu(E_n) \in \mathcal{Y}\). Hence, we have \(\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(E) \in \mathcal{Y}\). This completes the proof of the proposition.

Proposition 11.104 Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued pre-measure space, where \(\mathcal{Y}\) 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, \(\mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}\).

Proof Let \(A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}\). Note that \(A = A_1 \cup (\bigcup_{n=1}^{\infty} (A_{n+1} \setminus A_n))\), where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, \(\mu(A) = \mu(A_1) + \sum_{n=1}^{\infty} \mu(A_{n+1} \setminus A_n) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{n=1}^{\infty} \mu(A_{n+1} \setminus A_n)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}\), where the third equality follows from Fact 11.102. This completes the proof of the proposition.

Proposition 11.105 Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued pre-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 \(\mu(E_1) = \mu(E_2) \in \mathcal{Y}\).

Proof By Proposition 11.100, \(0 = \mathcal{P} \circ \mu(E_1 \Delta E_2) = \mathcal{P} \circ \mu(E_1 \setminus E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1)\). Then, \(\mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(E_2 \setminus E_1) = 0\) and \(\mu(E_1 \setminus E_2) = \emptyset_Y = \mu(E_2 \setminus E_1)\). Hence, we have \(\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_2),\) where the first and last equality follows from Fact 11.102. This completes the proof of the proposition.

Proposition 11.106 Let \((X, \mathcal{B}, \mu)\) be a \(\mathcal{Y}\)-valued pre-measure space, where \(\mathcal{Y}\) is a normed linear space. Then, there is a unique complete \(\mathcal{Y}\)-valued pre-measure space \((X, \mathcal{B}_0, \mu_0)\) such that

1. \(\mathcal{B} \subseteq \mathcal{B}_0;\)
2. \(\forall E \in \mathcal{B}, \mu(E) = \mu_0(E)\) and \(\mathcal{P} \circ \mu(E) = \mathcal{P} \circ \mu_0(E);\)
3. \(\forall E \subseteq X, E \in \mathcal{B}_0\) 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 \(\mathcal{Y}\)-valued pre-measure space that satisfies (i) and (ii), then, \(\mathcal{B}_0 \subseteq \mathcal{B}_1\), \(\mu_0(E) = \mu_1(E)\), and \(\mathcal{P} \circ \mu_0(E) = \mathcal{P} \circ \mu_1(E)\), \(\forall E \in \mathcal{B}_0\).
The first inequality follows from Definition 11.99 and the fact that $\nu_0(E) = P \circ \mu(E)$, $\forall E \in B$. Define $\mu_0 : B \to Y$ 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 $P \circ \mu(C) = 0$. $\mu_0(E) := \mu(B) \in Y$.

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, $0 \leq P \circ \mu(B_1 \Delta B_2) = P \circ \mu(B_1 \setminus B_2) + P \circ \mu(B_2 \setminus B_1) \leq P \circ \mu(C_1) = 0$. By Proposition 11.105, we have $\mu(B_1) = P \circ \mu(B_2)$. 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 $Y$-valued pre-measure on the measurable space $(X, B_0)$. Since $\emptyset \in B$, then $\mu_0(\emptyset) = \mu(\emptyset) = 0$. 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 $P \circ \mu(C_n) = 0$. 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 $E_n = (\bigcup_{n=1}^{\infty} A_n) \cup (\bigcup_{n=1}^{\infty} B_n)$, $E_n \subseteq \bigcup_{n=1}^{\infty} C_n$, $0 \leq P \circ \mu\left(\bigcup_{n=1}^{\infty} C_n\right) \leq \sum_{n=1}^{\infty} P \circ \mu(C_n) = 0$, and $\sum_{n=1}^{\infty} \mu(B_n) \subseteq \bigcup_{n=1}^{\infty} C_n \subseteq \bigcup_{n=1}^{\infty} A_n = E_n$.

Also, $\sum_{n=1}^{\infty} \|\mu_0(E_n)\| = \sum_{n=1}^{\infty} \|\mu(B_n)\| < +\infty$. This proves that $\mu_0$ is countably additive. Hence, $(X, B_0, \mu_0)$ is a $Y$-valued pre-measure space.

Next, we will show that $P \circ \mu_0 = \nu_0$. $\forall E \in B_0$, $\forall n \in \mathbb{N}_+$, a 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 $P \circ \mu(C_i) = 0$. Then, $(B_i)_{i=1}^{n} \subseteq B$ is pairwise disjoint, $E = (\bigcup_{i=1}^{n} A_i) \cup (\bigcup_{i=1}^{n} B_i) = A \cup B, A \subseteq \bigcup_{i=1}^{n} C_i = C \subseteq B, B \in B$, and $0 \leq P \circ \mu(C) \leq \sum_{i=1}^{n} P \circ \mu(B_i) = 0$, where the last inequality follows from Proposition 11.6. By Proposition 11.12 and its proof, $\nu_0(E) = P \circ \mu(B) = \sum_{i=1}^{n} P \circ \mu(B_i) \geq \sum_{i=1}^{n} \mu(B_i) \| = \sum_{i=1}^{n} \|\mu_0(E_i)\|$. Hence, $\nu_0(E) \geq P \circ \mu_0(E)$.

On the other hand, $\nu_0(E) = P \circ \mu(B) \leq P \circ \mu_0(B) \leq P \circ \mu_0(E)$, where the first inequality follows from Definition 11.99 and the fact that $\mu = \mu_0|_B$; and the second inequality follows from Proposition 11.4. Therefore, $P \circ \mu_0 = \nu_0$. $\forall E \in B, P \circ \mu_0(E) = \nu_0(E) = P \circ \mu(E)$. This shows that (ii) is satisfied. This proves that $(X, B_0, \mu_0)$ is a complete $Y$-valued pre-measure space satisfying (i), (ii), and (iii) and $(X, B_0, P \circ \mu_0)$ is the completion of $(X, B, P \circ \mu)$.
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 \((\mathcal{P} \circ \mu_1)|_{\mathcal{B}_0} = \mathcal{P} \circ \mu_0\). For every \(E \in \mathcal{B}_0\), there exists \(A, B, C \subseteq X\) such that \(E = A \cup B, A \subseteq C, B, C \subseteq B, \text{ and } \mathcal{P} \circ \mu(C) = 0\). Then, \(\mathcal{P} \circ \mu_1(C) = 0, \ E = B \cup (E \setminus B), \ B \not\subseteq E \setminus B = A \cup B \subseteq C, \text{ and } 0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0\). This leads to \(\mu_1(E \setminus B) = 0\) 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.

\begin{proposition}
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) \((X, \tilde{\mathcal{B}}, \mu|_{\tilde{\mathcal{B}}})\) 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 \(\tilde{\mathcal{X}} := (X, \tilde{\mathcal{B}}, \tilde{\mu})\) be the completion of \(\mathcal{X}\) as defined in Proposition 11.106, and \(\tilde{\mathcal{A}} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)\) be the \(\mathcal{Y}\)-valued pre-measure subspace of \(\tilde{\mathcal{X}}\). Then, \(\tilde{\mathcal{A}}\) is the completion of \(\mathcal{A}\).

\end{proposition}

\begin{proof}
(i) Since \(\tilde{\mathcal{B}}\) is a \(\sigma\)-algebra on \(X\), then \((X, \tilde{\mathcal{B}})\) is a measurable space. \(\forall B \in \mathcal{B} \subseteq \mathcal{B}, \mu|_{\tilde{\mathcal{B}}}(B) = \mu(B) \in \mathcal{Y}\). Then, \(\mu|_{\tilde{\mathcal{B}}} : \tilde{\mathcal{B}} \rightarrow \mathcal{Y}\). Clearly, \(\mu|_{\tilde{\mathcal{B}}}(|\tilde{\mathcal{B}}|) = 0\). \(\forall\) pairwise disjoint sequence \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \subseteq \tilde{\mathcal{B}}\), we have \(E := \bigcup_{i=1}^{\infty} E_i \in \tilde{\mathcal{B}}, \sum_{i=1}^{\infty} \|\mu|_{\tilde{\mathcal{B}}}(E_i)\| = \sum_{i=1}^{\infty} \|\mu(E_i)\| < +\infty\), and \(\mu|_{\tilde{\mathcal{B}}}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu|_{\tilde{\mathcal{B}}}(E_i) = \sum_{i=1}^{\infty} \|\mu(E_i)\|\). Hence, \((X, \tilde{\mathcal{B}}, \mu|_{\tilde{\mathcal{B}}})\) 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 \(\mu|_{\mathcal{B}_A}(\emptyset) = 0\). \(\forall\) 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 \(\mu|_{\mathcal{B}_A}(E) = \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \|\mu(E_i)\|\). 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 \(\mathcal{P} \circ \mu(B) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(B) = 0\).
\end{proof}
11.6. BANACH SPACE VALUED MEASURES

\[ \mathcal{P} \circ (\mu|_{\mathcal{B}_A})(B) = \mathcal{P} \circ \mu_A(B) = 0. \]
Then, \( E_A \in \mathcal{B} \) by the completeness of \( (X, \mathcal{B}, \mathcal{P} \circ \mu) \). Thus, \( E_A \in \mathcal{B}_A \). By the arbitrariness of \( E_A \), \( (A, \mathcal{B}_A, \mathcal{P} \circ \mu_A) \) is complete. Hence, \( \mathcal{A} \) is complete.

(iii) Note that \( \mathcal{B}_A = \{ C \subseteq A \mid C \in \mathcal{B} \} \), \( \mu_A = \tilde{\mu}|_{\mathcal{B}_A} \), and \( \mathcal{P} \circ \tilde{\mu}_A = (\mathcal{P} \circ \tilde{\mu})|_{\mathcal{B}_A} \). By (ii), \( \mathcal{A} \) is a complete \( \mathcal{Y} \)-valued pre-measure space. Since \( \mathcal{B} \subseteq \mathcal{B} \), then \( \mathcal{B}_A \subseteq \mathcal{B}_A \). \( \forall E \in \mathcal{B}_A, \mu_A(E) = \mu_A|_{\mathcal{B}_A}(E) = \mu(E) = \tilde{\mu}(E) = \mu_A(E) \) and, by (i), \( \mathcal{P} \circ \mu_A(E) = (\mathcal{P} \circ \mu)|_{\mathcal{B}_A}(E) = \mathcal{P} \circ \tilde{\mu}(E) = (\mathcal{P} \circ \tilde{\mu})|_{\mathcal{B}_A}(E) = \mathcal{P} \circ \tilde{\mu}_A(E) \), where the third equality follows from Proposition 11.106. \( \forall E \subseteq A \), assume that \( E = E_A \cup E_B \) with \( E_B, E_C \in \mathcal{B}_A \), \( E_A \subseteq E_C \), and \( \mathcal{P} \circ \mu_A(E_C) \neq 0 \). Then, \( \mathcal{P} \circ \tilde{\mu}_A(E_C) \neq 0 \) and \( E_A \in \mathcal{B}_A \), since \( \mathcal{A} \) is complete. Then, \( E \in \mathcal{B}_A \). \( \forall E \subseteq A \), assume that \( E \in \mathcal{B}_A \). Then, \( E \subseteq A \) and \( E \in \mathcal{B} \). By Proposition 11.106, \( E = E_A \cup E_B \) with \( E_B, E_C \subseteq \mathcal{B} \), \( E_A \subseteq E_C \), and \( \mathcal{P} \circ \mu(E_C) = 0 \). Then, \( E_B, E_C \subseteq \mathcal{B}_A \), \( E_A \subseteq E_C \), and \( 0 \leq \mathcal{P} \circ \mu_A(E_C \cap A) = \mathcal{P} \circ \tilde{\mu}(E_C \cap A) \leq \mathcal{P} \circ \mu(E_C) = 0 \), where the last inequality follows form Proposition 11.4. Hence, (i) — (iii) of Proposition 11.106 are satisfied and \( \mathcal{A} \) 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

\[
\nu(E) = \sup_{n \in \mathbb{Z}_+, \bigcap_{i=1}^{n} E_i \subseteq \mathcal{B}, \bigcup_{i=1}^{n} E_i = E, E \cap E_i = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(E_i) \|.
\]

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 \( \tilde{\mathcal{X}} \).

We will say that \( \mathcal{X} \) is complete if the measure space \( \tilde{\mathcal{X}} \) 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 \( \tilde{\mathcal{X}} \).
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 \((E_i)_{i=1}^n \subseteq \mathcal{B}\), \(\mu(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \mu(E_i)\), 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, \(\mu(E) = \mu\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n \mu(E_i)\). This completes the proof of the fact.

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, \(\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(\bigcap_{n=1}^\infty E_n) \in \mathcal{Y}\).

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 \(E_1 \subseteq E \cup (\bigcup_{n=1}^\infty (E_n \setminus E_{n+1}))\), and the sets in the union are pairwise disjoint and measurable. Therefore, \(\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}\) and \(\|\mu(E)\| + \sum_{n=1}^\infty \|\mu(E_n \setminus E_{n+1})\| < \infty\). Note that \(E_n = E_{n+1} \cup (E_n \setminus E_{n+1})\), \(n = 1, 2, \ldots\), which implies that \(\mu(E_n) = \mu(E_{n+1}) + \mu(E_n \setminus E_{n+1}) \in \mathcal{Y}\), by Fact 11.110. This leads to \(\mu(E_1) = \mu(E) + \sum_{n=1}^\infty \mu(E_n) - \mu(E_{n+1}) = \mu(E) + \mu(E_1) - \lim_{n \in \mathbb{N}} \mu(E_n) \in \mathcal{Y}\). Hence, we have \(\lim_{n \in \mathbb{N}} \mu(E_n) = \mu(E) \in \mathcal{Y}\). This completes the proof of the proposition.

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, \(\mu(A) = \lim_{n \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}\).

Proof \(A \in \text{dom}(\mu)\) implies that \(\mathcal{P} \circ \mu(A) \subset +\infty\). 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 \(A = A_1 \cup (\bigcup_{i=1}^\infty (A_{i+1} \setminus A_{i}))\), where the sets in the union are pairwise disjoint and measurable. Then, by countable additivity, \(\mu(A) = \mu(A_1) + \sum_{i=1}^\infty \mu(A_{i+1} \setminus A_i) = \lim_{n \in \mathbb{N}} (\mu(A_1) + \sum_{i=1}^n \mu(A_{i+1} \setminus A_i)) = \lim_{n \in \mathbb{N}} \mu(A_{n+1}) = \lim_{i \in \mathbb{N}} \mu(A_i) \in \mathcal{Y}\), where the third equality follows from Fact 11.110. This completes the proof of the proposition.

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 \(\mathcal{P} \circ \mu(E_1 \triangle E_2) = 0\), we have \(\mu(E_1) = \mu(E_2)\) whenever one of them makes sense.
11.6. BANACH SPACE VALUED MEASURES

Proposition 11.113, we have $\mu_B \ni \mu_B$.

Proof By Fact 11.3, $0 = \mathcal{P} \circ \mu(E_1 \triangle E_2) = \mathcal{P} \circ \mu(E_1 \setminus E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1)$. Then, $\mathcal{P} \circ \mu(E_1 \setminus E_2) = \mathcal{P} \circ \mu(E_2 \setminus E_1) = 0$ and $\mu(E_1 \setminus E_2) = \emptyset_y = \mu(E_2 \setminus E_1)$. Without loss of generality, assume $E_1 \in \text{dom} (\mu)$. Then, $\mathcal{P} \circ \mu(E_1) < +\infty$. This implies that $\mathcal{P} \circ \mu(E_2) = \mathcal{P} \circ \mu(E_1 \cap E_2) + \mathcal{P} \circ \mu(E_2 \setminus E_1) = \mathcal{P} \circ \mu(E_1 \cap E_2) < +\infty$ and $E_2 \in \text{dom} (\mu)$. By Fact 11.110, we have $\mu(E_1) = \mu(E_1 \cap E_2) + \mu(E_1 \setminus E_2) = \mu(E_1 \cap E_2) + \mu(E_2 \setminus E_1) = \mu(E_2)$. This completes the proof of the proposition. □

Proposition 11.114 Let $(X, B, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space. Then, there is a unique complete $\mathcal{Y}$-valued measure space $(X, B_0, \mu_0)$ such that

(i) $B \subseteq B_0$;

(ii) $(\mathcal{P} \circ \mu_0)|_B = \mathcal{P} \circ \mu$ and $\mu_0|_B = \mu$;

(iii) $\forall E \subseteq X, E \in B_0$ 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, $(X, B_0, \mathcal{P} \circ \mu_0)$ is the completion of $(X, B, \mathcal{P} \circ \mu)$.

Furthermore, let $(X, B_1, \mu_1)$ be another complete $\mathcal{Y}$-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 $\mathcal{Y}$ 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 $\mu_0(E) := \mu(B) \in \mathcal{Y}$.

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

$$E = A_1 \cup B_1 = A_2 \cup B_2 \quad B_1, C_1, B_2, C_2 \in B\quad \mathcal{P} \circ \mu(C_1) = \mathcal{P} \circ \mu(C_2) = 0; \quad A_1 \subseteq C_1; \quad A_2 \subseteq C_2$$

Then, $B \ni B_1 \setminus B_2 \ni E \setminus B_2 = A_2 \setminus B_2 \ni C_2$ and $B \ni B_2 \setminus B_1 \ni E \setminus B_1 = A_1 \setminus B_1 \subseteq C_1$. 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 $\mu(B_1) = \mu(B_2) \in \mathcal{Y}$. 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 $\mu_0(E) = \mu(E) \in \mathcal{Y}$. 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 \(\mu_0(\emptyset) = \mu(\emptyset) = \emptyset_\mathcal{Y}\). Hence, (i) of Definition 11.10 is satisfied. (ii) of Definition 11.10 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 \((E_n)_{n=1}^\infty \subseteq \mathcal{B}_0\) 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 \(E_n = A_n \cup B_n, B_n, C_n \in \mathcal{B}, A_n \subseteq C_n\), and \(\mathcal{P} \circ \mu(C_n) = 0\). Since \(B_n \subseteq E_n\), \forall \(n \in \mathbb{N}\), then, \((B_n)_{n=1}^\infty \subseteq \mathcal{P}\) 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 11.10 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 \((E_i)_{i=1}^n \subseteq \mathcal{B}_0\) 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 \(E_i = A_i \cup B_i, B_i, C_i \in \mathcal{B}, \forall i \subseteq C_i,\) and \(\mathcal{P} \circ \mu(C_i) = 0\). Then, \((B_i)_{i=1}^n \subseteq \mathcal{P}\) 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,

\[\nu_0(E) \geq \sup_{E \in \mathcal{B}_0, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_0(E_i)\| =: s_E\]

On the other hand,

\[+\infty > \nu_0(E) = \mathcal{P} \circ \mu(B)\]

\[= \sup_{E \in \mathcal{B}_0, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(B_i)\|\]

\[\leq \sup_{E \in \mathcal{B}_0, \bigcup_{i=1}^n E_i = E, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu_0(E_i)\| \leq s_E\]

Hence, we have \(\nu_0(E) = s_E\) and (iv) of Definition 11.10 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 \(\nu_0|_\mathcal{B} = \mathcal{P} \circ \mu,\) then \((\mathcal{P} \circ \mu_0)|_\mathcal{B} = \mathcal{P} \circ \mu.\) Hence, (i) – (iii) are satisfied by \((X, \mathcal{B}_0, \mu_0)\). Clearly, \((X, \mathcal{B}_0, \mathcal{P} \circ \mu_0)\) 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). \((X, \mathcal{B}_1, \mathcal{P} \circ \mu_1)\) is a complete \(\mathcal{Y}\)-valued 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 \( \mathcal{P} \circ \mu_1(C) = \mathcal{P} \circ \mu(C) = 0 \), \( E = B \cup (E \setminus B) \), \( \mathcal{B}_0 \ni E \setminus B = A \setminus B \subseteq C \), and \( 0 \leq \mathcal{P} \circ \mu_1(E \setminus B) \leq \mathcal{P} \circ \mu_1(C) = 0 \). This leads to \( \mu_1(E \setminus B) = 0 \), 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, \( \mu_0 = \mu_1|\mathcal{B}_0 \).

Therefore, \((X, \mathcal{B}_0, \mu_0)\) is a complete \(\mathcal{Y}\)-valued measure space that satisfies (i), (ii), and (iii). It is the unique such space since \(\mathcal{B}_0\) is unique by (iii), and \(\mu_0\) is also unique since any complete \(\mathcal{Y}\)-valued measure space \((X, \mathcal{B}_0, \mu_1)\) satisfying (i) and (ii) will be such that \(\mu_1 = \mu_0\) and \(\mathcal{P} \circ \mu = \mathcal{P} \circ \mu_1\).

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

**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, \(A \in \mathcal{B}\), and \(\mathcal{B}_A := \{ C \subseteq A \; | \; C \in \mathcal{B} \}\). 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 \(\bar{\mathcal{X}} := (X, \bar{\mathcal{B}}, \bar{\mu})\) be the completion of \(\mathcal{X}\) as defined in Proposition 11.114, and \(\bar{A} := (A, \bar{\mathcal{B}}_A, \bar{\mu}_A)\) be the \(\mathcal{Y}\)-valued measure subspace of \(\bar{\mathcal{X}}\). Then, \(\bar{A}\) 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, \(\nu_A := (\mathcal{P} \circ \mu)|\mathcal{B}_A : \mathcal{B}_A \rightarrow [0, +\infty) \subseteq \mathbb{R}_+\) is a measure on \((A, \mathcal{B}_A)\). Clearly, \(\emptyset \in \mathcal{B}_A\) and \(\mu_A(\emptyset) = \mu|\mathcal{B}_A (\emptyset) = \mu(\emptyset) = \nu_Y\). \(\forall E \in \mathcal{B}_A\) 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. \(\forall E \in \mathcal{B}_A\) with \(\nu_A(E) < +\infty\), we have \(\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty\). Then, \(\mu_A(E) = \mu(E) \in \mathcal{Y}\). \(\forall i\) pairwise disjoint \((E_i)_{i=1}^n \subseteq \mathcal{B}_A \subseteq \mathcal{B}\) with \(E = \bigcup_{i=1}^n E_i\), we have \(\sum_{i=1}^n \| \mu_A(E_i) \| = \sum_{i=1}^n \| \mu(E_i) \| \leq +\infty\), and \(\mu_A(E) = \mu(E) = \sum_{i=1}^n \mu_A(E_i) \in \mathcal{Y}\). \(\forall E \in \mathcal{B}_A\) with \(\nu_A(E) < +\infty\), we have \(\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty\). Then, \(\nu_A(E) = \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{Z}_+} (E_i)_{i=1}^n \subseteq \mathcal{B}_A, \bigcup_{i=1}^n E_i = E, \forall 1 \leq i < j \leq n \sum_{i=1}^n \| \mu(E_i) \| = \sup_{n \in \mathbb{Z}_+} (E_i)_{i=1}^n \subseteq \mathcal{B}_A, \bigcup_{i=1}^n E_i = E, \forall 1 \leq i < j \leq n \sum_{i=1}^n \| \mu(E_i) \| = \mathcal{P} \circ \mu_A = (\mathcal{P} \circ \mu)|\mathcal{B}_A = \nu_A = \mathcal{P} \circ \mu_A\). \(\forall i\) pairwise disjoint \((E_i)_{i=1}^n \subseteq \mathcal{B}_A \subseteq \mathcal{B}\) with \(E = \bigcup_{i=1}^n E_i\), we have \(\sum_{i=1}^n \| \mu_A(E_i) \| = \sum_{i=1}^n \| \mu(E_i) \| \leq +\infty\), and \(\mu_A(E) = \mu(E) = \sum_{i=1}^n \mu_A(E_i) \in \mathcal{Y}\). \(\forall E \in \mathcal{B}_A\) with \(\nu_A(E) < +\infty\), we have \(\mathcal{P} \circ \mu(E) = \nu_A(E) < +\infty\). Then, \(\nu_A(E) = \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{Z}_+} (E_i)_{i=1}^n \subseteq \mathcal{B}_A, \bigcup_{i=1}^n E_i = E, \forall 1 \leq i < j \leq n \sum_{i=1}^n \| \mu(E_i) \| = \sup_{n \in \mathbb{Z}_+} (E_i)_{i=1}^n \subseteq \mathcal{B}_A, \bigcup_{i=1}^n E_i = E, \forall 1 \leq i < j \leq n \sum_{i=1}^n \| \mu(E_i) \| = \mathcal{P} \circ \mu_A = (\mathcal{P} \circ \mu)|\mathcal{B}_A = \nu_A = \mathcal{P} \circ \mu_A\).

(ii) If, in addition, \(\mathcal{X}\) is complete, then \((X, \mathcal{B}, \mathcal{P} \circ \mu)\) is a complete measure space. 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 \(\mathcal{B}_A = \{ C \subseteq A \; | \; C \in \mathcal{B} \}\), \(\bar{\mu}_A = \bar{\mu}|\mathcal{B}_A\), and \(\mathcal{P} \circ \bar{\mu}_A = (\mathcal{P} \circ \bar{\mu})|\mathcal{B}_A\). By (ii), \(\bar{A}\) is a complete \(\mathcal{Y}\)-valued measure space. Since \(\mathcal{B} \subseteq \mathcal{B}_A\), then \(\mathcal{B} \subseteq \mathcal{B}_A\). Note that \(\mathcal{P} \circ \mu_A = (\mathcal{P} \circ \mu)|\mathcal{B}_A = (\mathcal{P} \circ \bar{\mu})|\mathcal{B}_A = (\mathcal{P} \circ \bar{\mu}_A)|\mathcal{B}_A = (\mathcal{P} \circ \bar{\mu}_A)|\mathcal{B}_A\), 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 \( B_A \subseteq \bar{B}_A \); and the last equality follows from (i). Note also that 
\[
\mu_A = \mu|_{B_A} = (\bar{\mu}|_{B_A})|_{B_A} = (\bar{\mu}|_{\bar{B}_A})|_{B_A} = \bar{\mu}_A|_{B_A},
\]
where the first equality follows from Proposition 11.1.4. Hence, (i) – (iii) is undefined, (iii) and \( E \) is complete. Then, \( E \in \bar{B}_A \). On the other hand, \( \forall E \subseteq A \), 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, \( \mathcal{P} \circ \bar{\mu}_A(E_C) = 0 \) and \( E_A \subseteq \bar{B}_A \) since \( \bar{A} \) is complete. Then, \( E \in \bar{B}_A \). On the other hand, \( \forall E \subseteq A \), assume that \( E \subseteq B_A \). Then, \( E \subseteq A \) and \( E \in \bar{B} \). 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, \( E_B, E_C \subseteq A \in B_A \), \( E_A \subseteq E_C \cap A \), 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 form Proposition 11.4. Hence, (i) — (iii) of Proposition 11.114 are satisfied and \( \bar{A} \) is the completion of \( A \).

This completes the proof of the proposition.

Proposition 11.116 Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a Banach space, \( \mathcal{W} \) be a separable subspace of \( Y \), and \( f : X \rightarrow \mathcal{W} \) be \( B \)-measurable. Define \( \bar{\nu} : B \rightarrow [0, +\infty] \subseteq \mathbb{R}_c \) by \( \bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mu, \forall E \in \mathcal{B} \). Then, \( \bar{\nu} \) is a measure on \((X, \mathcal{B})\). Define a function \( \nu \) from \( B \) to \( Y \) by \( \nu(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( \bar{\nu}(E) = +\infty \), and \( \nu(E) = \int_E f \, d\mu, \forall E \in \mathcal{B} \) with \( \bar{\nu}(E) < +\infty \). Then, \((X, \mathcal{B}, \nu)\) is a \( Y \)-valued measure space with \( \mathcal{P} \circ \nu = \bar{\nu} \).

We will call \( \nu \) the \( Y \)-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 \( Y \)-valued measure space.

Proof Since \( f \) is \( B \)-measurable, by Propositions 7.21 and 11.38, \( \mathcal{P} \circ f \) is \( B \)-measurable. Then, \( \bar{\nu} : B \rightarrow [0, +\infty] \subseteq \mathbb{R}_c \) 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 \mathcal{P} \circ \nu \leq +\infty, \nu(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( \bar{\nu}(E) < +\infty \), by Proposition 11.92, \( \nu(E) = \int_E f \, d\mu = \int_E f |_E \, d\mu_E \subseteq \mathcal{Y} \), where \( \mathcal{Y} := (E, \mathcal{B}_E, \mu_E) \) 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 equality follows from the fact that \( E_i \subseteq E \) and \( 0 \leq \bar{\nu}(E_i) \leq \bar{\nu}(E) < +\infty \); the first inequality follows from Proposition 11.92; and the second equality follows from Proposition 11.83. Then, \( \sum_{i=1}^\infty \nu(E_i) = \)
\[ \sum_{i=1}^{\infty} \int_{E_i} f \, d\mu = \sum_{i=1}^{\infty} \int_{E_i} f|_E \, d\mu_E = \int_E f|_E \, d\mu_E = \int_E f \, d\mu = \nu(E), \]
where the second, the third, and the fourth equalities follow from Proposition 11.92.

\( \forall E \in \mathcal{B} \) with \( \tilde{\nu}(E) < +\infty \), \( \forall n \in \mathbb{Z}_+ \), \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{n} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{n} E_i \), \( f|_E \) is absolutely integrable over \( E \). \( \sum_{i=1}^{n} \| \nu(E_i) \| = \sum_{i=1}^{n} \left\| \int_{E_i} f \, d\mu \right\| \leq \sum_{i=1}^{n} \int_{E_i} \mathcal{P} \circ f \, d\mu = \int_{E} \mathcal{P} \circ f \, d\mu = \tilde{\nu}(E) < +\infty \), where the first equality follows from the fact that \( E_i \subseteq E \) and \( 0 \leq \tilde{\nu}(E_i) \leq \tilde{\nu}(E) < +\infty \); the first inequality follows from Proposition 11.92; the third equality follows from Proposition 11.92; and the last equality follows from Proposition 11.83. Hence, \( s_E := \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E=\bigcup_{i=1}^{n} E_i, E \cap E_i = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \nu(E_i) \| \leq \tilde{\nu}(E). \)

On the other hand, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), by Proposition 11.95, \( \exists \) a simple function \( \phi : \mathcal{E} \rightarrow \mathcal{W} \) 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 \( 0 \leq \int_{E} (\mathcal{P} \circ (\phi - f)|_E) \, d\mu_E < \epsilon/2 \). Let \( \phi \) admit canonical representation \( \phi = \sum_{i=1}^{n} w_i \chi_{E_i} \), where \( n \in \mathbb{Z}_+ \), \( w_1, \ldots, w_n \in \mathcal{W} \) are distinct and none equals to \( v_y \), \( E_1, \ldots, E_n \in \mathcal{B}_E \) are nonempty, pairwise disjoint, and \( \mu_E(E_i) < +\infty \), \( \forall i \in \{1, \ldots, n\} \). Then, we have

\[
\tilde{\nu}(E) = \int_{E} (\mathcal{P} \circ f) \, d\mu = \int_{E} (\mathcal{P} \circ f)|_E \, d\mu_E < \int_{E} (\mathcal{P} \circ \phi) \, d\mu_E + \epsilon/2
\]

\[
= \sum_{i=1}^{n} \| \nu(E_i) \| + \sum_{i=1}^{n} \| w_i \| \mu_E(E_i) + \epsilon/2 - \sum_{i=1}^{n} \| \nu(E_i) \|
\]

\[
= \sum_{i=1}^{n} \| \nu(E_i) \| + \epsilon/2 + \sum_{i=1}^{n} \| \int_{E_i} \phi \, d\mu_E \| - \sum_{i=1}^{n} \| \int_{E_i} f|_E \, d\mu_E \|
\]

\[
\leq s_E + \epsilon/2 + \sum_{i=1}^{n} \| \int_{E_i} \phi \, d\mu_E - \int_{E_i} f|_E \, d\mu_E \|
\]

\[
= s_E + \epsilon/2 + \sum_{i=1}^{n} \| \int_{E_i} (\phi - f|_E) \, d\mu_E \|
\]

\[
\leq s_E + \epsilon/2 + \int_{\bigcup_{i=1}^{n} E_i} \mathcal{P} \circ (\phi - f|_E) \, d\mu_E
\]

\[
= s_E + \epsilon/2 + \int_{\bigcup_{i=1}^{n} E_i} \mathcal{P} \circ (\phi - f|_E) \, d\mu_E < s_E + \epsilon
\]

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 \( 0 \leq \tilde{\nu}(E_i) \leq \tilde{\nu}(E) < +\infty \); the fifth equality and the third inequality follow from Proposition 11.92; and the last equality and the last inequality follows from Proposition 11.83. Hence, by the arbitrariness of \( \epsilon \), we have \( \tilde{\nu}(E) \leq s_E \). Therefore, \( \tilde{\nu}(E) = s_E \).

The above shows that \( (X, \mathcal{B}, \nu) \) is a \( \mathbb{Y} \)-valued measure space with \( \mathcal{P} \circ \nu = \tilde{\nu} \).
Proposition 11.117 Let \( \mathcal{X} := (X, B, \mu) \) be a \( \mathcal{Y} \)-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 \( \mathcal{P} \circ \mu(X_n) < +\infty, \forall n \in \mathbb{N} \). Let \( X_n := X_n \cap \{ x \in X | \| f(x) \| \leq n \} \in \mathcal{B} \). Clearly, \( X = \bigcup_{m=1}^{\infty} X_i = \bigcup_{m=1}^{\infty} \bigcup_{i=1}^{m} X_{i,m} \). By Bounded Convergence Theorem 11.77, we have \( \nu(X_{i,m}) = \int_{X_{i,m}} (\mathcal{P} \circ f) \, d\mu \leq n \mu(X_{i,m}) \leq n \mu(X_i) < +\infty \). Then, \( \nu \) is \( \sigma \)-finite. Hence, \( (X, B, \nu) \) is a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure space.

(ii) Let \( f \) be absolutely integrable over \( \mathcal{X} \). Then, \( \nu(X) = \int_X (\mathcal{P} \circ f) \, d\mu < +\infty \). Hence, \( (X, B, \nu) \) is a \( \sigma \)-finite \( \mathcal{Y} \)-valued (pre)measure space. On the other hand, let \( \nu \) be finite. Then, \( \nu(X) = \int_X (\mathcal{P} \circ f) \, d\mu < +\infty \). Hence, \( f \) is absolutely integrable over \( \mathcal{X} \).

(iii) By (i), \( (X, B, \nu) \) is \( \sigma \)-finite. \( \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(E) = 0 \), then \( \int_E (\mathcal{P} \circ f) \, d\mu = 0 \). By Proposition 11.96, \( \mathcal{P} \circ f = 0 \) a.e. in \( E \). Since \( \mathcal{P} \circ f > 0 \) a.e. in \( X \), then, \( \mu(E) = 0 \). \( \forall A \subseteq E \), by the completeness of \( \mathcal{X} \), \( A \in \mathcal{B} \).

This completes the proof of the proposition.

Proposition 11.117 Let \( \mathcal{X} := (X, B, \mu) \) 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 \( \mathcal{P} \circ \mu(X_n) < +\infty, \forall n \in \mathbb{N} \), and \( X_n := (X_n, B_n, \mu_n) \) be the finite \( \mathcal{Y} \)-valued measure subspace of \( \mathcal{X} \). Then, any \( \mathcal{Y} \)-valued measure space \( \mathcal{X} := (X, B, \nu) \) with \( \mathcal{X}_n \) as its finite \( \mathcal{Y} \)-valued measure subspace, \( \forall n \in \mathbb{N} \), must be identical to \( \mathcal{X} \).

Proof \( \forall E \in \mathcal{B} \), \( E \cap X_n \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( E = \bigcup_{n=1}^{\infty} (E \cap X_n) \in \mathcal{B} \). Then, \( \mathcal{B} \subseteq \mathcal{B} \). On the other hand, \( \forall E \in \mathcal{B} \), \( E \cap X_n \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( E = \bigcup_{n=1}^{\infty} (E \cap X_n) \in \mathcal{B} \). Then, \( \mathcal{B} \subseteq \mathcal{B} \). Hence, \( \mathcal{B} = \mathcal{B} \).

\( \forall E \in \mathcal{B} \) with \( E \in \text{dom} (\mu) \), we have \( \mathcal{P} \circ \mu(E) < +\infty \) and \( \mu(E) \in \mathcal{Y} \). Let \( E_n := E \cap (X_n \setminus \bigcup_{i=1}^{n-1} X_i) \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \mu(E_n) = \mu_n(E_n) = \tilde{\mu}(E_n) \), \( \forall n \in \mathbb{N} \). Note that \( \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \tilde{\mu}(E_n) = \mathcal{P} \circ \tilde{\mu}(E_n) < +\infty \), where the second and third equalities follow from Proposition 11.115. Then, \( E_n \in \text{dom} (\tilde{\mu}) \) and \( \mu(E_n) = \sum_{n=1}^{\infty} \mu_n(E_n) = \sum_{n=1}^{\infty} \tilde{\mu}(E_n) = \mu(\tilde{\mu}(E_n)) \in \mathcal{Y} \).

On the other hand, \( \forall E \in \mathcal{B} \) with \( E \in \text{dom} (\tilde{\mu}) \), we have \( \mathcal{P} \circ \tilde{\mu}(E) < +\infty \) and \( \tilde{\mu}(E) \in \mathcal{Y} \). Let \( E_n := E \cap (X_n \setminus \bigcup_{i=1}^{n-1} X_i) \in \mathcal{B} \), \( \forall n \in \mathbb{N} \). Then, \( \mu(E_n) = \mu_n(E_n) = \tilde{\mu}(E_n), \forall n \in \mathbb{N} \). Note that \( \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \tilde{\mu}(E_n) = \mu(\mathcal{P} \circ \tilde{\mu}(E_n)) < +\infty \), where the second and third equalities follow from Proposition 11.115. Then, \( E_n \in \text{dom} (\mu) \) and \( \mu(E_n) = \sum_{n=1}^{\infty} \mu_n(E_n) = \sum_{n=1}^{\infty} \tilde{\mu}(E_n) = \tilde{\mu}(E) \in \mathcal{Y} \).

Therefore, \( \text{dom} (\mu) = \text{dom} (\tilde{\mu}) \) and \( \mu = \tilde{\mu} \). Hence, \( \mathcal{X} = \mathcal{X} \). This completes the proof of the proposition.

Proposition 11.118 Let \( \mathcal{X}_n := (X_n, B_n, \mu_n) \) be a finite \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a Banach space, \( \forall n \in \mathbb{N} \). \( \forall n, m \in \mathbb{N} \), let \( (X_n \cap X_m, B_{n,m}, \mu_n,m) \) be the finite \( \mathcal{Y} \)-valued measure subspace of \( \mathcal{X}_n \). Assume that \( (\mathcal{X}_n)_{n=1}^{\infty} \) is consistent, that is \( B_n,m = B_{n,m} \) and \( \mu_n,m = \mu_{n,m} \),
∀n, m ∈ N. Let $X := \bigcup_{n=1}^{\infty} X_n$, $B := \{ B \subseteq X \mid B \cap X_n \in \mathcal{B}_n, \forall n \in \mathbb{N} \}$, $\nu : B \rightarrow [0, \infty] \subset \mathbb{R}$ be defined by $\nu(E) := \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E_n)$, where $E_1 := E \cap X_1$ and $E_{n+1} := E \cap (X_{n+1} \setminus \bigcup_{j=1}^{n} X_j)$, $\forall \in \mathbb{N}$, and $\mu$ be a function from $B$ to $\mathcal{Y}$ defined by $\mu(E)$ is undefined, $\forall E \in B$ with $\nu(E) = +\infty$, and $\mu(E) := \sum_{n=1}^{\infty} \mu_n(E_n) \in \mathcal{Y}$, $\forall E \in B$ 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 $X_n$ 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 $B$ is a $\sigma$-algebra on $X$. Clearly, $\emptyset \in B$ and $X \in B$. $\forall (E_i)_{i=1}^{\infty} \subseteq B$, 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 $\forall i \in \mathbb{N}$, $E_i \in B$ 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 $E \in B$. $\forall E \in B$, $E \cap X_n \in \mathcal{B}_n$, $\forall n \in \mathbb{N}$. Then, $(X \setminus E) \cap X_n = X_n \setminus E = X_n \setminus (X \cap E) \in \mathcal{B}_n$. By the arbitrariness of $n$, we have $X \setminus E \in B$. Hence, $B$ is a $\sigma$-algebra on $X$. $\forall n \in \mathbb{N}$, $\forall \in \mathbb{N}$, $X_n \cap X_m \in \mathcal{B}_{n,m} = \mathcal{B}_{m,n} \subseteq \mathcal{B}_m$. By the arbitrariness of $m$, we have $X_n \in B$.

Next, we will show that $\nu$ is a $\sigma$-finite measure on $(X, \mathcal{B})$. Clearly, $\nu(\emptyset) = 0$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq B$, let $E := \bigcup_{i=1}^{\infty} E_i$, $\forall i \in \mathbb{N}$, let $E_{i,1} := E_i \cap X_1$ and $E_{i,n+1} := E_i \cap (X_{n+1} \setminus \bigcup_{j=1}^{n} X_j) \in B_{n+1}$, $\forall n \in \mathbb{N}$. Let $B_1 := E \cap X_1 = \bigcup_{i=1}^{\infty} E_i \cap X_1 \in B_1$ and $B_{n+1} := E \cap (X_{n+1} \setminus \bigcup_{j=1}^{n} X_j) = \bigcup_{i=1}^{\infty} E_{i,n+1} \in B_{n+1}$, $\forall n \in \mathbb{N}$. Then, $\nu(E) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_i(E_i,n) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_i(E_i,n) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n(E_{i,n}) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)$, where the first equality follows from the definition of $\nu$; the second equality follows from the fact that $\mathcal{P} \circ \mu$ 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 $(X, \mathcal{B})$. $\forall n \in \mathbb{N}$, let $X_{n,1} := X_n \cap X_1$ and $X_{n,i+1} := X_n \cap (X_{n+1} \setminus \bigcup_{j=1}^{i} X_j)$, $\forall i \in \mathbb{N}$. Then, $\nu(X_n) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_i(X_{n,i}) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_i((X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{i} X_j) + \mathcal{P} \circ \mu_1(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_{i+1,n}((X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{i} X_j) + \mathcal{P} \circ \mu_n(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_{n+1}(X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{i} X_j) + \mathcal{P} \circ \mu_n(X_n \cap X_1) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n((X_n \cap X_{n+1}) \setminus \bigcup_{j=1}^{i} X_j) + \mathcal{P} \circ \mu_n(X_n \cap X_1) = \mathcal{P} \circ \mu_n(X_n) < +\infty$, 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 $\mathcal{P} \circ \mu$ is a measure on $(X_n, \mathcal{B}_n)$. Hence, $\nu$ is $\sigma$-finite.

Next, we will show that $(X, \mathcal{B}, \mu)$ is a $\sigma$-finite $\mathcal{Y}$-valued measure space with $\mathcal{P} \circ \mu = \nu$. 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, $+\infty > \nu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(B_n) \geq \sum_{n=1}^{\infty} \| \mu_n(B_n) \|$. This implies that $\mu(E) = \sum_{n=1}^{\infty} \mu_n(B_n) \in \mathcal{Y}$ by Propo-
sition 7.27. ∀ pairwise disjoint \((E_i)_{i=1}^{\infty} \subseteq \mathcal{B}\) with \(E = \bigcup_{i=1}^{\infty} E_i, \forall i \in \mathbb{N}\), let \((E_{n,i})_{i=1}^{\infty} \subseteq \mathcal{B}\) be as defined in the previous paragraph. Then, we have
\[
\sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_{n,i}) \leq \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \nu(E_i) = \nu(E) < +\infty,
\]
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 \(\sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_{n,i}) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(B_n) = \mu(E),\)
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.

∀ \(E \in \mathcal{B}\) with \(\nu(E) < +\infty\), let \((B_n)_{n=1}^{\infty}\) be as defined in the previous paragraph. ∀ \(m \in \mathbb{Z}_+\), ∀ pairwise disjoint \((E^m_i)_{i=1}^{\infty} \subseteq \mathcal{B}\) with \(E = \bigcup_{i=1}^{\infty} E_i, \forall i \in \{1, \ldots, m\}\), let \((E_{n,i})_{i=1}^{\infty} \subseteq \mathcal{B}\) be as defined in the previous paragraph. Then, \(\sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_{n,i}) \leq \sum_{i=1}^{\infty} \sum_{n=1}^{\infty} \nu(E_{n,i}) = \nu(E) < +\infty\).

Hence,
\[
s_E := \max_{m \in \mathbb{Z}_+, \ (E^m_i)_{i=1}^{\infty} \subseteq \mathcal{B}, \ E = \bigcup_{i=1}^{\infty} E_i, \ E \cap E_j = \emptyset, \forall 1 \leq i < j \leq m} \sum_{i=1}^{m} \mu(E_i) \leq \nu(E).
\]
On the other hand, \(\nu(E) = \sum_{n=1}^{\infty} \mu(B_n) < +\infty\), 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, ∀ \(\epsilon \in (0, +\infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \exists m_n \in \mathbb{Z}_+, \exists \text{ pairwise disjoint} \ (B_{n,i})_{i=1}^{\infty} \subseteq \mathcal{B}_n\) such that \(\mu(B_{n,i}) = 2^{-n}\epsilon < \sum_{i=1}^{\infty} \mu_n(B_{n,i})\). It is easy to see that \(\mu_n(B_{n,i}) = \mu_n(B_{n,i})\), ∀ \(i \in \{1, \ldots, m_n\}\). Then, \(\nu(E) - \epsilon = \sum_{n=1}^{\infty} \mu_n(B_{n,i}) < \sum_{n=1}^{\infty} \sum_{i=1}^{m_n} \mu_n(B_{n,i}) < \sum_{n=1}^{\infty} \sum_{i=1}^{m_n} \mu(B_{n,i}) \leq \nu(E) < +\infty\).

Then, \(\exists N \in \mathbb{N}\) such that \(\nu(E) \leq \sum_{n=1}^{N} \sum_{i=1}^{m_n} \mu_n(B_{n,i}) < \epsilon \leq s_E + \epsilon\). By the arbitrariness of \(\epsilon\), we have \(\nu(E) \leq s_E\). Therefore, \(\nu(E) = s_E\) and (iv) of Definition 11.108 is satisfied. This shows that \((X, \mathcal{B}, \mu)\) is a σ-finite \(\mathcal{Y}\)-valued measure space with \(\mathcal{P} \circ \mu = \nu\).

Finally, we will show that \(X_n\) is the finite \(\mathcal{Y}\)-valued measure subspace of \(X\), ∀ \(n \in \mathbb{N}\). Fix any \(n \in \mathbb{N}\). ∀ \(E \in \mathcal{B}\) with \(E \subseteq X_n\), we have \(E = E \cap X_n \in \mathcal{B}_n\). On the other hand, ∀ \(E \in \mathcal{B}_n\), we have \(E \subseteq X_n\) and ∀ \(i \in \mathbb{N}\), \(E \cap X_i = E \cap X_i \cap X_n \in \mathcal{B}_{n,i} = B_{n,i} \subseteq B_i\). Then, \(E \in \mathcal{B}\). Hence, \(\mathcal{B}_n = \{E \in \mathcal{B} \mid E \subseteq X_n\}\).

∀ \(E \in \mathcal{B}_n\), let \(B_1 := E \cap X_1 \in \mathcal{B}_1\) and \(B_{i+1} := E \cap (X_{i+1} \setminus \bigcup_{j=1}^{i} X_j) \in B_{n,i} \), ∀ \(i \in \mathbb{N}\). Then, \(\nu(E) = \sum_{i=1}^{\infty} \mu(B_i)\). Note that \(B_i \subseteq X_n\), ∀ \(i \in \mathbb{N}\), this implies that \(B_i \subseteq X_n \subseteq B_{n,i} \subseteq B_i\). Then, we have \(\nu(E) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n(B_i) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu_n(B_i) \leq \mathcal{P} \circ \mu_n(X_n) < +\infty\), where the first equality follows from the fact that \(\mu_{i,n} = \mu_{i|B_{n,i}}\); the second equality follows from the consistency
assumption; the third equality follows from the fact that \( \mu_{n,t} = \mu_n|_{B_{n,t}} \); the fourth equality follows from countable additivity of \( \mathcal{P} \cup \mu_n \); 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 \( X \) 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 \( \mathbb{K} \), \( W \subseteq B \{ Y, Z \} \) be a subspace, \( f : X \to W \) be \( B \)-measurable, and \( \mathcal{I}(W) := (\mathcal{R}(W), \preceq) \) be the integration system on \( W \). \( \forall \alpha \in A \) \( \in \mathcal{R}(W) \), define \( F_R := \sum_{\alpha \in A} \omega \mu(f_{\alpha}(U_{\alpha})) \in Z \). This defines a net \( (F_R)_{R \in \mathcal{I}(W)} \); \( f \) is said to be integrable over \( X \) if the net admits a limit in \( Z \). In this case, \( \lim_{R \in \mathcal{I}(W)} F_R \in Z \) 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 \( \int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{R \in \mathcal{I}(W)} F_R \in \mathbb{R}_e \), whenever the limit exists.

**Definition 11.120** Let \( X := (X, \mathcal{B}, \mu) \) be a \( Y \)-valued measure space with \( \mathcal{P} \cup \mu(X) = +\infty \), where \( Y \) is a normed linear space over \( \mathbb{K} \), \( Z \) be a normed linear space over \( \mathbb{K} \), \( \hat{X} := (X, \mathcal{B}, \mathcal{P} \cup \mu) \) be the total variation, and \( f : X \to W := B \{ Y, Z \} \) be \( B \)-measurable. Define \( M(X) \) to be the directed system, \( M(X) \), of subsets of \( X \) as defined in Definition 11.71. \( \forall A \in M(X) \), let \( (A, B_A, \mu_A) \) be the finite \( Y \)-valued measure subspace of \( X \) and define \( F_A := \int_A f \, d\mu \in Z \), if the integral exists. This defines a net \( (F_A)_{A \in M(X)} \) 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, \( \lim_{A \in M(X)} F_A \in Z \) 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 \( (F_A)_{A \in M(X)} \subseteq \mathbb{R}_e \) and denote \( \int_X f \, d\mu := \int_X f(x) \, d\mu(x) := \lim_{A \in M(X)} F_A \in \mathbb{R}_e \), 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 \( \mathbb{K} \), \( Z \) be a normed linear space over \( \mathbb{K} \), \( f : X \to W := B \{ Y, Z \} \) be \( B \)-measurable, and \( \hat{X} := (X, \mathcal{B}, \hat{\mu}) \) be a finite \( Y \)-valued measure space that satisfies \( \mathcal{B} \subseteq \mathcal{B} \) and \( \mu = \hat{\mu}|_{\mathcal{B}} \). Then, \( \int_X f \, d\mu = \int_X f \, d\hat{\mu} \), 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, $(F_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f \, d\mu$, and $(\tilde{F}_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f \, d\tilde{\mu}$, as defined in Definition 11.119.\forall R := \{ (w_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in \mathcal{J}(\mathcal{W}), \forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \in \mathcal{B} \subseteq \tilde{\mathcal{B}}$ since $f$ is $\mathcal{B}$-measurable, and $\mu(f_{\text{inv}}(U_\alpha)) = \tilde{\mu}(f_{\text{inv}}(U_\alpha)) \in \mathcal{Y}$ since $\mu = \tilde{\mu}|_{\mathcal{B}}$. Then, $F_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} w_\alpha \tilde{\mu}(f_{\text{inv}}(U_\alpha)) = \tilde{F}_R \in \mathcal{Z}$. Hence, $\int_X f \, d\mu = \lim_{R \in \mathcal{J}(\mathcal{W})} F_R = \lim_{R \in \mathcal{J}(\mathcal{W})} \tilde{F}_R = \int_X f \, d\tilde{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. \qed

Lemma 11.122 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a finite $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $K$, $Z$ be a normed linear space over $K$, $f : X \to W := B(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$-measurable, $\hat{X} \in \mathcal{B}$, and $\hat{X} := (X, \mathcal{B}, \hat{\mu})$ be the finite $\mathcal{Y}$-valued measure subspace of $\mathcal{X}$ as defined in Proposition 11.115. Assume that $\mathcal{P} \circ \mu(\hat{X} \setminus \hat{X}) = 0$. 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, $(F_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f \, d\mu$, and $(\tilde{F}_R)_{R \in \mathcal{J}(\mathcal{W})}$ be the net for $\int_X f|_{\hat{X}} \, d\hat{\mu}$, as defined in Definition 11.119. $\forall R := \{ (w_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in \mathcal{J}(\mathcal{W}), \forall \alpha \in \Lambda$, we have $f_{\text{inv}}(U_\alpha) \in \mathcal{B} \subseteq \tilde{\mathcal{B}}$ since $f$ is $\mathcal{B}$-measurable, and $0 \leq \mathcal{P} \circ \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) \leq \mathcal{P} \circ \mu(X \setminus \hat{X}) = 0$. This implies that $\mu(f_{\text{inv}}(U_\alpha)) = \mu(f_{\text{inv}}(U_\alpha) \cap \hat{X}) + \mu(f_{\text{inv}}(U_\alpha) \cap (X \setminus \hat{X})) = \mu(f|_{\hat{X}} \, f_{\text{inv}}(U_\alpha)) = \tilde{\mu}(f|_{\hat{X}} \, f_{\text{inv}}(U_\alpha))$ by Fact 11.110. Then, $F_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(f_{\text{inv}}(U_\alpha)) = \sum_{\alpha \in \Lambda} w_\alpha \tilde{\mu}(f|_{\hat{X}} \, f_{\text{inv}}(U_\alpha)) = \tilde{F}_R \in \mathcal{Z}$. Hence, $\int_X f \, d\mu = \lim_{R \in \mathcal{J}(\mathcal{W})} F_R = \lim_{R \in \mathcal{J}(\mathcal{W})} \tilde{F}_R = \int_X f|_{\hat{X}} \, d\hat{\mu}$, whenever one of the integrals exists. This completes the proof of the proposition. \qed

Proposition 11.123 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space with $\mathcal{P} \circ \mu(X) = +\infty$, where $\mathcal{Y}$ is a normed linear space over $K$, $Z$ be a normed linear space over $K$, $f : X \to W := B(\mathcal{Y}, \mathcal{Z})$ be $\mathcal{B}$-measurable, and $\hat{\mathcal{X}} := (X, \mathcal{B}, \hat{\mu})$ be the completion of $\mathcal{X}$. Then, $\int_{\hat{\mathcal{X}}} f \, d\hat{\mu} = \int_X f \, d\mu$, whenever one of the integrals exists.

Proof Let $\hat{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu =: \nu)$ and $\hat{\mathcal{X}} := (\hat{X}, \mathcal{B}, \mathcal{P} \circ \tilde{\mu} =: \nu)$. 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 $(F_A)_{A \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f \, d\mu$ and $(\tilde{F}_A)_{\tilde{A} \in \mathcal{M}(\mathcal{X})}$ be the net for $\int_X f \, d\tilde{\mu}$ as defined in Definition 11.120. Clearly, $\mathcal{M}(\mathcal{X}) \subseteq \mathcal{M}(\hat{\mathcal{X}})$, since $\hat{\mathcal{X}}$ is the completion of $\mathcal{X}$. We will distinguish two exhaustive cases: Case 1: $\int_X f \, d\tilde{\mu}$ exists; Case 2: $\int_X f \, d\mu$ exists.

Case 1: $\int_X f \, d\tilde{\mu}$ exists. Then, the net $(\tilde{F}_A)_{\tilde{A} \in \mathcal{M}(\hat{\mathcal{X}})}$ is well defined. $\forall A \in \mathcal{M}(\mathcal{X})$, we have $A \in \mathcal{B}$ and $\nu(A) < +\infty$. Then, $A \in \mathcal{M}(\hat{\mathcal{X}})$. By Propositions 11.121 and 11.114, $F_A = \int_A f \, d\mu_A = \int_A f \, d\tilde{\mu}_A = \tilde{F}_A$, where $(A, \mathcal{B}_A, \mu_A)$ is the finite $\mathcal{Y}$-valued measure subspace of $\mathcal{X}$ and
11.6. BANACH SPACE VALUED MEASURES

$(A, \mathcal{B}_A, \mu_A)$ is the finite complete $\mathcal{Y}$-valued measure subspace of $\tilde{X}$. Then, $(F_A)_{A \in \mathcal{M}(X)}$ is well defined. $\forall A \in \mathcal{M}(X)$, we have $A \in \mathcal{B}$ with $\bar{\nu}(A) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\tilde{A} = A \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\tilde{A}) = \nu(\tilde{B})$. Then, $\tilde{A} \subseteq \tilde{B} \cup \tilde{C} =: A \in \mathcal{B}$ and $\nu(A) \leq \nu(B) + \nu(\tilde{C}) = \nu(B) = \bar{\nu}(\tilde{A}) < +\infty$. This shows that $A \in \mathcal{M}(X)$ and $A \subseteq A$. Hence, $(F_A)_{A \in \mathcal{M}(X)}$ is a subnet 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. $\forall A \in \mathcal{M}(X)$, we have $A \in \mathcal{B}$ with $\nu(A) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\tilde{A} = A \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\tilde{A}) = \nu(\tilde{B})$. Let $\tilde{B} =: A \in \mathcal{B}$. Then, $0 \leq \nu(\tilde{A} \setminus A) = \bar{\nu}(\tilde{A} \setminus B) = \nu(\tilde{C}) = 0$ and $\nu(A) = \bar{\nu}(\tilde{A}) < +\infty$. This shows that $A \in \mathcal{M}(X)$ and $A \subseteq A$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$-valued measure subspace of $X$, $\mathcal{A} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)$ be the finite complete $\mathcal{Y}$-valued measure subspace of $\tilde{X}$, and $\tilde{\mathcal{A}} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)$ be the finite complete $\mathcal{Y}$-valued measure subspace of $\tilde{A}$. Since $\tilde{X}$ is the completion of $X$, then, by Proposition 11.115, $\tilde{\mathcal{A}}$ is the completion of $\mathcal{A}$. By Proposition 11.121, we have $\int_A f \, d\mu_A = \int_A f \, d\tilde{\mu}_A = F_A$. By Lemma 11.122, we have $\tilde{F}_A = \int_A f \, d\tilde{\mu}_A = \int_A f \, d\mu_A = F_A$. 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 $U (U \subseteq \mathbb{R}_+)$ if $\mathcal{Z} = \mathbb{R}$ or $U \subseteq \mathcal{Z}$ if $\mathcal{Z} \neq \mathbb{R}$ with $\int_X f \, d\mu \in U$. 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 $\nu(A_0) = \bar{\nu}(A_0) < +\infty$. Then, $A_0 \in \mathcal{M}(\tilde{X})$. $\forall \tilde{A} \in \mathcal{M}(\tilde{X})$ with $A_0 \subseteq \tilde{A}$, then $\tilde{A} \in \mathcal{B}$ and $\bar{\nu}(\tilde{A}) < +\infty$. By Proposition 11.114 and its proof, $\exists \tilde{A}, \tilde{B}, \tilde{C} \subseteq X$ with $\tilde{A} = A \cup \tilde{B}$, $\tilde{B}, \tilde{C} \in \mathcal{B}$, $\tilde{A} \subseteq \tilde{C}$, $\nu(\tilde{C}) = 0$, and $\bar{\nu}(\tilde{A}) = \nu(\tilde{B})$. Let $\tilde{A} := A \cup B = \tilde{A} \cup B$ be $\mathcal{M}(\tilde{X})$. Then, $A_0 \subseteq A \subseteq A$, $0 \leq \nu(\tilde{A} \setminus A) = \nu(\tilde{A} \setminus B) \leq \nu(\tilde{A} \setminus B) \leq \nu(\tilde{A} \setminus \tilde{B}) = \nu(\tilde{C}) = 0$, and $\nu(A) \leq \nu(\tilde{B}) + \nu(A_0) = \nu(\tilde{A}) + \nu(A_0) < +\infty$. This shows that $A \in \mathcal{M}(\tilde{X})$. Let $\mathcal{A} := (A, \mathcal{B}_A, \mu_A)$ be the finite $\mathcal{Y}$-valued measure subspace of $X$, $\mathcal{A} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)$ be the finite complete $\mathcal{Y}$-valued measure subspace of $\tilde{X}$, and $\tilde{\mathcal{A}} := (A, \tilde{\mathcal{B}}_A, \tilde{\mu}_A)$ be the finite complete $\mathcal{Y}$-valued measure subspace of $\tilde{A}$. Since $\tilde{X}$ is the completion of $X$, then, by Proposition 11.115, $\tilde{\mathcal{A}}$ is the completion of $\mathcal{A}$. By Proposition 11.121, we have $\int_A f \, d\mu_A = \int_A f \, d\tilde{\mu}_A = F_A$. By Lemma 11.122, we have $\tilde{F}_A = \int_A f \, d\tilde{\mu}_A = \int_A f \, d\mu_A = F_A$. Therefore, we have $\int_X f \, d\mu = \lim_{A \in \mathcal{M}(\tilde{X})} \tilde{F}_A = \int_X f \, d\tilde{\mu}$.

Hence, in both cases, we have $\int_X f \, d\mu = \int_X f \, d\tilde{\mu}$. This completes the proof of the proposition. □

**Definition 11.124** Let $X := (X, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, and $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$. $\phi : X \to \mathcal{Z}$ is said to be a simple function if $\exists n \in \mathbb{Z}_+, \exists z_1, \ldots, z_n \in \mathcal{Z}$.
are nonempty and pairwise disjoint, and \( \tilde{\omega}(x) = \sum_{i=1}^{n} z_i \lambda_{A_i, X}(x) \), \( \forall x \in X \). We will say that a simple function \( \phi \) is in canonical representation if \( z_1, \ldots, z_n \) are distinct and none equals to \( \tilde{\omega}_Z \), and \( A_1, \ldots, A_n \) are nonempty and pairwise disjoint.

Clearly, every simple function admits a unique canonical representation.

**Proposition 11.1.25** 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 \( \mathbb{K} \), \( \phi : X \to B(\mathcal{Y}, Z) =: W \) be a simple function in canonical representation, i.e., \( \exists n \in \mathbb{Z}_+, \exists w_1, \ldots, w_n \in W \), which are distinct and none equals to \( \tilde{\omega}_W \), \( \forall A_1, \ldots, A_n \in \mathcal{B} \), which are nonempty, pairwise disjoint, and \( \mathcal{P} \circ \mu(A_i) < + \infty \), \( i = 1, \ldots, n \), such that \( \phi(x) = \sum_{i=1}^{n} w_i \lambda_{A_i, X}(x) \), \( \forall x \in X \). Then, \( \int_X \phi \, d\mu = \sum_{i=1}^{n} w_i \mu(A_i) =: I \in Z \) and \( 0 \leq \| \int_X \phi \, d\mu \| \leq \int_X \mathcal{P} \circ \phi \, d\mathcal{P} \circ \mu < + \infty \).

**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: \( \mathcal{P} \circ \mu(X) = + \infty \). Let \( w_{n+1} := \tilde{\omega}_W \) and \( A_{n+1} := X \setminus (\bigcup_{i=1}^{n} A_i) \in \mathcal{B} \). Let \( \epsilon_0 := \min(1, \min_{1 \leq i < j \leq n+1} \| w_i - w_j \|) > 0 \), by the assumption. Let \( (\Phi_R)_{R \in \mathcal{P}(W)} \) be the net as defined in Definition 11.119 for \( \int_X \phi \, d\mu \).

\( \forall \epsilon \in (0, \epsilon_0/2) \subset \mathbb{R} \), let \( U_i := B_W(w_i, \epsilon) \in B_B(W), i = 1, \ldots, n+1 \), 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 \( W = B_W(\tilde{\omega}_W, \delta) \). Then, \( W \) must be the trivial normed linear space, i.e., \( W \) is a singleton set containing \( \tilde{\omega}_W \). There is a single representation of \( W \), which is \( R_0 := \{ (\tilde{\omega}_W, W) \} \). Then, \( \Phi_{R_0} = \tilde{\omega}_Z \). Clearly, we must have \( n = 0 \), since \( w_1, \ldots, w_n \in \mathcal{Y} \) are distinct and none equals to \( \tilde{\omega}_W \). Then, \( I = \tilde{\omega}_Z \). Hence, \( \forall R \in \mathcal{P}(W) \) with \( R_0 \preceq R \), \( \| \Phi_R - I \| = \| \Phi_{R_0} - I \| = 0 \).

Case 1b: \( W \supset \bigcup_{i=1}^{n+1} U_i \). Let \( U_{n+2} := W \setminus (\bigcup_{i=1}^{n+1} U_i) \in B_B(W) \), 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 \( R_0 \in \mathcal{P}(W) \).

\( \forall \tilde{R} \in \mathcal{P}(W) \) with \( R_0 \preceq \tilde{R} \), then \( \tilde{R} = \{ (\tilde{w}_\alpha, \tilde{U}_\alpha) \mid \alpha \in \Lambda \} \), where \( \Lambda \) is a finite index set, \( \left( \tilde{U}_\alpha \right)_{\alpha \in \Lambda} \subseteq B_B(W) \) 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 \( w_i = \tilde{w}_\alpha \), and, by \( \tilde{R} \in \mathcal{P}(W) \), \( \exists \tilde{\alpha}_i \in \Lambda \) such that \( w_i \in \tilde{U}_{\tilde{\alpha}_i} \). Since \( \left( \tilde{U}_\alpha \right)_{\alpha \in \Lambda} \subseteq B_B(W) \) are nonempty and pairwise disjoint and \( \tilde{w}_\alpha \in \tilde{U}_\alpha \), \( \forall \alpha \in \Lambda \), then \( \alpha_i = \tilde{\alpha}_i \).

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 \( \Lambda := \{ \alpha_1, \ldots, \alpha_{n+2} \} \). \( \forall \alpha \in \Lambda \setminus \alpha \), \( \phi_{\text{inv}}(\tilde{U}_\alpha) = 0 \). Then, \( \tilde{\Phi}_R = \sum_{\alpha \in \Lambda} \tilde{w}_{\alpha} \mu(\phi_{\text{inv}}(\tilde{U}_\alpha)) = \sum_{i=1}^{n+2} \tilde{w}_{\alpha_i} \mu(\phi_{\text{inv}}(\tilde{U}_{\alpha_i})) = \sum_{i=1}^{n+2} w_i \mu(\phi_{\text{inv}}(U_{\alpha_i})). \)
where \( Y \) is measurable, \( \forall \alpha \), and \( \mu(A_\alpha) \in \mathcal{Y} \). Furthermore, \( \phi_{\text{meas}}(\bar{U}_{\alpha+2}) = 0 \). Hence, \( \Phi_R = \sum_{i=1}^n w_i \mu(A_i) = I \in \mathcal{Z} \). Then, we have \( \| \Phi_R - I \| = 0 \).

In both subcases, we have \( \exists R_0 \in \mathcal{J}(W), \forall R \in \mathcal{J}(W) \) with \( R_0 \leq R \), we have \( \| \Phi_R - I \| = 0 < \epsilon \). Hence, \( \int_X \phi \, d\mu = \lim_{R \in \mathcal{J}(W)} \Phi_R = I \). Note that \( 0 \leq \| \int_X \phi \, d\mu \| \leq \sum_{i=1}^n \| w_i \| \| \mu(A_i) \| \leq \sum_{i=1}^n \| w_i \| \| \mathcal{P} \circ \mu(A_i) \| = \int_X \mathcal{P} \circ \phi \, d\mathcal{P} \circ \mu < +\infty \), 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: \( \mathcal{P} \circ \mu(X) = +\infty \). \( \forall A \in \mathcal{M}(\mathcal{X}), \mathcal{P} \circ \mu(A) < +\infty \). \( \phi|_A : A \to W \) is a simple function. By Case 1, \( \Phi_A = \int_A \phi|_A \, d\mu_A \in \mathcal{Z} \) is well-defined, where \( (A, \mathcal{B}_A, \mu_A) \) is the finite \( \mathcal{Y} \)-valued measure subspace of \( \mathcal{X} \). Hence, the net \( (\Phi_A)_{A \in \mathcal{M}(\mathcal{X})} \) is well-defined. Take \( A_0 := \bigcup_{i=1}^n A_i \in \mathcal{B} \).

Then, \( \mathcal{P} \circ \mu(A_0) = \sum_{i=1}^n \mathcal{P} \circ \mu(A_i) < +\infty \). Therefore, \( A_0 \in \mathcal{M}(\mathcal{X}) \), \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_0 \subseteq A \), by Case 1, \( \Phi_A = \sum_{i=1}^n w_i \mu(A_i) = I \in \mathcal{Z} \). Then, \( \int_X \phi \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} \Phi_A = I \in \mathcal{Z} \). Note that \( 0 \leq \| \int_X \phi \, d\mu \| \leq \sum_{i=1}^n \| w_i \| \| \mu(A_i) \| \leq \sum_{i=1}^n \| w_i \| \| \mathcal{P} \circ \mu(A_i) \| = \int_X \mathcal{P} \circ \phi \, d\mathcal{P} \circ \mu < +\infty \), 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.

Clearly, the above result holds for simple functions given in any form, not necessarily in the canonical representation. It also shows that the integrals of simple functions are linear.

**Lemma 11.126** 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 \( \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \), \( \phi_n : X \to \mathcal{W} \) 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, \( \mathcal{W} := \text{span}\left(\left\{ z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \exists z = wy \right\}\right) \subseteq \mathcal{Z} \) 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 \( \bar{\mathcal{X}} := (X, \mathcal{B}, \mathcal{P} \circ \mu) \) be the finite measure space as defined in Definition 11.108. Let \( E := \{ x \in X \mid (\phi_n(x))_{n=1}^\infty \) does not converge to \( f(x) \) or \( (g_n(x))_{n=1}^\infty \) does not converge to \( g(x) \} \). Then, by (i), \( E \in \mathcal{B} \) and
Let \( F \in R \in X(\mathcal{W}) \) be the net for \( \int_X f \, d\mu \) as defined in Definition 11.119.

For every \( \varepsilon \in (0, \infty) \), there exists \( \delta > 0 \) such that for all \( \forall A \in \mathcal{B} \) with \( P \circ \mu(A) < \delta \), \( \forall n \in \mathbb{N} \) with \( n \geq n_0 \), we have

\[
0 \leq \int_A g \, dP \circ \mu < \varepsilon/6 \quad \text{and} \quad 0 \leq \int_A g_n \, dP \circ \mu < \varepsilon/6.
\]

By Egoroff’s Theorem 11.55, \( \exists E \in \mathcal{B} \) with \( P \circ \mu(E) < \delta \) such that \( \left( \phi_n|_{X \setminus E} \right)_{n=1}^{\infty} \) converges uniformly on \( f|_{X \setminus E} \). Let \( E_1 := E \cup E \). Then, \( E_1 \in \mathcal{B} \) with

\[
P \circ \mu(E_1) = P \circ \mu(E) < \delta \quad \text{such that} \quad \left( \phi_n|_{X \setminus E_1} \right)_{n=1}^{\infty} \text{ converges uniformly on } f|_{X \setminus E_1}.
\]

Then, \( \exists \gamma \in \mathbb{N} \) with \( n_1 \geq n_0 \), \( \forall n \in \mathbb{N} \) with \( n \geq n_1 \), \( \forall x \in X \setminus E_1 \),

\[
\| \phi_n(x) - f(x) \| < \frac{\varepsilon}{4 + \hat{g}(x) + 1} =: \hat{\varepsilon} \in (0, \infty) \subset \mathbb{R}.
\]

Fix any \( n \geq n_1 \). Let \( \phi_n \) admit the canonical representation

\[
\phi_n = \sum_{i=1}^{\hat{n}} w_i \chi_{A_i} =: \sum_{i=1}^{\hat{n}} \phi_i \in B(\mathcal{W}), \quad w_i, \ldots, w_\hat{n} \in \mathcal{W}
\]

are distinct and nonequal to \( \phi \), and \( A_1, \ldots, A_\hat{n} \) are nonempty, pairwise disjoint, and \( P \circ \mu(A_i) < +\infty, i = 1, \ldots, \hat{n} \). Let \( w_{i+1} := \phi_i \) and \( A_{i+1} := X \setminus \left( \bigcup_{j=1}^{i} A_i \right) \in \mathcal{B} \). Then, \( X = \bigcup_{i=1}^{\hat{n}+1} A_i \) and \( \sum_{i=1}^{\hat{n}} a_i \) are nonempty, pairwise disjoint and of finite total variation. Define \( \hat{V}_i := B_\mathcal{W}(w_i, \hat{\varepsilon}), i = 1, \ldots, \hat{n} + 1 \). Define \( V_i := V_1 \in \mathcal{B}(\mathcal{W}) \), \( V_{i+1} := V_{i+1} \setminus \left( \bigcup_{j=1}^{i} V_j \right) \in \mathcal{B}(\mathcal{W}), i = 1, \ldots, \hat{n} \). Let \( V_{\hat{n}+2} := \mathcal{W} \setminus \left( \bigcup_{j=1}^{\hat{n}+1} V_j \right) \in \mathcal{B}(\mathcal{W}) \). Clearly, we may form a representation \( \hat{R} := \{(\hat{w}_i, V_i) \mid i = 1, \ldots, \hat{n} + 2, V_i \neq \emptyset\} \in \mathcal{J}(\mathcal{W}) \), where \( \hat{w}_i \in V_i \) are any vectors that satisfy the assumption of Proposition 11.69. \( \forall R \in \mathcal{J}(\mathcal{W}) \) with \( \hat{R} \preceq R \). Let \( R := \{ (\hat{w}_\gamma, \hat{U}_\gamma) \mid \gamma \in \Gamma \} \). \( \forall \gamma \in \Gamma \), by \( \hat{R} \preceq R \), \( \exists \hat{i}_\gamma \in \{1, \ldots, \hat{n} + 2\} \) such that \( \hat{U}_\gamma \subseteq V_{\hat{i}_\gamma} \). Define \( \hat{\gamma} := \{ \gamma \in \Gamma \mid \hat{i}_\gamma = \hat{n} + 2 \} \). Let \( A_{\hat{n}+1} := A_{\hat{n}+1} \cap f_{\mathcal{W}}(\hat{U}_\gamma) \cap (X \setminus E_1) \in \mathcal{B}, i = 1, \ldots, \hat{n} + 1 \). Let \( \hat{\lambda}_\gamma := f_{\mathcal{W}}(\hat{U}_\gamma) \cap E_1 \). \( \forall \gamma \in \Gamma \), \( \hat{\gamma} \subseteq V_{\hat{n}+2} = \mathcal{W} \setminus \left( \bigcup_{j=1}^{\hat{n}+1} V_j \right) \).

Then, \( \mathcal{B} \ni f_{\mathcal{W}}(\hat{U}_\gamma) \subseteq E_1 \).

For any \( \gamma \in \Gamma \), \( \hat{U}_\gamma \subseteq V_{\hat{n}+2} = \mathcal{W} \setminus \left( \bigcup_{j=1}^{\hat{n}+1} V_j \right) \). Then, \( \mathcal{B} \ni f_{\mathcal{W}}(\hat{U}_\gamma) \subseteq E_1 \).

For any \( \gamma \in \Gamma \), \( \hat{A}_\gamma = f_{\mathcal{W}}(\hat{U}_\gamma) \cap (E_1 \setminus E) \neq \emptyset \). \( \forall x \in \hat{A}_\gamma \), we have \( f(x) \in \hat{U}_\gamma \setminus E \), \( \lim_{m \to \infty} \phi_m(x) = \hat{g}(x) \), \( \| \phi_m(x) \| \leq \hat{g}(x) \), \( \forall m \in \mathbb{N} \), and \( \lim_{m \to \infty} g_m(x) = g(x) \). Then, by Propositions 7.21 and 3.66, we have

\[
\| f(x) \| = \lim_{m \to \infty} \| \phi_m(x) \| \leq \lim_{m \to \infty} g_m(x) = g(x).
\]

Then, by \( R \in \mathcal{J}(\mathcal{W}), \| \hat{w}_\gamma \| < \inf_{x \in \hat{U}_\gamma} \| w \| + 1 \leq \| f(x) \| + 1 \leq g(x) + 1 \). By the arbitrariness of \( x \), we have \( \| \hat{w}_\gamma \| \leq \inf_{x \in \hat{A}_\gamma} \| g(x) + 1 \). Therefore, \( \| \hat{w}_\gamma \| \mu(\hat{A}_\gamma) \| \leq (1 + \inf_{x \in \hat{A}_\gamma} \| g(x) \|) \mu(\hat{A}_\gamma) \| \leq (1 + \inf_{x \in \hat{A}_\gamma} \| g(x) \|) \mu(\hat{A}_\gamma) \| \leq (1 + \inf_{x \in \hat{A}_\gamma} \| g(x) \|) \mu(\hat{A}_\gamma) \| \leq (1 + \inf_{x \in \hat{A}_\gamma} \| g(x) \|) \mu(\hat{A}_\gamma) \|
\]

Note that

\[
\left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(\hat{U}_\gamma) \cap E_1 \right\| = \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma (\mu(\hat{U}_\gamma) \cap E_1) \right\|
\]
Claim 11.126.3 \[ \forall \gamma \in \Gamma, \forall i \in \{1, \ldots, \bar{n} + 1 \}, \text{ we have } \| \hat{w}_\gamma - w_i \| \cdot \mu(A_{\gamma,i}) \leq \frac{3 \epsilon}{6 \beta \rho g(X) + 1} = \mu(A_{\gamma,i}). \]

Proof of claim: Since \( \| \phi_n(x) \| \leq g_n(x), \forall x \in X, \) the result is trivial if \( A_{\gamma,i} = \emptyset. \) If \( A_{\gamma,i} \neq \emptyset, \) then \( \exists x \in A_{\gamma,i} \cap f_{\text{inv}}(\hat{U}_\gamma) \cap (X \setminus E_1). \) This implies that \( \phi_n(x) = w_i \) and \( f(x) \in B_W(w_i, \epsilon) \cap \hat{U}_\gamma \subseteq B_W(w_i, \epsilon) \cap V_{i_\gamma} \subseteq B_W(w_i, \epsilon) \cap \hat{V}_{i_\gamma} = B_W(w_i, \epsilon) \cap \hat{B}_W(w_i, \epsilon). \) Note that \( \hat{w}_\gamma \in \hat{U}_\gamma \subseteq B_W(w_i, \epsilon). \) Therefore, we have \( \| \hat{w}_\gamma - w_i \| \leq \| \hat{w}_\gamma - w_i \| + \| w_i - f(x) \| + \| f(x) - w_i \| < 3 \epsilon = \frac{3 \epsilon}{6 \beta \rho g(X) + 1}. \) Hence, the result holds. \( \square \)

Note that \( F_R = \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) \) 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,

\[ \| F_R - \int_X \phi_n \, d\mu \| = \left\| \sum_{\gamma \in \Gamma} \hat{w}_\gamma \mu(f_{\text{inv}}(\hat{U}_\gamma)) - \sum_{i=1}^{\bar{n}} w_i \mu(A_i) \right\| \]
Claim 11.126.1.
\[ \forall \gamma \in \Gamma \]

The equality follows from Claim 11.126.1 and the fact that 
\[ n = 436 \]

In summary, we have shown that 
\[ \exists \bar{\Gamma}, \forall \gamma \in \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\}; \] and the third inequality follows from Claim 11.126.3.

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 \( A_{\gamma,i} = \emptyset \), \( \forall \gamma \in \bar{\Gamma}, \forall i \in \{1, \ldots, \bar{n} + 1\} \); 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
11.6. BANACH SPACE VALUED MEASURES

\[ \| F_R - \int_X \phi_n \, d\mu \| < \epsilon. \]
Then, the net \((F_R)_{R \in \mathbb{Z}(W)} \subseteq \mathcal{Z}\) 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 \(\| \int_X f \, d\mu - \int_X \phi_n \, d\mu \| = \lim_{R \to \mathbb{Z}(W)} \| F_R - \int_X \phi_n \, d\mu \| \leq \epsilon, \forall n \geq n_1\). Hence, we have \(\int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu\) 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 \(0 \leq \| \int_X f \, d\mu \| = \lim_{n \to \infty} \| \int_X \phi_n \, d\mu \| \leq \sup_{n \in \mathbb{N}} \int_X \mathcal{P} \circ \phi_n \, d\mu \leq \lim_{n \to \infty} \int_X g_n \, d\mu = \int_X g \, d\mu < +\infty\), 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 \(\mathcal{B}(\mathcal{Y}, \mathcal{Z})\), \(f : X \to \mathcal{W}\) be \(\mathcal{B}\)-measurable, and \(g : X \to [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that \(\| f(x) \| \leq g(x), \forall x \in \mathcal{X},\) and \(g\) is integrable over \(\mathcal{X} := (X, \mathcal{B}, \mathcal{P} \circ \mu)\). Then, there exists a sequence of simple functions \(\{ \varphi_k \}_{k=1}^{\infty}\) such that \(\lip_{k \in \mathbb{N}} \varphi_k = f\) a.e. in \(\mathcal{X}\), \(\| \varphi_k(x) \| \leq g(x), \forall x \in \mathcal{X}, \forall k \in \mathbb{N}\), \(f\) is integrable over \(\mathcal{X}\) with \(\int_X f \, d\mu = \lim_{k \to \infty} \int_X \varphi_k \, d\mu \in \mathcal{Z}\), \(\lim_{k \to \infty} \int_X \mathcal{P} \circ \varphi_k \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \in [0, \infty) \subseteq \mathbb{R}\), and \(\lim_{k \to \infty} \int_X \mathcal{P} \circ (\varphi_k - f) \, d\mu = 0\).

**Proof** By Proposition 11.66, there exists a sequence of simple functions \(\{ \varphi_k \}_{k=1}^{\infty}\) such that \(\| \varphi_k(x) \| \leq \| f(x) \| \leq g(x), \forall x \in \mathcal{X}, \forall k \in \mathbb{N}\), and \(\lim_{k \to \infty} \varphi_k = f\) a.e. in \(\mathcal{X}\). By Lemma 11.126, \(f\) is integrable over \(\mathcal{X}\) and \(\int_X f \, d\mu = \lim_{k \to \infty} \int_X \varphi_k \, d\mu \in \mathcal{Z}\).

Note that \(0 \leq \mathcal{P} \circ \varphi_k(x) \leq g(x), \forall x \in \mathcal{X}, \forall k \in \mathbb{N}\), \(\mathcal{P} \circ f\) is \(\mathcal{B}\)-measurable, and \(\lim_{k \to \infty} \mathcal{P} \circ \varphi_k = \mathcal{P} \circ f\) a.e. in \(\mathcal{X}\), by Propositions 7.21 and 11.52. Then, by Lebesgue Dominated Convergence Theorem 11.91, we have \(0 \leq \lim_{k \to \infty} \int_X \mathcal{P} \circ \varphi_k \, d\mu = \int_X \mathcal{P} \circ f \, d\mu \leq \int_X g \, d\mu < +\infty\), 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, \(\lim_{k \to \infty} \mathcal{P} \circ (\varphi_k - f) = 0\) a.e. in \(\mathcal{X}\). Note that \(0 \leq \mathcal{P} \circ (\varphi_k - f)(x) \leq 2g(x), \forall x \in \mathcal{X}, \forall k \in \mathbb{N}\). By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have \(\lim_{k \to \infty} \int_X \mathcal{P} \circ (\varphi_k - f) \, d\mu = 0\). 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 \(\mathcal{B}(\mathcal{Y}, \mathcal{Z})\), \(\mathcal{P} = \text{span}\{ \{ z \in \mathcal{Z} | \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} : z = wy \} \} \subseteq \mathcal{Z}\) 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, \(\mu_n : X \to [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable, \(\forall n \in \mathbb{N}\), and \(g : X \to [0, \infty) \subseteq \mathbb{R}\) be \(\mathcal{B}\)-measurable. Assume that

(i) \(\lim_{n \to \infty} f_n = f\) a.e. in \(\mathcal{X}\), \(\lim_{n \to \infty} g_n = g\) a.e. in \(\mathcal{X}\);
In the following statements hold.

Subspace of $\mathbb{Z}_4\mathbb{Z}$

**Proof**

By Lemma 11.43, $(\mathbb{P} \circ \mu)_n(x) = f_n(x)$ a.e. in $\mathcal{X}$, $\|\varphi_n(x)\| \leq g_n(x), \forall x \in \mathcal{X}$, for all $n \in \mathbb{N}$, $f_n$ is integrable over $\mathcal{X}$, and $\int_X f_n \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu = \varphi_n \in \mathbb{Z}$. By Lemma 11.126, $\int_X f_n \, d\mu = \varphi_n$, $\forall n \in \mathbb{N}$. By Proposition 11.58, $\lim_{n \to \infty} \varphi_n = f_n$ in measure in $\mathcal{X}$. Then, $\exists k_n \in \mathbb{N}$ such that $\mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{x \in \mathcal{X} \mid \|\varphi_n(x) - f_n(x)\| \geq 2^{-n}\}) < 2^{-n}$ and $\|\int_X f_n \, d\mu - \int_X \varphi_n \, d\mu\| < 2^{-n}$. Denote $\psi_n := \varphi_n$.

By Proposition 11.58, $\lim_{n \to \infty} f_n = f$ in measure in $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$, we have $2^{-n} < \epsilon/2$ and $\mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{x \in \mathcal{X} \mid \|f_n(x) - f(x)\| \geq \epsilon/2\}) < \epsilon/2$. Note that, by Lemma 11.43, $B \ni E_n := \{x \in \mathcal{X} \mid \|\psi_n(x) - f(x)\| \geq \epsilon\} \subseteq \{x \in \mathcal{X} \mid \|f_n(x) - f(x)\| \geq \epsilon\} \subseteq E_n \cup E_n$. Then, we have $\|\int_X f_n \, d\mu - \int_X \psi_n \, d\mu\| < \epsilon$. This implies that $\lim_{n \to \infty} \psi_n = f$ in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\psi_{n_i})_{i=1}^\infty$ of $(\psi_n)_{n=1}^\infty$ that converges to $f$ a.e. in $\mathcal{X}$. By Lemma 11.126, $f$ is integrable over $\mathcal{X}$, $\int_X f \, d\mu = \lim_{n \to \infty} \int_X \psi_n \, d\mu \in \mathbb{Z}$, and $0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty$. By the fact that $\|\int_X f_n \, d\mu - \int_X \psi_n \, d\mu\| < 2^{-n}$, $\forall i \in \mathbb{N}$, 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 $\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \in \mathbb{Z}$. 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 $\mathbb{W}$, we have $\int_X f \, d\mu \in \mathbb{W} \subseteq \mathbb{Z}$.

This completes the proof of the theorem. □

**Definition 11.129** Let $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ be a $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a normed linear space over $\mathbb{K}$, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $\mathcal{W} := \mathcal{B}([\mathcal{Y}, \mathcal{Z}], \mu)$, and $f : \mathcal{X} \to \mathcal{W}$ be $\mathcal{B}$-measurable. $f$ is said to be absolutely integrable if $\mathcal{P} \circ f$ is integrable over $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mathcal{P} \circ \mu)$, that is, $0 \leq \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < +\infty$.

**Proposition 11.130** Let $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ 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 $\mathcal{B}([\mathcal{Y}, \mathcal{Z}])$, $\mathcal{U}$ be a Banach space over $\mathbb{K}$, $\mathcal{W} := \text{span}\{\{x \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \cdot x = wy\}\} \subseteq \mathcal{Z}$ be the Banach subspace of $\mathcal{Z}$, $f_i : \mathcal{X} \to \mathcal{W}$ be absolutely integrable over $\mathcal{X}$, $i = 1, 2$. Then, the following statements hold.

(i) $f_i$ is integrable over $\mathcal{X}$ and $\int_X f_i \, d\mu \in \mathcal{W} \subseteq \mathcal{Z}$, $i = 1, 2$. 

(ii) $\int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu = \varphi_n \in \mathbb{Z}$. By Lemma 11.126, $\int_X f_n \, d\mu = \varphi_n$, $\forall n \in \mathbb{N}$. By Proposition 11.58, $\lim_{n \to \infty} \varphi_n = f_n$ in measure in $\mathcal{X}$. Then, $\exists k_n \in \mathbb{N}$ such that $\mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{x \in \mathcal{X} \mid \|\varphi_n(x) - f_n(x)\| \geq 2^{-n}\}) < 2^{-n}$ and $\|\int_X f_n \, d\mu - \int_X \varphi_n \, d\mu\| < 2^{-n}$. Denote $\psi_n := \varphi_n$.

By Proposition 11.58, $\lim_{n \to \infty} f_n = f$ in measure in $\mathcal{X}$. $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$, we have $2^{-n} < \epsilon/2$ and $\mathcal{P} \circ \mu(E_n) := \mathcal{P} \circ \mu(\{x \in \mathcal{X} \mid \|f_n(x) - f(x)\| \geq \epsilon/2\}) < \epsilon/2$. Note that, by Lemma 11.43, $B \ni E_n := \{x \in \mathcal{X} \mid \|\psi_n(x) - f(x)\| \geq \epsilon\} \subseteq \{x \in \mathcal{X} \mid \|f_n(x) - f(x)\| \geq \epsilon\} \subseteq E_n \cup E_n$. Then, we have $\|\int_X f_n \, d\mu - \int_X \psi_n \, d\mu\| < \epsilon$. This implies that $\lim_{n \to \infty} \psi_n = f$ in measure in $\mathcal{X}$. By Proposition 11.57, there exists a subsequence $(\psi_{n_i})_{i=1}^\infty$ of $(\psi_n)_{n=1}^\infty$ that converges to $f$ a.e. in $\mathcal{X}$. By Lemma 11.126, $f$ is integrable over $\mathcal{X}$, $\int_X f \, d\mu = \lim_{n \to \infty} \int_X \psi_n \, d\mu \in \mathbb{Z}$, and $0 \leq \|\int_X f \, d\mu\| \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty$. By the fact that $\|\int_X f_n \, d\mu - \int_X \psi_n \, d\mu\| < 2^{-n}$, $\forall i \in \mathbb{N}$, 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 $\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu \in \mathbb{Z}$. 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 $\mathbb{W}$, we have $\int_X f \, d\mu \in \mathbb{W} \subseteq \mathcal{Z}$.

This completes the proof of the theorem. □
(ii) \( f_1 + f_2 \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (f_1 + f_2) \, d\mu = \int_{\mathcal{X}} f_1 \, d\mu + \int_{\mathcal{X}} f_2 \, d\mu \in \mathcal{W} \).

(iii) \( \forall A \in B(\mathcal{Z}, \mathcal{U}), A f_1 \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (A f_1) \, d\mu = A \int_{\mathcal{X}} f_1 \, d\mu \in \mathcal{U} \).

(iv) \( \forall c \in \mathbb{K}, c f_1 \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (c f_1) \, d\mu = c \int_{\mathcal{X}} f_1 \, d\mu \in \mathcal{W} \).

(v) \( \forall H \in B, \) let \( \mathcal{H} := (H, B_H, \mu_H) \) 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 \( \int_{\mathcal{H}} (f_1|_H) \, d\mu = \int_H f_1 \, d\mu_H \in \mathcal{W}. \) We will henceforth denote \( \int_H f_1 \, d\mu_H \) by \( \int_H f_1 \, d\mu \).

(vi) If \( f_1 = f_2 \) a.e. in \( \mathcal{X} \) then \( \int_{\mathcal{X}} f_1 \, d\mu = \int_{\mathcal{X}} f_2 \, d\mu \in \mathcal{W} \).

(vii) \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq B, \sum_{i=1}^\infty \int_{E_i} f_1 \, d\mu = \int_{\bigcup_{i=1}^\infty E_i} f_1 \, d\mu \in \mathcal{W} \).

(viii) \( 0 \leq \| \int_{\mathcal{X}} f_1 \, d\mu \| \leq \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty. \)

**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 \( \int_{\mathcal{X}} f_i \, d\mu = \lim_{n \to \infty} \int_{\mathcal{X}} \phi_{i,n} \, d\mu \in \mathbb{Z}. \) By Lebesgue Dominated Convergence Theorem 11.128, \( \int_{\mathcal{X}} f_i \, d\mu \in \mathcal{W}, i = 1, 2. \)

(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 \( B \)-measurable. Note that \( \mathcal{P} \circ (f_1 + f_2)(x) = \| f_1(x) + f_2(x) \| = \| f_1(x) \| + \| f_2(x) \| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x), \forall x \in X. \) By Proposition 11.83, we have \( 0 \leq \int_X \mathcal{P} \circ (f_1 + f_2) \, d\mathcal{P} \circ \mu \leq \int_X (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2) \, d\mathcal{P} \circ \mu = \int_X (\mathcal{P} \circ f_2 \circ f_2) \, d\mathcal{P} \circ \mu + \int_X \mathcal{P} \circ f_2 \, d\mathcal{P} \circ \mu < +\infty. \) 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 \( \| \phi_{1,n}(x) + \phi_{2,n}(x) \| \leq \| \phi_{1,n}(x) \| + \| \phi_{2,n}(x) \| \leq (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x), \forall x \in X, \forall n \in \mathbb{N}. \) 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 \( \mathcal{X}. \) By Lebesgue Dominated Convergence Theorem 11.128, \( f_1 + f_2 \) is integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} (f_1 + f_2) \, d\mu = \lim_{n \to \infty} \int_{\mathcal{X}} (\phi_{1,n} + \phi_{2,n}) \, d\mu = \lim_{n \to \infty} \int_{\mathcal{X}} \phi_{1,n} \, d\mu + \int_{\mathcal{X}} \phi_{2,n} \, d\mu \leq \int_{\mathcal{X}} f_1 \, d\mu + \int_{\mathcal{X}} f_2 \, d\mu \in \mathcal{W}, \) 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 \( A \in B(\mathcal{Z}, \mathcal{U}). \) By Propositions 11.38, 7.21, and 7.62, \( A \phi_{1,n}, A f_1, \) and \( \mathcal{P} \circ (A f_1) \) are \( B \)-measurable, \( \forall n \in \mathbb{N}. \) Note that \( \mathcal{P} \circ (A f_1)(x) = \| A f_1(x) \| \leq \| A \| \| f_1(x) \| = \| A \| \| f_1 \|, \forall x \in X. \) By Proposition 11.83 and the fact that \( 0 \leq \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty, \) we have \( 0 \leq \int_{\mathcal{X}} (\mathcal{P} \circ f_1) \, d\mathcal{P} \circ \mu \leq \int_{\mathcal{X}} (\| A \| \| f_1 \|) \, d\mathcal{P} \circ \mu = \| A \| \int_{\mathcal{X}} \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu < +\infty. \) 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 \( \mathcal{W} := \{ F \in B(\mathcal{Y}, \mathcal{U}) \mid F = AF_1, F_1 \in \mathcal{W} \} \) is a separable subspace of \( B(\mathcal{Y}, \mathcal{U}). \)
By Propositions 11.52 and 7.62, \( \lim_{n \to \infty} A\varphi_{1,n} = Af \) a.e. in \( X \). By Proposition 11.125, we have \( \int_X (A\varphi_{1,n}) \, d\mu = A \int_X \varphi_{1,n} \, d\mu \in \mathcal{U} \), \( \forall n \in \mathbb{N} \). Note also that \( \| A\varphi_{1,n}(x) \| \leq \| A \| \| \varphi_{1,n}(x) \| \leq \| A \| P \circ f_1(x) \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \).

By Lebesgue Dominated Convergence Theorem 11.128, \( Af_1 \) is integrable over \( X \) and \( \int_X (Af_1) \, d\mu = \lim_{n \to \infty} \int_X (A\varphi_{1,n}) \, d\mu = \lim_{n \to \infty} A \int_X \varphi_{1,n} \, d\mu = A \int_X f_1 \, d\mu \in \mathcal{U} \), 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 \( (P \circ f_1)_{|H} = P \circ (f_1)_{|H} \). Then, by Proposition 11.115, we have \( 0 \leq \int_H P \circ (f_1)_{|H} \, d\mu_{|H} = \int_H (P \circ f_1)_{|H} \, d\mu_{|H} \leq \int_X P \circ f_1 \, d\mu < +\infty \), where the second equality and the second inequality follow from Proposition 11.83. Hence, \( f_1 \) is absolutely integrable over \( H \).

\( \{ \varphi_{1,n} \}_{n=1}^{\infty} \) is a sequence of simple functions that converges to \( f_1 \) a.e. in \( H \). Note that \( \| \varphi_{1,n} \|_H (x) \leq (P \circ f_1)_{|H} (x) = P \circ (f_1)_{|H} (x), \forall x \in H, \forall n \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_H \varphi_{1,n} \, d\mu_{|H} = \lim_{n \to \infty} \int_H \varphi_{1,n} \, d\mu_{|H} \in \mathbb{W} \). Note that \( \{ \varphi_{1,n} \}_{n=1}^{\infty} \) is a sequence of simple functions that converges to \( f_1 \) a.e. in \( X \), and \( \| \varphi_{1,n}(x) \|_{X,H} \leq (P \circ f_1)_{|H} (x), \forall x \in X, \forall n \in \mathbb{N} \).

Again by Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_X (f_1 \varphi_{1,n}) \, d\mu = \lim_{n \to \infty} \int_X (\varphi_{1,n}) \, d\mu = \lim_{n \to \infty} \int_H \varphi_{1,n} \, d\mu_{|H} = \int_X f_1 \, d\mu_{|H} \in \mathbb{W} \), where the second equality follows from Proposition 11.125.

(vi) If \( f_1 = f_2 \) a.e. in \( X \), then, by Proposition 11.54, \( \lim_{n \to \infty} \varphi_{1,n} = f_1 \) a.e. in \( X \). By Lebesgue Dominated Convergence Theorem 11.128, we have \( \int_X f_2 \, d\mu = \lim_{n \to \infty} \int_X \varphi_{1,n} \, d\mu = \int_X f_1 \, d\mu \).

(vii) \( \forall \) pairwise disjoint \( (E_i)_{i=1}^{\infty} \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \) and \( E_n := \bigcup_{i=1}^{n} E_i \in \mathcal{B}, \forall n \in \mathbb{N} \). Then, \( \lim_{n \to \infty} \int_{E_n} f_1 \, d\mu = \lim_{n \to \infty} \int_{E_n} f_1 \, d\mu = \int_{\bigcup_{i=1}^{\infty} E_i} f_1 \, d\mu = \int_{\bigcup_{i=1}^{n} E_i} f_1 \, d\mu = \int_X f_1 \, d\mu \in \mathbb{W} \), where the first equality follows from (v); the second equality follows from (ii); the third equality follows from the 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 the Lebesgue Dominated Convergence Theorem 11.128.

This completes the proof of the proposition.

\\[ \square \]

**Theorem 11.131 (Lebesgue Dominated Convergence)** Let \( (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over \( \mathcal{K} \), \( \mathcal{Z} \) be a Banach space over \( \mathcal{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B} \) \( \{ \mathcal{Y}, \mathcal{Z} \} \), \( \mathcal{W} := \text{span} \{ \{ z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} : z = wy \} \} \subseteq \mathcal{Z} \) 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, \( g_n : X \to [0, \infty) \subseteq \mathcal{R} \) be \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \), and \( g : X \to [0, \infty) \subseteq \mathcal{R} \) be \( \mathcal{B} \)-measurable. Assume that

(i) \( \lim_{n \to \infty} f_n = f \) a.e. in \( X \), \( \lim_{n \to \infty} g_n = g \) a.e. in \( X \);
(ii) \( \| f_n(x) \| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}; \)

(iii) \( \int_X g_n \, d\mathcal{P} \circ \mu = \mathbb{R}, \forall n \in \mathbb{N} \) and \( \int_X g \, d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_X g_n \, d\mathcal{P} \circ \mu \in \mathbb{R}. \)

Then, \( f \) is integrable over \( X, \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \tilde{\mathcal{W}} \subseteq \mathbb{Z}, \) 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 \( X, \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{W}, \) 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 \( X, \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f_n \, d\mu \in \mathcal{W}, \) 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 \( (F_n,A)_{A \in \mathcal{M}(X)} \) 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 \( A := (A, B, \mathcal{B}, \mathcal{P} \circ \mu) \) 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 \( A := (A, B, \mathcal{P} \circ \mu), \) which is the finite measure subspace of \( X := (X, B, \mathcal{P} \circ \mu) \) as shown in Proposition 11.115. By Proposition 11.83, \( \lim_{n \in \mathbb{N}} g_{\mid A} = g_{\mid A} \) a.e. in \( A \) and \( \int_A g_{\mid A} \, d\mathcal{P} \circ \mu = \lim_{n \in \mathbb{N}} \int_A g_{n_{\mid A}} \, d\mathcal{P} \circ \mu \in \mathbb{R}. \) By Lebesgue Dominated Convergence Theorem 11.128, we have \( f_{\mid A} \) and \( f_{n_{\mid A}} \)'s are integrable over \( A, \int_A f_{\mid A} \, d\mu_A = \lim_{n \in \mathbb{N}} \int_A f_{n_{\mid A}} \, d\mu_A \in \mathcal{W}, \) and \( 0 \leq \| \int_A f_{\mid A} \, d\mu_A \| \leq \int_A g_{\mid A} \, d\mathcal{P} \circ \mu < +\infty. \) 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 \( (F_{n,A})_{A \in \mathcal{M}(X)} \)'s are well defined.

\( \forall \varepsilon \in (0, \infty) \subset \mathbb{R}, \) by Definition 11.79, \( \exists \) a simple function \( \phi : X \rightarrow [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi(x) \leq g(x), \forall x \in X, \) and \( \int_X g \, d\mathcal{P} \circ \mu - \varepsilon/5 < \int_X \phi \, d\mathcal{P} \circ \mu \leq \int_X g \, d\mathcal{P} \circ \mu < +\infty. \) Let \( A_0 := \{ x \in X \mid \phi(x) > 0 \}. \) 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 \( A \subseteq A_0 , \) we have \( 0 \leq \int_{A \setminus A_0} g_{\mid A_0} \, d\mathcal{P} \circ \mu_{A \setminus A_0} = \int_A g_{\mid A} \, d\mathcal{P} \circ \mu - \int_{A_0} g_{\mid A_0} \, d\mathcal{P} \circ \mu_{A_0} \leq \int_X g \, d\mathcal{P} \circ \mu - \int_{A_0} \phi_{\mid A_0} \, d\mathcal{P} \circ \mu_{A_0} = \int_X g \, d\mathcal{P} \circ \mu - \int_X \phi \, d\mathcal{P} \circ \mu < \varepsilon/5, \) where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75. Note that \( \| F_A - F_{A_0} \| = \| \int_A f_{\mid A_0} \, d\mu_A - \int_{A_0} f_{\mid A_0} \, d\mu_{A_0} \| = \| \int_{A \setminus A_0} f_{\mid A_0} \, d\mu_{A \setminus A_0} \| \)

\( \leq \int_{A \setminus A_0} g_{\mid A \setminus A_0} \, d\mathcal{P} \circ \mu_{A \setminus A_0} < \varepsilon/5, \) 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, \( 0 \leq \| \int_X f \, d\mu \| = \lim_{A \in \mathcal{M}(X)} \| F_A \| \leq \lim_{A \in \mathcal{M}(X)} \int_A g_{\mid A} \, d\mathcal{P} \circ \mu = \int_X g \, d\mathcal{P} \circ \mu < +\infty. \)
∀n ∈ N, by Definition 11.79, \( \exists \) a simple function \( \phi_n : X \rightarrow [0, \infty) \subset \mathbb{R} \) such that \( 0 \leq \phi_n(x) \leq g_n(x), \forall x \in X \), and \( \int_X g_n \, d\mu < \int_X \phi_n \, d\mu < \int_X g_n \, d\mu + \epsilon/5 < \int_X \phi_n \, d\mu < \int_X g_n \, d\mu < \infty \). Let \( A_n := \{ x \in X \mid \phi_n(x) > 0 \} \).

Then, \( A_n \in \mathcal{F}, \mathcal{P} \circ \mu(A_n) < +\infty \), and \( A_n \in \mathcal{M}(\mathcal{X}) \). \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_n \subseteq A \), we have \( 0 \leq \int_{A \setminus A_n} g_n |_{A \setminus A_n} \, d\mu \circ \mu \circ A_n - \int_{A_n} g_n |_{A_n} \, d\mu \circ A_n \leq \int_X g_n \, d\mu < \int_X g_n \, d\mu < \epsilon/5 \), where the first equality and the second inequality follow from Proposition 11.83; and the second equality follows from Proposition 11.75.

Note that \( \| F_n \cdot A - F_n \cdot A_n \| = \int_{A \setminus A_n} g_n |_{A \setminus A_n} \, d\mu \circ A_n - \int_{A_n} g_n |_{A_n} \, d\mu \circ A_n < \epsilon/5 \), where the second equality and the first inequality follow from Proposition 11.130. Then, the net \( \{ F_n \cdot A \}_{A \in \mathcal{M}(\mathcal{X})} \) is a Cauchy net, which admits a limit, by Proposition 4.44. \( \int_X f \, d\mu = \lim_{A \in \mathcal{M}(\mathcal{X})} F_n \cdot A \in \mathcal{W} \).

Let \( \tilde{g}_n := g_n \wedge g, \forall n \in \mathbb{N} \). Then, \( 0 \leq \tilde{g}_n(x) \leq g(x), \forall x \in X \), and \( \tilde{g}_n \) is \( \mathcal{B} \)-measurable by Proposition 11.40, \( \forall n \in \mathbb{N} \). Then, \( \lim_{n \in \mathbb{N}} \tilde{g}_n = g \) a.e. in \( X \) by Proposition 11.50. By Fatou's Lemma 11.80, \( \int_X g \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X \tilde{g}_n \, d\mu \leq \int_X g \, d\mu \leq \limsup_{n \in \mathbb{N}} \int_X \tilde{g}_n \, d\mu \leq \int_X g \, d\mu < +\infty \). Therefore, by Proposition 3.83, \( \int_X g \, d\mu = \lim_{n \in \mathbb{N}} \int_X \tilde{g}_n \, d\mu \). This coupled with (iii) and the fact that \( F_n \cdot A = \lim_{n \in \mathbb{N}} F_n \cdot A_n \), implies that \( \exists \tilde{n}_0 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n_0 \leq n \), we have \( 0 \leq \int_X g \, d\mu - \int_X g \, d\mu < \epsilon/5 \), \( \int_X g \, d\mu - \int_X g \, d\mu < \epsilon/5 \), and \( \| F_n \cdot A - F_n \cdot A_n \| < \epsilon/5 \), \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_n \subseteq A \), we have \( 0 \leq \int_{A \setminus A_n} g_n |_{A \setminus A_n} \, d\mu \circ A_n - \int_{A_n} g_n |_{A_n} \, d\mu \circ A_n < \epsilon/5 \), and \( \| F_n \cdot A - F_n \cdot A_n \| < \epsilon/5 \), \( \forall A \in \mathcal{M}(\mathcal{X}) \) with \( A_n \subseteq A \), we have \( 0 \leq \int_{A \setminus A_n} g_n |_{A \setminus A_n} \, d\mu \circ A_n - \int_{A_n} g_n |_{A_n} \, d\mu \circ A_n < \epsilon/5 \), the second equality, the second inequality, and the third inequality follow from Proposition 11.83.

Furthermore, \( \| F_n \cdot A - F_n \cdot A_n \| = \int_{A \setminus A_n} g_n |_{A \setminus A_n} \, d\mu \circ A_n - \int_{A_n} g_n |_{A_n} \, d\mu \circ A_n < \epsilon/5 \), where the second equality and the first inequality follow from Proposition 11.130. This implies that \( \| F_n \cdot A - F_n \cdot A_n \| \leq \int_X f \, d\mu - \int_X f \, d\mu \leq \| F_n \cdot A - F_n \cdot A_n \| < \epsilon/5 \), \( \| F_n \cdot A - F_n \cdot A_n \| < \epsilon/5 \), and \( \| F_n \cdot A - F_n \cdot A_n \| < \epsilon/5 \), where the second inequality follows from Propositions 7.21 and 3.66. Hence, \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f \, d\mu \in \mathcal{W} \).

Then, in both cases, we have shown that \( f \) is integrable over \( \mathcal{X} \), \( \int_X f \, d\mu = \lim_{n \in \mathbb{N}} \int_X f \, d\mu \in \mathcal{W} \), and \( 0 \leq \| \int_X f \, d\mu \| \leq \int_X g \, d\mu < +\infty \).

Note that \( 0 \leq \mathcal{P} \circ f_n(x) \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N} \), \( \lim_{n \in \mathbb{N}} \mathcal{P} \circ f_n = \mathcal{P} \circ f \) a.e. in \( \mathcal{X} \) by Propositions 7.21 and 11.52, and \( \mathcal{P} \circ f \) and \( \mathcal{P} \circ f_n \)'s are \( \mathcal{B} \)-measurable by Propositions 7.21 and 11.38. Then, by Lebesgue Dominated Convergence Theorem 11.91, \( \int_X \mathcal{P} \circ f \, d\mu = \lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ f_n \, d\mu \in \mathcal{R} \). Hence, \( f \) is absolutely integrable over \( \mathcal{X} \).
This completes the proof of the theorem. 

\textbf{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 $\mathbb{K}$, $\tilde{\mathcal{W}} := \text{span}\{z \in \mathcal{Z} \mid \exists y \in \mathcal{Y}, \exists w \in \mathcal{W} \text{ s.t. } z = wy\} \subseteq \mathcal{Z}$ 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) $\forall \mathcal{H} \in \mathcal{B}$, 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, $\int_{H} f_1 d\mu_H$ is absolutely integrable over $\mathcal{H}$ and $\int_X f_1 d\mu_H = \int_H f_1 d\mu_H \in \mathcal{W}$. We will henceforth denote $\int_H f_1 d\mu_H$ 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) $0 \leq \left\| \int_X f_1 d\mu \right\| \leq \int_X \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < +\infty$.

\textbf{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 $(F_{i,E})_{E \in \mathcal{M}(X)}$ 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, $F_{i,E} = \int_E f_i d\mu_E \in \tilde{\mathcal{W}}$, $i = 1, 2$. Then, the net $(F_{i,E})_{E \in \mathcal{M}(X)}$ 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 $\lim_{E \in \mathcal{M}(X)} F_{i,E} = \int_X f_i d\mu \in \tilde{\mathcal{W}}$, $i = 1, 2$.

(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 $\mathcal{P} \circ (f_1 + f_2)(x) = \| f_1(x) + f_2(x) \| \leq \| f_1(x) \| + \| f_2(x) \| = (\mathcal{P} \circ f_1 + \mathcal{P} \circ f_2)(x)$, $\forall x \in X$. 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. For all $E \in \mathcal{M}(\mathcal{X})$, $\tilde{F}_E = \int_E (f_1 + f_2) \, d\mu_E = \int_E f_1 \, d\mu_E + \int_E f_2 \, d\mu_E = F_{1,E} + F_{2,E} \in \mathcal{W}$, 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, $A f_1$ and $\mathcal{P} \circ (A f_1)$ are $\mathcal{B}$-measurable. Note that $\mathcal{P} \circ (A f_1)(x) = \|A f_1(x)\| \leq \|A\| \|f_1(x)\| = \|A\| \mathcal{P} \circ f_1(x)$, for all $x \in X$. By Proposition 11.83 and the fact that $\int_X \mathcal{P} \circ f_1 \, d\mathcal{P} \circ \mu \in \mathbb{R}$, 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 also that $A f_1 : X \to \mathcal{W}$, 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 $B(\mathcal{Y}, \mathcal{U})$. By (i), $\int_X (A f_1) \, d\mu \in \mathcal{U}$. Let $(\tilde{F}_E)_{E \in \mathcal{M}(\mathcal{X})}$ be the net for $(A f_1) \, d\mu$ as defined in Definition 11.120. For all $E \in \mathcal{M}(\mathcal{X})$, $\tilde{F}_E = \int_E (A f_1) \, d\mu_E = A \int_E f_1 \, d\mu_E = A F_{1,E} \in \mathcal{U}$, where the second equality follows from Proposition 11.130. Then, by Propositions 3.66 and 7.62, we have $\int_X (A f_1) \, d\mu = \lim_{E \in \mathcal{M}(\mathcal{X})} \tilde{F}_E = \lim_{E \in \mathcal{M}(\mathcal{X})} A F_{1,E} = A \int_X f_1 \, d\mu \in \mathcal{U}$. Hence, $A f_1$ 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, $0 \leq \int_H \mathcal{P} \circ (f_1|_H) \, d\mathcal{P} \circ \mu_H \leq \int_X (\mathcal{P} \circ f_1) \, d\mathcal{P} \circ \mu < +\infty$. 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, $f_1|_{\mathcal{X}|_H}$ is $\mathcal{B}$-measurable. By Proposition 11.83, $f_1|_{\mathcal{X}|_H}$ is absolutely integrable over $\mathcal{X}$. Again, by (i), $f_1|_{\mathcal{X}|_H}$ is integrable over $\mathcal{X}$. For all $h \in (0, \infty) \subset \mathbb{R}$, $\exists A_0 \in \mathcal{M}(\mathcal{X})$ such that $\forall A \in \mathcal{M}(\mathcal{X})$ with $A_0 \subset A$, we have $\left\| \int_A (f_1|_{\mathcal{X}|_H}) \, d\mu_A - \int_X (f_1|_{\mathcal{X}|_H}) \, d\mu \right\| < \epsilon/2$. 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 $E_0 \subset E$, we have $\left\| \int_E f_1 \, d\mu_E - \int_H f_1|_H \, d\mu_H \right\| < \epsilon/2$. 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, $\int_{A_1} (f_1|_{\mathcal{X}|_H}) \, d\mu_{A_1} = \int_{H \cap A_1} f_1 \, d\mu_{H \cap A_1} \leq \int_{E_1} f_1 \, d\mu_{E_1}$, where the second equality follows from Proposition 11.130. Then, we have $\left\| \int_X (f_1|_{\mathcal{X}|_H}) \, d\mu - \int_H f_1|_H \, d\mu_H \right\| \leq \left\| \int_X (f_1|_{\mathcal{X}|_H}) \, d\mu - \int_{A_1} (f_1|_{\mathcal{X}|_H}) \, d\mu_{A_1} \right\| + \left\| \int_{E_1} f_1 \, d\mu_{E_1} - \int_H f_1|_H \, d\mu_H \right\| < \epsilon/2 + \epsilon/2 = \epsilon$.

By the arbitrariness of $\epsilon$, we have $\int_{A_1} (f_1|_{\mathcal{X}|_H}) \, d\mu_{A_1} = \int_H f_1|_H \, d\mu_H \in \mathcal{W}$.

Case 2b: $\mathcal{P} \circ \mu(H) < +\infty$. Let $A_1 := A_0 \cup H \in \mathcal{M}(\mathcal{X})$. By Proposition 11.130, $\int_{A_1} (f_1|_{\mathcal{X}|_H}) \, d\mu_{A_1} = \int_H f_1|_H \, d\mu_H$. Then, we have $\left\| \int_X (f_1|_{\mathcal{X}|_H}) \, d\mu - \int_H f_1|_H \, d\mu_H \right\| = \left\| \int_X (f_1|_{\mathcal{X}|_H}) \, d\mu - \int_{A_1} (f_1|_{\mathcal{X}|_H}) \, d\mu_{A_1} \right\| < \epsilon/2$. By the arbitrariness of $\epsilon$, we have $\int_X (f_1|_{\mathcal{X}|_H}) \, d\mu = \int_H f_1|_H \, d\mu_H \in \mathcal{W}$. 

444

CHAPTER 11. GENERAL MEASURE AND INTEGRATION
Hence, in both subcases, $\int_X (f_1 \chi_{E,x}) \, d\mu = \int_X f_1 \, d\mu = H \in \widetilde{W}$.

(vi) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

(vii) $∀$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq B$, let $E := \bigcup_{i=1}^{\infty} E_i \in B$ and $E_n := \bigcup_{i=1}^{n} E_i \in B$, $∀n \in \mathbb{N}$. Then, $\lim_{n \to \infty} \int_X E_n \chi_{E_n,X} = \int_X f_1 \chi_{E,X}$, $∀x \in X$. Then, we have $\sum_{i=1}^{\infty} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \int_{E_i} f_1 \, d\mu = \lim_{n \to \infty} \int_X (f_1 \chi_{E_n,X}) \, d\mu = \lim_{n \to \infty} \int_X (f_1 \chi_{E,X}) \, d\mu = \int_X f_1 \, d\mu = \int_X f \, d\mu$ in $\mathbb{W}$, where the first equality follows from (vii); and the last inclusion follows from (vii); the second equality follows from (vii); the third equality follows from Lebesgue Dominated Convergence Theorem 11.131; the last equality follows from (vii); and the last inclusion follows from (vii).

(viii) This follows immediately from Lebesgue Dominated Convergence Theorem 11.131.

This completes the proof of the proposition. □

**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 $\mathbb{K}$, $Z$ 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 absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}$, $\varphi_n : X \to \mathcal{W}$, $∀n \in \mathbb{N}$, such that $∥\varphi_n(x)∥ ≤ \varphi \circ f(x)$, $∀x \in X$, $∀n \in \mathbb{N}$, $\lim_{n \to \infty} \int_X \varphi_n \, d\mu = \int_X f \, d\mu$ in $\mathcal{Z}$, $\lim_{n \to \infty} \int_X \varphi_n \circ P \, dP \circ \mu = \int_X \varphi \circ f \, dP \circ \mu$ in $[0, \infty) \subset \mathbb{R}$, and $\lim_{n \to \infty} \int_X \varphi \circ (\varphi_n - f) \, dP \circ \mu = 0$.

**Proof** By Proposition 11.66, there exists a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}$, $\varphi_n : X \to \mathcal{W}$, $∀n \in \mathbb{N}$, such that $∥\varphi_n(x)∥ ≤ \varphi \circ f(x)$, $∀x \in X$, $∀n \in \mathbb{N}$, and $\lim_{n \to \infty} \varphi_n = f$ a.e. in $\mathcal{X}$, by Lemma 11.43, $\varphi \circ (\varphi_n - f)$ is $\mathcal{B}$-measurable. By Propositions 7.21, 7.23, 11.53, and 11.52, we have $\lim_{n \to \infty} \int_X (\varphi \circ (\varphi_n - f)) \, dP \circ \mu = 0$ a.e. in $\mathcal{X}$. Note that $0 ≤ P \circ (\varphi_n - f)(x) = ∥\varphi_n(x) - f(x)∥ ≤ ∥\varphi_n(x)∥ + ∥f(x)∥ ≤ 2\varphi \circ f(x)$, $∀x \in X$, $∀n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.91 and Proposition 11.75, we have $\lim_{n \to \infty} \int_X P \circ (\varphi_n - f) \, dP \circ \mu = 0$. This completes the proof of the proposition. □

**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 $\mathbb{K}$, $Z$ be a Banach space over $\mathbb{K}$, $W$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, and $f : X \to W$ be absolutely integrable over $\mathcal{X}$. Then, there exists a sequence of simple functions $(\phi_n)_{n=1}^{\infty}$, $\phi_n : X \to \mathcal{W}$, $∀n \in \mathbb{N}$, such that $∥\phi_n(x)∥ ≤ \varphi \circ f(x)$, $∀x \in X$, $∀n \in \mathbb{N}$, $\lim_{n \to \infty} \int_X \phi_n \, d\mu = \int_X f \, d\mu$ in $\mathcal{Z}$, $\lim_{n \to \infty} \int_X \phi_n \circ P \, dP \circ \mu = \int_X \varphi \circ f \, dP \circ \mu$ in $[0, \infty) \subset \mathbb{R}$, and $\lim_{n \to \infty} \int_X \phi \circ (\phi_n - f) \, dP \circ \mu = 0$.

**Proof** Let $\mathcal{X} := (X, \mathcal{B}, P \circ \mu)$, which is a measure space. By the assumption, $f$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.95, 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 \(\|\phi_n(x)\| \leq \mathcal{P} \circ f(x), \forall x \in X, \forall n \in \mathbb{N}\), \(\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \in [0, \infty) \subset \mathbb{R}\), and \(\lim_{n \in \mathbb{N}} \int_X \mathcal{P} \circ (\phi_n - f) \, d\mathcal{P} \circ \mu = 0\).

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, \(\lim_{n \in \mathbb{N}} \int_X \phi_n \, d\mu - \int_X f \, d\mu\) equals \(\lim_{n \in \mathbb{N}} \int_X (\phi_n - f) \, d\mu\), with \(\mathcal{P} \circ (\phi_n - f) \, d\mathcal{P} \circ \mu = 0\). 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, \(\mathcal{P} \circ \bar{\mu} = \mu\). 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 \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite \(\mathbb{R}\)-valued measure space with \(\nu := \mathcal{P} \circ \mu\) and \(\mu(E) \geq 0, \forall E \in \text{dom} (\mu)\). Then, any function \(\bar{\nu} : \mathcal{B} \to \mathbb{R}\) satisfying \(\bar{\nu}(E) = \sum_{i=1}^{\infty} \bar{\nu}(E_i), \forall\) 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

\[
\nu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \cup_{i=1}^{n} E_i = E, \forall i, j \geq n, E_i \cap E_j = \emptyset} \sum_{i=1}^{n} |\mu(E_i)| = \mu(E) = \bar{\nu}(E)
\]

Case 2: \(E \in \mathcal{B} \setminus \text{dom} (\mu)\). Then, \(\nu(E) = \infty\). Since \(\mathcal{X}\) 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, \(E = \bigcup_{n=1}^{\infty} (E \cap X_n), 0 \leq \nu(E \cap X_n) \leq \nu(X_n) < \infty, E \cap X_n \in \text{dom} (\mu), \forall n \in \mathbb{N}\). 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 \( \alpha \in \mathbb{K} \). Define \( \nu := \nu_1 + \nu_2 : \mathcal{B} \to \mathcal{Y} \) by \( \nu(E) = \nu_1(E) + \nu_2(E) \), \( \forall E \in \mathcal{B} \). Then, \( \nu \) is a finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \nu \leq \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu_2 \). 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 \( \mathcal{P} \circ \bar{\nu} = |\alpha| \mathcal{P} \circ \nu_1 \).

(iv) Let \( \mu \) be a finite \( \mathcal{K} \)-valued measure on \((X, \mathcal{B})\) and \( y \in \mathcal{Y} \). Define \( \nu := \mu y : \mathcal{B} \to \mathcal{Y} \) by \( \nu(E) = \mu(E)y \), \( \forall E \in \mathcal{B} \). Then, \( \nu \) is a finite \( \mathcal{Y} \) valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \nu = \|y\| \mathcal{P} \circ \mu \).

(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 \( \mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu \).

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, \( \sum_{i=1}^{\infty} a\mu_1(X_n) = \sum_{i=1}^{\infty} a\mu_1(E_i) = a \sum_{i=1}^{\infty} \mu_1(E_i) = a \mu(E) \), where the second equality follows from the fact that \( \sum_{i=1}^{\infty} \mu_1(E_i) = \mu(E) \leq \mu(X) < +\infty \). Hence, \( a\mu_1 \) is \( \sigma \)-finite.

(ii) Clearly, \( \bar{\mu}(\emptyset) = a\mu_1(\emptyset) = 0 \). \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^\infty E_i \in \mathcal{B} \). Then, \( \sum_{i=1}^{\infty} (\nu_1 + \nu_2)(E_i) = \sum_{i=1}^{\infty} (\nu_1(E_i) + \nu_2(E_i)) \leq \sum_{i=1}^{\infty} \nu_1(E_i) + \sum_{i=1}^{\infty} \nu_2(E_i) = +\infty \), where the last inequality follows from the fact that \( \nu_1 \) and \( \nu_2 \) are finite \( \mathcal{Y} \)-valued measures, and \( \sum_{i=1}^{\infty} (a\nu_1)(E_i) = \sum_{i=1}^{\infty} |\alpha| \nu_1(E_i) = |\alpha| \sum_{i=1}^{\infty} \nu_1(E_i) \leq +\infty \).

(iii) Clearly, \( (\nu_1 + \nu_2)(\emptyset) = \nu_1(\emptyset) + \nu_2(\emptyset) = \emptyset y \) and \( (a\nu_1)(\emptyset) = a\nu_1(\emptyset) = \emptyset y \). \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq \mathcal{B} \), let \( E := \bigcup_{i=1}^\infty E_i \in \mathcal{B} \). Then, \( \sum_{i=1}^{\infty} \nu_1(E_i) \leq \sum_{i=1}^{\infty} \nu_2(E_i) \) and \( \sum_{i=1}^{\infty} (a\nu_1)(E_i) \leq |\alpha| \sum_{i=1}^{\infty} \nu_1(E_i) \).
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 $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$, $\forall n \in \mathbb{Z}_+$, $\forall$ pairwise disjoint $(E_i)_{i=1}^{n} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{n} E_i$, we have $\sum_{i=1}^{n} \| (\nu_1 + \nu_2)(E_i) \| \leq \sum_{i=1}^{n} \| \nu_1(E_i) \| + \sum_{i=1}^{n} \| \nu_2(E_i) \| \leq P \circ \nu_1(E) + P \circ \nu_2(E)$. Hence, $P \circ (\nu_1 + \nu_2)(E) \leq P \circ \nu_1(E) + P \circ \nu_2(E)$. Then, $P \circ \nu \leq P \circ \nu_1 + P \circ \nu_2$. Since $\nu_1$ and $\nu_2$ are finite, then $P \circ \nu(X) \leq P \circ \nu_1(X) + P \circ \nu_2(X) < +\infty$. Therefore, $\nu_1 + \nu_2$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$, we have

$$P \circ (\alpha \nu_1)(E) = \sup_{n \in \mathbb{Z}_+, \ (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq n} \left| \alpha \right| \sum_{i=1}^{n} \| \nu_1(E_i) \| = \left| \alpha \right| P \circ \nu_1(E)$$

where the last equality follows from Proposition 3.81 and the fact that $0 \leq P \circ \nu_1(E) \leq P \circ \nu_1(X) < +\infty$. Hence, $P \circ (\alpha \nu_1) = \left| \alpha \right| P \circ \nu_1$. Therefore, $\alpha \nu_1$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$.

(iv) Clearly, $(\mu y)(\emptyset) = \mu(\emptyset) y = \emptyset y$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\sum_{i=1}^{\infty} \| (\mu y)(E_i) \| = \sum_{i=1}^{\infty} \| y \| \| \mu(E_i) \| = \| y \| \sum_{i=1}^{\infty} \| \mu(E_i) \| < +\infty$. Note that $\sum_{i=1}^{\infty} (\mu y)(E_i) = \sum_{i=1}^{\infty} \mu(E_i)y = \mu(E)y = (\mu y)(E) \in \mathcal{Y}$, 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 $(X, \mathcal{B})$.

$\forall E \in \mathcal{B}$,

$$P \circ (\mu y)(E) = \sup_{n \in \mathbb{Z}_+, \ (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq n} \left\| \sum_{i=1}^{n} (\mu y)(E_i) \right\| = \left\| y \right\| \left\| \sum_{i=1}^{\infty} \mu(E_i) \right\| = \left\| y \right\| P \circ \mu(E)$$

where the last equality follows from Proposition 3.81 and the fact that $0 \leq P \circ \mu(E) \leq P \circ \mu(X) < +\infty$. Hence, $P \circ (\mu y) = \| y \| P \circ \mu$. Therefore, $\mu y$ is a finite $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$.

(v) Clearly, $\bar{\mu}(\emptyset) = A \mu(\emptyset) = \emptyset \emptyset$. $\forall$ pairwise disjoint $(E_i)_{i=1}^{\infty} \subseteq \mathcal{B}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \in \mathcal{Y}$, and $\sum_{i=1}^{\infty} \| \mu(E_i) \| < +\infty$. This implies that $\bar{\mu}(E) = A \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} A \mu(E_i) = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \in \mathcal{Z}$, 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 \( \mathbb{Z} \)-valued pre-measure on \( (X, \mathcal{B}) \). \( \forall E \in \mathcal{B}, \)

\[
P \circ \tilde{\mu}(E) = \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| A \| \| \mu(E_i) \|
\]

\[
\leq \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| A \| \| \mu(E_i) \|
\]

\[
= \| A \| \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(E_i) \|
\]

\[
= \| A \| P \circ \mu(E) \leq \| A \| P \circ \mu(X) < \infty
\]

where the second equality follows from Proposition 3.81. Hence, \( P \circ \tilde{\mu} \leq \| A \| P \circ \mu \) and \( \tilde{\mu} \) is a finite \( \mathbb{Z} \)-valued pre-measure on \( (X, \mathcal{B}) \). Then, \( \tilde{\mu} \) is a finite \( \mathbb{Z} \)-valued measure on \( (X, \mathcal{B}) \).

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

**Proposition 11.137** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathcal{Y} \) be a normed linear space over \( \mathbb{K} \), and \( \nu_1 \) and \( \nu_2 \) be \( \mathcal{Y} \)-valued measures on \( (X, \mathcal{B}) \). Assume that, \( \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \), we have \( \nu_1(E) = \nu_2(E) \in \mathcal{Y} \). 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 \( P \circ \nu_1(X_n) < \infty \), \( i = 1, 2 \). Hence, \( \nu_1 \) and \( \nu_2 \) are \( \sigma \)-finite. Let \( \mathcal{X}_n := (X_n, \mathcal{B}_n, \nu_{n,i}) \) be the finite measure subspace of \( \mathcal{X} \), \( (X_n, \mathcal{B}_n, \nu_{n,i}) \) be the finite \( \mathcal{Y} \)-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, \( \nu_{n,1}(E) = \nu_1(E) = \nu_{n,2}(E) \). Thus, \( \nu_{n,1} = \nu_{n,2} \). By Proposition 11.117, we have \( \nu_1 = \nu_2 \). \( \square \)

**Proposition 11.138** Let \( (X, \mathcal{B}) \) be a measurable space, \( \mathcal{Y} \) 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, \( P \circ \mu \leq P \circ \mu_1 + P \circ \mu_2 \).

(ii) Let \( \mu \) be a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \) and \( \alpha \in [0, \infty) \subset \mathbb{R} \). 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 $\bar{\mu}(E) = \emptyset y$, $\forall E \in \mathcal{B}$; and $\mathcal{P} \circ (\alpha \mu) = |\alpha| \mathcal{P} \circ \mu$.

(iv) Let $\mu$ be a $\sigma$-finite $K$-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 $y \neq \emptyset y$, 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 $y = \emptyset y$, then $\bar{\mu}(E) = \emptyset y$, $\forall E \in \mathcal{B}$; and $\mathcal{P} \circ (\mu y) = \|y\| \mathcal{P} \circ \mu$.

(v) Let $\mu$ be a $\sigma$-finite $\mathcal{Y}$-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 $\mathcal{P} \circ \bar{\mu} \leq \|A\| \mathcal{P} \circ \mu$. 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 $\mathcal{P} \circ \mu_i(X_n) < \infty$, $\forall n \in \mathbb{N}$, $i = 1, 2$. 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 $(X, B, \mu_i)$, $\forall n \in \mathbb{N}$, $i = 1, 2$. 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 $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$, $\forall E \in \mathcal{B}$, $\mathcal{P} \circ \mu(E) \in [0, \infty) \subset \mathbb{R}$.

We will show that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E)$ by distinguishing two exhaustive and mutually exclusive cases: Case 1: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$; Case 2: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$. Case 1: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$. Then, it is trivially true that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) = \infty$. Case 2: $\mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$. Then, $\mathcal{P} \circ \mu(E) \in [0, \infty) \subset \mathbb{R}$, $i = 1, 2$.

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, $\mathcal{P} \circ \mu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ (\mu_{1,n} + \mu_{2,n})(E_n) \leq \sum_{n=1}^{\infty} (\mathcal{P} \circ \mu_{1,n}(E_n))^+ + \mathcal{P} \circ \mu_{2,n}(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_1(E_n) + \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_2(E_n) = \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$, 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, $\mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2$. 

$\forall E \in \text{dom}(\mu_1) \cap \text{dom}(\mu_2)$. Then, $\mathcal{P} \circ \mu_1(E) < \infty$ and $\mu_1(E) \in \mathcal{Z}$, $i = 1, 2$. This implies that $\mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu_1(E) + \mathcal{P} \circ \mu_2(E) < \infty$, $E \in \text{dom}(\mu)$, and $\mu(E) \in \mathcal{Z}$. Let $E_n := E \cap X_n \in B_n$, $\forall n \in \mathbb{N}$. Then, $\mu(E) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} (\mu_{1,n} + \mu_{2,n})(E_n) = \sum_{n=1}^{\infty} (\mu_1(E_n))^+$.
\( \mu_2(E_n) = \sum_{n=1}^{\infty} \mu_1(E_n) + \sum_{n=1}^{\infty} \mu_2(E_n) = \mu_1(E) + \mu_2(E) \in \mathcal{Z}, \) 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 \( \bar{\mu} : \mathcal{B} \to [0, \infty] \subset \mathbb{R} \) by \( \bar{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B}, \) if \( \alpha > 0 \)

and \( \bar{\mu}(E) = 0, \forall E \in \mathcal{B}, \) if \( \alpha = 0 \). Then, \( \bar{\mu} \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \) by Proposition 11.136. When \( \alpha = 0 \), clearly \( \bar{\mu} \) is a finite

measure on \( (X, \mathcal{B}) \). Hence, \( \bar{\mu} \) is the \( \sigma \)-finite measure we seek. Let \( \bar{\mu} \) be any

measure on \( (X, \mathcal{B}) \) such that \( \bar{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B} \) with \( \alpha \mu(E) < \infty \).

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. \( \forall E, B \in \mathcal{B}, \) 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 \( \bar{\mu}(E) = \sum_{n=1}^{\infty} \bar{\mu}(E_n) = \sum_{n=1}^{\infty} \mu(E_n) = \sum_{n=1}^{\infty} \mu(E_n) = \bar{\mu}(E) \in [0, \infty] \subset \mathbb{R} \). Hence, \( \bar{\mu} \) is unique.

(iii) Define \( \nu := \alpha \mathcal{P} \circ \mu \) as prescribed in (ii), which is a \( \mu \)-\( \sigma \)-finite measure on \( (X, \mathcal{B}) \). Define \( \bar{\mu} \) to be a function from \( \mathcal{B} \) to \( \mathcal{Y} \) by, if \( \alpha \neq 0 \), then \( \bar{\mu}(E) = \alpha \mu(E) \in \mathcal{Y}, \forall E \in \mathcal{B} \) \( \text{dom} (\mu) \), and \( \bar{\mu} \) is undefined, \( \forall E \in \mathcal{B} \setminus \text{dom} (\mu) \);

if \( \alpha = 0 \), then \( \bar{\mu}(E) = \bar{\nu}_Y, \forall E \in \mathcal{B} \). We will show that \( \bar{\mu} \) is a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) satisfying: \( \bar{\mu}(E) = \alpha \mu(E), \forall E \in \mathcal{B} \) \( \text{dom} (\mu) \), and \( \mathcal{P} \circ \bar{\mu} = \nu \) by distinguishing two exhaustive and mutually exclusive cases: Case 2: \( \alpha \neq 0 \).

Case 1: \( \alpha = 0 \). Then, \( \bar{\mu}(E) = \bar{\nu}_Y \) and \( \nu(E) = 0, \forall E \in \mathcal{B} \). It is easy to show that \( \bar{\mu} \) is a finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}) \) with \( \mathcal{P} \circ \bar{\mu} = \nu \). It is straightforward to check that \( \bar{\mu}(E) = 0 \mu(E) = \bar{\nu}_Y, \forall E \in \mathcal{B} \). 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 \( \bar{\mu}(E) \) is also undefined. \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty \), then \( \mathcal{P} \circ \mu(E) = \nu(E) \) \( |\alpha| < \infty \) and \( \bar{\mu}(E) \in \mathcal{B} \).

Then, \( \bar{\mu}(E) = \alpha \mu(E) \in \mathcal{Y}, \forall \mathcal{P} \)-pairwise disjoint \( \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^{\infty} E_i \), we have \( \bar{\mu}(E) = \alpha \mu(E) = \alpha \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \alpha \mu(E_i) = \sum_{i=1}^{\infty} \bar{\mu}(E_i) \in \mathcal{Y} \) and \( \sum_{i=1}^{\infty} \| \bar{\mu}(E_i) \| = \sum_{i=1}^{\infty} |\alpha| \| \mu(E_i) \| \leq \sum_{i=1}^{\infty} |\alpha| \| \mathcal{P} \circ \mu(E_i) \| = |\alpha| \| \mathcal{P} \circ \mu(E) \| = \nu(E) < \infty \). \( \forall E \in \mathcal{B} \) with \( \nu(E) < \infty \),

\[
\nu(E) = |\alpha| \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{Z}_+} \left( \sum_{i=1}^{n} \| \mu(E_i) \| \right) = \sup_{n \in \mathbb{Z}_+} \left( \sum_{i=1}^{n} \| \bar{\mu}(E_i) \| \right) = \sup_{n \in \mathbb{Z}_+} \left( \sum_{i=1}^{n} \| \bar{\mu}(E_i) \| \right)
\]
where the third equality follows from Proposition 3.81. Hence, \( \bar{\mu} \) is a \( \sigma \)-finite \( Y \)-valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \bar{\mu} = \nu = |\alpha| \mathcal{P} \circ \mu \). This case is proved.

Hence, \( \bar{\mu} \) is the \( \sigma \)-finite \( Y \)-valued measure on \((X, \mathcal{B})\) we seek. Let \( \bar{\mu} \) be any \( Y \)-valued measure on \((X, \mathcal{B})\) such that \( \bar{\mu}(E) = \alpha \mu(E), \forall E \in \text{dom}(\mu), \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(E) < \infty \), then \( E \in \text{dom}(\mu) \) and \( \bar{\mu}(E) = \bar{\mu}(E) \in Y \). By Proposition 11.137, \( \bar{\mu} = \bar{\mu} \). Hence, \( \bar{\mu} \) is unique.

(iv) Define \( \nu := \| y \| \mathcal{P} \circ \mu \) as prescribed in (ii), which is a \( \sigma \)-finite measure on \((X, \mathcal{B})\). Define \( \bar{\mu} \) to be a function from \( \mathcal{B} \) to \( Y \) by, if \( y \neq \emptyset \), then \( \bar{\mu}(E) = \mu(E)y \in Y, \forall E \in \text{dom}(\mu), \) and \( \bar{\mu}(E) \) is undefined, \( \forall E \in \mathcal{B} \setminus \text{dom}(\mu) \); if \( y = \emptyset \), then \( \bar{\mu}(E) = \emptyset, \forall E \in \mathcal{B} \). We will show that \( \bar{\mu} \) is a \( \sigma \)-finite \( Y \)-valued measure on \((X, \mathcal{B})\) satisfying: \( \bar{\mu}(E) = \mu(E)y, \forall E \in \text{dom}(\mu), \) and \( \mathcal{P} \circ \bar{\mu} = \nu \) by distinguishing two exhaustive and mutually exclusive cases: 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 finite \( Y \)-valued measure on \((X, \mathcal{B})\) with \( \mathcal{P} \circ \bar{\mu} = \nu \). It is straightforward to check that \( \bar{\mu}(E) = \mu(E)y = \emptyset, \forall E \in \text{dom}(\mu) \). 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 \( E \in \text{dom}(\mu) \). Then, \( \bar{\mu}(E) = \mu(E)y \in Y, \forall \text{pairwise disjoint } (E_i)_{i=1}^\infty \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^\infty E_i \), we have \( \bar{\mu}(E) = \mu(E)y = \left( \sum_{i=1}^\infty \mu(E_i) \right)y = \sum_{i=1}^\infty \mu(E_i)y = \sum_{i=1}^\infty \| \mu(E_i) \| \leq \sum_{i=1}^\infty \| y \| \| \mu(E_i) \| \leq \| y \| \mathcal{P} \circ \mu(E_i) = \| y \| \mathcal{P} \circ \mu(E_i) = \nu(E) < \infty, \forall E \in \mathcal{B} \) with \( \nu(E) < \infty \),

\[
\nu(E) = \| y \| \mathcal{P} \circ \mu(E) = \sup_{n \in \mathbb{N}, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| \mu(E_i) \| = \sup_{n \in \mathbb{N}, (E_i)_{i=1}^n \subseteq \mathcal{B}, E = \bigcup_{i=1}^n E_i, E \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| \mu(E_i) \| \]

where the third equality follows from Proposition 3.81. Hence, \( \bar{\mu} \) is a \( \sigma \)-finite \( Y \)-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 \( \sigma \)-finite \( Y \)-valued measure on \((X, \mathcal{B})\) we seek. Let \( \bar{\mu} \) be any \( Y \)-valued measure on \((X, \mathcal{B})\) such that \( \bar{\mu}(E) = \mu(E)y, \forall E \in \text{dom}(\mu), \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(E) < \infty \), then \( E \in \text{dom}(\mu) \) and \( \bar{\mu}(E) = \bar{\mu}(E) \in Y \). By Proposition 11.137, \( \bar{\mu} = \bar{\mu} \). Hence, \( \bar{\mu} \) is unique.

(v) 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 \( \mathcal{P} \circ \mu(X_n) < \infty, \forall n \in \mathbb{N} \). 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 $P \circ \mu_n \leq \|A\| P \circ \mu_n$. By Proposition 11.118, the generation process on $(X_n)_{n=1}^\infty$ yields a unique $\sigma$-finite $\mathcal{Z}$-valued measure space $\mathcal{X} := (X, \mathcal{B}, \mu)$ on the set $X$.

$\forall E \in \mathcal{B}, P \circ \bar{\mu}(E) = \sum_{n=1}^\infty P \circ \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty P \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty \|A\| P \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty 0 = 0 = (\|A\| P \circ \mu)(E)$. If $\|A\| > 0$, then $P \circ \bar{\mu}(E) \leq \sum_{n=1}^\infty \|A\| P \circ \mu_n(E \cap X_n) = \|A\| \sum_{n=1}^\infty P \circ \mu_n(E \cap X_n) = \|A\| P \circ \mu(E) = (\|A\| P \circ \mu)(E)$. Hence, we have $P \circ \bar{\mu} \leq \|A\| P \circ \mu$.

$\forall E \in \text{dom}(\mu), P \circ \mu(E) < \infty$. Then, $P \circ \bar{\mu}(E) \leq (\|A\| P \circ \mu)(E) < \infty$ and $E \in \text{dom}(\bar{\mu})$. This leads to $\bar{\mu}(E) = \sum_{n=1}^\infty \bar{\mu}(E \cap X_n) = \sum_{n=1}^\infty \mu_n(E \cap X_n) = \sum_{n=1}^\infty A\mu_n(E \cap X_n) = A \sum_{n=1}^\infty \mu(E \cap X_n) = A\mu(E)$, 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, $\bar{\mu}$ is the $\sigma$-finite $\mathcal{Z}$-valued measure we seek.

This completes the proof of the proposition.

\[ \] 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 $\mathbb{K}$, $j = 1, \ldots, n$, $\mathcal{Y} := \prod_{i=1}^m \mathcal{Y}_i$, $\mathcal{Z} := \prod_{j=1}^n \mathcal{Z}_j$, $h : \prod_{i=1}^m \prod_{j=1}^n \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \rightarrow \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be defined by,

\[
\begin{bmatrix}
A_{1,1} & \cdots & A_{1,m} \\
\vdots & \ddots & \vdots \\
A_{n,1} & \cdots & A_{n,m}
\end{bmatrix}
\]

$\forall A_{j,i} \in \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j), i = 1, \ldots, m, j = 1, \ldots, n$, and $p_{j,i} : \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \rightarrow \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$ defined by

\[
p_{j,i} \left( \begin{bmatrix}
A_{1,1} & \cdots & A_{1,m} \\
\vdots & \ddots & \vdots \\
A_{n,1} & \cdots & A_{n,m}
\end{bmatrix} \right) = A_{j,i}
\]

$\forall \begin{bmatrix}
A_{1,1} & \cdots & A_{1,m} \\
\vdots & \ddots & \vdots \\
A_{n,1} & \cdots & A_{n,m}
\end{bmatrix} \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}), i = 1, \ldots, m, j = 1, \ldots, n$. Then, $h$
CHAPTER 11. GENERAL MEASURE AND INTEGRATION

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

\[
\| h \| = \sup_{\| A_{j,i} \| \leq 1} \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \left[ \begin{array}{c} y_1 \\ \vdots \\ y_m \end{array} \right] 
\]

\[
= \sup_{\| A_{j,i} \| \leq 1} \sup_{\| y_i \| \leq 1} \left( \sum_{j=1}^{n} \left( \sum_{i=1}^{m} \| A_{j,i} \| \| y_i \| \right)^2 \right)^{1/2}
\]

\[
\leq \sup_{\| A_{j,i} \| \leq 1} \sup_{\| y_i \| \leq 1} \left( \sum_{j=1}^{n} \left( \sum_{i=1}^{m} \| A_{j,i} \|^2 \right) \left( \sum_{i=1}^{m} \| y_i \|^2 \right) \right)^{1/2}
\]

\[
\leq \sup_{\| A_{j,i} \| \leq 1} \left( \sum_{j=1}^{n} \sum_{i=1}^{m} \| A_{j,i} \|^2 \right)^{1/2} \leq 1
\]

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

\[
\| p_{j,i} \| = \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \left[ \begin{array}{c} y_i \\ \vdots \\ y_m \end{array} \right] 
\]

\[
= \sup_{\| y_i \| \leq 1} \left\| \begin{bmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{bmatrix} \right\| \left[ \begin{array}{c} \vartheta y_1 \\ \vdots \\ \vartheta y_m \end{array} \right]
\]

\[
\leq \sup_{\| y_i \| \leq 1} \left( \sum_{j=1}^{n} \sum_{i=1}^{m} \| A_{j,i} \|^2 \right)^{1/2} \leq 1
\]
Hence, 
\[ \int_X A \text{ as defined in Proposition 11.139.} \]
By Propositions 7.21 and 11.38, 
\[ i \in Y \]
\[ A \in \mathcal{Y} \]
\[ a \]
By Proposition 11.39, 
\[ A \text{ is absolutely integrable over } Y \text{ if, and only if, } A_{j,i} \text{ is absolutely integrable over } X. \]
In this case, we have 
\[ \int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \int_X A_{2,1} \, d\mu \int_X A_{2,2} \, d\mu \right] \in \mathcal{B}(Y_1 \times Y_2, Z_1 \times Z_2). \]

**Proof**
Let \( A_{j,i}, i = 1, 2, j = 1, 2, \) be absolutely integrable over \( X. \)
By Proposition 11.39, 
\[ a : X \to \mathcal{B}(Y_1 \times Y_2, Z_1 \times Z_2) \]
de defined by 
\[ a(x) = (A_{1,1}(x), A_{1,2}(x), A_{2,1}(x), A_{2,2}(x)), \forall x \in X, \text{ is } \mathcal{B}\text{-measurable.} \]
By Propositions 11.139 and 11.38, 
\[ A = ha \text{ is } \mathcal{B}\text{-measurable, where } h \text{ is the bounded linear function defined in Proposition 11.139 for } m = n = 2. \]
By Propositions 11.38 and 7.21, 
\[ \mathcal{P} \circ A \text{ is } \mathcal{B}\text{-measurable.} \]
\[ \|a(x)\| \leq \|h\| \|a(x)\| \leq \|A_{1,1}(x)\| + \|A_{1,2}(x)\| + \|A_{2,1}(x)\| + \|A_{2,2}(x)\| = \sum_{i=1}^{2} \sum_{j=1}^{2} \mathcal{P} \circ A_{j,i}(x). \]
By Proposition 11.83, 
\[ 0 \leq \int_X (\mathcal{P} \circ A) \, d\mathcal{P} \circ \mu \leq \int_X (\sum_{i=1}^{2} \sum_{j=1}^{2} \mathcal{P} \circ A_{j,i}) \, d\mathcal{P} \circ \mu = \sum_{i=1}^{2} \sum_{j=1}^{2} \int_X (\mathcal{P} \circ A_{j,i}) \, d\mathcal{P} \circ \mu < \infty. \]
Hence, \( A \) is absolutely integrable over \( X. \)
By Propositions 4.31, 7.66, and 7.22, 
\[ \mathcal{B}(Y_1, Z_{j}), i = 1, 2, j = 1, 2, \text{ and } \mathcal{B}(Y_1 \times Y_2, Z_1 \times Z_2) \text{ is a Banach space over } \mathcal{K}. \]
By Propositions 11.132, 
\[ A_{j,i}, i = 1, 2, j = 1, 2, a, \text{ and } A \text{ are integrable over } X \text{ and } \int_X A \, d\mu = h \int_X a \, d\mu \in \mathcal{B}(Y_1 \times Y_2, Z_1 \times Z_2). \]
On the other hand, let \( A \) be absolutely integrable over \( X. \) Fix any \( i \in \{1, 2\} \) and any \( j \in \{1, 2\}. \)
By Propositions 11.39, 11.38, 
\[ A_{j,i} = p_{j,i} A \text{ is } \mathcal{B}\text{-measurable, where } p_{j,i} \text{ is the bounded linear function as defined in Proposition 11.139.} \]
By Propositions 7.21 and 11.38, 
\[ \mathcal{P} \circ A_{j,i} \text{ is } \mathcal{B}\text{-measurable.} \]
\[ \forall x \in X, \mathcal{P} \circ A_{j,i}(x) \leq \|p_{j,i}\| \mathcal{P} \circ A(x) \leq \mathcal{P} \circ A(x). \]
By Proposition 11.83, 
\[ 0 \leq \int_X \mathcal{P} \circ A_{j,i} \, d\mathcal{P} \circ \mu \leq \int_X \mathcal{P} \circ A \, d\mathcal{P} \circ \mu < \infty. \]
Hence, \( A_{j,i} \) is absolutely integrable over \( X. \)
By Proposition 11.132, 
\[ A_{j,i} \text{ is integrable over } X \text{ and } \int_X A_{j,i} \, d\mu \in \mathcal{B}(Y_j, Z_j). \]
Hence, \( A \) is absolutely integrable over \( X \) if, and only if, \( A_{j,i}, i = 1, 2, j = 1, 2, \) are absolutely integrable over \( X. \)
Let $A, A_{j,i}$, $i = 1, 2$, $j = 1, 2$, be absolutely integrable over $X$. We will show that 
\[
\int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \right] \int_X A_{1,2} \, d\mu \left[ \int_X A_{2,1} \, d\mu \right] \int_X A_{2,2} \, d\mu \in B(\mathbb{R}^2)
\] 
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: $\mathcal{P} \circ \mu(X) < \infty$. By Proposition 11.133, $\forall n \in \mathbb{N}$, such that $\lim_{n \to \infty} \varphi_{j,n} = A_{j,i}$ a.e. in $X$, $\lim_{n \to \infty} \int_X \varphi_{j,n} \, d\mu = \lim_{n \to \infty} \hat{\varphi}_{j,n}(x) \leq \mathcal{P} \circ A_{j,i}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\int_X A_{j,i} \, d\mu = \lim_{n \to \infty} \int_X \varphi_{j,n} \, d\mu \in B(\mathbb{Z})$. Then, by Proposition 11.75, we have $\lim_{n \to \infty} \psi_n = a$ a.e. in $X$, where $\psi_n : X \to \prod_{j=1}^n W_{j,i}$ is defined by $\psi_n(x) = (\varphi_{j,1,n}(x), \varphi_{j,2,n}(x), \varphi_{j,3,n}(x), \varphi_{j,4,n}(x))$, $\forall x \in X$, $\forall n \in \mathbb{N}$. Clearly, $\psi_n$ is a simple function, $\forall n \in \mathbb{N}$. By Lebesgue Dominated Convergence Theorem 11.131, $\int_X a \, d\mu = \lim_{n \to \infty} \int_X \psi_n \, d\mu = \lim_{n \to \infty} \left( \int_X \varphi_{j,1,n} \, d\mu, \int_X \varphi_{j,2,n} \, d\mu, \int_X \varphi_{j,3,n} \, d\mu, \int_X \varphi_{j,4,n} \, d\mu \right) = \left( \int_X A_{j,1} \, d\mu, \int_X A_{j,2} \, d\mu, \int_X A_{j,3} \, d\mu, \int_X A_{j,4} \, d\mu \right) \in B(\mathbb{Z})$, where the second equality follows from Proposition 11.75; and the third equality follows from Proposition 3.67. Then, by Proposition 11.132, $\int_X A \, d\mu = h \int_X a \, d\mu = \left[ \int_X A_{1,1} \, d\mu \right] \int_X A_{1,2} \, d\mu \left[ \int_X A_{2,1} \, d\mu \right] \int_X A_{2,2} \, d\mu \in B(\mathbb{Z})$.

Case 2: $\mathcal{P} \circ \mu(X) = \infty$. Let $(A_E)_{E \in \mathcal{M}(X)}$ and $(A_{j,i,E})_{E \in \mathcal{M}(X)}$ be the nets for $\int_X A \, d\mu$ and $\int_X A_{j,i} \, d\mu$ as defined in Definition 11.71, respectively, $i = 1, 2$, $j = 1, 2$. $\forall E \in \mathcal{M}(X)$, by Case 1, $A_E = \left[ \int_E A_{1,1} \, d\mu_E \right] \left[ \int_E A_{1,2} \, d\mu_E \right] = \left[ \int_E A_{2,1} \, d\mu_E \right] \left[ \int_E A_{2,2} \, d\mu_E \right] \in B(\mathbb{R}^2)$, where $E := (E, B_E, \mu_E)$ is the finite $\mathbb{K}$-valued measure space of $X$. Since $A, A_{j,i}$, $i = 1, 2$, $j = 1, 2$ are integrable over $X$, we have $\int_X A \, d\mu = \lim_{E \in \mathcal{M}(X)} A_E = \lim_{E \in \mathcal{M}(X)} \left[ A_{1,1,E} A_{1,2,E} \right] = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \right] = \left[ \int_X A_{2,1} \, d\mu \int_X A_{2,2} \, d\mu \right] \in B(\mathbb{Z})$. Hence, in both cases, we have $\int_X A \, d\mu = \left[ \int_X A_{1,1} \, d\mu \int_X A_{1,2} \, d\mu \right] = \left[ \int_X A_{2,1} \, d\mu \int_X A_{2,2} \, d\mu \right] \in B(\mathbb{Z})$. This completes the proof of the proposition.

**Proposition 11.141** Let $(X, \mathcal{B})$ be a measurable space, $\mathbb{Y}$ be a Banach space over $\mathbb{K}$, and $Z$ be the set of $\sigma$-finite $\mathbb{Y}$-valued measures on $(X, \mathcal{B})$. Define vector addition and scalar multiplication on $Z$ according to Proposition 11.138 and $\mathcal{P} \subset Z$ by $\mathcal{P}(E) = \mathcal{P}(E)$, $\forall E \in \mathcal{B}$. Then, $(Z, \mathcal{P}, +, \cdot, \mathcal{P}) =: \mathcal{M}_\sigma(X, \mathcal{B}, \mathbb{Y})$ is a vector space over $\mathbb{K}$.
11.7. Calculation with Measures

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 $\vartheta(E) = \vartheta_y$, $\forall E \in \mathcal{B}$. Then, $\mathcal{Z} := (Z, +, \vartheta, \vartheta)$ 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, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) := (Z, || \cdot ||)$ 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 $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ and is a Banach space over $\mathbb{K}$.

Proof We will first show that $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})$ 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) By Proposition 11.136, $\mu_1 + \mu_2 + \mu_1 \in Z$ 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, $\forall E \in \mathcal{B}$, ($(\mu_1 + \mu_2) + \mu_3)(E) = (\mu_1 + \mu_2)(E) + \mu_3(E) = (\mu_1(E) + \mu_2(E)) + \mu_3(E) = (\mu_1(E) + \mu_3(E)) + \mu_2(E) = (\mu_1 + \mu_3)(E) + \mu_2(E) = (\mu_1 + \mu_3)(E) + \mu_2(E) = (\mu_1(E) + \mu_3(E)) + \mu_2(E) = (\mu_1 + \mu_3)(E) + \mu_2(E) = (\mu_1 + \mu_3)(E)$. Therefore, $\mathcal{M}_f(X, \mathcal{B}, \mathcal{Y}) = (Z, +, \vartheta)$ is a vector space over $\mathbb{K}$. This completes the proof of the proposition. □
uniform Cauchy sequence uniform in $\mathbb{E}$ and the second inequality follows from Definition 11.108; and the last step
\[
\mu(X) + (\mu_2(E) + \mu_3(E)) = \mu_1(E) + (\mu_2(E) + \mu_3(E)) = \mu_1(E) + (\mu_2 + \mu_3)(E) = \mu_1 + (\mu_2 + \mu_3).
\]
Hence, $(\mu_1 + \mu_2) + \mu_3 = \mu_1 + (\mu_2 + \mu_3)$.

(iii) By Proposition 11.136, $\mu_1 + \vartheta \in \mathbb{E}$ and $\forall E \in \mathcal{B}$, $(\mu_1 + \vartheta)(E) = \mu_1(E) + \vartheta(E) = \mu_1(E) \in \mathbb{Y}$. Hence, $\mu_1 + \vartheta = \mu_1$.

(iv) By Proposition 11.136, $\alpha(\mu_1 + \mu_2), \alpha \mu_1 + \alpha \mu_2 \in \mathbb{E}$ and $\forall E \in \mathcal{B}$,
\[
(\alpha(\mu_1 + \mu_2))(E) = \alpha(\mu_1 + \mu_2)(E) = \alpha(\mu_1(E) + \mu_2(E)) = \alpha \mu_1(E) + \alpha \mu_2(E) = (\alpha \mu_1)(E) + (\alpha \mu_2)(E) = (\alpha \mu_1 + \alpha \mu_2)(E) \in \mathbb{Y}.
\]
Hence, $\alpha(\mu_1 + \mu_2) = \alpha \mu_1 + \alpha \mu_2$.

(v) By Proposition 11.136, $(\alpha + \beta)\mu_1, \alpha \mu_1 + \beta \mu_1 \in \mathbb{E}$ and $\forall E \in \mathcal{B}$,
\[
((\alpha + \beta)\mu_1)(E) = (\alpha + \beta)\mu_1(\mathbb{E}) = (\alpha \mu_1 + \beta \mu_1)(E) = (\alpha \mu_1)(E) + (\beta \mu_1)(E) = (\alpha \mu_1 + \beta \mu_1)(E) \in \mathbb{Y}.
\]
Hence, $(\alpha + \beta)\mu_1 = \alpha \mu_1 + \beta \mu_1$.

Let $f : \mathcal{E} \to \mathbb{K}$ be a Banach space. Clearly,
\[
P \circ \mu_1(X) \in [0, \infty) \subset \mathbb{R}
\]
and $(\mu_1 + \mu_2)(X) = \mu_1(X) + \mu_2(X) \in \mathbb{E}$. Hence, $P \circ (\alpha \mu_1)(X) = (|\alpha| P \circ \mu_1)(X) = |\alpha| P \circ \mu_1(X) = |\alpha| \mu_1(X)$.

Hence, $M_f(X, \mathbb{K}, \mathbb{Y})$ is a normed linear space.

Let $\mathbb{Y}$ be a Banach space. Clearly, $M_f(X, \mathbb{K}, \mathbb{Y}) \subseteq M_f(X, \mathbb{K}, \mathbb{Y})$. Therefore, $M_f(X, \mathbb{K}, \mathbb{Y})$ is a subspace of $M_f(X, \mathbb{K}, \mathbb{Y})$.

Finally, we show that $M_f(X, \mathbb{K}, \mathbb{Y})$ is a Banach space when $\mathbb{Y}$ is a Banach space. Let $(\mu_n)_{n=1}^\infty \subseteq M_f(X, \mathbb{K}, \mathbb{Y})$ be a Cauchy sequence. $\forall \epsilon > 0, \exists N \in \mathbb{N}$, $\forall m, n \in \mathbb{N}$ with $N \leq m$ and $N \leq n$, we have $|\mu_n(\mu_m(\mathbb{E})) - \mu_m(\mu_n(\mathbb{E}))| = P \circ (\mu_m(\mu_m(\mathbb{E})) - \mu_n(\mu_n(\mathbb{E})))) < \epsilon$. $\forall E \in \mathcal{E}$, $\|\mu(E) - \mu_m(\mathbb{E})\| = \|\mu_n(\mu_n(\mathbb{E}))) - \mu_m(\mu_n(\mathbb{E})))\| = P \circ (\mu_n(\mu_n(\mathbb{E}))) - \mu_m(\mu_n(\mathbb{E}))) < \epsilon$. Hence, $(\mu_n)_{n=1}^\infty$ is a uniform Cauchy sequence uniform in $E \in \mathcal{E}$. By $\mathbb{Y}$ being complete, $\forall E \in \mathcal{E}$, there exists a unique $\mu(\mathbb{E}) \in \mathbb{Y}$ such that $\lim_{n \to \infty} \mu_n(\mathbb{E}) = \mu(\mathbb{E})$. This defines a function $\mu : \mathcal{E} \to \mathbb{Y}$. Clearly, $\mu(E) = \lim_{n \to \infty} \mu_n(E) = \vartheta_E$, $\forall$ pairwise disjoint $(E_i)_{i=1}^\infty \subseteq \mathcal{E}$, let $E := \bigcup_{i=1}^\infty E_i \in \mathcal{E}$. Then,
\[
\sum_{i=1}^\infty \|\mu(E_i)\| = \sum_{i=1}^\infty \lim_{n \to \infty} \mu_n(E_i) \leq \liminf_{n \to \infty} \sum_{i=1}^\infty \|\mu_n(E_i)\| \leq \liminf_{n \to \infty} P \circ \mu_n(E) \leq P \circ \mu_n(E) = \sum_{i=1}^\infty \|\mu(E_i)\| \in \mathbb{R}
\]
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 \(\sum_{i=1}^{\infty} \mu(E_i) = \lim_{n \to \infty} \sum_{i=1}^{m} \mu_n(E_i) = \lim_{n \to \infty} \lim_{m \to \infty} \mu_n(\bigcup_{i=1}^{m} E_i) = \lim_{m \to \infty} \lim_{n \to \infty} \mu_n(E_i) = 0\), where fourth equality follows from Iterated Limit Theorem 4.56. Hence, \(\mu\) is a \(\mathcal{F}\)-valued measure on \((X, \mathcal{B})\).

Note that

\[
P \circ \mu(X) = \sup_{n \in \mathbb{Z}^+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, \ X = \bigcup_{i=1}^{n} E_i, \ E_i \cap E_j = \emptyset, \ \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(E_i) \|\]

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 \(\mathcal{F}\)-valued measure on \((X, \mathcal{B}, \mathcal{Y})\) and \(\mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\).

Note that, \(\forall n \in \mathbb{N}\) with \(N \leq n\),

\[
\| \mu_n - \mu \| = P \circ (\mu_n - \mu)(X)
\]

where the fourth equality follows from Propositions 7.21, 7.23, 3.66, and 3.67. Hence, \(\lim_{n \to \infty} \| \mu_n - \mu \| = 0\) and \(\lim_{n \to \infty} \mu_n = \mu \in \mathcal{M}_f(X, \mathcal{B}, \mathcal{Y})\).
Thus, we have shown that every Cauchy sequence in $\mathcal{M}_f(X, B, \gamma)$ converges in $\mathcal{M}_f(X, B, \gamma)$. Hence, $\mathcal{M}_f(X, B, \gamma)$ is a Banach space.

This completes the proof of the proposition.

In the above proposition, $\mathcal{M}_f(X, B, \gamma)$ admits a topology $\mathcal{O}_1$ that is induced by the norm. We may define a topology on the vector space $\mathcal{M}_\sigma(X, B, \gamma)$ when $\gamma$ is a Banach space, whose subset topology on $\mathcal{M}_f(X, B, \gamma)$ is exactly $\mathcal{O}_1$. Let $\mathcal{O}_B$ be a collection of subsets of $\mathcal{M}_\sigma(X, B, \gamma)$ such that, $\forall B \in \mathcal{O}_B$, $\exists E \in B$, let $B_E := \{ A \in B \mid A \subseteq E \}$, $\exists \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \nu_E \in \mathcal{M}_f(E, B_E, \gamma)$ such that

$$B = \{ \mu \in \mathcal{M}_\sigma(X, B, \gamma) \mid (E, B_E, \mu_E) \text{ is the finite } \gamma\text{-valued measure subspace of } (X, B, \mu) \text{ and } \|\mu_E - \nu_E\|_{\mathcal{M}_f(E, B_E, \gamma)} < \epsilon \}$$

It is easy to show that $\mathcal{O}_B$ is a basis for a topology $\mathcal{O}$ on $\mathcal{M}_\sigma(X, B, \gamma)$ by Proposition 3.18. We will abuse the notation and denote the topological space $(\mathcal{M}_\sigma(X, B, \gamma), \mathcal{O})$ by $\mathcal{M}_\sigma(X, B, \gamma)$. With this topology in place, we may define convergence for nets in $\mathcal{M}_\sigma(X, B, \gamma)$ according to Section 3.9. For clarity, we summarize the concept of convergence for sequences in $\mathcal{M}_\sigma(X, B, \gamma)$ in the following definition.

**Definition 11.143** Let $(X, B)$ be a measurable space, $\gamma$ be a Banach space over $\mathbb{K}$, $(\mu_n)_{n=1}^\infty \subseteq \mathcal{M}_\sigma(X, B, \gamma)$, and $\nu \in \mathcal{M}_\sigma(X, B, \gamma)$. 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 $(E, B_E, \mu_n)$ of $(X, B, \mu_n)$ is finite for sufficiently large $n \in \mathbb{N}$ and $\lim_{n \in \mathbb{N}} \mu_n(E) = \nu(E)$ in $\mathcal{M}_f(E, B_E, \gamma)$, where $(E, B_E, \nu_E)$ is the finite $\gamma$-valued measure subspace of $(X, B, \nu)$. 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 $(X, B)$ be a measurable space, $\mu_1$ and $\mu_2$ be $\sigma$-finite measures on $(X, B)$, $a \in (0, \infty) \subseteq \mathbb{R}$, $f : X \to [0, \infty) \subseteq \mathbb{R}$ be $B$-measurable, and $\mathcal{X}_i := (X, B, \mu_i), i = 1, 2$. Then, the following statements hold.

(i) If $\mu_1 \leq \mu_2$ (that is $\mu_1(E) \leq \mu_2(E), \forall E \in B$), 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: \( \int_X f \, d\mu_2 < \infty \). The result holds trivially in this case. Case 2: \( \int_X f \, d\mu_2 < \infty \). By Definition 11.79, \( 0 \leq \int_X f \, d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi \, d\mu_1 \leq \infty \), where the supremum is over all simple functions \( \phi : \mathcal{X}_1 \to [0, \infty) \subseteq \mathbb{R} \) 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 \( n \in \mathbb{Z}_+ \), \( a_1, \ldots, a_n \in (0, \infty) \subseteq \mathbb{R} \) 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 11.75, \( \int_X \phi \, d\mu_1 = \sum_{i=1}^{n} a_i \mu_1(A_i) \subseteq [0, +\infty) \subseteq \mathbb{R} \). 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 \( \mathcal{X}_2 \) is \( \sigma \)-finite, then, by Proposition 11.7, \( \exists E \in \mathcal{B} \) with \( E \subseteq A_{i_0} \) such that \( \frac{1}{\mu_0} \int_X f \, d\mu_2 < \mu_2(E) < \infty \). Choose \( \phi : \mathcal{X}_2 \to [0, \infty) \subseteq \mathbb{R} \) by \( \phi = a_{i_0} \chi_{E,X} \). Clearly, \( \phi \) is a simple function of \( \mathcal{X}_2 \) 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 \( \mathcal{X}_2 \). 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 \( (\phi_n)_{n=1}^{\infty} \subseteq \mathcal{X} \to [0, \infty) \subseteq \mathbb{R} \), \( \forall n \in \mathbb{N} \), such that \( 0 \leq \phi_n(x) \leq f(x) < \infty, \forall x \in X \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n \phi = f \text{ a.e. in } \mathcal{X} \), where \( \mathcal{X} := (X, \mathcal{B}, \mu) \). Since \( \mu = \mu_1 + \mu_2 \), then \( \mu_i \leq \mu, i = 1, 2 \). This implies that \( (\phi_n)_{n=1}^{\infty} \) is a sequence of simple functions of \( \mathcal{X} \) and \( \lim_{n \to \infty} \phi_n = f \text{ a.e. in } \mathcal{X}_i, i = 1, 2 \). By Fatou's Lemma 11.80 and Propositions 3.83 and 11.83, we have \( 0 \leq \int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X \phi_n \, d\mu \leq \limsup_{n \to \infty} \int_X \phi_n \, d\mu \leq \int_X f \, d\mu \leq \int_X f \, d\mu \leq \infty \). Then, \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu \subseteq [0, \infty) \subseteq \mathbb{R} \). By an argument that is similar to the above, we have \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X \phi_n \, d\mu = \lim_{n \to \infty} \int_X f \, d\mu \). Hence, the result holds.

(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.
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 \( X := (X, B, \nu) \) with \( 0 \leq \phi(x)/a \leq f(x), \forall x \in X \). Then, \( \int_X \phi d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X (\phi) d\mu_1 = \sup_{0 \leq \phi \leq f} \int_X \phi d\nu = \int_X f d\nu \in [0, \infty] \subset \mathbb{R} \), 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. \( \square \)

**Proposition 11.145** Let \( (X, B) \) be a measurable space, \( Y \) be a Banach space over \( \mathbb{K} \), \( \mu_1, \mu_2 \in \mathcal{M}_\sigma(X, B, Y) \), \( Z \) be a Banach space over \( \mathbb{K} \), \( W \) be a separable subspace of \( B(Y, Z) \), \( f : X \to W \) be \( B \)-measurable, and \( \mu := \mu_1 + \mu_2 \in \mathcal{M}_\sigma(X, B, Y) \). Assume that \( f \) is absolutely integrable over \( X := (X, B, \mu) \) and \( \int_X f d\mu = \int_X f d\mu_1 + \int_X f d\mu_2 \in Z \).

**Proof** By Propositions 11.38 and 11.36, \( \mathcal{P} \circ \mu \leq \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \mu_2 =: \nu \) and \( \nu \) is \( \sigma \)-finite. By Proposition 11.144, \( 0 \leq \int_X \mathcal{P} \circ f d\mu \leq \int_X \mathcal{P} \circ f d\nu = \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu_1 + \int_X \mathcal{P} \circ f d\mathcal{P} \circ \mu_2 < \infty \). Hence, \( f \) is absolutely integrable over \( X \).

By Proposition 11.66, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^\infty, \varphi_n : \tilde{X} \to W, \forall n \in \mathbb{N} \), 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 \( \tilde{X} \), where \( \tilde{X} := (X, B, \nu) \). Since \( \mathcal{P} \circ \mu \leq \nu \), \( \mathcal{P} \circ \mu_1 \leq \nu \), \( i = 1, 2 \), then \( \lim_{n \in \mathbb{N}} \varphi_n = f \) a.e. in \( \tilde{X} \), \( \lim_{n \in \mathbb{N}} \varphi_n = f \) a.e. in \( X_i \), \( i = 1, 2 \), and \( \varphi_n \) is a simple function of \( X_i \), \( i = 1, 2 \), or \( \mathcal{X}_2 \), \( \forall n \in \mathbb{N} \). By Lebesgue Dominated Convergence Theorem 11.131, we have \( \int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu_1 + \int_X \varphi_n d\mu_2 \in Z \), \( \forall n \in \mathbb{N} \), let \( \varphi_n \) admit canonical representation \( \varphi_n = \sum_{i=1}^{m_n} \varphi_{i,n} x_{A_{i,n}, X} \).

By Proposition 11.125, we have \( \int_X \varphi_n d\mu = \sum_{i=1}^{m_n} \varphi_{i,n} \mu(A_{i,n}) = \sum_{i=1}^{m_1} \varphi_{i,n} (\mu_1(A_{i,n}) + \mu_2(A_{i,n})) = \sum_{i=1}^{m_1} \varphi_{i,n} \mu_1 + \sum_{i=1}^{m_2} \varphi_{i,n} \mu_2 \in \mathbb{Z} \). Then, by Propositions 7.23, 3.66, and 3.67, \( \int_X f d\mu = \lim_{n \in \mathbb{N}} \int_X \varphi_n d\mu = \lim_{n \in \mathbb{N}} (\int_X \varphi_n d\mu_1 + \int_X \varphi_n d\mu_2) = \lim_{n \in \mathbb{N}} (\int_X \varphi_n d\mu_1 + \int_X \varphi_n d\mu_2) = \int_X f d\mu_1 + \int_X f d\mu_2 \in \mathbb{Z} \). This completes the proof of the proposition. \( \square \)

**Proposition 11.146** Let \( X := (X, B, \mu) \) be a \( \sigma \)-finite \( Y \)-valued measure space, where \( Y_1 \) is a normed linear space over \( \mathbb{K} \), \( Y_2 \) and \( Z \) be Banach spaces over \( \mathbb{K} \), \( A \in B(Y_1, Y_2) \), \( \nu := A\mu \in \mathcal{M}_\sigma(X, B, Y_2) \) as defined in Proposition 11.138, \( W_2 \) be a separable subspace of \( B(Y_2, Z) \), and \( f : X \to W_2 \) be \( B \)-measurable. Assume that \( \| A \| \mathcal{P} \circ f \) is integrable over \( (X, B, \mathcal{P} \circ \mu) \). Then, \( f \) is absolutely integrable over \( X := (X, B, \nu) \), \( f A \) is absolutely integrable over \( X \), and \( \int_X f d\nu = \int_X (f A) d\mu \in \mathbb{Z} \).

**Proof** By Propositions 11.38, 7.64, and 7.21, \( f A \) and \( \mathcal{P} \circ (fA) \) are \( B \)-measurable. By Proposition 11.83, \( 0 \leq \int_X \mathcal{P} \circ (fA) d\mathcal{P} \circ \mu \leq \int_X f d\nu \).
\[ \int_X (\|A\| \mathcal{P} \circ f) \, d\mathcal{P} \circ \mu < \infty. \]

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 \( \|\varphi_n(x)\| \leq \|f(x)\| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \), 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, \( \|\varphi_n(x)A\| \leq \|A\| \|f(x)\| \), \( \forall x \in X \), \( \forall n \in \mathbb{N} \). Note that \( fA : X \to \mathcal{W}_1 \) and \( \varphi_nA : X \to \mathcal{W}_1 \), \( \forall n \in \mathbb{N} \), where \( \mathcal{W}_1 = \{ F \in \mathcal{B}(\mathcal{Y}_1, \mathcal{Z}) \mid F = F_1A, F_1 \in \mathcal{W}_2 \} \) is a separable subspace of \( \mathcal{B}(\mathcal{Y}_1, \mathcal{Z}) \).

By Proposition 11.52, we have \( \lim_{n \to \infty} \varphi_nA = fA \) a.e. in \( \mathcal{X} \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X (fA) \, d\mu = \lim_{n \to \infty} \int_X \varphi_nA \, d\mu \in \mathcal{Z} \). By Proposition 11.125, we have \( \int_X (fA) \, d\mu = \int_X f \, d\nu \in \mathcal{Z} \), \( \forall n \in \mathbb{N} \). Then, \( \int_X (fA) \, d\mu = \int_X f \, d\nu \in \mathcal{Z} \).

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

**Proposition 11.147** Let \( n, m \in \mathbb{Z}_+ \), \( (X, \mathcal{B}) \) be a measurable space, and \( \mu_{j,i} \) be a \( \sigma \)-finite \( \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \)-valued measure on \( (X, \mathcal{B}) \), 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 \( (X, \mathcal{B}) \) such that \( \mu(E) = \left[ \begin{array}{ccc} \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 \cap_{j=1}^n \text{dom}(\mu_{j,i}), \text{ where } \mathcal{Y} := \prod_{i=1}^m \mathcal{Y}_i \text{ and } \mathcal{Z} := \prod_{j=1}^n \mathcal{Z}_j. \text{ Furthermore, } \max_{i,j} \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 \text{ will be called the vector measure and denoted by } \left[ \begin{array}{ccc} \mu_{1,1} & \cdots & \mu_{1,m} \\ \vdots & \ddots & \vdots \\ \mu_{n,1} & \cdots & \mu_{n,m} \end{array} \right]. \]

**Proof** Consider the special case where \( \mu_{j,i} \)'s are finite. Then, \( \mu \) is uniquely defined since \( \text{dom}(\mu_{j,i}) = \mathcal{B}, i = 1, \ldots, m, j = 1, \ldots, n \).

\[ \mu(\emptyset) = \left[ \begin{array}{ccc} \mu_{1,1}(\emptyset) & \cdots & \mu_{1,m}(\emptyset) \\ \vdots & \ddots & \vdots \\ \mu_{n,1}(\emptyset) & \cdots & \mu_{n,m}(\emptyset) \end{array} \right] = \left[ \begin{array}{ccc} \vartheta_{B(\mathcal{Y}_1, \mathcal{Z}_1)} & \cdots & \vartheta_{B(\mathcal{Y}_m, \mathcal{Z}_1)} \\ \vdots & \ddots & \vdots \\ \vartheta_{B(\mathcal{Y}_1, \mathcal{Z}_n)} & \cdots & \vartheta_{B(\mathcal{Y}_m, \mathcal{Z}_n)} \end{array} \right]. \]

\( \forall \) pairwise disjoint \( \{E_k\}_{k=1}^\infty \subseteq \mathcal{B}, \) we have \( \sum_{k=1}^\infty \|\mu(E_k)\| \leq \sum_{k=1}^\infty \|h\| \left( \sum_{j=1}^m \sum_{i=1}^n \|\mu_{j,i}(E_k)\|^2 \right)^{1/2} \leq \sum_{k=1}^\infty \sum_{j=1}^m \sum_{i=1}^n \|\mu_{j,i}(E_k)\| = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^\infty \|\mu_{j,i}(E_k)\| < \infty \), 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 \( (X, \mathcal{B}) \) and therefore is a finite \( \mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j) \)-valued pre-measure on \( (X, \mathcal{B}) \), \( i = 1, \ldots, m, j = 1, \ldots, n \). We
also have \( \sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} \left[ \begin{array}{c} \mu_1(E_k) \\ \vdots \\ \mu_n(E_k) \end{array} \right] = \left[ \begin{array}{c} \sum_{k=1}^{\infty} \mu_1(E_k) \\ \vdots \\ \sum_{k=1}^{\infty} \mu_n(E_k) \end{array} \right] \)

\[ \leq \left[ \begin{array}{c} \sum_{k=1}^{\infty} \mu_n(E_k) \end{array} \right] = \mu(E) \in B(\mathcal{Y}, \mathcal{Z}), \]

where the first 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 also that, for all \( E \in \mathcal{B} \),

\[ \mathcal{P} \circ \mu_{t,s}(E) = \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \| \mu_{t,s}(E_i) \| \]

\[ = \mathcal{P} \circ \mu(E) \]

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, for all \( E \in \mathcal{B} \),

\[ \mathcal{P} \circ \mu(E) = \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \| \mu(E_i) \| \]

\[ \leq \sup_{k \in \mathbb{Z}_+, (E_i)_{i=1}^k \subseteq \mathcal{B}, E = \bigcup_{i=1}^k E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq k} \sum_{i=1}^k \sum_{j=1}^m \sum_{i=1}^n \| \mu_{t,s}(E_i) \| \]

\[ \leq \sum_{s=1}^m \sum_{t=1}^n \| \mu_{t,s}(E) \| \]

\[ = \sum_{s=1}^m \sum_{t=1}^n \mathcal{P} \circ \mu_{t,s}(X) \leq \sum_{s=1}^m \sum_{t=1}^n \mathcal{P} \circ \mu_{t,s}(X) < \infty \]

where the first equality follows from Definition 11.99; the first inequality follows from Propositions 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 $\mu_{j,i}$'s are $\sigma$-finite. Then, $\exists \ (X_k)_{k=1}^\infty \subseteq \mathcal{B}$ such that $X = \bigcup_{k=1}^\infty X_k$ and $\mathcal{P} \circ \mu_{j,i}(X_k) < \infty, \ \forall i \in \{1, ..., m\}, \ \forall j \in \{1, ..., n\}, \ \forall k \in \mathbb{N}$. 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 $\mathcal{B}(\mathcal{Y}_i, \mathcal{Z}_j)$-valued measure subspace of $(X, \mathcal{B}, \mu_{j,i}), \ \forall k \in \mathbb{N}, \ \forall i \in \{1, ..., m\}, \ \forall j \in \{1, ..., n\}$. Then, by the special case, there exists a unique finite $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$-valued measure $\mu_k$ on $(X_k, \mathcal{B}_k)$ such that $\mu_k(E) = \left[ \begin{array}{ccc} \mu_{1,1,k}(E) & \cdots & \mu_{1,m,k}(E) \\ \vdots & & \vdots \\ \mu_{n,1,k}(E) & \cdots & \mu_{n,m,k}(E) \end{array} \right] \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}), \ \forall E \in \mathcal{B}_k$. Let $\tilde{\mathcal{X}}_k := (X_k, \mathcal{B}_k, \tilde{\mu}_k)$ be the finite $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$-valued measure space, $\forall k \in \mathbb{N}$. By Propositions 4.31, 7.22, and 7.66, $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$ is a Banach space over $\mathbb{K}$. By Proposition 11.118, there is a unique $\sigma$-finite $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$-valued measure space $\mathcal{X} := (X, \mathcal{B}, \mu)$ on $X$ such that $\tilde{\mathcal{X}}_k$ is the finite $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$-valued measure subspace of $\mathcal{X}, \ \forall k \in \mathbb{N}$. $\forall s \in \{1, ..., m\}, \ \forall t \in \{1, ..., n\}, \ \forall E \in \mathcal{B}$, we have $\mathcal{P} \circ \mu_{j,i,s}(E) = \sum_{k=1}^\infty \mathcal{P} \circ \mu_{j,i,k}(E \cap X_k) = \sum_{k=1}^\infty P \circ \mu_{j,i,k}(E \cap X_k) = \mathcal{P} \circ \mu_{j,i}(E) \leq \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i,k}(E \cap X_k) = \sum_{k=1}^\infty \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E) = P \circ \mu(E) \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}(E), \ \forall E \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}), \ \forall \mu \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. Then, max$_{s \in \{1, ..., m\}} \mathcal{P} \circ \mu_{j,i,s} \leq \mathcal{P} \circ \mu \leq \sum_{i=1}^m \sum_{j=1}^n \mathcal{P} \circ \mu_{j,i}, \ \forall E \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}), \ \forall \mu \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$. This implies that $\mu(E) = \sum_{k=1}^\infty \mu(E \cap X_k) = \sum_{k=1}^\infty \tilde{\mu}_k(E \cap X_k) = \left[ \begin{array}{ccc} \mu_{1,1}(E \cap X_k) & \cdots & \mu_{1,m}(E \cap X_k) \\ \vdots & & \vdots \\ \mu_{n,1}(E \cap X_k) & \cdots & \mu_{n,m}(E \cap X_k) \end{array} \right] = \left[ \begin{array}{ccc} \sum_{k=1}^\infty \mu_{1,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{1,m}(E \cap X_k) \\ \vdots & & \vdots \\ \sum_{k=1}^\infty \mu_{n,1}(E \cap X_k) & \cdots & \sum_{k=1}^\infty \mu_{n,m}(E \cap X_k) \end{array} \right] = \left[ \begin{array}{ccc} \mu_{1,1}(E) & \cdots & \mu_{1,m}(E) \\ \vdots & & \vdots \\ \mu_{n,1}(E) & \cdots & \mu_{n,m}(E) \end{array} \right]$, 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 \( \mathcal{X}_i := (X, \mathcal{B}, \mu_i) \) be a \( \sigma \)-finite \( y_i \)-valued measure space, where \( y_i \) is a Banach space over \( K \), \( Z_j \) be a Banach space over \( K \), \( W_{j,i} \subseteq B(y_i, Z_j) \) be a separable subspace, \( f_{j,i} : X \to W_{j,i}, \ i = 1, 2, \ j = 1, 2, \mu := \left[ \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right] \) be the vector measure as defined in Proposition 11.147, \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be the \( \sigma \)-finite \( y_1 \times y_2 \)-valued measure space, and \( F : X \to B(y_1 \times y_2, Z_1 \times Z_2) \) be defined by \( F(x) = \left[ \begin{array}{c} f_{1,1}(x) \\ f_{1,2}(x) \\ f_{2,1}(x) \\ f_{2,2}(x) \end{array} \right], \forall x \in X. \) Assume that \( F \) is absolutely integrable over \( \mathcal{X} \). Then, \( f_{j,i}, j = 1, 2 \) are absolutely integrable over \( \mathcal{X}_i, i = 1, 2 \). In this case, we have \( \int_X F \, d\mu = \left[ \int f_{1,1} \, d\mu_1 + \int f_{1,2} \, d\mu_2 \right] \in Z_1 \times Z_2. \)

**Proof** Let \( F \) be absolutely integrable over \( \mathcal{X} \). Fix any \( i \in \{1, 2\} \) and any \( j \in \{1, 2\} \). Then, by Propositions 11.139 and 11.38, \( f_{j,i} = p_{j,i} F \) is \( B \)-measurable, where \( p_{j,i} \) is the bounded linear function as defined in Proposition 11.139. By Propositions 7.21 and 11.38, \( \mathcal{P} \circ f_{j,i} \) is \( B \)-measurable, \( \forall x \in X, \mathcal{P} \circ f_{j,i}(x) \leq \|p_{j,i}\| \mathcal{P} \circ F(x) \leq \mathcal{P} \circ F(x). \) Then, by Propositions 11.83, 11.147, and 11.144, \( 0 \leq \int_X \mathcal{P} \circ f_{j,i} \, d\mu \leq \|x\| \int_X \mathcal{P} \circ f_{j,i} \, d\mu \leq \int \mathcal{P} \circ F \, d\mu < \infty. \) Hence, \( f_{j,i} \) is absolutely integrable over \( \mathcal{X}_i \). By Proposition 11.132, \( f_{j,i} \) is integrable over \( \mathcal{X}_i \) and \( \int f_{j,i} \, d\mu = B(y_i, Z_j). \)

By Propositions 7.22, 4.4, and 3.28, \( \prod_{i=1}^2 \prod_{j=1}^2 B(y_i, Z_j) \subseteq \prod_{i=1}^2 \prod_{j=1}^2 B(y_i, Z_j) \) is a separable subspace. Let \( h \) be the bounded linear function as defined in Proposition 11.139 for \( n = m = 2 \). Then, \( \mathcal{W} := h(\prod_{i=1}^2 \prod_{j=1}^2 W_{j,i}) \subseteq B(y_1 \times y_2, Z_1 \times Z_2) \) is a separable subspace. By Proposition 11.66, 3 sequence of simple functions \( \{\psi_n\}_{n=1}^\infty, \psi_n : \mathcal{X} \to \mathcal{W}, \forall n \in \mathbb{N}, \) such that \( \lim_{n \to \infty} \psi_n = F \) a.e. in \( \mathcal{X}, \|\psi_n(x)\| \leq \|F(x)\|, \forall x \in X, \forall n \in \mathbb{N}. \) Fix any \( i \in \{1, 2\} \) and any \( j \in \{1, 2\} \). Let \( \varphi_{j,i,n} := p_{j,i} \psi_n. \) Then, \( \varphi_{j,i,n} : \mathcal{X} \to B(y_i, Z_j) \) is a simple function. By Proposition 11.52, \( \lim_{n \to \infty} \varphi_{j,i,n} = F \) a.e. in \( \mathcal{X} \). Since \( \mathcal{P} \circ \mu \geq \mathcal{P} \circ \mu_i \), we have \( \varphi_{j,i,n} : \mathcal{X}_i \to B(y_i, Z_j) \) 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 \( \|\varphi_{j,i,n}(x)\| \leq \|p_{j,i}\| \|\psi_n(x)\| \leq \|F(x)\|, \forall x \in X, \forall n \in \mathbb{N}. \) By Lebesgue Dominated Convergence Theorem 11.131, \( \int f_{j,i} \, d\mu = \lim_{n \to \infty} \int \varphi_{j,i,n} \, d\mu = \int \varphi_{1,1,i,n} \, d\mu_1 + \int \varphi_{1,2,i,n} \, d\mu_2 \)

\[
\int f_{1,1} \, d\mu_1 + \int f_{1,2} \, d\mu_2 \in Z_1 \times Z_2
\]

where the second equality follows from Proposition 11.125; and the third equality follows from Propositions 3.67, 3.66, and 7.23.
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 \(\lim_{n \to \infty} \mu_n(E) = \mu(E), \forall E \in \mathcal{B}\) 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 \(\mu(\{x \in X \mid \phi(x) > 0\}) < \infty\). By \(\lim_{n \to \infty} \mu_n(E) = \mu(E) < \infty, \forall \mu_n \in \mathbb{N}\) such that \(\mu_n(E) < \infty, \forall n \in \mathbb{N}\). Thus, \(\forall n \in \mathbb{N}\) with \(n_0 \leq n\), \(\phi\) is a simple function of \(X_n := (X, \mathcal{B}, \mu_n)\). By Proposition 11.75, \(\int_X \phi \, d\mu_n = c_1 \in [0, \infty) \subset \mathbb{R}\). Let \(\hat{E} := \{x \in X \mid (f_n(x))_{n=1}^{\infty} \text{ does not converge to } f(x)\}\). By the assumption \(\hat{E} \in \mathcal{B}\) and \(\mu(\hat{E}) = 0\).

Fix any \(\epsilon \in (0, 1) \subset \mathbb{R}\). Define \(E_n := \{x \in X \setminus \hat{E} \mid f_k(x) \geq (1 - \frac{\epsilon}{2+2c_k})\phi(x), \forall k \geq n\}, \forall n \in \mathbb{N}\). By Propositions 7.23, 11.38, and 11.39, \(E_n \in \mathcal{B}, \forall n \in \mathbb{N}\). Clearly, \(\forall x \in X \setminus (\hat{E} \cup \hat{E}), x \in E_n, \forall n \in \mathbb{N}\). Then, \(X \setminus E_n \subseteq \hat{E} \cup \hat{E}, \forall n \in \mathbb{N}\). \(\mu(X \setminus E_n) \leq \mu(\hat{E} \cup \hat{E}) \leq \mu(\hat{E}) < \infty, \forall n \in \mathbb{N}\). \(\forall x \in \hat{E} \setminus \hat{E}\), then \(\lim_{n \to \infty} f_n(x) = f(x) \geq \phi(x) > 0\) and there exists \(n_x \in \mathbb{N}\) such that \(x \in E_{n_x}\). Then, \(\bigcup_{n=1}^{\infty} E_n = X \setminus \hat{E}\). Clearly, \(E_n \subseteq E_{n+1}, \forall n \in \mathbb{N}\). Then, \(X \setminus E_n \supseteq X \setminus E_{n+1}, \forall n \in \mathbb{N}\), and \(\bigcap_{n=1}^{\infty} (X \setminus E_n) = X \setminus (\bigcup_{n=1}^{\infty} E_n) = \hat{E}\). 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 \(\mu(X \setminus E_{n_1}) < \frac{\epsilon}{2+2c_1}\). By the assumption that \(\lim_{n \to \infty} \mu(X \setminus E_n) = \mu(X \setminus E_{n_1}) < \frac{\epsilon}{2+2c_1}\). Then, there exists \(n_2 \in \mathbb{N}\) with \(n_1 \leq n_2, \forall n \in \mathbb{N}\) with \(n_2 \leq n\), we have \(\mu(X \setminus E_{n_1}) < \frac{\epsilon}{2+2c_1}\). Fix any \(n \in \mathbb{N}\) with \(n_2 \leq n\), \(X \setminus E_{n_1} \supseteq X \setminus E_n\), which implies that \(0 \leq \mu(X \setminus E_{n_1}) \leq \mu(X \setminus E_n) < \frac{\epsilon}{2+2c_1}\). Thus, by Proposition 11.83, \(\int_X f_n \, d\mu_n \geq \int_{E_n} f_n \, d\mu_n \geq \int_{E_n} (1 - \frac{\epsilon}{2+2c_1})\phi \, d\mu_n = (1 - \frac{\epsilon}{2+2c_1})(\int_X \phi \, d\mu_n - \int_{X \setminus E_n} \phi \, d\mu_n) \geq (1 - \frac{\epsilon}{2+2c_1})(\int_X \phi \, d\mu_n - M\mu_n(X \setminus E_n)) > (1 - \frac{\epsilon}{2+2c_1})\int_X \phi \, d\mu_n - \epsilon/2\). By the assumption that \(\lim_{n \to \infty} \mu_n(E) = \mu(E), \forall E \in \mathcal{B}\) with \(\mu(E) < \infty\), and Proposition 11.75, we have \(\lim_{n \to \infty} \int_X \phi \, d\mu_n = \int_X \phi \, d\mu = c_1\). Then, \(\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq (1 - \frac{\epsilon}{2+2c_1}) c_1 - \epsilon/2 > c_1 - \epsilon = \int_X \phi \, d\mu - \epsilon\). By the arbitrariness of \(\epsilon\), we have \(\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq \int_X \phi \, d\mu\). By the arbitrariness of \(\phi\), we have \(\liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \geq \int_X f \, d\mu\).

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

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 $B$-measurable, $g_n : X \to [0, \infty) \subset \mathbb{R}$ be $B$-measurable, $\forall n \in \mathbb{N}$, $g : X \to [0, \infty) \subset \mathbb{R}$ be $B$-measurable. Assume that

\[(i) \lim_{n \in \mathbb{N}} f_n = f \ a.e. \ in \ \mathcal{X} \ and \ \lim_{n \in \mathbb{N}} g_n = g \ a.e. \ in \ \mathcal{X}, \ where \ \mathcal{X} := (X, B, \mu); \]

\[(ii) |f_n(x)| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N}; \]

\[(iii) \int_X g_n \, d\mu_n \in \mathbb{R}, \forall n \in \mathbb{N}, \ and \ \int_X g \, d\mu = \lim_{n \in \mathbb{N}} \int_X g_n \, d\mu_n \in \mathbb{R}; \]

\[(iv) \lim_{n \in \mathbb{N}} \mu_n(E) = \mu(E), \forall E \in B \ with \ \mu(E) < \infty. \]

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 $\mathcal{X}_n := (X, B, \mu_n)$ 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 $0 \leq |\int_X f \, d\mu| \leq \int_X P \circ f \, d\mu \leq \int_X g \, d\mu < \infty$.

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, $\lim_{n \in \mathbb{N}} (g_n - f_n) = \lim_{n \in \mathbb{N}} (g_n - f_n) \vee 0 = (g - f) \vee 0 = g - f$ a.e. in $\mathcal{X}$ and $\lim_{n \in \mathbb{N}} (g_n + f_n) = \lim_{n \in \mathbb{N}} (g_n + f_n) \vee 0 = (g + f) \vee 0 = g + f$ a.e. in $\mathcal{X}$. By Fatou's Lemma 11.149 and Propositions 11.92 and 3.83, we have $0 \leq \int_X ((g - f) \vee 0) \, d\mu = \int_X (g - f) \, d\mu = \int_X g \, d\mu - \int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X (g_n - f_n) \, d\mu_n = \liminf_{n \in \mathbb{N}} \int_X g_n \, d\mu_n - \int_X f_n \, d\mu_n = \int_X g \, d\mu - \limsup_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \ and \ 0 \leq \int_X ((g + f) \vee 0) \, d\mu = \int_X (g + f) \, d\mu = \int_X g \, d\mu + \int_X f \, d\mu \leq \limsup_{n \in \mathbb{N}} \int_X (g_n + f_n) \, d\mu_n = \limsup_{n \in \mathbb{N}} \int_X g_n \, d\mu_n + \int_X f_n \, d\mu_n = \int_X g \, d\mu + \limsup_{n \in \mathbb{N}} \int_X f_n \, d\mu_n$. Then, by Proposition 3.83, we have $\int_X f \, d\mu \leq \liminf_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \leq \limsup_{n \in \mathbb{N}} \int_X f_n \, d\mu_n \leq \int_X f \, d\mu \in \mathbb{R}$. 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. \(\Box\)

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 $(X, B)$ be a measurable space, $Y$ be a Banach space over $K$, $(\mu_n)_{n=1}^{\infty} \subseteq M_\sigma(X, B, Y)$, $\mu \in M_\sigma(X, B, Y)$, $Z$ be a Banach space over $K$, $\mathcal{W}$ be a separable subspace of $B(Y, Z)$, $W := \text{span}\{\{z \in Z \mid \exists y \in Y, \exists w \in W \triangleright z = wy\}\} \subseteq Z$ be the Banach subspace of $Z$, $f_n : X \to \mathcal{W}$ 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, $\forall n \in \mathbb{N}$, $g : X \to [0, \infty) \subset \mathbb{R}$ be $B$-measurable. Assume that
(i) \( \lim_{n \to \infty} f_n = f \) a.e. in \( X \) and \( \lim_{n \to \infty} g_n = g \) a.e. in \( X \), where \( X := (X, \mathcal{B}, \mu) \);

(ii) \( \| f_n(x) \| \leq g_n(x), \forall x \in X, \forall n \in \mathbb{N} \);

(iii) \( \int_X g_n \, d\mathcal{P} \circ \mu_n \in \mathbb{R}, \forall n \in \mathbb{N}, \) and \( \int_X g \, d\mathcal{P} \circ \mu = \lim_{n \to \infty} \int_X g_n \, d\mathcal{P} \circ \mu_n \in \mathbb{R} \);

(iv) \( \lim_{n \to \infty} \mu_n = \mu \) as defined in Definition 11.143.

Then, \( f \) is absolutely integrable over \( X \), \( \int_X f \, d\mu = \lim_{n \to \infty} \int_X f_n \, d\mu_n \in \mathbb{W} \subseteq \mathbb{Z} \), and \( 0 \leq \| \int_X f \, d\mu \| \leq \int_X g \, d\mathcal{P} \circ \mu < \infty \).

**Proof**

By Proposition 11.132, \( f_n \) is absolutely integrable over \( X_n := (X, \mathcal{B}, \mu_n) \) and \( \int_X f_n \, d\mu_n \in \mathbb{W} \). Clearly, \( \| f(x) \| \leq g(x) \) a.e. in \( X \). By Proposition 11.83, \( f \) is absolutely integrable over \( X \). By Proposition 11.132, \( \int_X f \, d\mu \in \mathbb{W} \) and \( 0 \leq \| \int_X f \, d\mu \| \leq \int_X \mathcal{P} \circ f \, d\mu \leq \int_X g \, d\mathcal{P} \circ \mu < \infty \).

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: \( \mathcal{P} \circ \mu(X) < \infty \). By \( \lim_{n \to \infty} \mu_n = \mu \), we have \( \lim_{n \to \infty} \mathcal{P} \circ \mu_n(X) = \mathcal{P} \circ \mu(X) < \infty \), \( \forall n \in \mathbb{N} \). Fix any \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \). By Proposition 11.64, \( E \in \mathcal{B} \) such that \( \forall A \in \mathcal{B} \) with \( \mathcal{P} \circ \mu(A) < \delta \), we have \( \int_A g \, d\mathcal{P} \circ \mu < \epsilon/6 \).

By Egoroff's Theorem 11.59, \( \exists E \subseteq B \) with \( \mathcal{P} \circ \mu(E) < \delta/2 \) such that \( \left( f_n \right)_{n=1}^{\infty} \) converges uniformly to \( f \) on \( X \setminus E \). Then, \( \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( \| f_n(x) - f(x) \| < \frac{\epsilon}{4 \mathcal{P} \circ \mu(X) + 1} \), \( \forall x \in X \setminus E \). 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 \( E := E \cup E_1 \in \mathcal{B} \). Then, \( \mathcal{P} \circ \mu(E) < \delta \). Hence, \( \int_E g \, d\mathcal{P} \circ \mu < \epsilon/6 \). By Lebesgue Dominated Convergence Theorem 11.150, we have \( \lim_{n \to \infty} \int_X g_n \, d\mathcal{P} \circ \mu_n = \lim_{n \to \infty} \int_X g \, d\mathcal{P} \circ \mu_n = \int_X (g \chi_{X \setminus E} \chi_{X \setminus E}) \, d\mathcal{P} \circ \mu = \int_E g \, d\mathcal{P} \circ \mu < \epsilon/3 \). By the assumption (iv), \( \lim_{n \to \infty} \mathcal{P} \circ (\mu_n - \mu)(X) = 0 \). This implies that \( \exists n_1 \in \mathbb{N} \) with \( n_0 \leq n_1 \), \( \forall n \in \mathbb{N} \) with \( n_1 \leq n \), we have \( 0 \leq \int_E g_n \, d\mathcal{P} \circ \mu_n < \epsilon/3 \). 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 \( \mathcal{P} \circ (\mu_n - \mu)(X) < \mathcal{P} \circ \mu_n(X) < \mathcal{P} \circ \mu(X) + 1/4 \).

Fix any \( n \in \mathbb{N} \) with \( n_2 \leq n \). Clearly, \( f \chi_{X \setminus E} \) is absolutely integrable over \( E \), and \( \int_{X \setminus E} f \, d\mu_n \in \mathbb{Z} \), since \( \| f(x) \| \leq M \) a.e. in \( E \), where \( E := (X \setminus E, \mathcal{B}_X \setminus E, \mu_n, \chi_{X \setminus E}) \) is the finite \( \mathcal{P} \)-valued measure subspace of \( X_n \). We have

\[
\| \int_X f_n \, d\mu_n - \int_X f \, d\mu \| = \| \int_E f_n \, d\mu_n + \int_{X \setminus E} f_n \, d\mu_n - \int_{X \setminus E} f \, d\mu_n + \int_{X \setminus E} f \, d\mu_n - \int_{E} f \, d\mu - \int_{E} f \, d\mu \|.
\]
be a separable subspace of where

\[
\int F_n \, d\mu_n \leq \int F \, d\mu + \int (f_n - f) \, d\mu_n
\]

+ \left\| \int_{X \setminus E} f \, d(\mu_n - \mu) \right\|

\leq \int g_n \, d\mathcal{P} \circ \mu_n + \int g \, d\mathcal{P} \circ \mu + \int_{X \setminus E} \mathcal{P} \circ (f_n - f) \, d\mathcal{P} \circ \mu_n

+ \int_{X \setminus E} \mathcal{P} \circ f \, d\mathcal{P} \circ (\mu_n - \mu)

< \epsilon/3 + \epsilon/6 + \frac{\mathcal{P} \circ (X \setminus E) \epsilon}{4 \mathcal{P} \circ \mu(X) + 1} + M \mathcal{P} \circ (\mu_n - \mu)(X \setminus E) < \epsilon

where the first equality follows from Proposition 11.132; the first inequality follows from Propositions 11.132, 11.145, and 11.146; the second inequality follows from Proposition 11.132; and the third inequality follows from Proposition 11.83. Hence, we have \( \lim_{n \to \infty} \int f \, d\mu_n = \int f \, d\mu \in \mathcal{W} \). This case is proved.

Case 2: \( \mathcal{P} \circ \mu(X) = \infty \). Fix any \( \epsilon \in (0, \infty) \subset \mathbb{R} \). 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 \( \int_X g \, d\mathcal{P} \circ \mu < \epsilon/4 < \int_X \phi \, d\mathcal{P} \circ \mu \leq \int_X g \, d\mathcal{P} \circ \mu < \infty \). Let \( A_0 := \{ x \in X \mid \phi(x) > 0 \} \). Then, \( A_0 \in \mathcal{B} \) and \( \mathcal{P} \circ \mu(A_0) < \infty \). Note that, by Propositions 11.83 and 11.75, \( 0 \leq \int_X \phi \, d\mathcal{P} \circ \mu - \int_{X \setminus A_0} \phi \, d\mathcal{P} \circ \mu = \int_X \phi \, d\mathcal{P} \circ \mu - \int_{A_0} \phi \, d\mathcal{P} \circ \mu \leq \int_X g \, d\mathcal{P} \circ \mu < \epsilon/4 \). This implies that \( 0 \leq \left\| \int_{X \setminus A_0} f \, d\mu - \int_{A_0} f \, d\mu \right\| \leq \frac{\epsilon}{4} \). This leads to \( \left\| \int f \, d\mu - \int f_n \, d\mu_n \right\| \leq \epsilon/4 + \epsilon/4 + \epsilon/4 < \epsilon \). This completes the proof of the theorem.

**Proposition 11.152** Let \( X := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) is a normed linear space over the Banach space \( K \), \( \mathcal{W} \) 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, \( \mathcal{X} := (X, \mathcal{B}, \nu) \) is a finite \( \mathcal{Z} \)-valued measure space and \( 0 \leq \mathcal{P} \circ \nu(E) \leq \int f \, d\mathcal{P} \circ \mu \leq \epsilon/4 \).
\[ \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty, \forall E \in \mathcal{B}. \]

**Proof**

By Proposition 11.125, \( \nu(\emptyset) = \int_{\emptyset} f \, d\mu = 0 \). \( \forall E \in \mathcal{B} \), by Proposition 11.132, \( \nu(E) = \int_E f \, d\mu \in \mathcal{Z} \) and \( 0 \leq \|\nu(E)\| \leq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \). \( \forall E \in \mathcal{B} \). Then, \( \sum_{n=1}^{\infty} \|\nu(E_n)\| \leq \sum_{n=1}^{\infty} \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \), where the equality follows from Proposition 11.83. By Proposition 11.132, we have \( \nu(E) = \int_E f \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mu = \sum_{n=1}^{\infty} \nu(E_n) \in \mathcal{Z} \). Hence, \( \nu \) is a \( \mathcal{Z} \)-valued pre-measure on \( (X, \mathcal{B}) \). \( \forall E \in \mathcal{B} \), \( \mathcal{P} \circ \nu \). This completes the proof of the proposition. \( \square \)

**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 \( \mathcal{B} \), 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 \( \int_X \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty \). Furthermore, \( \nu \) is \( \sigma \)-finite and \( 0 \leq \mathcal{P} \circ \nu \). \( \forall E \in \mathcal{B} \).

**Proof**

Since \( \mathcal{X} \) is \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( \mathcal{P} \circ \mu(X_n) < \infty, \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( (X_n)_{n=1}^{\infty} \) is pairwise disjoint. \( \forall m \in \mathbb{N} \), let \( \bar{X}_m := \{ x \in X \mid m - 1 \leq \|f(x)\| < m \} \subseteq \mathcal{B} \). Clearly, \( (\bar{X}_m)_{m=1}^{\infty} \subseteq \mathcal{B} \) is pairwise disjoint and \( X = \bigcup_{m=1}^{\infty} \bar{X}_m \). \( \forall n, m \in \mathbb{N} \), let \( \mathcal{X}_{m,n} := (\bar{X}_m \cap X_n, \mathcal{B}_{m,n}, \mu_{m,n}) \) be the finite \( \mathcal{Y} \)-valued measure subspace of \( \mathcal{X} \). Then, \( f|_{\mathcal{X}_{m,n}} \) is absolutely integrable over \( \mathcal{X}_{m,n} \). By Propositions 11.152 and 11.132, we may define a finite \( \mathcal{Z} \)-valued measure space \( \bar{\mathcal{X}}_{m,n} := (\bar{X}_m \cap X_n, B_{m,n}, \nu_{m,n}) \) by \( \nu_{m,n}(E) = \int_{\bar{X}_m \cap X_n} f \, d\mu \in \mathcal{Z}, \forall E \in B_{m,n} \). 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 \( \mathcal{X}_{m,n} \) 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 \( \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \infty \). \( \forall m, n \in \mathbb{N} \), let \( E_{m,n} := E \cap \bar{X}_m \cap X_n \subseteq \mathcal{B} \). Then, \( \mathcal{P} \circ \nu(E) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mathcal{P} \circ \nu_{m,n}(E_{m,n}) \leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int_{E_{m,n}} f \, d\mu = \int_E f \, d\mu \in \mathcal{Z} \). 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 \( \mathbb{Z} \)-valued measure we seek.

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

**Proposition 11.154** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( \mathbb{Z} \) be a Banach space over \( \mathbb{K} \), \( \mathcal{W} \) be a separable subspace of \( \mathcal{B}(\mathbb{K}, \mathbb{Z}) = \mathbb{Z} \), 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 \( \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \in [0, \infty] \subset \mathbb{R}_+ \) for all \( E \in \mathcal{B} \), \( \forall \mu \in \mathcal{B} \), and \( \nu(E) = \int_E f \, d\mu \in \mathbb{Z} \), \( \forall \mathcal{P} \circ \nu(E) < \infty \). 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 \( \nu(E) = \int_E f \, d\mu \in \mathbb{Z} \), \( \forall \mathcal{P} \circ \nu(E) < \infty \). Furthermore, \( \nu \) is \( \sigma \)-finite and \( 0 \leq \mathcal{P} \circ \nu(E) \leq \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \), \( \forall \mathcal{P} \circ \nu(E) < \infty \).

Define \( \bar{\nu} : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+ \) by \( \bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mathcal{P} \circ \mu \), \( \forall \mathcal{P} \circ \nu(E) < \infty \). By Proposition 11.116, \( \bar{\nu} \) is a \( \sigma \)-finite measure on \((X, \mathcal{B})\). Fix any \( \mathcal{P} \in \mathcal{B} \). We will distinguish two exhaustive and mutually exclusives cases: Case 1: \( E \in \mathcal{B} \setminus \text{dom}(\bar{\nu}) \); Case 2: \( E \in \mathcal{B} \setminus \text{dom}(\nu) \). Clearly, we have \( \mathcal{P} \circ \nu(E) = \infty \geq \bar{\nu}(E) \). This case is proved. Case 2: \( E \in \text{dom}(\nu) \). Since \( \bar{\nu} \) is \( \sigma \)-finite, then \( \exists (X_n)_{n=1}^\infty \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^\infty X_n \) and \( \bar{\nu}(X_n) < \infty \), \( \forall n \in \mathbb{N} \). 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 \( \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \bar{\nu}(E_n) < \infty \). Fix any \( \mathcal{P} \in \mathcal{B} \), \( \forall F \subseteq E_n \), we have \( 0 \leq \mathcal{P} \circ \nu(F) \leq \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu = \bar{\nu}(E_n) \) and \( \nu(F) = \int_{E_n} f \, d\mu \). Fix any \( \epsilon \in (0, \infty) \subset \mathbb{R} \). Fix any \( n \in \mathbb{N} \).

Let \( \mathcal{E}_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n}) \) be the \( \sigma \)-finite \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \). By Proposition 11.94, \( \exists \) simple function \( \varphi_n : \mathcal{E}_n \to \mathcal{W} \) 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 \( \int_{E_n} \mathcal{P} \circ (f|_{E_n} - \varphi_n) \, d\mathcal{P} \circ \mu \leq 2^{-n-1} \epsilon \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m_n} w_{i,n} \chi_{E_{i,n},E_n} \), where \( m_n \in \mathbb{Z}_+ \), \( w_{i,n} \in \mathcal{W} \) are distinct and none equals to \( \vartheta_{\mathcal{Z}_+} \), and \( E_{1,n}, \ldots, E_{m_n,n} \in \mathcal{B} \) are subsets of \( E_n \), nonempty, pairwise disjoint, and \( \mathcal{P} \circ \nu(E_{i,n}) < \infty \), \( i = 1, \ldots, m_n \). \( \forall i \in \{1, \ldots, m_n\} \), \( \exists w_{i,n} \in \mathbb{Z}_+ \), \( \exists \) 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

\[
\bar{\nu}(E_n) = \int_{E_n} \mathcal{P} \circ f \, d\mathcal{P} \circ \mu < \int_{E_n} \mathcal{P} \circ \varphi_n \, d\mathcal{P} \circ \mu + 2^{-n-2} \epsilon
\]

\[
= \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \| \nu(E_{j,i,n}) \| + \sum_{i=1}^{m_n} \| w_{i,n} \| \mathcal{P} \circ \mu(E_{i,n}) + 2^{-n-2} \epsilon
\]

\[- \sum_{i=1}^{m_n} \sum_{j=1}^{m_{i,n}} \| \nu(E_{j,i,n}) \|.
\]
11.7. CALCULATION WITH MEASURES

\[ \mu_{\text{case, we will denote}} \]

\[ \text{Let } \mu \text{ be a measurable space, } \mu_1 \text{ and } \mu_2 \text{ be measures on } (X, \mathcal{B}). \text{ We will say that } \mu_1 \text{ is absolutely continuous with respect to } \mu_2 \text{ if } \forall A \in \mathcal{B} \text{ with } \mu_2(A) = 0, \text{ we have } \mu_1(A) = 0. \text{ In this case, we will denote } \mu_1 \ll \mu_2. \text{ } \mu_1 \text{ and } \mu_2 \text{ are said to be mutually singular if } \exists A, B \in \mathcal{B} \text{ with } A \cap B = \emptyset \text{ and } X = A \cup B, \text{ such that } \mu_1(A) = \mu_2(B) = 0. \text{ In this case, we will denote } \mu_1 \perp \mu_2.

**Proposition 11.156** Let } \mu, \lambda_1, \text{ and } \lambda_2 \text{ be measures on a measurable space } (X, \mathcal{B}), \mathcal{Y} \text{ be a normed linear space over } \mathbb{K}, \nu_1 \text{ and } \nu_2 \text{ be } \sigma\text{-finite } \mathcal{Y}\text{-valued measures on } (X, \mathcal{B}), a, b \in (0, \infty) \subset \mathbb{R}, \text{ and } \alpha, \beta \in \mathbb{K}. \text{ Then, the following statements hold.

(i) } \lambda_1 \ll \mu \text{ and } \lambda_2 \ll \mu \text{ 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 \( \lambda_1 \perp \mu \) and \( \lambda_2 \perp \mu \), 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, \( \lambda_1 \perp \lambda_2 \). “Necessity” Let \( \lambda_1 \perp \mu \) and \( \lambda_2 \perp \mu \). Then, \( \exists A, B \in \mathcal{B} \) with \( A = X \setminus B \) such that \( \lambda_1(A) = \mu(B) = 0 \), \( i = 1, 2 \).

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 \( 0 \leq \lambda_1(A) + b\lambda_2(A) = (a\lambda_1 + b\lambda_2)(A) \leq a\lambda_1(A_1) + b\lambda_2(A_2) = 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0 \). 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 \( \mathbb{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \) is well-defined by Proposition 11.142.

(vii) By Proposition 11.138, \( \nu_1 + \nu_2 \in \mathbb{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \) 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

\[ P \circ (\nu_1 + \nu_2)(E) \geq P \circ \nu_1(E) + P \circ \nu_2(E) \] by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( P \circ \nu_1(E) + P \circ \nu_2(E) < \infty \); Case 2: \( P \circ \nu_1(E) + P \circ \nu_2(E) = \infty \).

Case 1: \( P \circ \nu_1(E) + P \circ \nu_2(E) < \infty \). Then, \( P \circ \nu_1(E) < \infty \) and \( E \in \text{dom}(\nu_i) \), \( i = 1, 2 \). Without loss of generality, assume that \( X \) is the second 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 \( P \circ \nu_1(E) + P \circ \nu_2(E) \leq P \circ (\nu_1 + \nu_2)(E) \). This case is proved.

Case 2: \( P \circ \nu_1(E) + P \circ \nu_2(E) = \infty \). Without loss of generality, assume \( P \circ \nu_1(E) = \infty \). Since \( \nu_1 \) and \( \nu_2 \) are \( \sigma \)-finite, there exists \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{B} \) such that \( X = \bigcup_{n=1}^{\infty} X_n \) and \( P \circ \nu_1(X_n) < \infty \), \( \forall n \in \mathbb{N} \), \( i = 1, 2 \). Without loss of generality, we may assume that \( X_n \subseteq X_{n+1} \), \( \forall n \in \mathbb{N} \). Let \( E_n := X \cap X_n \), \( \forall n \in \mathbb{N} \). 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 \( \lim_{n \to \infty} P \circ \nu_1(E_n) = P \circ \nu_1(E) = \infty \). \( \forall M \in (0, \infty) \subseteq \mathbb{R} \), \( \exists M_0 \in \mathbb{N} \) such that \( P \circ \nu_1(E_{M_0}) > M + 1 \). Clearly, \( P \circ \nu_1(E_{M_0}) \leq P \circ \nu_1(X_{M_0}) < \infty \), \( i = 1, 2 \). Then, \( \exists n_1 \in \mathbb{Z}_+ \), \( \exists \) pairwise disjoint \( \{E_i\}_{i=1}^{n_1} \subseteq \mathcal{B} \) with \( E_{M_0} = \bigcup_{i=1}^{n_1} E_i \) such that \( P \circ \nu_1(E_{M_0}) < \sum_{i=1}^{n_1} \nu_1(E_i) \) + 1. Then,

\[ M < \sum_{i=1}^{n_1} \nu_1(E_i) = \sum_{i=1}^{n_1} \nu_1(E_i) \cap A + \nu_1(E_i \cap B) \]
\[ = \sum_{i=1}^{n_1} \| \nu_2(E_i \cap B) + \nu_1(\overline{E}_i \cap B) \| \]
\[ = \sum_{i=1}^{n_1} \| (\nu_1 + \nu_2)(E_i \cap B) \| \leq \sum_{i=1}^{n_1} \mathcal{P} \circ (\nu_1 + \nu_2)(E_i \cap B) \]
\[ = \mathcal{P} \circ (\nu_1 + \nu_2)(E_{n_0} \cap B) \leq \mathcal{P} \circ (\nu_1 + \nu_2)(E) \]

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 \( \nu_1(\overline{A}) = \theta_y \), and the fact that \( \forall A \in \mathcal{B} \) with \( \nu(A) \leq B \), we have \( \mathcal{P} \circ \nu_2(A) = 0 \) and \( \nu_2(B) = \theta_y \); 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, \( \alpha \nu_1, \beta \nu_2, \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \). 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, \( \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \). \( \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \), by \( \mathcal{P} \circ \nu_1 \ll \mu \), we have \( \mathcal{P} \circ \nu_1(E) = 0 \), \( i = 1, 2 \). 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, \( \alpha \nu_1 + \beta \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y}) \). By \( \mathcal{P} \circ \nu_1 \perp \mu \), \( \exists A_1, B_1 \in \mathcal{B} \) with \( A_1 = X \setminus B_1 \) such that \( \mathcal{P} \circ \nu_1(A_1) = \mu(B_1) = 0 \), \( i = 1, 2 \). 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 \( 0 \leq \mathcal{P} \circ (\alpha \nu_1 + \beta \nu_2)(A) \leq |\alpha| \mathcal{P} \circ \nu_1 + |\beta| \mathcal{P} \circ \nu_2(A) \leq |\alpha| \mathcal{P} \circ \nu_1(A_1) + |\beta| \mathcal{P} \circ \nu_2(A_2) = 0 \leq \mu(B) \leq \mu(B_1) + \mu(B_2) = 0 \), 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 \( \mathcal{Y} \)-valued measure space, where \( \mathcal{Y} \) 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 \( \mathcal{Y} \) by; \( \forall E \in \mathcal{B} \) with \( E \cap A \in \text{dom}(\mu) \), \( \mu_A(E) := \mu(E \cap A) \in \mathcal{Y} \); \( \forall E \in \mathcal{B} \) 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 \( \mathcal{Y} \) similarly as \( \mu_A \) with \( A \) replaced with \( B \). Then, the following statements hold.

(i) \( \mu_A \) and \( \mu_B \) are \( \mathcal{Y} \)-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 \( \mathcal{Y} \) 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 $\nu_A : \mathcal{B} \to [0, \infty] \subset \mathbb{R}$ 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 $\mu_A(E) = \mu(E \cap A) \in \mathcal{Y}$. $\forall$ pairwise disjoint $(E_n)_{n=1}^\infty \subseteq \mathcal{B}$ with $E = \bigcup_{n=1}^\infty E_n$, clearly, $0 \leq \mathcal{P} \circ \mu(E_n \cap A) \leq \mathcal{P} \circ \mu(E \cap A) < \infty$, $E_n \cap A \in \text{dom}(\mu)$, and $\mu_A(E_n) = \mu(E_n \cap A) \in \mathcal{Y}$, $\forall n \in \mathbb{N}$. By Definition 11.108, $\mu_A(E) = \mu(E \cap A) = \sum_{n=1}^\infty \mu(E_n \cap A) = \sum_{n=1}^\infty \mu_A(E_n)$ and $\sum_{n=1}^\infty \|\mu_A(E_n)\| = \sum_{n=1}^\infty \|\mu(E_n \cap A)\| < \infty$. $\forall E \in \mathcal{B}$ with $\nu_A(E) < \infty$,

$$\nu_A(E) = \mathcal{P} \circ \mu(E \cap A) = \sup_{n \in \mathbb{N}_+} \sum_{1 \leq i < j \leq n} \|\mu(A_i)\|$$

Hence, $\mu_A$ is a $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_A = \nu_A$. By symmetry, $\mu_B$ is a $\mathcal{Y}$-valued measure on $(X, \mathcal{B})$ with $\mathcal{P} \circ \mu_B = \mathcal{P} \circ \mu(E \cap B)$, $\forall E \in \mathcal{B}$.

(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 $\mu$-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. \qed

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
then, \( \mu \), which case the result holds, or it contains a subset of negative measure. In the latter case, let \( (0, \infty) \) thus, either the result holds or we have \( (n) \), let \( E_n := E \cap A_n \in B, \forall n \in \mathbb{N} \). then, \( E = \bigcup_{n=1}^{\infty} E_n \), \((E_n)_{n=1}^{\infty}\) is pairwise disjoint, \( E_n \in \text{dom}(\mu) \), \( E_n \subseteq A_n \), and \( \mu(E_n) \in [0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N} \). this leads to \( \mu(E) = \sum_{i=1}^{\infty} \mu(E_n) \in [0, \infty) \subseteq \mathbb{R} \). 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 \( \bar{E} \in \text{dom}(\mu) \) and \( \mu(\bar{E}) \in \mathbb{R} \). 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 \( E_j \subseteq E \setminus \left( \bigcup_{i=1}^{j-1} E_i \right) \), \( \mu(E_j) < -1/n_j \), and \( \mu(E \setminus \left( \bigcup_{i=1}^{j-1} E_i \right)) \in (0, +\infty) \subseteq \mathbb{R} \), \( \forall j \in \{1, \ldots, k\} \). either \( E \setminus \left( \bigcup_{j=1}^{k} E_j \right) \) 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 \( E_{k+1} \subseteq E \setminus \left( \bigcup_{j=1}^{k} E_j \right) \) with \( \mu(E_{k+1}) < -1/n_{k+1} \). then, \( \mu(E \setminus \left( \bigcup_{j=1}^{k+1} E_j \right)) = \mu(E \setminus \left( \bigcup_{j=1}^{k} E_j \right)) - \mu(E_{k+1}) \in (0, +\infty) \subseteq \mathbb{R} \). 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 \( A := E \setminus \left( \bigcup_{k=1}^{\infty} E_k \right) \in \mathcal{B} \). then, \( E = A \cup \left( \bigcup_{k=1}^{\infty} E_k \right) \), where the sets in the union are pairwise disjoint. then, \( \mu(E) = \mu(A) + \sum_{k=1}^{\infty} \mu(E_k) \), and \( |\mu(E)| - \sum_{k=1}^{\infty} \mu(E_k) < +\infty \). 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 \( A \subseteq E \setminus \left( \bigcup_{j=1}^{k_0-1} E_j \right) \), then \( A \) cannot 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 \( A \subseteq \mathcal{B} \) and a negative set \( B \subseteq \mathcal{B} \) 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 exclusives cases: Case 1: \( X \) is an \( \mathbb{R} \)-valued pre-measure space; Case 2: \( X \) is a \( \sigma \)-finite \( \mathbb{R} \)-valued measure space.

Case 1: \( X \) is an \( \mathbb{R} \)-valued pre-measure space. Let \( \lambda := \sup \{ \mu(A) \in \mathbb{R} \mid A \) 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: \( X \) is a \( \sigma \)-finite \( \mathbb{R} \)-valued 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 \( X_n := (X_n, B_n, \mu_n) \) be the finite \( \mathbb{R} \)-valued measure subspace of \( X \), \( \forall n \in \mathbb{N} \). Fix any \( n \in \mathbb{N} \). \( X_n \) is a finite \( \mathbb{R} \)-valued pre-measure space. By Case 1, there exists \( A_n, B_n \in B_n \subseteq \mathcal{B} \) 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 \( A \cap B = (\bigcup_{n=1}^\infty A_n) \cap (\bigcup_{n=1}^\infty B_n) = \bigcup_{n=1}^\infty \bigcap_{m=1}^\infty (A_n \cap B_m) = \emptyset \). This case is proved.

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

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 \( \mu_+: \mathcal{B} \to [0, \infty) \subseteq \mathbb{R} \) by \( \mu_+(E) = \mu(A \cap E), \forall E \in \mathcal{B} \) and \( \mu_-: \mathcal{B} \to [0, \infty) \subseteq \mathbb{R} \) 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 \( \hat{\mu}_+ \) and \( \hat{\mu}_- \) be any pair of mutually singular finite measures on \((X, \mathcal{B})\) such that \( \mu = \hat{\mu}_+ - \hat{\mu}_- \). Then, \( \exists A, B \in \mathcal{B} \) with \( A \cap B = \emptyset \) and \( A \cup B = X \) such that \( \hat{\mu}_+(B) = \hat{\mu}_-(\hat{A}) = 0 \). Note that \( \hat{\mu}_+(A \cap \hat{A}) = \hat{\mu}_+(A) - \hat{\mu}_-(A \setminus \hat{A}) \leq \hat{\mu}_+(A) = \hat{\mu}_+(A) - \hat{\mu}_-(A) = \hat{\mu}_+(A \cap \hat{A}) - \hat{\mu}_-(A \cap \hat{A}) \leq \hat{\mu}_+(A \cap \hat{A}) = \hat{\mu}_+(A) - \hat{\mu}_-(A \cap \hat{A}) \leq \hat{\mu}_+(A) = \mu_+(A) = \hat{\mu}_+(A \cap \hat{A}) = \hat{\mu}_+(A) = \mu_+(A) - \mu_-(A \cap \hat{A}) = \mu_+(A \cap \hat{A}) - \mu_-(A \cap \hat{A}) = \mu_+(A \cap \hat{A}) - \mu_-(A \cap \hat{A}) \in \mathbb{R} \). This implies that \( \mu_+(A \cap \hat{A}) = \mu_+(A) = \mu_+(A \cap \hat{A}) = \hat{\mu}_+(A) = \mu(A) \) and \( \mu_+(A \setminus \hat{A}) = \mu_+(A \setminus \hat{A}) = \hat{\mu}_+(A \setminus \hat{A}) = \mu_-(A \setminus \hat{A}) = 0 \). \( \forall E \in \mathcal{B} \), we have \( \mu_+(E) = \mu_+(E \cap A) = \mu_+(E \cap A) + \mu_-(E \cap A \setminus \hat{A}) = \mu_+(E \cap A) = \mu_+(E \cap A \setminus \hat{A}) = \mu_+(E \cap A \setminus \hat{A}) = \mu_+(E \cap A) \). \hfill \Box

**Theorem 11.163 (Jordan Decomposition Theorem)** Let \((X, \mathcal{B}, \mu)\) be an \( \sigma \)-finite \( \mathbb{R} \)-valued measure space. Then, there exists a unique pair of mutually singular \( \sigma \)-finite measures \( \mu_+ \) and \( \mu_- \) on \((X, \mathcal{B})\) such that \( \mu = \mu_+ - \mu_- \). Furthermore, \( \mathcal{P} \circ \mu = \mu_+ + \mu_- \).

**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: \( \mu_A(E) = \mu(E \cap A) \). Define \( \mu_B \) to be a function from \( \mathcal{B} \) to \( \mathbb{R} \) by: \( \mu_B(E) = \mu(E \cap B) \). Then, by Proposition 11.157, \( \mu_A \) and \( \mu_B \) are \( \sigma \)-finite \( \mathcal{R} \)-valued measures on \((X, \mathcal{B})\) such that \( \mu = \mu_+ + \mu_- \), \( \mathcal{P} \circ \mu = \mathcal{P} \circ \mu_A + \mathcal{P} \circ \mu_B \), and \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \). Since \( \mu_A(E) \in [0, \infty) \subset \mathbb{R} \) and \( \mathcal{P} \circ \mu_A \) is defined for \( E \in \mathcal{B} \), \( \mathcal{P} \circ \mu_B \) is defined for \( E \in \mathcal{B} \), \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \). Since \( \mu_A(E) \in [0, \infty) \subset \mathbb{R} \) and \( \mathcal{P} \circ \mu_A \) is defined for \( E \in \mathcal{B} \), \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \). Since \( \mu_A(E) \in [0, \infty) \subset \mathbb{R} \) and \( \mathcal{P} \circ \mu_A \) is defined for \( E \in \mathcal{B} \), \( \mathcal{P} \circ \mu_A \perp \mathcal{P} \circ \mu_B \). The pair \( \mu_+ \) and \( \mu_- \) is the pair of mutually singular measure spaces 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 \( \mu_+ \), \( \mu_- \), \( \hat{\mu}_+ \), 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 \( \mu_+(X_n) \in \mathbb{R} \), \( \mu_-(X_n) \in \mathbb{R} \). Then, \( \mathcal{P} \circ
$\mu(X_n) < \infty, \forall n \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. Let $X_n := (X_n, \mathcal{B}_n, \mu_n)$ be the finite $\mathbb{R}$-valued measure subspace of $X$, $(X_n, \mathcal{B}_n, \mu_{n,+})$ be the finite measure subspace of $(X, \mathcal{B}, \mu_+)$, $(X_n, \mathcal{B}_n, \mu_{n,-})$ be the finite measure subspace of $(X, \mathcal{B}, \mu_-)$, $(X_n, \mathcal{B}_n, \tilde{\mu}_{n,+})$ be the finite measure subspace of $(X_n, \mathcal{B}_n, \mu_{n,+})$, and $(X_n, \mathcal{B}_n, \tilde{\mu}_{n,-})$ be the finite measure subspace of $(X_n, \mathcal{B}_n, \mu_{n,-})$. Then, it is easy to see that $\mu_n = \mu_{n,+} - \mu_{n,-} = \tilde{\mu}_{n,+} + \tilde{\mu}_{n,-}, \mu_{n,+} \perp \mu_{n,-}, \tilde{\mu}_{n,+} \perp \mu_{n,-}, \mu_{n,+} \Rightarrow \mu_{n,-}$, and $\mathcal{P} \circ \mu_n = \mu_{n,+} + \mu_{n,-} = \tilde{\mu}_{n,+} + \tilde{\mu}_{n,-}$. By Jordan Decomposition Theorem 11.162, we have $\mu_{n,+} = \tilde{\mu}_{n,+}$ and $\mu_{n,-} = \tilde{\mu}_{n,-}$. By Proposition 11.117, we have $\mu_{n,+} = \tilde{\mu}_{n,+}$ and $\mu_{n,-} = \tilde{\mu}_{n,-}$. Hence, the pair is unique.

This completes the proof of the theorem.

**Proposition 11.164** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\mathbb{C}$-valued pre-measure space. Then, $\mu = \mu_r + i\mu_i$, where $\mu_r : \mathcal{B} \to \mathbb{R}$ and $\mu_i : \mathcal{B} \to \mathbb{R}$ be 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 $\mathbb{C}$-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 $\mu_i(\emptyset) = 0$. Any 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 $\mathbb{C}$-valued pre-measure. We also have $\sum_{n=1}^\infty \mu_r(E_n) = \sum_{n=1}^\infty \text{Re}(\mu(E_n)) = \text{Re}(\sum_{n=1}^\infty \mu(E_n)) = \text{Re}(\mu(\bigcup_{n=1}^\infty E_n)) = \mu_r(\bigcup_{n=1}^\infty E_n)$. 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 $\sum_{n=1}^\infty |\mu_i(E_n)| = \sum_{n=1}^\infty |\mu_i(E_n)| = \text{Im}(\mu(\bigcup_{n=1}^\infty E_n)) = \text{Im}(\mu(\bigcup_{n=1}^\infty E_n))$. 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 $(X, \mathcal{B})$.

Then, $\mathcal{P} \circ \mu(X) < +\infty$. Hence, $\mathcal{X}$ is a finite $\mathbb{C}$-valued measure space.

This completes the proof of the proposition.

**Proposition 11.165** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $\mathbb{C}$-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 $\exists (X_n)_{n=1}^\infty \subseteq \mathcal{B}$ 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. Fix any \(n \in \mathbb{N}\). Let \(X_n := (X_n, \mathcal{B}_n, \mu_n)\) be the finite \(\mathbb{C}\)-valued measure subspace of \(X\). Then, \(X_n\) 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 \(0 \leq \max\{\mathcal{P} \circ \mu_n, r(E), \mathcal{P} \circ \mu_n, i(E)\} \leq \mathcal{P} \circ \mu_n(E) \leq \mathcal{P} \circ \mu_n, r(E) + \mathcal{P} \circ \mu_n, i(E) < \infty, \forall E \in \mathcal{B}_n\). 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 \((X, \mathcal{B}, \mu)\) on \(X\) and the generation process on \((X_n, \mathcal{B}, \mu_n)\) yields a unique \(\sigma\)-finite \(\mathbb{R}\)-valued measure space \((X, \mathcal{B}, \mu_i)\) on \(X\).

\(\forall E \in \mathcal{B}\), let \(E_n := E \cap X_n \in \mathcal{B}_n\), \(\forall n \in \mathbb{N}\). Then, \(\mathcal{P} \circ \mu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E_n)\), where the first equality follows from the fact that \(\mathcal{P} \circ \mu\) is a measure; and the second equality follows from Propositions 11.118 and 11.115. By Propositions 11.118 and 11.115, we have \(\mathcal{P} \circ \mu(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n, r(E_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n, i(E_n) = \mathcal{P} \circ \mu_n(E)\). Similarly, we have \(\mathcal{P} \circ \mu(E) \geq \mathcal{P} \circ \mu_i(E)\). By Propositions 11.118 and 11.115, \(\mathcal{P} \circ \mu(E) \leq \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n, r(E_n) + \mathcal{P} \circ \mu_n, i(E_n) = \mathcal{P} \circ \mu_n, r(E) + \mathcal{P} \circ \mu_n, i(E)\). Hence, we have \(0 \leq \mathcal{P} \circ \mu_n, r(E) + \mathcal{P} \circ \mu_n, i(E) \leq \mathcal{P} \circ \mu(E) \leq \mathcal{P} \circ \mu(E) + \mathcal{P} \circ \mu_i(E) \leq \infty, \forall E \in \mathcal{B}\).

By Proposition 11.138, \(\mu_n, r + i \mu_n, i\) 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 \(\mathcal{P} \circ \mu_n, r(E) + \mathcal{P} \circ \mu_n, i(E) \leq \mathcal{P} \circ \mu(E) < \infty\) and \(E \in \text{dom}(\mu_n, r) \cap \text{dom}(\mu_n, i)\). Then, \(\mu_n, r + i \mu_n, i(E) = \mu_n, r(E) + i \mu_n, i(E) = \mu_n, r(E) + \mu_n, i(E) = \mu_n(E) = \mu(E)\), 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, \(X_n\) is the 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 \(\mu_n = \mu_n, r + i \mu_n, i\).

This completes the proof of the proposition.

\(\Box\)

**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 \(\mathbf{K}\)-valued measure on \((X, \mathcal{B})\), \(\mathcal{Y}\) be a normed linear space over \(\mathbf{K}\), \(\nu\) be a \(\mathcal{Y}\)-valued measure on \((X, \mathcal{B})\), and \(\tilde{f} : X \to \mathcal{Y}\) be \(\mathcal{B}\)-measurable. We will say that \(\tilde{f}\) is a Radon-Nikodym derivative of \(\nu\) with respect to \(\mu\) if \(\mathcal{P} \circ \tilde{f}\) is a Radon-Nikodym derivative of \(\mathcal{P} \circ \nu\) with respect to \(\mathcal{P} \circ \mu\) (that is, \(\mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ \tilde{f} \, d\mathcal{P} \circ \mu \in [0, \infty] \subset \mathbf{R}_e, \forall E \in \mathcal{B}\) and \(\nu(E) = \int_E \tilde{f} \, d\mu \in \mathcal{Y}, \forall E \in \text{dom}(\nu)\)).

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 11.115 guarantees that any measurable function \(\tilde{f} : X \to \mathcal{W}\)
Proposition 11.167 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $\mathbb{K}$-valued measure space, $\nu$ be a $\sigma$-finite $\mathbb{Y}$-valued measure on $(X, \mathcal{B})$, $f : X \to \mathbb{Y}$ be a Radon-Nikodym derivative of $\nu$ with respect to $\mu$, and $g : X \to \mathbb{Y}$ be $\mathcal{B}$-measurable. Then, $g$ is a Radon-Nikodym derivative of $\nu$ with respect to $\mu$ if, and only if, $g = f$ a.e. in $\mathcal{X}$. In this case, we will write $\frac{d \nu}{d \mu} = f$ 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 $E := (E, \mathcal{B}_E, \mu_E)$, which is the $\mathbb{K}$-valued measure subspace of $\mathcal{X}$. This implies that, by Proposition 11.132, $\nu(E) = \int_E f \ d\mu = \int_E g \ d\mu \in \mathbb{Y}$. Hence, $g$ is a Radon-Nikodym derivative of $\nu$ with respect to $\mu$.

“Necessity” Let $g : X \to \mathbb{Y}$ be another Radon-Nikodym derivative of $\nu$ with respect to $\mu$. Define $\lambda := \nu - \nu$, which is a $\sigma$-finite $\mathbb{Y}$-valued measure on $(X, \mathcal{B})$, by Proposition 11.138. Then, by Proposition 11.137, $\lambda(E) = \theta_\mathbb{Y}$ and $\mathcal{P} \circ \lambda(E) = 0$, $\forall E \in \mathcal{B}$. By Proposition 11.154, we may define a $\sigma$-finite $\mathbb{Y}$-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 $E$. By Proposition 11.132, $(f - g)|_E$ is absolutely integrable over $E$. Then, we have $\lambda(E) = \nu(E) - \nu(E) = \int_E f \ d\mu - \int_E g \ d\mu = \int_E (f - g) \ d\mu = \hat{\lambda}(E)$. By Proposition 11.137, we have $\lambda = \hat{\lambda}$. Then, $\forall E \in \mathcal{B}$, $E \in \text{dom}(\lambda)$, $0 = \mathcal{P} \circ \lambda(E) = \mathcal{P} \circ \hat{\lambda}(E) = \int_E \mathcal{P} \circ (f - g) \ d\mathcal{P} \circ \mu$. 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.

Proposition 11.168 Let $(X, \mathcal{B})$ be a measurable space, $\lambda$ and $\lambda_1$ be $\sigma$-finite measures on $(X, \mathcal{B})$, $\hat{\lambda} := (X, \mathcal{B}, \lambda)$, $\lambda_1 := (X, \mathcal{B}, \lambda_1)$, $\mu, \mu_1 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathbb{K})$, $\mathcal{X} := (X, \mathcal{B}, \mu)$, $\mathcal{X}_1 := (X, \mathcal{B}, \mu_1)$, $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$, $\mathcal{Z}$ be a Banach space over $\mathbb{K}$, $\mathcal{W}$ be a separable subspace of $\mathcal{B}(\mathbb{Y}, \mathbb{Z})$, $\nu_1, \nu_2 \in \mathcal{M}_\sigma(X, \mathcal{B}, \mathcal{Y})$, $\hat{\mathcal{X}} := (X, \mathcal{B}, \nu_1)$, $\frac{d \nu_1}{d \nu_2} := f : X \to \mathbb{Y}$ a.e. in $\mathcal{X}$, $i = 1, 2$, $\frac{d \lambda_1}{d \lambda} := f : X \to [0, \infty) \subset \mathbb{R}$ a.e. in $\mathcal{X}_1$, $\frac{d \nu}{d \lambda} := f : X \to \mathbb{K}$ a.e. in $\mathcal{X}$, $g : X \to \mathcal{W}$ be $\mathcal{B}$-measurable, $\hat{f} : X \to [0, \infty) \subset \mathbb{R}$ be $\mathcal{B}$-measurable, $\hat{f} : X \to \mathcal{K}$ be $\mathcal{B}$-measurable, $\alpha_0 \in \mathbb{K}$, $A \in \mathcal{B}(\mathbb{Y}, \mathbb{Z})$, and $y_0 \in \mathbb{Y}$. Then, the following statements hold.

(i) $\int_X f \ d\lambda = \int_X (\hat{f} \hat{f}) \ d\lambda_1 \in [0, \infty) \subset \mathbb{R}$.
(ii) $g$ is absolutely integrable over $X_1$ if, and only if, $(P \circ g)(P \circ f_1)$ is integrable over $(X, \mathcal{B}, P \circ \mu)$. In which case, $\int_X g \, d\nu_1 = \int_X (g f_1) \, d\mu \in \mathbb{R}$.

(iii) $\frac{d(\nu_1 + \nu_2)}{d\mu} = \tilde{f}_1 + \tilde{f}_2 = \frac{d\nu_1}{d\mu} + \frac{d\nu_2}{d\mu}$ a.e. in $X$.

(iv) $\hat{f}_1$ is a Radon-Nikodym derivative of $A \nu_1$ with respect to $\mu$.

(v) $\frac{d(\alpha \nu)}{d\mu} = \alpha \tilde{f}_1 = \alpha_0 \frac{d\nu}{d\mu}$ a.e. in $X$.

(vi) $\frac{d\alpha x}{d\mu} = f y_0 = \left( \frac{d\alpha}{d\mu} \right) y_0$ a.e. in $X_1$.

(vii) $\frac{d\alpha x}{d\mu} = \tilde{f}_1 f = \left( \frac{d\alpha}{d\mu} \right) \left( \frac{df}{d\mu} \right)$ a.e. in $X_1$.

(viii) $\frac{d\alpha x}{d\mu} = \hat{f}$ 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 $(\tilde{\phi}_n)_{n=1}^{\infty}$, $\tilde{\phi}_n : \tilde{X} \to [0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, such that $0 \leq \tilde{\phi}_n(x) \leq \tilde{f}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \to \infty} \tilde{\phi}_n = \tilde{f}$ a.e. in $\tilde{X}$, where $\tilde{X} := (X, \mathcal{B}, \lambda + \lambda_1)$ is a $\sigma$-finite measure space by Proposition 11.136. $\forall n \in \mathbb{N}$, define $\phi_n := \bigvee_{i=1}^{n} \phi_i$. Then, $\phi_n$ is a simple function of $\tilde{X}$ to $[0, \infty) \subset \mathbb{R}$, $0 \leq \phi_n(x) \leq \phi_n(x) \leq \tilde{f}(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. By Proposition 11.50, $\lim_{n \to \infty} \phi_n = \tilde{f}$ a.e. in $\tilde{X}$. Since $\lambda \leq \lambda + \lambda_1$, then $\phi_n$ is a simple function of $\tilde{X}$, $\forall n \in \mathbb{N}$, and $\lim_{n \to \infty} \phi_n = \tilde{f}$ a.e. in $\tilde{X}$. By Monotone Convergence Theorem 11.81, $\int_X f \, d\lambda = \lim_{n \to \infty} \int_X \phi_n \, d\lambda \in \mathbb{R}_+$. Fix any $n \in \mathbb{N}$. Let $\phi_n$ admit the canonical representation $\phi_n = \sum_{i=1}^{n} a_i \chi_{A_i}$. Then, by Propositions 11.75 and 11.83, $\int_X f \, d\lambda = \sum_{i=1}^{\infty} a_i \mu_i = \sum_{i=1}^{n} a_i \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1 = \sum_{i=1}^{n} \int_{A_i} f \, d\lambda_1$.

(ii) Since $\frac{d\alpha x}{d\mu} = \tilde{f}_1$ a.e. in $X$, then $\frac{d(\alpha \nu)}{d\mu} = \alpha \frac{d\nu}{d\mu}$ a.e. in $X$. By (i), $\int_X P \circ g \, dP \circ \nu_1 = \int_X (P \circ g)(P \circ f_1) \, dP \circ \mu \in \mathbb{R}_+$. Hence, $g$ is absolutely integrable over $X_1$ if, and only if, $(P \circ g)(P \circ f_1)$ is integrable over $(X, \mathcal{B}, P \circ \mu)$. Let $g$ be absolutely integrable over $X_1$, then $\int_X P \circ g \, dP \circ \nu_1 = \int_X (P \circ g)(P \circ f_1) \, dP \circ \mu < \infty$. By Proposition 11.66, there exists a sequence of simple functions $(\phi_n)_{n=1}^{\infty}$, $\phi_n : \tilde{X} \to \mathcal{W}$, $\forall n \in \mathbb{N}$, such that $\| \phi_n(x) \| \leq \| g(x) \|$, $\forall x \in X$, $\forall n \in \mathbb{N}$, and $\lim_{n \to \infty} \phi_n = g$ a.e. in $\tilde{X}$, where $\tilde{X} := (X, \mathcal{B}, P \circ \mu + P \circ \nu_1)$ is a $\sigma$-finite measure space by Proposition 11.136. Since $P \circ \nu_1 \leq P \circ \mu + P \circ \nu_1$, then $\phi_n$ is a simple function.
of \( \bar{X}_1 \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \phi_n = g \) a.e. in \( \bar{X}_1 \). By Lebesgue Dominated Convergence Theorem 11.131, \( \int_X g \, d\nu_1 = lim_{n \to \infty} \int_X \phi_n \, d\nu_1 \in \mathbb{Z} \). 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 \( \mathcal{W} := span \{ \{ z \in \mathbb{Z} \mid z = wy, w \in \mathcal{W}, y \in \mathcal{Y} \} \) is a separable subspace of \( \mathbb{Z} \). 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 \( \mathcal{P} \circ (\phi_n f_1)(x) = \| \phi_n(x) f_1(x) \| \leq \| \phi_n(x) \| \| f_1(x) \| \leq (\mathcal{P} \circ g(x))(\mathcal{P} \circ f_1(x)), \forall x \in X, \forall n \in \mathbb{N} \). Note also that \( \phi_n f_1 : X \to \mathcal{W} \subseteq \mathbb{Z} \) and \( g f_1 : X \to \mathcal{W}, \forall n \in \mathbb{N} \). 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 \( \bar{X}_1 \). 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, \( \int_X (g f_1) \, d\mu = \lim_{n \to \infty} \int_X (\phi_n f_1) \, d\mu = \lim_{n \to \infty} \int_X \phi_n f_1 \, d\nu_1 = \int_X g \, d\nu_1 \in \mathbb{Z} \).

(iii) By Proposition 11.154, \( f_1 + f_2 \) is a Radon-Nikodym derivative of a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure \( \nu \) on \( (X, \mathcal{B}) \) 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 \( \mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < \infty \), \( i = 1, 2 \). This implies that \( f_i |_E \), \( i = 1, 2 \), are absolutely integrable over \( E := (E, \mathcal{B}_E, \mu_E) \), which is the \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \). Then, \( \nu(E) = \int_E f_1 + f_2 \, d\mu = \int_E f_1 \, d\mu + \int_E f_2 \, d\mu = \nu_1(E) + \nu_2(E) = (\nu_1 + \nu_2)(E) \in \mathcal{Y} \), 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 \( (X, \mathcal{B}) \) 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 \( A f_1 : X \to \mathcal{R}(A) \subseteq \mathbb{Z} \), where \( \mathcal{R}(A) \) is a separable subspace of \( \mathbb{Z} \) 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 \( (X, \mathcal{B}) \) with respect to \( \mu \). \( \forall E \in \mathcal{B} \) with \( (\mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu)(E) < \infty \), then \( \mathcal{P} \circ \nu_1(E) = \int_E \mathcal{P} \circ f_1 d\mathcal{P} \circ \mu < \infty \), which is the \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \). Then, \( \nu(E) = \int_E (A f_1) \, d\mu = A \int_E f_1 \, d\mu = A \nu_1(E) = (A \nu)(E) \in \mathbb{Z} \), where the second equality follows from Proposition 11.132; and the last equality follows from Proposition 11.138. Note that \( \mathcal{P} \circ \nu_1 + \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \) by Proposition 11.136. By Proposition 11.137, we have \( \nu = A \nu_1 \). 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 \( (X, \mathcal{B}) \) with respect to \( \mu_1 \). \( \forall E \in \mathcal{B} \) with \( (\mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu + \mathcal{P} \circ \nu)(E) < \infty \), 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 \( \mathcal{P} \circ \mu + \mathcal{P} \circ \mu_1 + \mathcal{P} \circ \nu + \mathcal{P} \circ \nu \) is a \( \sigma \)-finite measure on \( (X, \mathcal{B}) \) by Proposition 11.136. By Proposition 11.137, we have
\( \tilde{\nu} = \nu_1 \). Hence, by Proposition 11.167, \( \frac{d\tilde{\mu}}{d\mu} = \tilde{f}f \) a.e. in \( \mathcal{X}_1 \).

(viii) “Necessity” Let \( \frac{d\mu}{d\mu} = \tilde{f}f \) a.e. in \( \mathcal{X} \). By (vii), \( \frac{d\mu}{d\nu} = \tilde{f}f \) a.e. in \( \mathcal{X}_1 \). 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 \( \mathcal{X}_1 \).

“Sufficiency” Let \( \tilde{f} \) be such that \( \tilde{f}f = 1 \) a.e. in \( \mathcal{X}_1 \). 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 \( \mu \). \( \forall \mathcal{E} \in \mathcal{B} \) with \((\mathcal{P} \circ \mu_1 + \mathcal{P} \circ \tilde{\mu})(\mathcal{E}) < \infty \),

\[ \tilde{\mu}(\mathcal{E}) = \int_{\mathcal{E}} \tilde{f}d\mu = \int_{\mathcal{E}} (\tilde{f}f) d\mu_1 = \int_{\mathcal{E}} 1 d\mu_1 = \mu_1(\mathcal{E}) \in \mathbb{K}, \]

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, \( \frac{d\mu}{d\mu} = \tilde{f}f \) a.e. in \( \mathcal{X} \).

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

**Theorem 11.169 (Radon-Nikodym)** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) 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 \( \mathcal{X} \).

**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 \( \mathcal{F} \neq \emptyset \).

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, \( f_1 \leq f_2 \) is \( \mathcal{B} \)-measurable. Then, \( B := \{ x \in X \mid f_1(x) - f_2(x) \geq 0 \} \in \mathcal{B} \) 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, \( f_1 \lor f_2 \) is \( \mathcal{B} \)-measurable. \( \forall A \in \mathcal{B} \), by Proposition 11.83, \( \int_A (f_1 \lor f_2) d\mu = \int_{A \cap B} (f_1 \lor f_2) d\mu + \int_{A \cap C} (f_1 \lor f_2) d\mu = \int_{A \cap B} f_1 d\mu + \int_{A \cap C} f_2 d\mu \leq \nu(A \cap B) + \nu(A \setminus C) = \nu(A) \).

Therefore, \( f_1 \lor f_2 \in \mathcal{F} \). 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 \( \int_A f d\mu = \int_A (f \chi_{A,X}) d\mu = \lim_{n \to \infty} \int_A (f_n \chi_{A,X}) d\mu = \lim_{n \to \infty} \int_A f_n d\mu \leq \nu(A) \). 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 \( \tilde{\nu} : \mathcal{B} \to \mathbb{R} \) 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 \( \tilde{\nu} \) with \( \tilde{\nu}(A_0) \in (0, \infty) \subset \mathbb{R}. \) Clearly, \( 0 < \tilde{\nu}(A_0) \leq \nu(A_0). \) By \( \nu \ll \mu \) and \( \mu \) is finite, we have \( \mu(A_0) \in (0, \infty) \subset \mathbb{R}. \) \( \forall E \in \mathcal{B}, \nu(E) = \nu(E \cap A_0) + \nu(E \setminus A_0) \geq \tilde{\nu}(E \cap A_0) + \epsilon_0 \mu(E \cap A_0) + \int_{E \setminus A_0} f \, d\mu + \int_E f \, d\mu \geq \epsilon_0 \mu(E \cap A_0) + \int_E f \, d\mu, \) where the last inequality follows from Proposition 11.83. Let \( f_1 := \epsilon_0 \chi_{A_0}. \) By Propositions 11.38, 11.39, and 7.23, \( f + f_1 \) is \( \mathcal{B} \)-measurable. \( \forall E \in \mathcal{B}, \) \( \int_E (f + f_1) \, d\mu = \int_E f \, d\mu + \int_E f_1 \, d\mu = \int_E f \, d\mu + \epsilon_0 \mu(E \cap A_0) \leq \nu(E), \) where the first equality follows from Proposition 11.83; and the second equality follows from Proposition 11.75. Hence, \( f + f_1 \in \mathcal{F}. \) Then, \( \int_X (f + f_1) \, d\mu = \int_X f \, d\mu + \int_X f_1 \, d\mu = \lambda + \epsilon_0 \mu(A_0) > \lambda. \) This contradicts with the definition of \( \lambda. \) Therefore, the claim holds.

Hence, \( f \) is a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu. \) This special case is proved.

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. \( \forall E \in \mathcal{B}, \) \( \nu(E) = \sum_{n=1}^{\infty} \nu_n(E \cap X_n) = \sum_{n=1}^{\infty} \int_{E \cap X_n} f_n \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f_n \, d\mu_n = \sum_{n=1}^{\infty} \int_{E \cap X_n} f_n \, d\mu = \int_E f \, d\mu, \) 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 \( \mathcal{P} \circ \nu = \nu_+ + \nu_- \). By \( \mathcal{P} \circ \nu \ll \mu, \) 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 \( f := f_+ - f_- \). \( \forall E \in \text{dom}(\nu), \) 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, \( f_+|_E \) and \( f_-|_E \) are (absolutely) integrable over \( \mathcal{E} := (\mathcal{B}, \mathcal{B}_E, \mu_E) \), 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 \( \nu = \nu \). 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 \( \mathcal{X} := (X, \mathcal{B}) \) be a measurable space, \( \mathcal{Y} \) be a separable normed linear space over \( \mathbb{K} \), and \( f : X \to \mathcal{Y} \). Assume that \( \mathcal{Y}^* \) is separable, \( \mathcal{D}_s \subseteq \mathcal{Y}^* \) is a countable dense set, and, \( \forall y* \in \mathcal{D}_s, f_{y*} : X \to \mathbb{K} \) 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 \( D \subseteq \mathcal{Y} \) be a countable dense set. By Proposition 4.4, \( \mathcal{Y} \) is second countable with a countable basis \( \mathcal{E} := \{ \mathcal{B}(y_0, r) \mid y_0 \in D, r \in (\mathbb{Q}, r > 0) \} \). By Proposition 11.34, all we need to show is that \( \forall \mathcal{B}_0(y_0, r) \in \mathcal{E}, V := f_{\text{inv}}(\mathcal{B}_0(y_0, r)) \in \mathcal{B} \). \( \exists N \in \mathbb{N} \) such that \( r > 1/N \). We will show that \( V = U := \bigcup_{n=N}^{\infty} \bigcap_{y* \in \mathcal{D}_s, y* \neq y_0} f_{y*}(\mathcal{B}(\langle \langle y*, y_0 \rangle \rangle, \|y*\|(r - 1/n))) \in \mathcal{B} \). \( \forall n \in \mathbb{N} \) with \( n \geq N \), \( \forall y* \in \mathcal{D}_s \) with \( y* \neq y_0 \), \( f_{y*} \) is \( \mathcal{B} \)-measurable implies that \( f_{y*}(\mathcal{B}(\langle \langle y*, y_0 \rangle \rangle, \|y*\|(r - 1/n))) \in \mathcal{B} \). Hence, \( U \subseteq \mathcal{B} \).

\( \forall x \in V \), we have \( \|f(x) - y_0\| < r \). Then, \( \exists n \in \mathbb{N} \) with \( n \geq N \) such that \( \|f(x) - y_0\| < r - 1/n \). \( \forall y* \in \mathcal{D}_s \) with \( y* \neq y_0 \), we have \( f_{y*}(x) - \langle \langle y*, y_0 \rangle \rangle < \|y*\|(r - 1/n) \). Hence, \( x \in f_{y*}(\mathcal{B}(\langle \langle y*, y_0 \rangle \rangle, \|y*\|(r - 1/n))) \). Then, \( x \in U \). By the arbitrariness of \( x \), we have \( U \subseteq \mathcal{V} \).

On the other hand, \( \forall x \in U \), \( \exists n \in \mathbb{N} \) with \( n \geq N \), \( \forall y* \in \mathcal{D}_s \) with \( y* \neq y_0 \), we have \( |f_{y*}(x) - \langle \langle y*, y_0 \rangle \rangle| < \|y*\|(r - 1/n) \). Then, \( |f_{y*}(x) - \langle \langle y*, y_0 \rangle \rangle| < \|y*\|(r - 1/n) \). If \( f(x) = y_0 \), then \( x \in V \). If \( f(x) \neq y_0 \), by Proposition 7.85, \( \exists y* \in \mathcal{Y} \) with \( \|y_0\| = 1 \) such that \( \|f(x) - y_0\| = \langle \langle y*, f(x) - y_0 \rangle \rangle \). Since \( \mathcal{D}_s \) is dense in \( \mathcal{Y}^* \), then, by Proposition 4.13, \( \exists (y_{n})_{n=1}^{\infty} \subseteq \mathcal{D}_s \) that converges to \( y_0 \). Without loss of generality, assume \( y_{n} \neq y_{n}, \forall i \in \mathbb{N} \). By Propositions 3.66, 3.67, 7.21, and 7.72, we have \( \|f(x) - y_0\| = \lim_{n \in \mathbb{N}} |\langle \langle y_{n}, f(x) - y_0 \rangle \rangle| \leq \lim_{n \in \mathbb{N}} \|y_{n}\|(r - 1/n) = r - 1/n < r \). 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 \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite \( \mathbb{K} \)-valued measure space, \( \mathcal{Y} \) be a separable reflexive Banach space over \( \mathbb{K} \), \( \nu \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on \((X, \mathcal{B})\). Assume that \( \mathcal{P} \circ \nu \ll \mathcal{P} \circ \mu \) and \( \mathcal{Y}^* \) is separable. Then, there exists \( f : X \to \mathcal{Y} \) such that \( f \) is the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \). The function \( f \) is unique in the sense that \( g : X \to \mathcal{Y} \) 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 \( \mathcal{X} \).
11.8. The Radon-Nikodym Theorem

Proof. First, we consider the special case when $X$ is a $\sigma$-finite measure space and $\mathcal{Y} = \mathbb{C}$. By Proposition 11.165, $\nu = \nu_r + \nu_i$, where $\nu_r$ and $\nu_i$ are $\sigma$-finite $\mathbb{R}$-valued measures on $(X, \mathcal{B})$, and $\max\{\mathcal{P} \circ \nu_r(E), \mathcal{P} \circ \nu_i(E)\} \leq \mathcal{P} \circ \nu(E) \leq \mathcal{P} \circ \nu_r(E) + \mathcal{P} \circ \nu_i(E)$, $\forall E \in \mathcal{B}$. 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 $X$ and \( \frac{d\nu_i}{d\mu} = f_i \) a.e. in $X$. By Proposition 11.168, we have \( \frac{d\nu}{d\mu} = \frac{d(\nu_r + \nu_i)}{d\mu} = \frac{d\nu_r}{d\mu} + \frac{i d\nu_i}{d\mu} = f_r + if_i = f : X \to \mathbb{C} \) a.e. in $X$. This completes the proof of the special case.

Next, we consider the special case when $X$ is a $\sigma$-finite measure space, $\forall y_s \in \mathcal{Y}^*$, define $\nu_y$ to be the $\sigma$-finite $\mathbb{K}$-valued measure on $(X, \mathcal{B})$ such that $\nu_y(E) = \langle \langle y_s, \nu(E) \rangle \rangle$, $\forall E \in \text{dom}(\nu)$, as shown in Proposition 11.138. Then, by Proposition 11.138, $\mathcal{P} \circ \nu_y \ll \|y_s\| \mathcal{P} \circ \nu$. 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, $\exists f_{y_s} : X \to \mathbb{K}$ such that \( \frac{d\nu_y}{d\mu} = f_{y_s} \) a.e. in $X$. $\forall y_{s_1}, y_{s_2} \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}$, $\forall E \in \text{dom}(\nu)$, we have $E \in \text{dom}(\nu_{y_{s_1}}) \cap \text{dom}(\nu_{y_{s_2}})$.

Then, $\mathbb{K} \ni \nu_{y_{s_1} + y_{s_2}}(E) = \int_E f_{y_{s_1} + y_{s_2}} \ dm = \langle \langle \alpha y_{s_1} + \beta y_{s_2}, \nu(E) \rangle \rangle = \alpha \langle \langle y_{s_1}, \nu(E) \rangle \rangle + \beta \langle \langle y_{s_2}, \nu(E) \rangle \rangle = \alpha f_{y_{s_1}}(E) + \beta f_{y_{s_2}}(E) = \alpha \int_E f_{y_{s_1}} \ dm + \beta \int_E f_{y_{s_2}} \ dm \leq \int_E \mathcal{P} \circ \nu_{y_{s_1}} \ dm \leq \infty$, $\int_E \mathcal{P} \circ \nu_{y_{s_2}} \ dm \leq \infty$, and $\int_E \mathcal{P} \circ f_{y_{s_1}} \ dm \leq \infty$, and $\int_E \mathcal{P} \circ f_{y_{s_2}} \ dm \leq \infty$.

Then, $f_{y_{s_1}} \in \mathcal{E}$ and $f_{y_{s_2}} \in \mathcal{E}$ are absolutely integrable over $\mathcal{E} := (E, B_E, \mu_E)$, which is the $\sigma$-finite measure subspace of $X$. By Proposition 11.92, we have $\nu_{y_{s_1} + y_{s_2}}(E) = \int_E f_{y_{s_1} + y_{s_2}} \ dm = \int_E (\alpha f_{y_{s_1}} + \beta f_{y_{s_2}}) \ dm \in \mathbb{K}$, $\forall E \in \mathcal{B}$ with $\mathcal{P} \circ \nu(E) < \infty$. By Propositions 11.116 and 11.137, $\alpha f_{y_{s_1}} + \beta f_{y_{s_2}}$ is a Radon-Nikodym derivative of $\nu_{y_{s_1} + y_{s_2}}$ with respect to $\mu$. Then, by Proposition 11.167,

$$f_{y_{s_1} + y_{s_2}} = \alpha f_{y_{s_1}} + \beta f_{y_{s_2}} \ a.e. \ in \ X; \ \forall y_{s_1}, y_{s_2} \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}$$

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, $\exists f : X \to [0, \infty) \subset \mathbb{R}$ such that \( \frac{d\mathcal{P} \circ \nu}{d\mu} = f \) a.e. in $X$. This yields $\mathcal{P} \circ \nu(E) = \int_E f \ dm$, $\forall E \in \mathcal{B}$. $\forall y_s \in \mathcal{Y}^*$, by Definition 11.166, $\mathcal{P} \circ \nu_y(E) = \int_E (\mathcal{P} \circ f_y) \ dm (\mathcal{P} \circ \nu_y)(E)$, $\forall E \in \mathcal{B}$. When $\|y_s\| > 0$, $\mathcal{P} \circ \nu_y(E) \leq (\|y_s\| \mathcal{P} \circ \nu)(E) = \|y_s\| \int_E f \ dm = \int_E (\|y_s\| f) \ dm$, where the first equality follows from Proposition 11.136; and the third equality follows from Proposition 11.83. When $\|y_s\| = 0$, $\mathcal{P} \circ \nu_y(E) \leq 0 = \int_E (\|y_s\| f) \ dm$, where the inequality follows from Proposition 11.138; and the equality follows from Proposition 11.75. Hence, $\mathcal{P} \circ \nu_y(E) = \int_E (\mathcal{P} \circ f_y) \ dm \leq \int_E (\|y_s\| f) \ dm$, $\forall E \in \mathcal{B}$, $\forall y_s \in \mathcal{Y}^*$.

Then, by Propositions 11.96,

$$\mathcal{P} \circ f_y \leq \|y_s\| \hat{f} \ a.e. \ in \ X; \ \forall y_s \in \mathcal{Y}^*$$

Let $\mathbb{K}_Q := \mathbb{Q}$ if $\mathbb{K} = \mathbb{R}$ and $\mathbb{K}_Q := \{\alpha + i\beta \in \mathbb{C} \mid \alpha, \beta \in \mathbb{Q}\}$ if $\mathbb{K} = \mathbb{C}$. Clearly, $\mathbb{K}_Q$ is a countable dense set in $\mathbb{K}$. 


Let $D \subseteq \mathcal{Y}^*$ be a countable dense set and \( \hat{D} := \\{ \sum_{i=1}^n \alpha_i y_i \in \mathcal{Y}^* \mid n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in \mathbb{K}, y_1, \ldots, y_n \in D \} \). Then, \( \hat{D} \subseteq \mathcal{Y}^* \) is also a countable dense set. Note that, \( \forall y_1, y_2 \in \hat{D}, \forall \alpha, \beta \in \mathbb{K} \), we have \( \alpha y_1 + \beta y_2 \in \hat{D} \).

Define

\[
E_0 := \{ x \in X \mid \exists y_1, y_2 \in \hat{D}, \exists \alpha, \beta \in \mathbb{K} : f_{\alpha y_1 + \beta y_2}(x) \neq \alpha f_{y_1}(x) + \beta f_{y_2}(x) \}
\]

Then, \( E_0 \in \mathcal{B} \) and \( \mu(E_0) = 0 \).

We have the following results.

\[
f_{\alpha y_1 + \beta y_2}(x) = \alpha f_{y_1}(x) + \beta f_{y_2}(x);
\]

\[
\forall x \in X \setminus E_0, \forall y_1, y_2 \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}
\]

and

\[
|f_y(x)| \leq \|y\| \hat{f}(x); \quad \forall x \in X \setminus E_0, \forall y \in \hat{D}
\]

\[
\forall x \in X \setminus E_0, \forall y \in \mathcal{Y}^*, \forall \alpha, \beta \in \mathbb{K}, \exists (\alpha_i, \beta_i, i)_{i=1}^\infty \subseteq \hat{D} \text{ that converges to } y_i, \exists (\alpha_i)_{i=1}^\infty \subseteq \mathbb{K} \text{ that converges to } \alpha, \text{ and } \exists (\beta_i)_{i=1}^\infty \subseteq \mathbb{K} \text{ that converges to } \beta.
\]

By Propositions 7.23, 3.66, and 3.67, we have

\[
\lim_{i \to \infty} (\alpha_i \gamma_{y_1, i} + \beta_i \gamma_{y_2, i}) = \alpha y_1 + \beta y_2, \text{ where the sequence on the left-hand-side is contained in } \hat{D}.
\]

This implies that

\[
F(x, \alpha y_1 + \beta y_2) = \lim_{i \to \infty} f_{\alpha_i \gamma_{y_1, i} + \beta_i \gamma_{y_2, i}}(x) = \lim_{i \to \infty} (\alpha f_{y_1, i}(x) + \beta f_{y_2, i}(x)) = \alpha F(x, y_1) + \beta F(x, y_2), \text{ where the second equality follows from (11.4); and the last equality follows from Propositions 7.23, 3.66, and 3.67.}
\]

Furthermore, \( |F(x, y_1)| = \lim_{i \to \infty} |f_{y_1, i}(x)| \leq \lim_{i \to \infty} \|y_1, i\| \hat{f}(x) = \|y_1\| \hat{f}(x), \text{ where the inequality follows from (11.5); and the last equality follows from Propositions 3.66 and 7.21.} \)

Hence,

\[
\forall x \in X \setminus E_0, F(x, \cdot) : \mathcal{Y}^* \to \mathbb{K} \text{ is a bounded linear functional on } \mathcal{Y}^*.
\]

Since \( \hat{Y} \) is reflexive, then \( \exists f(x) \in \mathcal{Y}^* \text{ such that } F(x, y) = \langle y, f(x) \rangle, \forall y \in \mathcal{Y}^* \) and \( \|f(x)\| \leq \hat{f}(x) \).

Therefore, we may define \( f : X \to \mathcal{Y} \) by assigning

\[
f(x) = \hat{y}, \forall x \in E_0.
\]

\[\forall y \in \mathcal{Y}^*, \text{ define } F_y : X \to \mathbb{K} \text{ by } F_y(x) = \langle y, f(x) \rangle, \forall x \in X. \text{ Fix any } y \in \hat{D}. \text{ Then, } \forall x \in X \setminus E_0, F_{\hat{y}}(x) = \langle \hat{y}, f(x) \rangle = F(x, y) = f_y(x) \text{ and } \forall x \in E_0, F_{\hat{y}}(x) = \langle \hat{y}, f(x) \rangle = 0. \text{ By Proposition 11.41, } F_y \text{ is } \mathcal{B}\text{-measurable. By the arbitrariness of } \hat{y} \text{ and Proposition 11.170, } f \text{ is } \mathcal{B}\text{-measurable.} \]
Claim 11.171.1 \( \forall y_0 \in \mathcal{Y}^\ast, f_{y_0}(x) = \langle \langle y_0, f(x) \rangle \rangle \) a.e. \( x \in \mathcal{X} \).

Proof of claim: Let \( \hat{D} := D \cup \{ y_0 \} \) and \( \hat{D} := \{ \sum_{n=1}^\infty \alpha_i y_i \in \mathcal{Y}^\ast \mid n \in \mathbb{Z}_+, \alpha_1, \ldots, \alpha_n \in \mathbb{K}_Q, y_1, \ldots, y_n \in \hat{D} \} \). Then, \( \hat{D} \subseteq \mathcal{Y}^\ast \) is also a countable dense set and \( \hat{D} \subseteq \hat{D} \). Define

\[
E_1 := \{ x \in \mathcal{X} \mid \exists y_1, y_2 \in \hat{D}, \exists \alpha, \beta \in \mathbb{K}_Q \ni f_{\alpha y_1 + \beta y_2} \neq \alpha f_{y_1}(x) + \beta f_{y_2}(x) \text{ or } |f_{y_1}(x)| > \|y_1\| \|f(x)\| \}
\]

Then, \( E_1 \in \mathcal{B}, \mu(E_1) = 0 \) and \( E_0 \subseteq E_1 \).

\( \forall x \in \mathcal{X} \setminus E_1, \forall y_1, y_2 \in \hat{D}, \forall \alpha, \beta \in \mathbb{K}_Q \), we have \( f_{\alpha y_1 + \beta y_2} = \alpha f_{y_1} + \beta f_{y_2} \) and \( |f_{y_1}(x)| \leq \|y_1\| \|f(x)\| \). 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 limit \( \hat{F}(x, y_0) \in \mathbb{K} \). 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 \( (\hat{y}_i)_{i=1}^\infty \subseteq \hat{D} \) converging to \( y_0 \) and \( (y_0, y_0, \ldots) \subseteq \hat{D} \) leads to \( f_{\hat{y}_0}(x) = \hat{F}(x, y_0) = F(x, y_0) \). Therefore, \( f_{\hat{y}_0}(x) = F(x, y_0) = \langle \langle y_0, f(x) \rangle \rangle = F_{y_0}(x) \), \( \forall x \in \mathcal{X} \setminus E_1 \). Note that \( f_{y_0} \in \mathcal{B} \) measurable and, by Propositions 7.2 and 11.38, \( F_{y_0} \) is \( \mathcal{B} \)-measurable. Then, \( E_1 \supseteq \{ x \in \mathcal{X} \mid f_{y_0}(x) \neq F_{y_0}(x) \} = \{ x \in \mathcal{X} \mid |f_{y_0}(x) - f_{\hat{y}_0}(x)| > 0 \} \in \mathcal{B} \). 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. \( \forall E \in \text{dom}(\nu), \mathcal{P} \circ \nu(E) = \int_E f \, d\mu < \infty \). This implies that \( \int_E \mathcal{P} \circ f \, d\mu < \infty \). By Proposition 11.92, we have \( \int_E f \, d\mu \in \mathcal{Y} \). \( \forall y \in \mathcal{Y} \), \( \mathcal{P} \circ \nu_\mathcal{Y}(E) \leq \langle \langle y, \nu \rangle \rangle \), \( \mathcal{P} \circ \nu_{\mathcal{Y}}(E) \leq (\|y\| \|\mathcal{P} \circ \nu\|(E) < \infty \), \( E \in \text{dom}(\nu_{\mathcal{Y}}), \|y\| \|\mathcal{P} \circ \nu\|(E) = \int_E f \, d\mu = \int_E f \, d\nu = \int_E \langle \langle y, f(x) \rangle \rangle \, d\mu(x) = \langle \langle y, \int_E f \, d\mu \rangle \rangle \), where the last three equalities follow from Proposition 11.92. Hence, by Proposition 7.85, \( \nu(E) = \int_E f \, d\mu \in \mathcal{Y} \), \( \forall E \in \text{dom}(\nu) \). 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 \( \hat{\mathcal{X}} := (\mathcal{X}, \mathcal{B}, \mathcal{P} \circ \mu) \), which is a \( \sigma \)-finite measure space. By the second special case, \( \exists f_1 : X \to \mathcal{Y} \) such that \( \frac{d\nu}{d\mu_{\mathcal{P} \circ \mu}} = f_1 \) a.e. in \( \hat{\mathcal{X}} \). 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 \( \hat{\mathcal{X}} \). By Definition 11.166, \( \frac{d\nu}{d\mu_{\mathcal{P} \circ \mu}} = \mathcal{P} \circ f_2 \) a.e. in \( \hat{\mathcal{X}} \). 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 \( \hat{\mathcal{X}} \) and \( \exists f_3 : X \to \mathcal{K} \) such that \( f_3 f_2 = 1 \) a.e. in \( \hat{\mathcal{X}} \). By Proposition 11.168, \( \frac{d\nu}{d\mu} = f_3 \) a.e. in \( \hat{\mathcal{X}} \) and \( \frac{d\nu}{d\mu} = \left( \frac{d\nu}{d\mu_{\mathcal{P} \circ \mu}} \right) \left( \frac{d\nu}{d\mu} \right) = f_1 f_3 =: f : X \to \mathcal{Y} \) a.e. in \( \mathcal{X} \).

This completes the proof of the theorem. \( \square \)
Theorem 11.172 (Lebesgue Decomposition) Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( Y \) be 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 \( A := \{ x \in X \mid f(x) > 0 \} \subseteq \mathcal{B} \) and \( B := \{ x \in X \mid f(x) = 0 \} \subseteq \mathcal{B} \). 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 \( \mathbb{R} \) by: \( \nu_0(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( E \cap B \subseteq \mathcal{B} \setminus \text{dom}(\nu) \); and \( \nu_0(E) = \nu(E \cap B) \in \mathbb{R} \), \( \forall E \in \mathcal{B} \) with \( E \cap B \subseteq \text{dom}(\nu) \). Define \( \nu_1 \) to be a function from \( \mathcal{B} \) to \( \mathbb{R} \) by: \( \nu_1(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( E \setminus A \subseteq \mathcal{B} \setminus \text{dom}(\nu) \); and \( \nu_1(E) = \nu(E \setminus A) \in \mathbb{R} \), \( \forall E \in \mathcal{B} \) with \( E \setminus A \subseteq \text{dom}(\nu) \). By Proposition 11.157, \( \nu_0 \) and \( \nu_1 \) are \( \sigma \)-finite \( Y \)-valued measures on \((X, \mathcal{B})\) such that \( \nu = \nu_0 + \nu_1 \), \( \mathcal{P} \circ \nu = \mathcal{P} \circ \nu_0 + \mathcal{P} \circ \nu_1 \), \( \mathcal{P} \circ \nu_0(E) = \mathcal{P} \circ \nu(E \cap B) \), and \( \mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \nu(E \setminus A) \), \( \forall E \in \mathcal{B} \). This implies that \( \mathcal{P} \circ \nu_0(A) = \mathcal{P} \circ \nu(A \cap B) = 0 \). 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 \( \mathcal{E} := (E, \mathcal{B}_E, \lambda_E) \) 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 \( \mathcal{E} \). Then, \( \lambda(E \cap A) = 0 \) and \( \mathcal{P} \circ \nu_0(E) = \mathcal{P} \circ \nu(E \cap A) = 0 \), 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, \( \exists \tilde{A}, \tilde{B} \subseteq \mathcal{B} \) with \( \tilde{A} = X \setminus \tilde{B} \) such that \( \mathcal{P} \circ \tilde{\nu}_0(\tilde{A}) = \mu(\tilde{B}) = 0 \). \( \forall E \in \mathcal{B} \) with \( \lambda(E) < \infty \), we have \( \mu(E) < \infty \), \( \mathcal{P} \circ \nu(E) = \mathcal{P} \circ \nu_0(E) + \mathcal{P} \circ \nu_1(E) = \mathcal{P} \circ \tilde{\nu}_0(E) + \mathcal{P} \circ \tilde{\nu}_1(E) < \infty \). Then, \( \tilde{\nu}_0(E) = \tilde{\nu}_0(E \cap A) + \tilde{\nu}_0(E \cap B) = \tilde{\nu}_0(E \cap \tilde{B}) = \tilde{\nu}_0(E \cap \tilde{B}) + \tilde{\nu}_1(E \cap \tilde{B}) = \nu(E \cap \tilde{B}) = \nu_0(E \cap \tilde{B}) + \nu_1(E \cap \tilde{B}) + \nu_0(E \cap A \cap B) + \nu_1(E \cap A \cap B) = \nu_0(E \cap \tilde{B}) + \nu_0(E \cap A \cap B) = \nu_0(E \cap \tilde{B}) + \nu_0(E \cap \tilde{A}) = \nu_0(E) \in \mathbb{R} \), where 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 \( \mathcal{P} \circ \tilde{\nu}_0(\tilde{A}) = 0 \); \( \mu(B) = 0 \) and \( \mathcal{P} \circ \tilde{\nu}_1 \ll \mu \); and the sixth equality
follows from the fact that $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \nu_1 \ll \mu$. We also have $\hat{\nu}_1(E) = \nu_1(E \cap \hat{A}) + \nu_1(E \cap \hat{B}) = \nu_1(E \cap \hat{A}) + \nu_1(E \cap \hat{B}) = \nu_1(E \cap \hat{A}) = \nu_0(E \cap \hat{A} \cap B) + \nu_1(E \cap \hat{A}) + \nu_1(E \cap \hat{B}) = \nu_1(E \cap \hat{A}) = \nu_1(E \cap \hat{B}) = \nu_1(E) \in \mathcal{Y}$ where the second equality follows from the fact that $\mu(\hat{B}) = 0$ and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$; the third equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$; the eighth equality follows from the fact that $\mathcal{P} \circ \hat{\nu}_0(\hat{A}) = 0$, $\mu(\hat{B}) = 0$, and $\mathcal{P} \circ \hat{\nu}_1 \ll \mu$. 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 $\hat{\nu}$ on $(X, \mathcal{B})$ such that $\hat{\nu}(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\hat{\nu})$. By the general case, there exists a pair of $\sigma$-finite $\mathbb{R}$-valued measures $\hat{\nu}_0$ and $\hat{\nu}_1$ on $(X, \mathcal{B})$ that satisfies the desired properties. By the proof of the general case, $\hat{\nu}_0(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\hat{\nu}_0)$, and $\hat{\nu}_1(E) \in [0, \infty) \subset \mathbb{R}$, $\forall E \in \text{dom}(\hat{\nu}_1)$. By Proposition 11.135, $\hat{\nu}_0$ and $\hat{\nu}_1$ are identified with $\sigma$-finite measures $\nu_0$ and $\nu_1$, respectively. Then, $\nu_0 = \mathcal{P} \circ \hat{\nu}_0 \ll \mu$, $\nu_1 = \mathcal{P} \circ \hat{\nu}_1 \ll \mu$, and $\nu = \mathcal{P} \circ \hat{\nu} = \mathcal{P} \circ \hat{\nu}_0 + \mathcal{P} \circ \hat{\nu}_1 = \nu_0 + \nu_1$. 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 $\vartheta$. 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 $Z_p := \{ f \in \mathcal{M}(\mathcal{X}, \mathcal{Y}) \mid f$ is $\mathcal{B}$-measurable and $\mathcal{P}_p \circ f$ is integrable over $\mathcal{X} \}$. Define $\| \cdot \|_p : Z_p \to [0, \infty) \subset \mathbb{R}$ by $\| f \|_p := \left( \int_{\mathcal{X}} (\mathcal{P}_p \circ f) \, d\mu \right)^{1/p}$, $\forall f \in Z_p$. 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, $\vartheta \in Z_p \neq \emptyset$. $\forall f, g \in Z_p$, $\forall \alpha \in \mathbb{K}$, 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 $\| f + g \|_p \leq \| f \|_p + \| g \|_p$. $\| \alpha f \|_p = \left( \int_{\mathcal{X}} (\mathcal{P}_p \circ (\alpha f)) \, d\mu \right)^{1/p} = | \alpha | \left( \int_{\mathcal{X}} (\mathcal{P}_p \circ f) \, d\mu \right)^{1/p} = | \alpha \| f \|_p \in \mathbb{R}$, where the second equality follows from Proposition 11.92. Hence, $\alpha f \in Z_p$. 

The above shows that $Z_p$ is a subspace of $(M(X, \mu), \mathbb{K})$. Hence, $(Z_p, \mathbb{K})$ is a vector space. Clearly, $\|\vartheta\|_p = 0$. 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 $L_p(X, Y)$. We will denote the vector space $(Z_p, \mathbb{K})$ with the pseudo-norm $\|\cdot\|_p$ by $\tilde{L}_p(X, Y)$.

$\forall f \in Z_p$ with $\|f\|_p = 0$, we have $\int_X P \circ f \ d\mu = 0$. By Proposition 11.96, we have $P \circ f = 0$ a.e. in $X$ and $f = \vartheta_Y$ a.e. in $X$. On the other hand, $\forall f \in M(X, Y)$ with $f = \vartheta_Y$ a.e. in $X$ and $f$ being $\mathcal{B}$-measurable, then, by Proposition 11.83, $\|f\|_p = 0$ and $f \in Z_p$. Hence, $\forall f \in M(X, Y)$, $f \in Z_p$ and $\|f\|_p = 0$ if, and only if, $f = \vartheta_Y$ a.e. in $X$ and $f$ is $\mathcal{B}$-measurable. We will denote the norm in $L_p(X, Y)$ by $\|\cdot\|_p$ and elements in $L_p(X, Y)$ by $[f]$, where $f \in \tilde{L}_p(X, Y)$.

**Theorem 11.174 (Minkowski’s Inequality)** 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 $L_p(X, \mathcal{Y})$ be the vector space over $\mathbb{K}$ with the pseudo-norm $\|\cdot\|_p$ as defined in Example 11.173. $\forall f, g \in L_p(X, \mathcal{Y})$, then, $f + g \in L_p(X, \mathcal{Y})$ and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

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 $\alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $X$ and $\mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $X$.

**Proof** Clearly, $f$ and $g$ are $\mathcal{B}$-measurable. By Propositions 7.23, 11.88, 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: $\|f\|_p, \|g\|_p > 0$. Without loss of generality, assume that $\|g\|_p = 0$. Then, $\int_X (P \circ f) \ d\mu = 0$. By Proposition 11.96, we have $P \circ g = 0$ a.e. in $X$. Thus, $g = \vartheta_Y$ a.e. in $X$ and $f + g$ a.e. in $X$. By Proposition 11.83, the integral $\int_X (P \circ (f + g)) \ d\mu = \int_X (P \circ f) \ d\mu \in \mathbb{R}$. Hence, $\|f + g\|_p = \|f\|_p + \|g\|_p \in \mathbb{R}$ and $f + g \in L_p(X, \mathcal{Y})$. Clearly, equality holds implies $\alpha = 0, \beta = 1, \alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g$ a.e. in $X$ and $\mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g$ a.e. in $X$.

Case 2: $\|f\|_p, \|g\|_p > 0$. Let $\lambda := \frac{\|f\|_p}{\|f\|_p + \|g\|_p} \in (0, 1) \subset \mathbb{R}$. $\forall x \in X$, we have

$$\frac{\|f(x) + g(x)\|^p}{\|f\|_p + \|g\|_p} \leq \left( \frac{\|f(x)\| + \|g(x)\|}{\|f\|_p + \|g\|_p} \right)^p$$

$$= \left( \lambda \frac{\|f(x)\|}{\|f\|_p} + (1 - \lambda) \frac{\|g(x)\|}{\|g\|_p} \right)^p \leq \lambda \frac{\|f(x)\|^p}{\|f\|_p} + (1 - \lambda) \frac{\|g(x)\|^p}{\|g\|_p}$$

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 $\frac{\|f(x)\|}{\|f\|_p} = \frac{\|g(x)\|}{\|g\|_p}$. 


By Proposition 11.83, we have \( \int_X \left( \frac{(P_\beta \circ f + g)}{\|f\|_p + \|g\|_p} \right) \, d\mu \leq \int_X \left( \frac{(P_\alpha \circ f)}{\|f\|_p} \right) \, d\mu = \int_X \left( \frac{(P_\alpha \circ f)}{\|f\|_p} \right) \, d\mu + \frac{\lambda}{\|g\|_p} \int_X (P_\beta \circ f) \, d\mu = 1 \). This implies that \( \int_X (P_\beta \circ (f + g)) \, d\mu \leq (\|f\|_p + \|g\|_p) \lambda < +\infty \) and \( \|f + g\|_p \leq \|f\|_p + \|g\|_p \). Therefore, \( f + g \in L_p(\mathcal{X}, \mathcal{Y}) \).

Fix any \( p \in (1, +\infty) \subset \mathbb{R} \). By Proposition 11.97, equality holds if, and only if, \( \mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \) and \( \frac{\mathcal{P}_\beta}{\|\mathcal{P}\|_p} = \frac{\mathcal{P}_\alpha}{\|\mathcal{P}\|_p} \) a.e. in \( \mathcal{X} \). Equality implies that \( \alpha = 1/\|f\|_p \) and \( \beta = 1/\|g\|_p \) such that \( \alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \) and \( \mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g \) 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 \( \alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \) and \( \mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \). Without loss of generality, assume that \( \beta \neq 0 \). Then, \( \mathcal{P} \circ g = \alpha_1 \mathcal{P} \circ f \) 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 \( \alpha \mathcal{P} \circ f = \beta \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \) and \( \mathcal{P} \circ (f + g) = \mathcal{P} \circ f + \mathcal{P} \circ g \) a.e. in \( \mathcal{X} \).

Hence, the result holds in both cases. This completes the proof of the theorem. \( \square \)

**Definition 11.175** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space and \( f : X \to \mathbb{R} \) be \( \mathcal{B} \)-measurable. The essential supremum of \( f \) is

\[
es\sup_{x \in \mathcal{X}} f(x) := \inf \{ M \in \mathbb{R} \mid \mu(\{ x \in X \mid f(x) > M \}) = 0 \} \in \mathbb{R}
\]

**Proposition 11.176** Let \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a measure space and \( f : X \to \mathbb{R} \) and \( g : X \to \mathbb{R} \) be \( \mathcal{B} \)-measurable. Then,

(i) \( \es\sup_{x \in \mathcal{X}} (f(x) + g(x)) \leq \es\sup_{x \in \mathcal{X}} f(x) + \es\sup_{x \in \mathcal{X}} g(x) \);

(ii) if \( f \leq g \) a.e. in \( \mathcal{X} \), then \( \es\sup_{x \in \mathcal{X}} f(x) \leq \es\sup_{x \in \mathcal{X}} g(x) \);

(iii) \( \forall \alpha \in (0, \infty) \subset \mathbb{R} \), \( \es\sup_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \es\sup_{x \in \mathcal{X}} f(x) \); and \( \forall \alpha \in [0, \infty) \subset \mathbb{R} \), \( \es\sup_{x \in \mathcal{X}} (\alpha f(x)) = \alpha \es\sup_{x \in \mathcal{X}} f(x) \) if \( \es\sup_{x \in \mathcal{X}} f(x) \in \mathbb{R} \).

(iv) \( \lambda := \es\sup_{x \in \mathcal{X}} f(x) \in \mathbb{R} \). 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, \( \es\sup_{x \in \mathcal{X}} f(x) = -\infty \). Similarly, \( \es\sup_{x \in \mathcal{X}} g(x) = \es\sup_{x \in \mathcal{X}} (f(x) + g(x)) = -\infty \). 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 11.7, \( \exists n \in \mathbb{N} \) such that \( \mu(A_n) > 0 \). Then, \( \es\sup_{x \in \mathcal{X}} f(x) \geq -n > -\infty \). Similarly, \( \es\sup_{x \in \mathcal{X}} g(x) > -\infty \) and \( \es\sup_{x \in \mathcal{X}} (f(x) + g(x)) > -\infty \). Then, \( \lambda := \es\sup_{x \in \mathcal{X}} f(x) + \)

\[ \text{ess sup}_{x \in X} g(x) \in (-\infty, +\infty) \subset \mathbb{R}. \]

If \( \lambda = +\infty \), then (i) holds. On the other hand, if \( \lambda \in \mathbb{R}, \forall M \in \mathbb{R} \) with \( \lambda < M \), then \( \exists M_1, M_2 \in \mathbb{R} \) with \( \text{ess sup}_{x \in X} f(x) < M_1 \) and \( \text{ess sup}_{x \in X} g(x) < M_2 \) and \( M_1 + M_2 = M \).

Then, \( \mu(\{ x \in X \mid f(x) > M \}) = 0 = \mu(\{ x \in X \mid g(x) > M \}). \)

By Propositions 7.23, 11.38, and 11.39, \( f + g \) is \( \mathcal{B} \)-measurable. Hence, \( \mu(\{ x \in X \mid f(x) + g(x) > M \}) = 0 \) and \( \text{ess sup}_{x \in X} (f(x) + g(x)) \leq M \). 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 11.7, \( \mu(E) := \mu(\{ x \in X \mid f(x) > \lambda \}) = \mu(\bigcup_{n=1}^{\infty} E_n) = \lim_{n \to \infty} \mu(E_n) = 0 \). 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.

**Example 11.177** Let \( X := (X, \mathcal{B}, \mu) \) be a measure space, \( Y \) be a separable normed linear space over \( \mathbb{K} \), 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, \( \vartheta \in Z_{\infty} \neq \emptyset \). \( \forall f, g \in Z_{\infty}, \forall a \in \mathbb{K}, \) by Propositions 7.23, 11.38, and 11.39, \( f + g \) and \( \alpha f \) are \( \mathcal{B} \) measurable. By Proposition 11.176,

\[
\| f + g \|_{\infty} = \max \{ \text{ess sup}_{x \in X} \| f(x) + g(x) \|, 0 \} \\
\leq \max \{ \text{ess sup}_{x \in X} (\| f(x) \| + \| g(x) \|), 0 \}
\]
11.9. \( L_p \) SPACES

\[
\leq \max\{\text{ess sup}_{x \in X} \| f(x) \| + \text{ess sup}_{x \in X} \| g(x) \|, 0\}
= \begin{cases} 
\text{ess sup}_{x \in X} \| f(x) \| + \text{ess sup}_{x \in X} \| g(x) \| & \text{if } \mu(X) > 0 \\
0 & \text{if } \mu(X) = 0
\end{cases}
\leq \| f \|_{\infty} + \| g \|_{\infty} < +\infty
\]
and
\[
\| \alpha f \|_{\infty} = \max\{\text{ess sup}_{x \in X} \| \alpha f(x) \|, 0\} = \max\{\text{ess sup}_{x \in X} \| \alpha \| \| f(x) \|, 0\}
= \begin{cases} 
|\alpha| \text{ess sup}_{x \in X} \| f(x) \| & \text{if } \mu(X) > 0 \\
0 & \text{if } \mu(X) = 0
\end{cases} = |\alpha| \| f \|_{\infty} < +\infty
\]

Then, \( f + g, \alpha f \in Z_{\infty} \). Hence, \( Z_{\infty} \) is a subspace of \( (\mathcal{M}(X, Y), \mathbb{K}) \). This implies that \( (Z_{\infty}, \mathbb{K}) \) is a vector space. Clearly, \( \| \theta \|_{\infty} = 0 \). Therefore, \( \| \cdot \|_{\infty} \)
defines a pseudo-norm on \( (Z_{\infty}, \mathbb{K}) \). By Proposition 7.47, the quotient space of \( (Z_{\infty}, \mathbb{K}) \) modulo \( \| \cdot \|_{\infty} \) is a normed linear space, to be denoted \( L_{\infty}(X, Y) \).

We will denote the vector space \( (Z_{\infty}, \mathbb{K}) \) with the pseudo-norm \( \| \cdot \|_{\infty} \) by \( L_{\infty}(X, Y) \).

\( \forall f \in Z_{\infty} \) with \( \| f \|_{\infty} = 0 \), we have \( \text{ess sup}_{x \in X} \| f(x) \| \leq 0 \). Then, by Proposition 11.176, \( \mathcal{P} \circ f = 0 \) a.e. in \( X \). Hence, \( f = \vartheta_Y \) a.e. in \( X \). On the other hand, \( \forall f \in \mathcal{M}(X, Y) \) with \( f = \vartheta_Y \) a.e. in \( X \) and \( f \) being \( \mathcal{B} \)-measurable, we have \( \| f \|_{\infty} = 0 \) and \( f \in Z_{\infty} \). Hence, \( \forall f \in \mathcal{M}(X, Y) \), \( f \in Z_{\infty} \) and \( \| f \|_{\infty} = 0 \) if, and only if, \( f = \vartheta_Y \) a.e. in \( X \) and \( f \) is \( \mathcal{B} \)-measurable. We will denote the norm in \( L_{\infty}(X, Y) \) by \( \| \cdot \|_{\infty} \) and elements in \( L_{\infty}(X, Y) \) by \( [ f ] \), where \( f \in L_{\infty}(X, Y) \).

In the following, we will write \( \lim_{n \in \mathbb{N}} z_n \overset{\text{\( \circ \)}}{=} z \) in \( L_p(X, Y) \), when the sequence \( (z_n)_{n=1}^{\infty} \subseteq L_p(X, Y) \) converges to \( z \in L_p(X, Y) \) in \( L_p(X, Y) \) pseudo-norm. We will simply write \( \lim_{n \in \mathbb{N}} z_n \overset{\text{\( \circ \)}}{=} z \) if there is no confusion in which pseudo-norm convergence occurs. For \( z \in L_p(X, Y) \), we will denote the corresponding equivalence class in \( L_p(X, Y) \) by \( [ z ] \). Then, for \( (z_n)_{n=1}^{\infty} \subseteq L_p(X, Y) \), \( \lim_{n \in \mathbb{N}} z_n \overset{\text{\( \circ \)}}{=} z \) in \( L_p(X, Y) \) if, and only if, \( \lim_{n \in \mathbb{N}} [ z_n ] = [ z ] \) in \( L_p(X, Y) \) (or simply \( \lim_{n \in \mathbb{N}} [ z_n ] = [ z ] \) 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 \( q \in (1, +\infty] \subseteq \mathbb{R}_e \) with \( 1/p + 1/q = 1 \), \( X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, and \( Y \) be a separable normed linear space over \( \mathbb{K} \) with \( Y^* \) being separable. Then, \( \forall f \in \mathcal{L}_p(X, Y) \), \( \forall g = L_q(X, Y^*) \), the function \( r : X \to \mathbb{K} \), defined by \( r(x) = \langle \langle g(x), f(x) \rangle \rangle, \forall x \in X \), is absolutely integrable over \( X \) and

\[
\int_X \| \langle \langle g(x), f(x) \rangle \rangle \| \, d\mu(x) \leq \| f \|_p \| g \|_q
\]

When \( q < \infty \), equality holds if, and only if, \( \| \langle \langle g(x), f(x) \rangle \rangle \| = \| f(x) \| \cdot \| g(x) \| \) a.e. \( x \in X \) and \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \) a.e. in \( X \).
Lemma 7.7 with equality holding if, and only if, \( |\langle g(x), f(x) \rangle| = |g(x)\| |f(x)\| \) a.e. \( x \in X \). Then, by Propositions 11.83 and 11.92,

\[
\int_X |\langle g(x), f(x) \rangle| \, d\mu(x) \leq \|g\|_\infty \int_X \mathcal{P} \circ f \, d\mu = \|f\|_p \|g\|_q
\]

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: \( \|f\|_p \|g\|_q = 0 \). Without loss of generality, assume \( \|g\|_q = 0 \). Then, \( g = \varnothing \) a.e. in \( X \). This implies that \( |\langle g(x), f(x) \rangle| = 0 \) a.e. \( x \in X \) and, by Propositions 11.83 and 11.75,

\[
\int_X |\langle g(x), f(x) \rangle| \, d\mu(x) = 0 = \|f\|_p \|g\|_q
\]

Equality holds \( \Rightarrow \alpha = 0 \), \( \beta = 1 \), \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \) a.e. in \( X \), and \( |\langle g(x), f(x) \rangle| = 0 = \|f(x)\| \|g(x)\| \) a.e. \( x \in X \). This subcase is proved.

Case 2b: \( \|f\|_p \|g\|_q > 0 \). Then, \( \|f\|_p > 0 \) and \( \|g\|_q > 0 \). \( \forall x \in X \), 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

\[
\frac{|\langle g(x), f(x) \rangle|}{\|f\|_p \|g\|_q} \leq \frac{|g(x)\| \|f(x)\|}{\|g\|_q \|f\|_p} \leq \frac{1}{p} \|f(x)\|^p + \frac{1}{q} \|g(x)\|^q
\]

with equality holding if, and only if, \( |\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\| \) and \( \|f(x)\|^p = \frac{\|g(x)\|^q}{\|g\|_q^q} \). Integrating the above inequality over \( X \), we have, by Propositions 11.83 and 11.97,

\[
\int_X \frac{|\langle g(x), f(x) \rangle| \, d\mu(x)}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1
\]

with equality holding if, and only if, \( |\langle g(x), f(x) \rangle| = \|g(x)\| \cdot \|f(x)\| \) a.e. \( x \in X \) and \( \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \) a.e. in \( X \). Equality \( \Rightarrow \alpha = 1/\|f\|_p^p \), \( \beta = 1/\|g\|_q^q \), \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \) a.e. in \( X \) and \( |\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\| \) a.e. \( x \in X \). On the other hand, if \( |\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\| \) a.e. \( x \in X \) and \( \exists \alpha, \beta \in \mathbb{R} \), which are not both zeros, such that \( \alpha \mathcal{P}_p \circ f = \beta \mathcal{P}_q \circ g \) a.e. in \( X \), then, without loss of generality, assume \( \beta \neq 0 \). Let \( \alpha_1 = \alpha/\beta \). Then, \( \alpha_1 \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \) a.e. in \( X \). Hence, \( \alpha_1 \|f\|_p^p = \|g\|_q^q \), which further implies that \( \alpha_1 = \|g\|_q^q / \|f\|_p^p \). Hence, \( \mathcal{P}_p \circ f = \mathcal{P}_q \circ g \) a.e. in \( X \). This implies equality. Therefore, equality holds.
if, and only if, $|\langle g(x), f(x) \rangle| = \|g(x)\| \|f(x)\|$ a.e. $x \in X$ and $\exists \alpha, \beta \in \mathbb{R}$, which are not both zeros, such that $\alpha P_\beta \circ f = \beta P_\alpha \circ g$ a.e. in $X$. This subcase is proved.

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

When $p = 2 = q$, the Hölder’s inequality becomes the well-known Cauchy-Schwarz Inequality:

$$\int_X |\langle g(x), f(x) \rangle| \, \mu(x) \leq \left( \int_X P_2 \circ f \, d\mu \right)^{1/2} \left( \int_X P_2 \circ g \, d\mu \right)^{1/2}$$

**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 $L_p(X, Y)$ be 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 $(\{f_n\})_{n=1}^\infty \subset L_p(X, Y)$ with $f_n \in L_p(X, Y)$, $\forall n \in \mathbb{N}$, and $\sum_{n=1}^\infty \|f_n\|_p =: M < \infty$. Set $g_n : X \to [0, \infty) \subset \mathbb{R}$ 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 $P \circ f_n \in L_p(X, \mathbb{R})$. Then, $g_n \in L_p(X, \mathbb{R})$ and $\|g_n\|_p = \sum_{i=1}^n \|P \circ f_i\|_p = \sum_{i=1}^n \|f_i\|_p \leq M$. This implies that $\int_X (l \circ g_n) \, d\mu \leq M^p$. 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, $\exists \alpha : X \to [0, \infty) \subset \mathbb{R}$, which is $\mathcal{B}$-measurable, such that $\lim_{n \to \infty} \alpha g_n = \alpha g$ a.e. in $X$ and $\int_X (\log g_n) \, d\mu = \lim_{n \to \infty} \int_X (\log g_n) \, d\mu \leq M^p$.

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 $\mu(E) = 0$. $\forall x \in X \setminus E$, $\lim_{n \to \infty} g_n(x) = g(x) \in \mathbb{R}$ and $\lim_{n \to \infty} \sum_{i=1}^n \|f_i(x)\| = g(x) \in \mathbb{R}$. By Proposition 7.27 and the completeness of $Y$, we have $\lim_{n \to \infty} \sum_{i=1}^n f_i(x) =: \lim_{n \to \infty} s_n(x) =: s(x) \in Y$. Define $s(x) = s_n(x) = \emptyset_Y$, $\forall x \in E$ and $\forall n \in \mathbb{N}$. Then, by Proposition 11.41, $s_n$ is $\mathcal{B}$-measurable, $\forall n \in \mathbb{N}$. Clearly, $\lim_{n \to \infty} s_n(x) = s(x)$, $\forall x \in X$. By Proposition 11.48, $s$ is $\mathcal{B}$-measurable. By Lemma 11.43, $s_n = \sum_{i=1}^n f_i \text{ a.e. in } X$. 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, $\lim_{n \to \infty} \sum_{i=1}^n f_i = s \text{ a.e. in } X$. Note that $\lim_{n \to \infty} P_\beta \circ (s_n(x) - s(x)) = 0$, $\forall x \in X$. Note also that $\|s_n(x) - s(x)\|^p \leq \|s_n(x)\|^p + \|s(x)\|^p \leq (g_n(x) + g(x))^p \leq 2^p(g(x))^p$, $\forall x \in X \setminus E$, $\forall n \in \mathbb{N}$, 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, $\lim_{n \to \infty} \int_X (P_\beta \circ (s_n - s)) \, d\mu = 0$. This implies that $\lim_{n \to \infty} \|s_n - s\|_p = 0$. Hence, $s \in L_p(X, \mathbb{R})$ and $\lim_{n \to \infty} \sum_{i=1}^n f_i = \lim_{n \to \infty} s_n \equiv s$. 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. \hfill \diamond

**Example 11.180** Let $X := (X, \mathcal{B}, \mu)$ be a measurable space, $Y$ be a separable Banach space over $\mathbb{K}$, and $L_{\infty}(X, Y)$ be the normed linear space
over \( K \) 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 \( f_n \in L_\infty(\mathcal{X}, \mathcal{Y}), \forall n \in \mathbb{N}. \) ∀\( k \in \mathbb{N}, \exists N_k \in \mathbb{N}, \forall n, m \in \mathbb{N} \) with \( n \geq N_k \) and \( m \geq N_k, \) we have \( \|f_n - f_m\|_\infty < 1/k. \) Then, by Proposition 11.176, \( A_{n,m,k} := \{ x \in X \mid \|f_n(x) - f_m(x)\| \geq 1/k \} \in \mathcal{B} \) and \( \mu(A_{n,m,k}) = 0. \) Let \( A := \bigcup_{k=1}^\infty \bigcup_{n=N_k}^\infty \bigcup_{m=N_k}^\infty A_{n,m,k} \in \mathcal{B}. \) Clearly, \( \mu(A) = 0. \) ∀\( x \in X \setminus A, (f_n(x))_{n=1}^\infty \subseteq \mathcal{Y} \) is a Cauchy sequence, which converges to \( f(x) \in \mathcal{Y} \) by the completeness of \( \mathcal{Y}. \) Define \( f(x) = \psi_y, \forall x \in A. \) Then, \( f : X \to \mathcal{Y} \) is well defined and \( \lim_{n \to \infty} f_n(x) = f(x), \forall x \in X \setminus A. \) By Propositions 11.48 and 11.41, \( f \) is \( \mathcal{B} \)-measurable. ∀\( k \in \mathbb{N}, \forall n \in \mathbb{N} \) with \( n \geq N_k, \forall x \in X \setminus A, \) by Propositions 3.66, 3.67, 7.21, and 7.23, we have \( \|f_n(x) - f(x)\| = \lim_{m \to \infty} \|f_n(x) - f_m(x)\| \leq 1/k. \) Then, \( 0 \leq \mu(\{ x \in X \mid \|f(x) - f_n(x)\| > 1/k \}) \leq \mu(A) = 0. \) This shows that \( \|f_n - f\|_\infty \leq 1/k. \) Then, \( \lim_{n \to \infty} \|f_n - f\|_\infty = 0, \lim_{n \to \infty} f_n = f, \) and \( f \in L_\infty(\mathcal{X}, \mathcal{Y}). \) Hence, \( \lim_{n \to \infty} f_n = [f] \) in \( L_\infty(\mathcal{X}, \mathcal{Y}). \) Therefore, \( L_\infty(\mathcal{X}, \mathcal{Y}) \) is a Banach space when \( \mathcal{X} \) is a measure space and \( \mathcal{Y} \) is a separable Banach space.

**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 normal linear space over \( \mathbb{K}, \mathcal{L}_p(\mathcal{X}, \mathcal{Y}) \) be the normed linear space over \( \mathbb{K} \) as defined in Example 11.173, and \( f \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y}). \) Then, there exists a sequence of simple functions \((\phi_i)_{i=1}^\infty, \phi_i : \mathcal{X} \to \mathcal{Y}, \forall i \in \mathbb{N}, such that \( \lim_{i \to \infty} \phi_i = f \) a.e. in \( \mathcal{X}, \) \( \|\phi_i(x)\| \leq \|f(x)\|, \forall x \in X, \forall i \in \mathbb{N}, \lim_{i \to \infty} \int_X \mathcal{P}_p \circ (\phi_i - f) d\mu = 0, \) and \( \lim_{i \to \infty} \phi_i = f \) in \( \mathcal{L}_p(\mathcal{X}, \mathcal{Y}). \)

**Proof** Since \( f \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y}), \) 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} \to \mathcal{Y}, \forall i \in \mathbb{N}, such that \( \|\phi_i(x)\| \leq \|f(x)\|, \forall x \in X, \forall i \in \mathbb{N}, \) and \( \lim_{i \to \infty} \phi_i = f \) 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, \( \lim_{i \to \infty} \mathcal{P}_p \circ (\phi_i - f) = f \) a.e. in \( \mathcal{X} \) and \( \mathcal{P}_p \circ (\phi_i - f)(x) \leq 2^p \mathcal{P}_p \circ f(x), \forall x \in X, \forall i \in \mathbb{N}. \) By Lebesgue Dominated Convergence Theorem 11.91, we have \( \lim_{i \to \infty} \int_X \mathcal{P}_p \circ (\phi_i - f) d\mu = 0. \) Hence, \( \lim_{i \to \infty} \phi_i = f. \) This completes the proof of the proposition. \( \square \)

**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 \( f \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y}). \) Then, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \) a continuous function \( g : \mathcal{X} \to \mathcal{Y} \) such that \( g \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y}) \) and \( \|g - f\|_p < \epsilon. \)

**Proof** Fix \( \epsilon \in (0, \infty) \subseteq \mathbb{R}. \) By Proposition 11.181, there exists a simple function \( \phi : \mathcal{X} \to \mathcal{Y} \) 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 \(\emptyset\), \(A_1, \ldots, A_n \in \mathcal{B}\) 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 \(\mu(U_i \setminus F_i) = \mu(U_i \setminus A_i) + \mu(A_i \setminus F_i) < \frac{e^{-c\mu(Y)}}{2(n+1)\|\mu\|}\). 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, \(\|g_i - \chi_{A_i,x}\|_p \leq \frac{e^{-c\mu(Y)}}{2(n+1)\|\mu\|}\).

Define \(g : X \to Y\) 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, \(\|g - f\|_p \leq \|g - \phi\|_p + \|\phi - f\|_p \leq \sum_{i=1}^{n} \|y_i(g_i - \chi_{A_i,x})\|_p + \epsilon/2 = \sum_{i=1}^{n} \|y_i\| \|g_i - \chi_{A_i,x}\|_p + \epsilon/2 \leq \sum_{i=1}^{n} \frac{\|y_i\|}{2(n+1)\|\mu\|} + \epsilon/2 < \epsilon\). Hence, \(g \in L_p(X,Y)\).

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

**Proposition 11.183** Let \(X := (X, \mathcal{B}, \mu)\) be a \(\sigma\)-finite measure space, \(Y\) be a separable Banach space, \(f : X \to Y\) 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 \(X\), and \(M \subseteq [0, \infty) \subset \mathbb{R}\). Assume that, \(\forall E \in \mathcal{B}\) with \(0 \leq \mu(E) < \infty\), we have \(\|\frac{1}{\mu(E)} \int_E f \, d\mu\| \leq M\). Then, \(\mathcal{P} \circ f \leq M\) a.e. in \(X\).

**Proof** Consider the open set \(O := \{y \in Y \mid \|y\| > M\} \subseteq Y\). Since \(Y\) 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 \(\mu(f_{\text{inv}}(B_y(y,r))) = 0\), \(\forall \mathcal{B}_Y(y,r) \in \mathcal{M}\), 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 \(X\) 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 \(x \in E\) and \(\|f(x) - y\| < r\). Then, \(\|\frac{1}{\mu(E)} \int_E f \, d\mu - y\| = \|\frac{1}{\mu(E)} \int_E (f - y) \, d\mu\| = \frac{1}{\mu(E)} \int_E (f - y) \, d\mu(x)\| \leq \frac{1}{\mu(E)} \|\int_E f \, d\mu - y\| \leq \frac{1}{\mu(E)} \|\int_E f \, d\mu\| = \frac{1}{\mu(E)} \|\int_E f \, d\mu\| > M\). This contradicts the assumption. Hence, \(\mu(f_{\text{inv}}(B_y(y,r))) = 0\), \(\forall \mathcal{B}_Y(y,r) \in \mathcal{M}\).

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 \(X\). \(\square\)
Lemma 11.184 Let \( p, q \in (1, \infty) \subset \mathbb{R} \) with \( 1/p + 1/q = 1 \), \( \mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( Y \) be a separable normed linear space over \( \mathbb{K} \) with \( Y^* \) being separable, \( \mathcal{Z} := L_p(\mathcal{X}, Y) \) be the normed linear space over \( \mathbb{K} \) as defined in Example 11.173, and \( g : \mathcal{X} \to Y^* \) be \( \mathcal{B} \)-measurable. Assume 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 \( \mathcal{X} \);

(ii) \( \exists M \in [0, \infty) \subset \mathbb{R} \) such that \( \forall \) simple function \( \phi : \mathcal{X} \to Y \) (\( \phi \in L_p(\mathcal{X}, \mathcal{Y}) \)), the function \( \langle \langle g(\cdot), \phi(\cdot) \rangle \rangle : \mathcal{X} \to \mathbb{K} \) is absolutely integrable over \( \mathcal{X} \) and \( \int_{\mathcal{X}} |\langle \langle g(x), \phi(x) \rangle \rangle| \, d\mu(x) \leq M \| \phi \|_p \).

Then, \( g \in L_q(\mathcal{X}, Y^*) \) and \( \| g \|_q \leq M \).

Proof By Proposition 11.116, define a \( \sigma \)-finite \( Y^* \)-valued measure \( \nu \) on \( (\mathcal{X}, \mathcal{B}) \) by: \( \nu(E) \) is undefined, \( \forall E \in \mathcal{B} \) with \( \int_E P \circ g \, d\mu = \infty; \nu(E) = \int_E g \, d\mu \in Y^*, \forall E \in \mathcal{B} \) with \( \int_E P \circ g \, d\mu < \infty \). Then, \( P \circ \nu(E) = \int_E P \circ g \, d\mu, \forall E \in \mathcal{B} \). By (i), \( P \circ \nu(E) < \infty, \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \). By Proposition 11.92, \( \nu(E) = \emptyset_{Y^*}, \forall E \in \mathcal{B} \) with \( \mu(E) = 0 \). Fix any \( E \in \mathcal{B} \) with \( \mu(E) < \infty \). Define \( P_q \circ \nu(E) \in [0, \infty] \subset \mathbb{R} \) by

\[
P_q \circ \nu(E) := \sup_{\begin{array}{c}
\forall n \in \mathbb{Z}_+, \forall \text{pairwise disjoint } (E_i)_{i=1}^n \subseteq \mathcal{B} \\
\forall \epsilon \in (0, 1) \subset \mathbb{R}, \forall i \in \{1, \ldots, n\} \text{ with } \mu(E_i) > 0, \text{ by Lemma } 7.75,
\end{array}} \sum_{i=1}^n \frac{\|\nu(E_i)\|^{q-1}/(\mu(E_i))^{q-1} \leq (1 - \epsilon) \|\nu(E_i)\| \|y_i\| = (1 - \epsilon) \|\nu(E_i)\|^{q}/(\mu(E_i))^{q-1}}
\]

Then,

\[
\|\phi\|_p = \left( \sum_{i=1}^n \frac{\|\nu(E_i)\|^{(q-1)p} (\mu(E_i))^{(1-q)p} \mu(E_i)}{\mu(E_i)} \right)^{1/p}
\]

By (ii), we have

\[
\sum_{i=1}^n \frac{1}{\mu(E_i)} (1 - \epsilon) \|\nu(E_i)\|^{q} / (\mu(E_i))^{q-1} \leq \sum_{i=1}^n \langle \langle \nu(E_i), y_i \rangle \rangle \leq \sum_{i=1}^n \int_{E_i} \langle \langle g(x), y_i \rangle \rangle \, d\mu(x)
\]

\[
= \left| \sum_{i=1}^n \langle \int_{E_i} g \, d\mu, y_i \rangle \right| = \left| \sum_{i=1}^n \int_{E_i} \langle \langle g(x), y_i \rangle \rangle \, d\mu(x) \right|
\]
where the second and third equalities follow from Proposition 11.92. Hence, we have \(\sum_{i=1}^{n} \mu(E_i) > 0\) \(\|\nu(E_i)\|^q (\mu(E_i))^{-q} \leq M^q/(1-\epsilon)^q\). By the arbitrariness of \(\epsilon\), we have \(\sum_{i=1}^{n} \mu(E_i) > 0\) \(\|\nu(E_i)\|^q (\mu(E_i))^{-q} \leq M^q\). Then, \(\mathcal{P}_q \circ \nu(E) \leq M^q\).

Since \(\mathcal{X}\) 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 \((\psi_i)_{i=1}^{\infty}, \psi_i : \mathcal{X} \to \mathbb{Y}\), \(\forall i \in \mathbb{N}\), such that \(\|\psi_i(x)\| \leq \|\psi(x)\|, \forall x \in X, \forall i \in \mathbb{N}\), and \(\lim_{i \in \mathbb{N}} \psi_i = g\) a.e. in \(\mathcal{X}\). Fix any \(n \in \mathbb{N}\), let \(E_n := \{ x \in X_n | \|\psi(x)\| \leq n\} \in \mathcal{B}\) and \(E_n := (E_n, \mathcal{B}_{E_n}, \mu_{E_n})\) be the finite measure subspace of \(\mathcal{X}\). Then, \((\mathcal{P} \circ g)|_{E_n}\) and \((\mathcal{P}_q \circ g)|_{E_n}\) are integrable over \(E_n\). By Propositions 7.21, 7.23, 11.52, and 11.53, \(\lim_{i \in \mathbb{N}} (\mathcal{P} \circ (\psi_i - g))|_{E_n} = 0\) a.e. in \(E_n\). Note that \((\mathcal{P} \circ (\psi_i - g))|_{E_n}\) \(x) \leq 2 (\mathcal{P} \circ g)|_{E_n}\) \(x)\), \(\forall x \in E_n, \forall i \in \mathbb{N}\). By Lebesgue Dominated Convergence Theorem 11.91, we have \(\lim_{i \in \mathbb{N}} \int_{E_n} \mathcal{P} \circ (\psi_i - g) \, d\mu = 0\). Note also that \((\mathcal{P}_q \circ \psi_i)|_{E_n}(x) \leq (\mathcal{P}_q \circ g)|_{E_n}(x)\), \(\forall x \in E_n, \forall i \in \mathbb{N}\). By Propositions 7.21 and 11.52, \(\lim_{i \in \mathbb{N}} (\mathcal{P}_q \circ \psi_i)|_{E_n} = (\mathcal{P}_q \circ g)|_{E_n}\) a.e. in \(E_n\).

Again, by Lebesgue Dominated Convergence Theorem 11.91, we have \(\lim_{i \in \mathbb{N}} \int_{E_n} \mathcal{P}_q \circ \psi_i \, d\mu = \int_{E_n} \mathcal{P}_q \circ g \, d\mu\). \(\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists i_0 \in \mathbb{N}\) such that \(0 \leq \int_{E_n} \mathcal{P} \circ (\psi_{i_0} - g) \, d\mu < \frac{\epsilon}{2^{\frac{1}{2}} m_{E_n}}\) and \(\left| \int_{E_n} \mathcal{P}_q \circ \psi_{i_0} \, d\mu - \int_{E_n} \mathcal{P}_q \circ g \, d\mu \right| < \epsilon/2\). Let \(\psi_{i_0}\) admit the canonical representation \(\psi_{i_0} = \sum_{j=1}^{\bar{n}} y_{i_0} \chi_{A_j}\). Then,
\[
\begin{align*}
&\leq \mathcal{P}_q \circ \nu(E_n) - \sum_{n=1}^{\bar{n}} \left\| \nu(\tilde{A}_j) \right\|^q (\mu(\tilde{A}_j))^{1-q} \\
&\quad + \sum_{n=1}^{\bar{n}} \left\| \int_{\tilde{A}_j} \psi_{i_0} \, d\mu \right\|^q (\mu(\tilde{A}_j))^{1-q} + \epsilon/2 \\
&\leq M^q + \sum_{n=1}^{\bar{n}} (\mu(\tilde{A}_j))^{1-q} \left( \left\| \int_{\tilde{A}_j} \psi_{i_0} \, d\mu \right\| - \left\| \int_{\tilde{A}_j} g \, d\mu \right\| \right) q \\
&\quad \cdot \left( t_j \left\| \int_{\tilde{A}_j} \psi_{i_0} \, d\mu \right\| + (1 - t_j) \left\| \int_{\tilde{A}_j} g \, d\mu \right\| \right)^{q-1} + \epsilon/2 \\
&\leq M^q + \sum_{n=1}^{\bar{n}} q n^{q-1} \left\| \int_{\tilde{A}_j} (\psi_{i_0} - g) \, d\mu \right\| + \epsilon/2 \\
&\leq M^q + \sum_{n=1}^{\bar{n}} q n^{q-1} \int_{\tilde{A}_j} \mathcal{P} \circ (\psi_{i_0} - g) \, d\mu + \epsilon/2 \\
&\leq M^q + q n^{q-1} \int_{E_n} \mathcal{P} \circ (\psi_{i_0} - g) \, d\mu + \epsilon/2 < M^q + \epsilon
\end{align*}
\]

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 \( \left\| \psi_{i_0}(x) \right\| \leq \left\| g(x) \right\| \leq n, \forall x \in E_n \); 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 \( \int_X \mathcal{P}_q \circ g \, d\mu = \lim_{n \to \infty} \int_{E_n} (\mathcal{P}_q \circ g)_{|E_n} \, d\mu = \lim_{n \to \infty} \int_{E_n} \mathcal{P}_q \circ g \, d\mu \leq M^q \). Hence, \( \left\| g \right\|_{q} \leq M \) and \( g \in L_p(X, \mathbb{Y}^*) \).

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

**Lemma 11.185** Let \( p \in [1, \infty) \subset \mathbb{R}, \ X := (X, \mathcal{B}, \mu) \) be a \( \sigma \)-finite measure space, \( \mathbb{Y} \) be a separable reflexive Banach space over \( \mathbb{K} \) with \( \mathbb{Y}^* \) being separable, and \( Z := L_p(X, \mathbb{Y}) \) be the Banach space over \( \mathbb{K} \) as defined in
Example 11.179. Then, \( \forall f \in \mathcal{Z}^* \), \( \exists g : X \to \mathcal{Y}^* \), 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 \( \phi : X \to \mathcal{Y} \) (\( \phi \in \mathcal{Z} := \mathcal{L}_\mu(X, \mathcal{Y}) \)), the function 
\[
\langle \langle g(\cdot), \phi(\cdot) \rangle \rangle : X \to \mathcal{K}
\]
is absolutely integrable over \( X \), 
\[
f(\langle \phi \rangle) = \int_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \text{ and } \left| \left| \int_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \right| \right| \leq \| f \|_p \cdot \| \phi \|_p.
\]

Furthermore, \( g \) is unique in the sense that \( \tilde{g} : X \to \mathcal{Y}^* \) is another function with the above properties if, and only if, \( \tilde{g} \) is \( \mathcal{B} \)-measurable and \( g = \tilde{g} \) a.e. in \( X \).

Proof Fix any \( f \in \mathcal{Z}^* \). \( \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \), \( \forall y \in \mathcal{Y} \), let 
\[
z_{E, y} := y \chi_{E, X} \in \mathcal{Z}.
\]
Define \( f_E : \mathcal{Y} \to \mathcal{K} \) by \( f_E(y) = f([z_{E, y}]) \), \( \forall y \in \mathcal{Y} \).
Since \( f \in \mathcal{Z}^* \), then \( f_E \) is linear and continuous and \( f_E = y_{*E} \in \mathcal{Y}^* \). Note that 
\[
[z_{E, y}] = \vartheta_\mathcal{Z}, \forall y \in \mathcal{Y}, \text{ and } f([z_{E, y}]) = 0.
\]
Then, \( f_\vartheta = \vartheta_\mathcal{Y}^* \) and 
\[
y_{*E} = \vartheta_\mathcal{Y}^*.
\]

Claim 11.185.1 \( \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \), \( \forall \) pairwise disjoint \( (E_i)_{i=1}^\infty \subseteq \mathcal{B} \) with \( E = \bigcup_{i=1}^\infty E_i \). Then, 
\[
\sum_{i=1}^\infty \| y_{*E_i} \| \leq \| f \|_p \left( \mu(E) \right)^{1/p} < \infty.
\]

Proof of claim: \( \sum_{i=1}^\infty \| y_{*E_i} \| = \sum_{i=1}^\infty \sup_{y \in \mathcal{Y}, \| y \| \leq 1} | f_{E_i}(y) | \), \( \forall i \in \mathbb{N} \),
by Propositions 7.85 and 7.90, \( \exists y_i \in \mathcal{Y} \) with \( \| y_i \| \leq 1 \) such that 
\[
\| y_{*E_i} \| = f_{E_i}(y_i) = f([z_{E_i, y_i}]).
\]
Then, \( \forall n \in \mathbb{N} \), 
\[
\sum_{i=1}^n \| y_{*E_i} \| = \sum_{i=1}^n f([z_{E_i, y_i}]) = f((\sum_{i=1}^n z_{E_i, y_i})) \leq \| f \|_p \| \sum_{i=1}^n z_{E_i, y_i} \|_p = \| f \|_p \| \sum_{i=1}^n z_{E_i, y_i} \|_p \leq \| f \|_p \cdot \left( \mu(E) \right)^{1/p} < \infty,
\]
where the first inequality follows from Proposition 7.72; and the second inequality follows from Propositions 11.83 and 11.75. By the arbitrariness of \( n \), we have 
\[
\sum_{n=1}^\infty \| y_{*E_n} \| \leq \| f \|_p \left( \mu(E) \right)^{1/p}.
\]
This completes the proof of the claim.

Claim 11.185.2 \( \forall E \in \mathcal{B} \) with \( \mu(E) < \infty \), \( \forall \) pairwise disjoint \( (E_n)_{n=1}^\infty \subseteq \mathcal{B} \) with \( E = \bigcup_{n=1}^\infty E_n \), we have 
\( y_{*E} = \sum_{n=1}^\infty y_{*E_n} \in \mathcal{Y}^* \).

Proof of claim: By Claim 11.185.1 and Propositions 7.27 and 7.72, 
\[
\sum_{n=1}^\infty y_{*E_n} \in \mathcal{Y}^*, \forall y \in \mathcal{Y},
\]
\[
\langle \langle \sum_{n=1}^\infty y_{*E_n}, y \rangle \rangle = \lim_{n \to \infty} \langle \langle \sum_{i=1}^n y_{*E_i}, y \rangle \rangle = \lim_{n \to \infty} \sum_{i=1}^n \langle \langle y_{*E_i}, y \rangle \rangle
\]
\[
= \lim_{n \to \infty} \sum_{i=1}^n f([z_{E_i, y}]) = \lim_{n \to \infty} f((\sum_{i=1}^n z_{E_i, y}))
\]
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\in \mathbb{N}} \int_{\mathcal{P}_n} (\sum_{i=1}^n z_{E_i,y} - z_{E_i,y}) \, d\mu = 0 \quad \text{and} \quad \lim_{n\in \mathbb{N}} \sum_{i=1}^n z_{E_i,y} = z_{E,E,y} \in \mathcal{L}.
By Propositions 7.72 and 3.66, \lim_{n\in \mathbb{N}} f((\sum_{i=1}^n z_{E_i,y})) = f(z_{E,y}) = \langle \langle y_{E,y}, y \rangle \rangle. \quad \text{Hence,} \quad \langle \langle y_{E,y}, y \rangle \rangle = \langle \langle \sum_{i=1}^n y_{E,i,y} \rangle \rangle, \; \forall y \in \mathcal{Y}.
This implies that, by Proposition 7.85, \forall y_{E} \in \sum_{n=1}^\infty y_{E,n}.
\square

Since \mathcal{X} is \sigma\text{-finite, then} \exists (X_n)_{n=1}^\infty \subseteq \mathcal{B} \text{such that} X = \bigcup_{n=1}^\infty X_n \text{and} \mu(X_n) < \infty, \forall n \in \mathbb{N}. \text{Without loss of generality, we may assume that} (X_n)_{n=1}^\infty \text{is pairwise disjoint. Fix any} \; n \in \mathbb{N}, \text{let} \mathcal{X}_n := (X_n, \mathcal{B}_n, \mu_n) \text{be the finite measure subspace of} \mathcal{X}. \text{We may define a function} \nu_n : \mathcal{B}_n \to \mathcal{Y}^* \text{by} \nu_n(E) = f_E = y_{E,y}, \forall E \in \mathcal{B}_n. \text{Clearly,} \nu_n(\emptyset) = y_{\emptyset,\emptyset} = \emptyset \in \mathcal{Y}^*. \forall \text{pairwise disjoint} (E_i)_{i=1}^\infty \subseteq \mathcal{B}_n, \text{let} E := \bigcup_{i=1}^\infty E_i \in \mathcal{B}_n. \text{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.
By Claim 11.185.2, \nu_n(E) = \sum_{i=1}^\infty \nu_n(E_i) \in \mathcal{Y}^*. \text{This shows that} \nu_n \text{is a} \mathcal{Y}^*-\text{valued pre-measure on} (X_n, \mathcal{B}_n). \text{By Claim 11.185.1,} \mathcal{P} \circ \nu_n(X_n) \leq \|f\| (\mu(X_n))^{1/p} < +\infty. \text{Then,} \nu_n \text{is finite. Hence,} (X_n, \mathcal{B}_n, \nu_n) \text{is a} \mathcal{Y}^*-\text{valued measure space. By Proposition 11.118, the generation process on} ((X_n, \mathcal{B}_n, \nu_n))_{n=1}^\infty \text{yields a unique} \sigma\text{-finite} \mathcal{Y}^*-\text{valued measure space} (X, \mathcal{B}, \nu) \text{on} \mathcal{X}.

Next, we will show that \mathcal{P} \circ \nu(E) \leq \|f\| (\mu(E))^{1/p} < \infty \text{and} \nu(E) = y_{E,y}. \forall E \in \mathcal{B} \text{with} \mu(E) < \infty. \text{Fix any} \; E \in \mathcal{B} \text{with} \mu(E) < \infty. \text{Let} \mathcal{E}_n := X_n \cap E \in \mathcal{B}_n, \forall n \in \mathbb{N}. \text{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 \mathbb{N}, \exists m_n \in \mathbb{Z}_+, \exists \text{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) < \sum_{i=1}^{m_n} \|\nu_n(E_{n,i})\| + 2^{-n}\epsilon < \infty. \text{Then,}

\mathcal{P} \circ \nu(E) < \sum_{n=1}^\infty \sum_{i=1}^{m_n} \|\nu_n(E_{n,i})\| + 2^{-n}\epsilon \\
\leq \sum_{n=1}^\infty \sum_{i=1}^{m_n} \|y_{E,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, \text{we have} \mathcal{P} \circ \nu(E) \leq \|f\| (\mu(E))^{1/p} < \infty. \text{Then,} \; E \in \text{dom} (\nu) \text{and} \nu(E) = \sum_{n=1}^\infty \nu_n(E_n) \in \mathcal{Y}^*. \text{By Claim 11.185.2, we have} \nu(E) = \sum_{n=1}^\infty y_{E,n} = y_{E,y}. \text{Thus, by Proposition 11.137,} \nu \text{is uniquely defined independent of the choice of} (X_n)_{n=1}^\infty.

\forall E \in \mathcal{B} \text{with} \mu(E) = 0, \mathcal{P} \circ \nu(E) \leq 0. \text{Thus,} \mathcal{P} \circ \nu \ll \mu. \text{By Radon-Nikodym Theorem 11.171 and Proposition 7.90,} \exists g : \mathcal{X} \to \mathcal{Y}^* \text{such that} g \text{is the Radon-Nikodym derivative of} \nu \text{with respect to} \mu. \forall E \in \mathcal{B} \text{with} \mu(E) < \infty, \text{we have} \mathcal{P} \circ \nu(E) = \int_E \mathcal{P} \circ g \, d\mu < \infty. \text{Thus, (i) holds.} \forall \text{simple function} \phi : \mathcal{X} \to \mathcal{Y}, \exists M \in [0, \infty) \subseteq \mathbb{R} \text{such that} \|\phi(x)\| \leq M, \forall x \in \mathcal{X}, \text{and} \mu(E) := \mu(\{x \in \mathcal{X} \mid \phi(x) \neq \emptyset\}) < \infty. \text{Then, by Proposition 7.72,}
\|\langle \langle g, \phi \rangle \rangle\| \leq M \|g(x)\| \|\chi_{E,E}(x)\| = M \mathcal{P} \circ g(x) \chi_{E,E}(x), \forall x \in \mathcal{X}. \text{By Propositions 7.72, 11.38, and 11.39, the function} \langle \langle g, \phi \rangle \rangle : \mathcal{X} \to \mathcal{K} \text{is} \mathcal{B}\text{-measurable. Hence, the function} \langle \langle g, \phi \rangle \rangle : \mathcal{X} \to \mathcal{K} \text{is absolutely integrable over} \mathcal{X}. \text{Let} \phi \text{admit the canonical representation} \phi = \sum_{i=1}^n y_i \chi_{E_i,E_i}.
Then,

\[
f([\phi]) = f(\sum_{i=1}^{n} y_i \chi_{E_i}, X) = \sum_{i=1}^{n} f([y_i \chi_{E_i}, X]) = \sum_{i=1}^{n} f([z_{E_i}, y_i])
\]

\[
= \sum_{i=1}^{n} \langle \langle y_i \chi_{E_i}, y_i \rangle \rangle = \sum_{i=1}^{n} \langle \langle \mu(E_i), y_i \rangle \rangle = \sum_{i=1}^{n} \left\langle \left\langle \int_{E_i} g \, d\mu, y_i \right\rangle \right\rangle
\]

\[
= \sum_{i=1}^{n} \int_{E_i} \langle \langle g(x), y_i \rangle \rangle \, d\mu(x) = \int_{X} \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x) \in K
\]

where the last two equalities follow from Proposition 11.92. Then, \( |f_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x)| = |f([\phi])| \leq \|f\| \cdot \|[\phi]\|_p = \|f\| \cdot \|\phi\|_p \). 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 \( g \in L_q(X, Y^*) \) such that \( f([z]) = \langle \langle [g], [z] \rangle \rangle := \int_X \langle \langle g(x), z(x) \rangle \rangle \, d\mu(x) \) and the function \( \langle \langle g(\cdot), z(\cdot) \rangle \rangle : X \to K \) is absolutely integrable over \( X \), \( \forall \nu \in Z := L_p(X, Y) \), 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 \( \phi : X \to Y \) (\( \phi \in L_p(X, Y) \)), 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 \( |f_X \langle \langle g(x), \phi(x) \rangle \rangle \, d\mu(x)| \leq \|f\| \cdot \|\phi\|_p \).

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 \( g \in L_q(X, Y^*) \) 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 \( g \in L_q(X, Y^*) \) 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, \( y_{\mu E} := \int_E y \, d\mu \in Y^* \) and, \( \forall y \in Y \), \( |\langle \langle y_{\mu E}, y \rangle \rangle| = |\int_E \langle \langle g(x), y \rangle \rangle \, d\mu(x)| = |\int_X \langle \langle g(x), y \chi_{E}, x \rangle \rangle \, d\mu(x)| \leq \|f\| \|y \chi_{E}, x\|_1 = \|f\| \|y\| \mu(E) \), 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 \( P \circ g \leq \| f \| \) a.e. in \( X \). Hence \( \| g \|_\infty \leq \| f \| \) and \( g \in L_\infty(X, Y^*) \).

Hence, in both cases, we have \( g \in L_q(X, Y^*) \) and \( \| g \|_q \leq \| f \| < +\infty \).

\( \forall x \in \tilde{Z} \), by Proposition 11.181, there exists a sequence of simple functions \( (\varphi_i)_{i=1}^\infty \), \( \varphi_i : X \to Y \), \( \forall i \in \mathbb{N} \), such that \( \lim_{i \to \infty} \varphi_i = z \) a.e. in \( X \). \( \| \varphi_i(x) \| \leq \| z(x) \| \), \( \forall x \in X \), \( \forall i \in \mathbb{N} \) and \( \lim_{i \to \infty} \varphi_i = z \) in \( \tilde{Z} \). By Propositions 7.72 and 3.66, \( f([z]) = \lim_{i \to \infty} f(\langle \varphi_i \rangle) = \lim_{i \to \infty} \int_X \langle g(x), \varphi_i(x) \rangle \, d\mu(x) \). Note that, by Propositions 7.72, 11.52, and 11.53, \( \lim_{i \to \infty} \int_X \langle g(x), \varphi_i(x) \rangle \, d\mu(x) \) is absolutely integrable over \( X \). Hence, we may define a function \( \Phi : \tilde{Z} \to K \) by \( \Phi(z) = \int_X \langle g(x), z(x) \rangle \, d\mu(x) \) and the function \( \langle g(\cdot), z(\cdot) \rangle : X \to K \) is absolutely integrable over \( X \).

Thus, we have \( \| f \| = \| g \|_q \).

\( \| f \| = \sup_{|z| < Z, \|z\|_p \leq 1} |f([z])| = \sup_{|z| < Z, \|z\|_p \leq 1} \left| \int_X \langle g(x), z(x) \rangle \, d\mu(x) \right| \)

\leq \sup_{|z| < Z, \|z\|_p \leq 1} \int_X \| \langle g(x), z(x) \rangle \| \, d\mu(x)

\leq \sup_{|z| < 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 \( \Phi : \tilde{Z}^* \to L_q(X, Y^*) \) by \( \Phi(f) = [g] \), \( \forall f \in \tilde{Z}^* \). Clearly, \( \Phi \) is well defined and \( \| \Phi(f) \|_q = \| f \| \), \( \forall f \in \tilde{Z}^* \). \( \forall f_1, f_2 \in \tilde{Z}^* \), \( \forall \alpha, \beta \in \mathbb{K} \), let \( [g_1] = \Phi(f_1) \), \( i = 1, 2 \). \( \forall [z] \in Z \),

\( (\alpha f_1 + \beta f_2)([z]) = \alpha f_1([z]) + \beta f_2([z]) \)

\( = \alpha \int_X \langle g_1(x), z(x) \rangle \, d\mu(x) + \beta \int_X \langle g_2(x), z(x) \rangle \, d\mu(x) \)

\( = \int_X \langle \alpha g_1(x) + \beta g_2(x), z(x) \rangle \, d\mu(x) \)

where the last equality follows from Proposition 11.92. By the uniqueness of \( g \), we have \( \Phi(\alpha f_1 + \beta f_2) = \alpha g_1 + \beta g_2 = \alpha [g_1] + \beta [g_2] = \alpha \Phi(f_1) + \beta \Phi(f_2) \).

Hence, \( \Phi \) is linear. \( \forall [g] \in L_q(X, Y^*) \), we may define \( f : Z \to \mathbb{K} \) by \( f([z]) = \int_X \langle g(x), z(x) \rangle \, d\mu(x) \), \( \forall z \in \tilde{Z} \). By Hölder’s Inequality 11.178 and Proposition 11.92, we have \( f \in \tilde{Z}^* \). Then, \( \Phi(f) = [g] \) and \( \Phi \) is surjective. \( \forall f_1, f_2 \in \tilde{Z}^* \) with \( \Phi(f_1) = \Phi(f_2) \), 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, \( \Phi : \tilde{Z}^* \to L_q(X, Y^*) \) is a isometrical isomorphism. This completes the proof of the theorem. \( \square \)
11.10 Dual of $\mathcal{C}(\mathcal{X}, \mathcal{Y})$

Proposition 11.187 Let $\mathcal{X} := (\mathcal{X}, \rho)$ be a metric space and $\mathcal{X} := (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$ be a metric measure space. Then, the following statements hold.

(i) $\forall$ compact subset $K \subseteq \mathcal{X}$, we have $K \in \mathcal{B}(\mathcal{X})$.

(ii) If $\mathcal{X}$ is locally finite and $\sigma$-compact, then it is $\sigma$-finite.

(iii) If $\mathcal{X}$ is finite, then it is locally finite.

Proof

(i) Fix any compact subset $K \subseteq \mathcal{X}$. We will show that $K = \bigcap_{n=1}^{\infty} \bigcup_{x \in X_n} \mathcal{B}(x, 1/n) := E$, where $X_n$'s are finite subsets of $K$. Clearly, we have $E \in \mathcal{B}(\mathcal{X})$.

By Proposition 11.10, $\forall n \in \mathbb{N}$, $K \subseteq \bigcup_{x \in K} \mathcal{B}(x, 1/n)$. Since $K$ is compact, then there exists a finite set $X_n \subseteq K$ such that $K \subseteq \bigcup_{x \in X_n} \mathcal{B}(x, 1/n)$. Then, $K \subseteq E$. On the other hand, $\forall \bar{x} \in \bar{K}$, by Proposition 5.5, $K$ is closed in $\mathcal{X}$. By Proposition 4.10, $\text{dist}(\bar{x}, K) > 0$. Then, $\exists n_0 \in \mathbb{N}$ such that $\text{dist}(\bar{x}, K) > 1/n_0$. Then, $\bar{x} \notin \mathcal{B}(x, 1/n_0), \forall x \in K$, which implies that $\bar{x} \notin \bigcup_{x \in X_{n_0}} \mathcal{B}(x, 1/n_0)$, which further implies that $\bar{x} \notin E$. Hence, $K \subseteq E$ and $E \subseteq K$. Thus, we have $K = E \in \mathcal{B}(\mathcal{X})$.

(ii) By $\mathcal{X}$ being $\sigma$-compact, there exists $(K_n)_{n=1}^{\infty} \subseteq \mathcal{X}$ such that $K_n$ is compact, $\forall n \in \mathbb{N}$. By $\mathcal{X}$ being locally finite, we have $K_n \in \mathcal{B}(\mathcal{X})$ and $\mu(K_n) < \infty$. Then, $\mathcal{X}$ is $\sigma$-finite.

(iii) $\forall$ compact subset $K \subseteq \mathcal{X}$, by (i), $K \in \mathcal{B}(\mathcal{X})$. Since $\mathcal{X}$ is finite, then $\mu(K) \leq \mu(K) < \infty$. Then, $\mathcal{X}$ is locally finite.

This completes the proof of the proposition. ☐

Proposition 11.188 Let $\mathcal{X} := (\mathcal{X}, \mathcal{O})$ be a topological space, $(\mathcal{X}_n)_{n=1}^{\infty} \subseteq \mathcal{O}$, $\mathcal{X}_n := (\mathcal{X}_n, \mathcal{O}_n)$ be a topological subspace of $\mathcal{X}$, and $(\mathcal{X}_n, \mathcal{B}(\mathcal{X}_n), \mu_n) := (\mathcal{X}_n)$ 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 $\mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu)$ such that $\mathcal{X}_n$ is the finite measure subspace of $\mathcal{X}$, $\forall n \in \mathbb{N}$. 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 $E \cap X_n \in \mathcal{B}(\mathcal{X}_n), \forall n \in \mathbb{N}$. 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, $\bigcup_{n=1}^{\infty} E_{(n)} = E$, and $E_{(n)} \in \mathcal{B}(\mathcal{X}_n) \subseteq \mathcal{B}(\mathcal{X})$, $\forall n \in \mathbb{N}$. Define $\bar{\mu} : \mathcal{B}(\mathcal{X}) \to [0, \infty] \subseteq \mathbb{R}$ by $\bar{\mu}(E) = \sum_{n=1}^{\infty} \mu_n(E_{(n)}) \in [0, \infty] \subseteq \mathbb{R}$, $\forall E \in \mathcal{B}(\mathcal{X})$.

We will show that $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \bar{\mu})$ 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, \( E_{(n+1)} = (E \cap X_{n+1}) \setminus \bigcup_{i=1}^{n} E_{i,(i)} = \bigcup_{i=1}^{\infty} (E_i \cap X_{n+1}) \setminus \bigcup_{i=1}^{n} E_{i,(i)} = \bigcup_{i=1}^{\infty} ((E_i \cap X_{n+1}) \setminus \bigcup_{i=1}^{n} E_{i,(i)}) = \bigcup_{i=1}^{\infty} E_{i,(n+1)}, \forall n \in \mathbb{N}, \) where the second equality follows from the inductive assumption; and the third equality follows from the fact that \( (E_j)_{j=1}^{\infty} \) is pairwise disjoint and \( E_{j,(i)} \subseteq E_j, \forall i \in \mathbb{N}, \forall j \in \mathbb{N} \).

Then, 

\[
\tilde{\mu}(E) = \sum_{n=1}^{\infty} \mu_n(E_{(n)}) = \sum_{n=1}^{\infty} \mu_n\left( \bigcup_{j=1}^{\infty} E_{j,(n)} \right) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \mu_n(E_{j,(n)}) = \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} \mu_n(E_{j,(n)}) = \sum_{j=1}^{\infty} \tilde{\mu}(E_j)
\]

where the fourth equality follows from the fact that all summands are non-negative real numbers. Hence, \((X, B_{\mathcal{B}}(X), \tilde{\mu})\) is a measure space. It is \( \sigma \)-finite since \( \tilde{\mu}(X_n) = \sum_{i=1}^{\infty} \mu_i(X_{n,(i)}) = \sum_{i=1}^{\infty} \mu_n(X_{n,(i)}) = \mu_n(X_n) < \infty, \forall n \in \mathbb{N} \), where the second equality follows from the consistency assumption of Proposition 11.118 on \((X_n)_{n=1}^{\infty} \).

Fix any \( n \in \mathbb{N} \). \( \forall E \in B_{\mathcal{B}}(X_n) \subseteq B_{\mathcal{B}}(X), \tilde{\mu}(E) = \sum_{i=1}^{\infty} \mu_i(E_{(i)}) = \sum_{i=1}^{\infty} \mu_n(E_{(i)}) = \mu_n\left( \bigcup_{i=1}^{\infty} E_{(i)} \right) = \mu_n(E), \) 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 \((X, B_{\mathcal{B}}(X), \tilde{\mu}), \forall n \in \mathbb{N} \). By Proposition 11.118, we have \( B = B_{\mathcal{B}}(X) \) and \( \mu = \tilde{\mu} \). Then, \( X = (X, B_{\mathcal{B}}(X), \tilde{\mu}) \).

Finally, we will show that \( X \) is a topological measure space. Fix any \( E \in B_{\mathcal{B}}(X) \). Then, \( E = \bigcup_{n=1}^{\infty} E_{(n)} \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \forall n \in \mathbb{N} \), since \( X_n \) is a topological measure space and \( E_{(n)} \in B_{\mathcal{B}}(X_n) \), \( \exists O_n \in O_n \) 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 \( X \) and \( X_n \in O \), then \( O_n \in O_n \subseteq O \subseteq B_{\mathcal{B}}(X) \). 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

\[
\mu(O \setminus E) = \tilde{\mu}(O \setminus E) = \mu\left( \left( \bigcup_{n=1}^{\infty} O_n \right) \setminus \left( \bigcup_{n=1}^{\infty} E_{(n)} \right) \right) = \mu\left( \left( \bigcup_{n=1}^{\infty} O_n \right) \setminus \left( \bigcup_{j=1}^{\infty} \tilde{E}_{(j)} \right) \right) \leq \mu\left( \bigcup_{n=1}^{\infty} O_n \setminus \tilde{E}_{(n)} \right) = \mu\left( \bigcup_{n=1}^{\infty} O_n \setminus E_{(n)} \right) < \epsilon
\]

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 \( \mathcal{O}, (X_n)_{n=1}^{\infty} \subseteq \mathcal{O}, \mathcal{X}_n := (X_n, \rho) \) be a metric subspace of \( \mathcal{X} \),
Proposition 11.190

Let $I K$ be a finite metric measure space, $\forall n \in N$. Assume that $(X_n)_{n=1}^{\infty}$ satisfies the assumptions of Proposition 11.118. By Proposition 11.188, the generation process on $(X_n)_{n=1}^{\infty}$ yields a unique $\sigma$-finite metric measure space $X := (X, B_B(X), \mu)$ such that $X_n$ is the finite metric measure subspace of $X$, $\forall n \in N$. Then, $X$ is locally finite.

**Proof** Fix any compact subset $K \subseteq X$. By Proposition 11.187, $K \in B_B(X)$. Clearly, $K \subseteq \bigcup_{n=1}^{\infty} X_n = X$. By the compactness of $K$, $\exists N \in N$ 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, $X$ is locally finite.

**Proposition 11.190** Let $p \in [1, \infty) \subseteq R$, $X_1$ be a finite-dimensional Banach space, $X_2$ be a $\sigma$-compact metric space, $Y_1$ be a finite-dimensional Banach space over $K$, $Y_2$ be a separable normed linear space over $K$, $X \subseteq X_1$ be a subset with subset topology $O$, $X := (X, O)$, and $X := (X, B_B(X), \mu)$ 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 $C(X, Y_1)$ is a separable Banach space.

(iv) If $X$ is a locally finite $\sigma$-compact metric measure space, then $L_p(X, Y_1)$ is a separable Banach space.

(v) $X_2$ is separable.

**Proof** Let $Y_1$ be $m$-dimensional, where $m \in Z_+$, and $\{e_1, \ldots, e_m\} \subseteq Y_1$ be a basis of 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, $\forall \alpha := (\alpha_1, \ldots, \alpha_m) \in K^m$, $\| a \|_1 := \| \sum_{i=1}^{m} \alpha_i e_i \|_{Y_1} \in R$. Then, $Y_1$ is isometrically isomorphic to the normed linear space $(K^m, K, \| \cdot \|_1)$, 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 $K = R$; and $K_Q := \{ a + ib \in C \mid a, b \in Q \}$, if $K = C$. Clearly, $K_Q$ is a countable dense subset of $K$. Then, $K_Q^m$ is a countable dense subset of $(K^m, K, \| \cdot \|_1)$. Since $\| \cdot \|_1$ is equivalent to $| \cdot |$, then $K_Q^m$ is a countable dense subset of $(K^m, K, \| \cdot \|_1)$. Hence, $Y_1$ is separable, since $(K^m, K, \| \cdot \|_1)$ is separable.

(ii) Let $D \subseteq Y_2$ be a countable dense subset and, $\forall i \in N$, $D_i := \{ (y_1, \ldots, y_i, y_{i+1}, y_{i+2}, \ldots) \in Z_2 \mid y_j \in D, 1 \leq j \leq i \}$. Then, $D_i$ is countable, $\forall i \in N$. Let $D := \bigcup_{i=1}^{\infty} D_i \subseteq Z_2$, which is also countable. We will show that $D$ is dense in $Z_2 = L_p(Y_2)$. Let $e := (y_1, y_2, \ldots) \in Z_2$, $\forall \epsilon \in (0, +\infty) \subseteq R$, we have $\sum_{i=1}^{\infty} \| y_i \|_{Y_2} < \infty$. Then, $\exists N \in N$ such that $\sum_{i=1}^{N} \| y_i \|_{Y_2} < \epsilon^2$. $\forall i \in \{ 1, \ldots, N \}$, $\exists \tilde{y}_i \in D$ such that $\| y_i - \tilde{y}_i \|_{Y_2} < \epsilon/2$. Hence, $D \subseteq Z_2 = L_p(Y_2)$. Then, $\bigcup_{i=1}^{\infty} (y_i, \ldots, y_i, y_{i+1}, y_{i+2}, \ldots) \subseteq Z_2 = L_p(Y_2)$. Hence, $D$ is dense in $Z_2 = L_p(Y_2)$. Therefore, $(Y_2, L_p(Y_2), \| \cdot \|_1)$ is separable.
\[ \forall D \subseteq \mathcal{C}(\mathcal{X}, y_1) \text{ be a countable dense subset. By Proposition 4.11, } \mathcal{X} \text{ is a normal topological space. Then, by Proposition 11.182 and (i), } \forall f \in \tilde{L}_p(\mathcal{X}, y_1), \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}, \exists g \in \tilde{L}_p(\mathcal{X}, y_1) \text{ such that } g \in \mathcal{C}(\mathcal{X}, y_1), \text{ and } \mu(Y) < \epsilon/2. \]

Clearly, \( g \) is continuous and \( g \in \mathcal{C}(\mathcal{X}, y_1) \). By Proposition 11.37, \( h \) is measurable. Hence, \( h \in \tilde{L}_p(\mathcal{X}, y_1) \) is separable. Therefore, \( \tilde{L}_p(\mathcal{X}, y_1) \) is separable. By Example 11.179, \( L_p(\mathcal{X}, y_1) \) is a Banach space. Hence, \( L_p(\mathcal{X}, y_1) \) is a separable Banach space.

Now, consider the general case where \( \mathcal{X} \) is a locally finite \( \sigma \)-compact metric measure space. Then, \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) such that \( \mathcal{X} = \bigcup_{n=1}^{\infty} X_n \) 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 \( \mathcal{X} \) is locally finite, then \( X_n \in \mathcal{B}_\mathcal{X} \) and \( \mu(X_n) < \infty, \forall n \in \mathbb{N} \). Fix any \( n \in \mathbb{N} \). By Proposition 4.37, let \( \mathcal{X}_n \) be the metric subspace of \( \mathcal{X} \). Let \( X_n := (X_n, \mathcal{B}_n, \mu_n) \) be the finite measure subspace of \( \mathcal{X} \). By Proposition 11.29, \( B_n = \mathcal{B}_\mathcal{X}(X_n) \) and \( X_n \) is a finite compact metric measure space. By the special case, \( L_p(\mathcal{X}_n, y_1) \) is a separable Banach space. Let \( D_n \subseteq L_p(\mathcal{X}_n, y_1) \) be a countable dense subset.

\[ \forall f \in \tilde{L}_p(\mathcal{X}_n, y_1), \text{ define } \bar{f} \in L_p(\mathcal{X}, y_1) \text{ by } \bar{f}(x) = \begin{cases} f(x) & \forall x \in X_n \\ \overline{y}_n & \forall x \in X \setminus X_n \end{cases}, \forall x \in \mathcal{X} \text{.}
\]

Let \( D := \{ \bar{f} \in L_p(\mathcal{X}, y_1) \mid \exists n \in \mathbb{N}, [f] \in D_n \} \). Clearly, \( D \subseteq L_p(\mathcal{X}, y_1) \) is countable.

We will show that \( D \) is dense in \( L_p(\mathcal{X}, y_1) \). Fix any \( f \in L_p(\mathcal{X}, y_1) \). We have \( \int_X (P_p \circ f)\,d\mu < \infty \). By Monotone Convergence Theorem 11.81, the integral exists. Let \( f_n := f\chi_{X_n} \in \tilde{L}_p(\mathcal{X}, y_1), \forall n \in \mathbb{N} \). Then,

\[ \lim_{n \in \mathbb{N}} \|f_n - f\|_p = \lim_{n \in \mathbb{N}} \left( \int_X (P_p \circ f_n - f)\,d\mu \right)^{1/p} = \lim_{n \in \mathbb{N}} \left( \int_X (P_p \circ f_n - f - P_p \circ f)\,d\mu \right)^{1/p} = \lim_{n \in \mathbb{N}} \left( \int_X P_p \circ f\,d\mu - \int_X P_p \circ f_n\,d\mu \right)^{1/p} = 0. \]
11.10. DUAL OF $\mathcal{C}(\mathcal{X}, \mathcal{Y})$

where the fourth equality follows from Proposition 11.83. Hence, \( \lim_{n \to \infty} f_n = f \) in \( \mathcal{L}_p(\mathcal{X}, \mathcal{Y}) \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists n_0 \in \mathbb{N} \) such that \( \| f_{n_0} - f \|_p < \epsilon/2 \).

Let \( f_{n_0} := f_{n_0}|_{\mathcal{X}_{n_0}} \). Then, \( f_{n_0} \in \mathcal{L}_p(\mathcal{X}_{n_0}, \mathcal{Y}_1) \). This implies that \( \exists [g] \in D_{n_0} \) with \( g \in L_p(\mathcal{X}_{n_0}, \mathcal{Y}_1) \) such that \( \| f_{n_0} - g \|_p < \epsilon/2 \). Let \( \tilde{g} \in \mathcal{L}_p(\mathcal{X}, \mathcal{Y}_1) \) be defined as in the second paragraph from last. Then, \([\tilde{g}] \in D \) and

\[
\| f - \tilde{g} \|_p \leq \| f - f_{n_0} \|_p + \| f_{n_0} - \tilde{g} \|_p < \epsilon/2 + \left( \int_{\mathcal{X}} p_p \circ (f_{n_0} - \tilde{g}) \, d\mu \right)^{1/p} \]

\[
= \epsilon/2 + \left( \int_{\mathcal{X}} p_p \circ (f_{n_0} - \tilde{g}) |_{\mathcal{X}_{n_0}} \, d\mu_{n_0} \right)^{1/p} \]

\[
= \epsilon/2 + \left( \int_{\mathcal{X}_{n_0}} p \circ (f_{n_0} - g) \, d\mu_{n_0} \right)^{1/p} \]

\[
= \epsilon/2 + \| \tilde{f}_{n_0} - g \|_p < \epsilon
\]

where the second equality follows from Proposition 11.83. Hence, \( D \) is dense in \( L_p(\mathcal{X}, \mathcal{Y}_1) \). Then, by Example 11.179 and Proposition 11.187, \( L_p(\mathcal{X}, \mathcal{Y}_1) \) is a separable Banach space.

(a) Since \( \mathcal{X}_2 \) is \( \sigma \)-compact, then \( \exists \) compact sets \( (K_n)_{n=1}^\infty \) such that \( \mathcal{X}_2 = \bigcup_{n=1}^\infty K_n \), \( \forall n \in \mathbb{N} \), \( \forall i \in \mathbb{N} \), \( K_n \subseteq \bigcup_{x \in K} \mathcal{B}_{\mathcal{X}_2}(x, 1/i) \). By the compactness of \( K_n \), \( \exists \) finite subset \( D_{n,i} \subseteq K_n \) such that \( K_n \subseteq \bigcup_{x \in D_{n,i}} \mathcal{B}_{\mathcal{X}_2}(x, 1/i) \).

Then, \( D := \bigcup_{n=1}^\infty \bigcup_{i=1}^\infty D_{n,i} \subseteq \mathcal{X}_2 \) is a countable dense set. Hence, \( \mathcal{X}_2 \) is separable.

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

**Definition 11.191** Let \( \mathcal{X} := (\mathcal{X}, \mathcal{O}) \) be a topological space, \( \mathcal{Y} \) be a normed linear space, and \( (\mathcal{X}, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued measure space on the same set \( \mathcal{X} \). The triple \( \mathcal{X} := (\mathcal{X}, \mathcal{B}, \mu) \) is said to be a \( \mathcal{Y} \)-valued topological measure space if \( \mathcal{B} = \mathcal{B}_\mathcal{B}(\mathcal{X}) \) and \( \mathcal{X} := (\mathcal{X}, \mathcal{B}_\mathcal{B}(\mathcal{X}), \mathcal{P} \circ \mu) \) is a topological measure space. We will say that \( \mathcal{X} \) is finite, \( \sigma \)-finite, or locally finite if \( \mathcal{X} \) is so. We will say that \( \mathcal{X} \) 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 \( \mathcal{X} \) is so.

Let \( \mathcal{X} := (\mathcal{X}, \rho) \) be a metric space with the natural topology \( \mathcal{O} \), \( \mathcal{Y} \) be a normed linear space, and \( (\mathcal{X}, \mathcal{B}_\mathcal{B}(\mathcal{X}), \mu) \) be a \( \mathcal{Y} \)-valued measure space on the same set \( \mathcal{X} \). The triple \( \mathcal{X} := (\mathcal{X}, \mathcal{B}_\mathcal{B}(\mathcal{X}), \mu) \) is said to be a \( \mathcal{Y} \)-valued metric measure space if \( ((\mathcal{X}, \mathcal{O}), \mathcal{B}_\mathcal{B}(\mathcal{X}), \mu) \) is a \( \mathcal{Y} \)-valued topological measure space. \( \mathcal{X} \) is said to be complete or totally bounded if \( \mathcal{X} \) is so.

Let \( \mathcal{X} := (\mathcal{X}, \mathfrak{K}, \| \cdot \|) \) be a normed linear space over the field \( \mathfrak{K} \), \( \mathcal{O} \) be the natural topology on \( \mathcal{X} \) generated by the norm \( \| \cdot \| \), \( \mathcal{Y} \) be a normed linear space, and \( (\mathcal{X}, \mathcal{B}_\mathcal{B}(\mathcal{X}), \mu) \) be a \( \mathcal{Y} \)-valued measure space on the same set \( \mathcal{X} \). The triple \( \mathcal{X} := (\mathcal{X}, \mathcal{B}_\mathcal{B}(\mathcal{X}), \mu) \) is said to be a \( \mathcal{Y} \)-valued normed linear measure.
space if $((X, \mathcal{O}), \mathcal{B}_B(X), \mu)$ is a $\mathcal{Y}$-valued topological measure space. When $X$ is a Banach space, then $\mathcal{X}$ is said to be a $\mathcal{Y}$-valued Banach measure space. Depending on whether $\mathcal{K} = \mathbb{R}$ or $\mathcal{K} = \mathbb{C}$, we will say that $X$ is a $\mathcal{Y}$-valued real or complex Banach measure space.

**Proposition 11.192** Let $X := (X, \mathcal{O})$ be a topological space, $\mathcal{Y}$ be a normed linear space over $\mathcal{K}$. Define $\mathcal{Z} := \{ \mu \in \mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y}) \mid (X, \mathcal{B}_B(X), \mu)$ is a finite $\mathcal{Y}$-valued topological measure space $\}$. Then, $\mathcal{Z}$ is a closed subspace of $\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$ and $(\mathcal{Z}, \mathcal{K}, \| \cdot \|_{\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})}) =: \mathcal{M}_{f1}(X, \mathcal{Y})$ is a normed linear space.

If, in addition, $\mathcal{Y}$ is a Banach space. Then, $\mathcal{M}_{f1}(X, \mathcal{Y})$ is a Banach space. Furthermore, define $\mathcal{Z} := \{ \mu \in \mathcal{M}_s(X, \mathcal{B}_B(X), \mathcal{Y}) \mid (X, \mathcal{B}_B(X), \mu)$ is a $\sigma$-finite $\mathcal{Y}$-valued topological measure space $\} \subseteq \mathcal{M}_s(X, \mathcal{B}_B(X), \mathcal{Y})$. Let $\mathcal{Z}$ admit the subset topology $\mathcal{O}_Z$. Then, $\mathcal{Z} := \mathcal{M}_{s1}(X, \mathcal{Y})$ is a subspace of $\mathcal{M}_s(X, \mathcal{B}_B(X), \mathcal{Y})$. We will abuse the notation and denote the topological space $(\mathcal{Z}, \mathcal{O}_Z)$ by $\mathcal{M}_{s1}(X, \mathcal{Y})$.

**Proof** Let $\mathcal{Y}$ be a normed linear space. We will show that $\mathcal{Z}$ is a subspace of $\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$, $\forall \alpha_1, \alpha_2 \in \mathcal{K}$, $\forall \mu_1, \mu_2 \in \mathcal{Z}$, by Proposition 11.136, $\mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \in \mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$, $\forall E \in \mathcal{B}_B(X)$, $\forall \epsilon \in (0, +\infty) \subseteq \mathcal{R}$, $\forall i = 1, 2$, by $\mu_i \in \mathcal{Z}$, $\exists O_i \in \mathcal{O}$ with $E \subseteq O_i$ such that $P \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|)$. Let $O := O_1 \cap O_2 \in \mathcal{O}$. Clearly, $E \subseteq O$ and $P \circ \mu(O \setminus E) \leq \alpha_i|P \circ \mu_i(O \setminus E) = \sum_{i=1}^{2} |\alpha_i|P \circ \mu_i(O_i \setminus E) < \epsilon$, where the first inequality and the first equality follow from Proposition 11.136. This shows that $(X, \mathcal{B}_B(X), \mu)$ is a topological measure space. Then, $(X, \mathcal{B}_B(X), \mu)$ is a finite $\mathcal{Y}$-valued topological measure space and $\mu \in \mathcal{Z}$. Clearly, $\exists \mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y}) \ni \mathcal{Z} \neq \emptyset$. Hence, $\mathcal{Z}$ is a subspace of $\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$. Then, $\mathcal{M}_f(X, \mathcal{Y})$ is a normed linear space since $\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$ is a normed linear space.

Next, we will show that $\mathcal{Z}$ is closed. $\forall \mu \in \mathcal{Z} \subseteq \mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$, by Proposition 4.13, $\exists (\mu_n)_{n \in \mathbb{N}} \subseteq \mathcal{Z}$ such that $\lim_{n \in \mathbb{N}} \mu_n = \mu$. Then, $\lim_{n \in \mathbb{N}} \| \mu_n - \mu \|_{\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})} = \lim_{n \in \mathbb{N}} P \circ (\mu_n - \mu)(X) = 0$. $\forall E \in \mathcal{B}_B(X)$, $\forall \epsilon \in (0, +\infty) \subseteq \mathcal{R}$, $\exists \mu \in \mathcal{N}$ such that $P \circ (\mu_n - \mu)(X) < \epsilon/2$. By $\mu_n \in \mathcal{Z}$, $\exists O \in \mathcal{O}$ with $E \subseteq O$ such that $P \circ \mu_n(O \setminus E) < \epsilon/2$. Then, $P \circ \mu(O \setminus E) = P \circ (\mu_n - \mu)(O \setminus E) \leq P \circ (\mu_n + \mu)(O \setminus E) = P \circ \mu(O \setminus E) + P \circ (\mu_n - \mu)(O \setminus E) < \epsilon/2 + P \circ (\mu_n - \mu)(X) < \epsilon$, where the first equality and the first inequality follow from Proposition 11.136. Hence, $\mu \in \mathcal{Z}$. By the arbitrariness of $\mu$, we have $\mathcal{Z} = \overline{\mathcal{Z}}$ and $\mathcal{Z}$ is closed.

Let $\mathcal{Y}$ be a Banach space. By Proposition 11.142, $\mathcal{M}_f(X, \mathcal{B}_B(X), \mathcal{Y})$ is a Banach space. By Proposition 4.39, $\mathcal{M}_f(X, \mathcal{Y})$ is a Banach space.

Finally, we will show that $\mathcal{Z}$ is a subspace of $\mathcal{M}_s(X, \mathcal{B}_B(X), \mathcal{Y})$, $\forall \alpha_1, \alpha_2 \in \mathcal{K}$, $\forall \mu_1, \mu_2 \in \mathcal{Z}$, by Proposition 11.138, $\mu := \alpha_1 \mu_1 + \alpha_2 \mu_2 \in \mathcal{M}_s(X, \mathcal{B}_B(X), \mathcal{Y})$, $\forall E \in \mathcal{B}_B(X)$, $\forall \epsilon \in (0, +\infty) \subseteq \mathcal{R}$, $\forall i = 1, 2$, by $\mu_i \in \mathcal{Z}$, $\exists O_i \in \mathcal{O}$ with $E \subseteq O_i$ such that $P \circ \mu_i(O_i \setminus E) < \epsilon/(1 + 2|\alpha_i|)$. Let $O := O_1 \cap O_2 \in \mathcal{O}$. Clearly, $E \subseteq O$ and $P \circ \mu(O \setminus E) \leq \sum_{i=1}^{2} |\alpha_i|P \circ \mu(O \setminus E) < \epsilon$. Hence, $\mu \in \mathcal{Z}$.
Let Proposition 11.193

\[ \mu_i)(O \setminus E) = \sum_{i=1}^{2} |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) \leq \sum_{i=1}^{2} |\alpha_i| \mathcal{P} \circ \mu_i(O \setminus E) < \epsilon, \]

where the first inequality and the first equality follow from Proposition 11.138. 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 \(\sigma\)-finite \(\mathcal{Y}\)-valued topological measure space and \(\mu \in \mathcal{Y}\). Clearly, \(\partial \mathcal{M}_\sigma(X, \mathcal{B}_B(X), \mathcal{Y}) \in \mathcal{Y}\). Hence, \(\mathcal{Z}\) is a subspace of \(\mathcal{M}_\sigma(X, \mathcal{B}_B(X), \mathcal{Y})\).

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

A bit of notation to simplify our presentation. Let \(\mathcal{M}_\sigma(X, \mathcal{B})\) denote the set of \(\sigma\)-finite measures on the measurable space \((X, \mathcal{B})\); \(\mathcal{M}_f(X, \mathcal{B})\) denote the set of finite measures on the measurable space \((X, \mathcal{B})\); \(\mathcal{M}_\sigma(X)\) denote the set of \(\sigma\)-finite topological measures on the topological space \(X\); and \(\mathcal{M}_f(X)\) denote the set of finite topological measures on the topological space \(X\).

**Proposition 11.193** Let \(X := (X, \mathcal{O})\) be a topological space and \(\mu_o : \mathcal{O} \to [0, \infty) \subset \mathbb{R}\). Assume that

(i) \(\mu_o(\emptyset) = 0;\)

(ii) \(\mu_o(O_1) \leq \mu_o(O_2), \forall O_1, O_2 \in \mathcal{O} \text{ with } O_1 \subseteq O_2;\)

(iii) \(\mu_o(\bigcup_{i=1}^{\infty} O_i) \leq \sum_{i=1}^{\infty} \mu_o(O_i), \forall (O_i)_{i=1}^{\infty} \subseteq \mathcal{O};\)

(iv) \(\mu_o(O_1 \cup O_2) = \mu_o(O_1) + \mu_o(O_2), \forall O_1, O_2 \in \mathcal{O} \text{ with } O_1 \cap O_2 = \emptyset;\)

(v) \(\mu_o(O) = \sup_{U \in \mathcal{O}, U \subseteq \mathcal{T} \subseteq \mu_o(U), \forall O \in \mathcal{O}.}\)

Define \(\tilde{\mu}_o : X \to [0, \infty) \subset \mathbb{R}\) by \(\tilde{\mu}_o(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O), \forall E \subseteq X.\) Then, the following statements hold.

1. \(\tilde{\mu}_o\) is an outer measure. It induces a finite complete measure space \((X, \mathcal{B}, \tilde{\mu})\), where \(\mathcal{B} := \{E \subseteq X \mid E\ \text{is measurable with respect to } \tilde{\mu}_o\}\)
   and \(\tilde{\mu} := \tilde{\mu}_o|_{\mathcal{B}}.\)

2. \(\mathcal{B}_B(X) \subseteq \mathcal{B}\) and the triple \(X := (X, \mathcal{B}_B(X), \mu := \tilde{\mu}|_{\mathcal{B}_B(X)})\) is a finite topological measure space with \(\mu(O) = \mu_o(O), \forall O \in \mathcal{O}.\)

3. The measure \(\mu\) is unique in the sense that if \(\tilde{\mu}\) be another measure on \((X, \mathcal{B}_B(X))\) satisfying \(\tilde{\mu}(O) = \mu_o(O) = \mu(O), \forall O \in \mathcal{O},\) then \(\tilde{\mu} = \mu.\)

**Proof**

1. By (i), we have \(\tilde{\mu}_o(\emptyset) = 0. \forall A \subseteq B \subseteq X,\) we have \(0 \leq \tilde{\mu}_o(A) = \inf_{O \in \mathcal{O}, A \subseteq O} \mu_o(O) \leq \inf_{O \in \mathcal{O}, B \subseteq O} \mu_o(O) = \tilde{\mu}_o(B) \leq \mu_o(X) < \infty.\)

   \(\forall E \subseteq \bigcup_{i=1}^{\infty} E_i \subseteq X, \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \forall i \in \mathbb{N}, \exists O_i \in \mathcal{O} \text{ with } E_i \subseteq O_i \text{ such that } \mu_o(O_i) < \tilde{\mu}_o(E_i) + 2^{-i} \epsilon. \text{ Then, } \tilde{\mu}_o(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_o(O) \leq \mu_o(\bigcup_{i=1}^{\infty} O_i) \leq \sum_{i=1}^{\infty} \tilde{\mu}_o(E_i) < \sum_{i=1}^{\infty} \tilde{\mu}_o(E_i) + \epsilon,\) where the second inequality follows from (iii). The arbitrariness of \(\epsilon,\) we have \(\tilde{\mu}_o(E) \leq \sum_{i=1}^{\infty} \tilde{\mu}_o(E_i).\)

   Hence, \(\tilde{\mu}_o : X \to [0, \infty) \subset \mathbb{R}\) is an outer measure. It is easy to see that \(\tilde{\mu}_o(O) = \mu_o(O), \forall O \in \mathcal{O}.\) By Theorem 11.17, \((X, \mathcal{B}, \tilde{\mu})\) is a finite complete measure space.
CHAPTER 11. GENERAL MEASURE AND INTEGRATION

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 \( \mu_{\sigma}(O_{1}) < \bar{\mu}_{\sigma}(E) + \epsilon/2 \). By (v), \( \exists U \in \mathcal{O} \) with \( U \subseteq \overline{E} \subseteq O_{1} \cap O \) such that \( \mu_{\sigma}(O \cap O_{1}) < \mu_{\sigma}(U) + \epsilon/2 \). Then, \( \bar{\mu}_{\sigma}(E) > \mu_{\sigma}(O_{1}) - \epsilon/2 \geq \mu_{\sigma}((O_{1} \setminus U) \cup U) - \epsilon/2 = \mu_{\sigma}(O_{1} \setminus U) + \mu_{\sigma}(U) - \epsilon/2 = \bar{\mu}_{\sigma}(O_{1} \setminus U) + \mu_{\sigma}(U) - \epsilon/2 > \bar{\mu}_{\sigma}(O_{1} \setminus (O_{1} \cap O)) + \mu_{\sigma}(O \cap O_{1}) - \epsilon = \bar{\mu}_{\sigma}(O_{1} \setminus O) + \bar{\mu}_{\sigma}(O \cap O_{1}) - \epsilon > \bar{\mu}_{\sigma}(E \cap O) + \bar{\mu}_{\sigma}(E \cap O') - \epsilon \geq \bar{\mu}_{\sigma}(E) - \epsilon \), where the second inequality follows from (ii); the first equality follows from (iv); and the third, the fourth, and the last inequality follow from the fact that \( \bar{\mu}_{\sigma} \) is an outer measure. By the arbitrariness of \( \epsilon \), we have \( \bar{\mu}_{\sigma}(E) = \bar{\mu}_{\sigma}(E \cap O) + \bar{\mu}_{\sigma}(E \cap O') \). By the arbitrariness of \( O \), we have \( \mathcal{O} \subseteq \mathcal{B} \). Since \( \mathcal{B} \) is a \( \sigma \)-algebra on \( X \), then \( \mathcal{B}(X) \subseteq \mathcal{B} \). By Proposition 11.13, \( (X, \mathcal{B}(X), \mu) \) is a measure space. Clearly, \( \mu(O) = \bar{\mu}(O) = \bar{\mu}_{\sigma}(O), \forall O \in \mathcal{O} \). This coupled with the definition of \( \bar{\mu}_{\sigma} \) leads to the conclusion that \( X \) is a topological measure space. Clearly, \( \mu(X) = \bar{\mu}_{\sigma}(X) < \infty \). Hence, \( X \) is a finite topological measure space.

3. Let \( \tilde{\mu} \) be another measure on \( (X, \mathcal{B}(X)) \) satisfying \( \hat{\mu}(O) = \bar{\mu}_{\sigma}(O) = \mu(O), \forall O \in \mathcal{O} \). Then, \( \tilde{\mu} \) is finite. Suppose \( \mu \neq \tilde{\mu} \). Then, \( \exists E \in \mathcal{B}(X) \) such that \( \hat{\mu}(E) \neq \bar{\mu}(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \mu_{\sigma}(O) = \inf_{O \in \mathcal{O}, E \subseteq O} \bar{\mu}(O) \). Since \( \hat{\mu}(E) < \inf \mu(E) \), \( \forall O \in \mathcal{O} \) with \( E \subseteq O \). Then, we must have \( \bar{\mu}(E) < \mu(E) \). Since \( X \) is a topological measure space, then, by Proposition 11.27, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \mu(E \setminus F) < (\mu(E) - \bar{\mu}(E))/2 \). This implies that \( \mu(F) = \mu(E) - \mu(E \setminus F) > \mu(E)/2 + \bar{\mu}(E)/2 > \bar{\mu}(E) \geq \hat{\mu}(F) \) and \( \mu_{\sigma}(X \setminus F) = \mu(X \setminus F) = \mu(X) - \mu(F) < \hat{\mu}(X) - \bar{\mu}(F) = \tilde{\mu}(X \setminus F) = \mu_{\sigma}(X \setminus F) \), where the first equality follows from \( X \setminus F \in \mathcal{O} \); the second equality follows from \( \mu \) being a finite measure; the inequality follows from \( X \in \mathcal{O} \); the third equality follows from \( \mu \) being a measure; and the last equality follows from \( X \setminus F \in \mathcal{O} \). This is a contradiction. Hence, we have \( \bar{\mu} = \mu \) and \( \mu \) is unique.

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

**Proposition 11.194** Let \( (X, O) \) be a compact Hausdorff topological space, \( \mathcal{Y} \) be a Banach space, \( \tilde{\mu} \) be a function that assigns a vector \( \tilde{\mu}(F) \in \mathcal{Y} \) for each closed subset \( F \subseteq X \), \( \bar{\mathcal{X}} := (X, \mathcal{B}(X), \nu) \) be a finite topological measure space. Assume that

(\( i \)) \( \tilde{\mu}(F_{1} \cup F_{2}) = \tilde{\mu}(F_{1}) + \tilde{\mu}(F_{2}), \forall X \setminus F_{1}, X \setminus F_{2} \in \mathcal{O} \) with \( F_{1} \cap F_{2} = \emptyset \);

(\( ii \)) \( \| \tilde{\mu}(F_{1}) - \tilde{\mu}(F_{2}) \| \leq \nu(O), \forall X \setminus F_{1}, X \setminus F_{2}, O \in \mathcal{O} \) with \( F_{1} \triangle F_{2} \subseteq O \).

Then, there exists a unique \( \mu \in \mathcal{M}_{f_{1}}(\bar{\mathcal{X}}, \mathcal{Y}) \) such that \( \mu(F) = \tilde{\mu}(F), \forall X \setminus F \in \mathcal{O} \). Furthermore, \( \mathcal{P} \circ \mu \leq \nu \).

**Proof** Fix any \( E \in \mathcal{B}(X) \). Let \( \tilde{\mathcal{A}}_{E} := \{ F \subseteq E \mid X \setminus F \in \mathcal{O} \} \) and \( \bar{\mathcal{A}}_{E} := (\tilde{\mathcal{A}}_{E}, \subseteq) \). Clearly, \( \bar{\mathcal{A}}_{E} \) is a directed system.

**Claim 11.194.1** The net \( (\tilde{\mu}(F))_{F \in \bar{\mathcal{A}}_{E}} \) is Cauchy.
Proof of claim: \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R} \), by \( \mathcal{X} \) being a topological measure space, \( \mathbf{F} \in \mathcal{O} \) with \( E \subseteq V \) such that \( \nu(V \setminus E) < \epsilon/2 \). By Proposition 11.27, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \nu(E \setminus F) < \epsilon/2 \). \( \forall F_1 \in \mathcal{A}_E \) with \( F \subseteq F_1 \), we have \( F \setminus F_1 \subseteq E \subseteq V \) and \( F \Delta F_1 = F_1 \setminus F \subseteq V \setminus F \in \mathcal{O} \). By (ii), \( \| \mu(F_1) - \mu(F) \| \leq \nu(V \setminus F) = \nu(V \setminus E) + \nu(E \setminus F) < \epsilon \). Hence, the net is Cauchy. This completes the proof of the claim. \( \square \)

By Proposition 4.44, we may define \( \mu(E) = \lim_{F \in \mathcal{A}_E} \mu(F) \in \mathcal{Y} \). Thus, we have defined a function \( \mu : \mathcal{B}_E \rightarrow \mathcal{Y} \). We will show that \( \mu \) is the \( \mathcal{Y} \)-valued measure we seek. Clearly, \( \mu(F) = \bar{\mu}(F), \forall X \setminus F \in \mathcal{O} \). By (i), \( \bar{\mu}(\emptyset) + \bar{\mu}(\emptyset) = \bar{\mu}(\emptyset) \) and then \( \mu(\emptyset) = \bar{\mu}(\emptyset) = \partial_y \).

\[ \forall E \in \mathcal{B}_E(X), \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \exists V \in \mathcal{O} \text{ with } E \subseteq V \text{ such that } \nu(V \setminus E) < \epsilon/2. \text{ By } \mu(E) = \lim_{F \in \mathcal{A}_E} \bar{\mu}(F), \exists F \in \mathcal{A}_E \text{ such that } \| \mu(E) - \bar{\mu}(F) \| < \epsilon/2. \] Clearly, \( F \subseteq V \subseteq V \). Then, \( \| \mu(E) \| < \| \bar{\mu}(F) \| + \epsilon/2 < \nu(V) + \epsilon/2 < \nu(E) + \epsilon \), where the second inequality follows from (ii). By the arbitrariness of \( \epsilon \), we have \( \| \mu(E) \| \leq \nu(V) \).

Fix any pairwise disjoint \( (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_E(X) \), let \( E := \bigcup_{n=1}^{\infty} E_n \subseteq \mathcal{B}_E(X) \). \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \exists V_n \in \mathcal{O} \) with \( E_n \subseteq V_n \) such that \( \nu(V_n \setminus E_n) < 2^{-n-1}\epsilon/5 \). By Proposition 11.27, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \nu(V_n \setminus F_n) < 2^{-n-1}\epsilon/5 \). By \( \mu(E_n) = \lim_{F \in \mathcal{A}_E} \bar{\mu}(F_n), \exists F_n \in \mathcal{A}_E \) with \( F_n \subseteq F \) such that \( \| \mu(E_n) - \bar{\mu}(F_n) \| < 2^{-n}\epsilon/5 \). Clearly, \( F_n \subseteq F_n \subseteq E_n \subseteq V_n \), \( \exists V \in \mathcal{O} \) with \( E \subseteq V \) such that \( \nu(V \setminus E) < \epsilon/5 \). By Proposition 11.27, \( \exists X \setminus F \in \mathcal{O} \) with \( F \subseteq E \) such that \( \nu(E \setminus F) < \epsilon/5 \). By \( \mu(E) = \lim_{F \in \mathcal{A}_E} \bar{\mu}(F), \exists F \in \mathcal{A}_E \) with \( F \subseteq F \) such that \( \| \mu(E) - \bar{\mu}(F) \| < \epsilon/5 \). Clearly, \( F \subseteq F \subseteq E \subseteq V \). By Proposition 5.5, \( F \) is compact. Since \( F \subseteq E = \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} V_n \), then \( \exists n_0 \in \mathbb{N} \) such that \( F \subseteq \bigcup_{i=1}^{n_0} V_i \).

\( \forall n \in \mathbb{N} \) with \( n_0 \leq n \), \( \| \mu(E) - \sum_{i=1}^{n} \mu(E_i) \| \leq \| \mu(E) - \bar{\mu}(F) \| + \| \bar{\mu}(F) - \sum_{i=1}^{n} \bar{\mu}(F_i) \| + \sum_{i=1}^{n} \| \bar{\mu}(F_i) - \mu(E_i) \| < \epsilon/5 + \| \bar{\mu}(F) - \mu(E) \| < \epsilon/5 + \| \bar{\mu}(F) - \mu(E) \| < \epsilon/5 + \| \bar{\mu}(F) - \mu(E) \| < \epsilon/5 \), where the second inequality follows from (i). Clearly, \( F \) and \( \bigcup_{i=1}^{n_0} F_i \) are closed sets. \( \hat{F} \setminus \left( \bigcup_{i=1}^{n_0} F_i \right) = (\hat{F} \setminus \left( \bigcup_{i=1}^{n_0} F_i \right)) \cup (\left( \bigcup_{i=1}^{n_0} F_i \right) \setminus \hat{F}) \subseteq ((\bigcup_{i=1}^{n_0} V_i) \setminus \left( \bigcup_{i=1}^{n_0} F_i \right)) \cup (V \setminus F) \subseteq (\bigcup_{i=1}^{n_0} V_i) \setminus (\bigcup_{i=1}^{n_0} F_i) \subseteq \mathcal{O} \). By (ii), \( \| \mu(E) - \sum_{i=1}^{n} \mu(E_i) \| < 2\epsilon/5 + \nu(\bigcup_{i=1}^{n} (V_i \setminus F_i)) \leq 2\epsilon/5 + \sum_{i=1}^{n} \nu(V_i \setminus F_i) + \nu(V \setminus F) = 2\epsilon/5 + \sum_{i=1}^{n} \nu(V_i \setminus E_i) + \nu(E_i \setminus F_i) + \nu(V \setminus E) + \nu(E \setminus F) < \epsilon \). Then, \( \mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \in \mathcal{Y} \).

\( \forall \) pairwise disjoint \( (E_n)_{n=1}^{\infty} \subseteq \mathcal{B}_E(X), \sum_{n=1}^{\infty} \| \mu(E_n) \| \leq \sum_{n=1}^{\infty} \nu(E_n) = \nu(\bigcup_{n=1}^{\infty} E_n) \leq \nu(X) < \infty \). This shows that \( \mu \) is a \( \mathcal{Y} \)-valued pre-measure on \( (X, \mathcal{B}_E(X)) \).

\( \forall E \in \mathcal{B}_E(X), \forall n \in \mathbb{Z}_+, \forall \) pairwise disjoint \( (E_i)_{i=1}^{n} \subseteq \mathcal{B}_E(X) \) with \( E = \bigcup_{i=1}^{n} E_i \), \( \sum_{i=1}^{n} \| \mu(E_i) \| \leq \sum_{i=1}^{n} \nu(E_i) = \nu(E) \). Hence, \( \mathcal{P} \circ \mu(E) \leq \nu(E) \). By the arbitrariness of \( E \), we have \( \mathcal{P} \circ \mu \leq \nu \). Hence, \( \mu \) is a finite \( \mathcal{Y} \)-valued measure on \( (X, \mathcal{B}_E(X)) \). It is easy to show that \( (X, \mathcal{B}_E(X), \mu) \) is a \( \mathcal{Y} \)-valued topological measure space. Then, \( \mu \in \mathcal{M}_f(X, \mathcal{Y}) \).
Finally, we need to show that $\mu \in \mathcal{M}_f(\mathcal{X}, \mathcal{Y})$ is unique. Let $\hat{\mu} \in \mathcal{M}_f(\mathcal{X}, \mathcal{Y})$ be such that: $\hat{\mu}(F) = \hat{\mu}(F) = \mu(F), \forall X \subseteq F \in \mathcal{O}$. By Proposition 11.27, $\exists X \subseteq F \in \mathcal{O}$ such that $\mathcal{P} \circ \hat{\mu}(E \setminus F) < \epsilon/2$. Again by Proposition 11.27, $\exists X \subseteq F \in \mathcal{O}$ with $F_1 \subseteq E$ such that $\mathcal{P} \circ \hat{\mu}(E \setminus F) < \epsilon/2$. Let $\hat{F} := F_1 \subseteq E$, which is clearly closed. Then, $\|\mu(E) - \hat{\mu}(E)\| = |\|\mu(E) - \hat{\mu}(F)\| + \|\hat{\mu}(F) - \hat{\mu}(E)\| = |\|\mu(E \setminus F)\| + 0 + \|\hat{\mu}(E \setminus \hat{F})\| \leq \mathcal{P} \circ \mu(E \setminus \hat{F}) + \mathcal{P} \circ \hat{\mu}(E \setminus \hat{F}) < \epsilon$. By the arbitrariness of $\epsilon$, we have $\mu(E) = \hat{\mu}(E)$. Hence, $\mu = \hat{\mu}$.

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

**Definition 11.195** Let $\mathcal{X} := (X, \mathcal{O})$ be a topological space. $E \subseteq X$ is said to be a Gδ set 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 Fσ set if exists $\left(\bar{F}_i\right)_{i=1}^{\infty} \subseteq \mathcal{O}$ such that $\bar{E} = \bigcup_{i=1}^{\infty} F_i$.

**Proposition 11.196** Let $\mathcal{X} := (X, \mathcal{O})$ be a locally compact separable metric space with the natural topology $\mathcal{O}$, $\mathcal{E} := \{ E \subseteq X \mid E \text{ is a compact } G_\delta \}$, and $\mathcal{B}_a(\mathcal{X})$ be the $\sigma$-algebra generated by $\mathcal{E}$. Then, $\mathcal{B}_a(\mathcal{X}) = \mathcal{B}_B(\mathcal{X})$.

**Proof** $\forall E \in \mathcal{E}$, by Proposition 5.5, $E$ is closed. Then, $E \in \mathcal{B}_B(\mathcal{X})$. Hence, $\mathcal{E} \subseteq \mathcal{B}_B(\mathcal{X})$. Since $\mathcal{B}_B(\mathcal{X})$ is a $\sigma$-algebra, then $\mathcal{B}_a(\mathcal{X}) \subseteq \mathcal{B}_B(\mathcal{X})$.

$\forall x \in \bar{E}$, by Definition 5.49, $\exists O_x \in \mathcal{O}$ such that $x \in O_x$ and $\overline{O_x}$ is compact. Then, $X = \bigcup_{x \in \bar{E}} O_x$. By Propositions 4.4 and 3.24, $\exists$ a countable set $D \subseteq X$ such that $X = \bigcup_{x \in D} O_x = \bigcup_{x \in D} \overline{O_x}$.

$\forall \bar{F} \in \mathcal{O}$, $\forall x \in D$, $F_x := F \cap O_x$ is compact and closed by Propositions 5.5 and 3.5. Then, $F_x = \bigcap_{i=1}^{\infty} \bigcup_{x \in F_x} B_x(x, 1/n)$, by Proposition 4.10. This implies that $F_x$ is a $G_\delta$ in addition to being compact. Then, $F_x \in \mathcal{E} \subseteq \mathcal{B}_a(\mathcal{X})$. This implies that $F = \bigcup_{x \in D} F_x \in \mathcal{B}_a(\mathcal{X})$. Hence, $\bar{F} \in \mathcal{B}_a(\mathcal{X})$. By the arbitrariness of $\bar{F}$, we have $\mathcal{O} \subseteq \mathcal{B}_a(\mathcal{X})$. By $\mathcal{B}_a(\mathcal{X})$ being a $\sigma$-algebra, we have $\mathcal{B}_B(\mathcal{X}) \subseteq \mathcal{B}_a(\mathcal{X})$.

Therefore, $\mathcal{B}_a(\mathcal{X}) = \mathcal{B}_B(\mathcal{X})$. This completes the proof of the proposition. \hfill $\Box$

**Lemma 11.197** Let $\mathcal{X} := (X, \mathcal{O})$ be a normal topological space. Then, $\forall E \in \mathcal{O}$ 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 $\phi_i|_{\bar{O}_i} = 1$ and $\phi_i|_{\partial O_i} = 0$. 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 $\phi(x) < 1$, $\forall x \in \bigcup_{i=1}^{\infty} \overline{O_i} = \overline{E}$. Hence, the result holds. This completes the proof of the lemma. \hfill $\Box$
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 $(X, \mathcal{B}_B(\mathcal{X}))$. Then, $\mathcal{X} := (\mathcal{X}, \mathcal{B}_B(\mathcal{X}), \mu)$ is a finite metric measure space. Thus, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X})) = \mathcal{M}_{f\mu}(\mathcal{X})$. As a consequence, $\mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mu) = \mathcal{M}_{f\mu}(\mathcal{X}, \mathcal{Y})$, where $\mathcal{Y}$ is any normed linear space.

Proof By Propositions 4.4, 3.24, and 5.72, $\mathcal{X}$ is $\sigma$-compact. Let $\mathcal{R} \subseteq \mathcal{B}_B(\mathcal{X})$ be such that $E \in \mathcal{R}$ if $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, we have
(i) $\exists \delta > 0$ with $\sigma$-compact and $E \subseteq \delta$ such that $\mu(O \cap E) < \epsilon$;
(ii) $\exists F \in \mathcal{R}$ with $F$ being a compact $G_{\delta}$ and $F \subseteq E$ such that $\mu(E \cap F) < \epsilon$.
We will show that $\mathcal{R} = \mathcal{B}_B(\mathcal{X})$.

Claim 11.198.1 $\forall F \in \mathcal{O}$, $F = \bigcap_{n=1}^{\infty} \bigcup_{x \in F} \mathcal{B}_\mathcal{X}(x, 1/n) =: \tilde{F}$ and is a $G_{\delta}$.

Proof of claim: Clearly, $F \subseteq \tilde{F}$. $\forall x_0 \in \tilde{F}$, by Proposition 4.10, $\epsilon_0 := \text{dist}(x_0, F) > 0$. Then, $\forall n \in \mathbb{N}$ with $1/n < \epsilon_0$, $x_0 \notin \bigcup_{x \in F} \mathcal{B}_\mathcal{X}(x, 1/n)$ and $x_0 \in \tilde{F}$. Hence, $\tilde{F} \subseteq F$ and $F = \tilde{F}$. Clearly, $F$ is a $G_{\delta}$. This completes the proof of the claim. □

Claim 11.198.2 $\forall \mathcal{O} \in \mathcal{O}$, $\mathcal{O}$ is $\sigma$-compact.

Proof of claim: By Claim 11.198.1, $\tilde{O}$ is a $G_{\delta}$. Then, $\exists (O_i)_{i=1}^{\infty} \subseteq \mathcal{O}$ such that $\tilde{O} = \bigcap_{i=1}^{\infty} O_i$ and $O = \bigcup_{i=1}^{\infty} O_i$. By $\mathcal{X}$ being $\sigma$-compact, $\exists$ compact sets $(K_i)_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} K_i$. Then, $\tilde{O} = O \cap X = \bigcup_{i=1}^{\infty} K_i = \bigcup_{i=1}^{\infty} (\tilde{O} \cap K_i)$. $\forall i, j \in \mathbb{N}$, by Propositions 3.5 and 5.5, $\tilde{O} \cap K_j$ is compact. Hence, $\tilde{O}$ is $\sigma$-compact.

$\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{R}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}_B(\mathcal{X})$. $\forall \epsilon \in (0, +\infty) \subset \mathbb{R}$, $\forall i \in \mathbb{N}$, $\exists O_i \in \mathcal{O}$ with $O_i$ being $\sigma$-compact and $E_i \subseteq O_i$ such that $\mu(O_i \cap E_i) < 2^{-i-1}\epsilon$. $\exists \tilde{F} \subseteq O$ with $F_i$ being a compact $G_{\delta}$ and $F_i \subseteq E_i$ such that $\mu(E_i \cap F_i) < 2^{-i-1}\epsilon$. $\forall \epsilon > 0$ and $\mu(O \cap E) = \mu((\bigcup_{i=1}^{\infty} O_i) \setminus (\bigcup_{i=1}^{\infty} E_i)) \leq \mu(\bigcup_{i=1}^{\infty} (O_i \setminus E_i)) \leq \sum_{i=1}^{\infty} \mu(O_i \setminus E_i) < \epsilon$. Hence, $E$ satisfies (i). By Proposition 11.7, $\mu(E) = \lim_{n \to \infty} \mu(\bigcup_{i=1}^{n} E_i)$. Then, $\exists n_0 \in \mathbb{N}$ such that $\mu(E) < \mu(\bigcup_{i=1}^{n_0} E_i) + \epsilon/2$. Let $F := \bigcup_{i=1}^{n_0} F_i$. Clearly, $F$ is compact and closed. By Claim 11.198.1, $F$ is a $G_{\delta}$. Then, $F \subseteq E$ and $\mu(E \cap F) = (\mu(E) - \mu(F)) < \mu(\bigcup_{i=1}^{n_0} E_i) - \mu(F) < \mu(\bigcup_{i=1}^{n_0} E_i) - \mu(F) < \epsilon/2$. $\forall \epsilon > 0$ and $\mu(O \cap E) = \mu((\bigcup_{i=1}^{\infty} O_i) \setminus (\bigcup_{i=1}^{\infty} E_i)) \leq \mu((\bigcup_{i=1}^{\infty} O_i \setminus E_i) \setminus \bigcup_{i=1}^{\infty} (O_i \setminus E_i)) \leq \sum_{i=1}^{\infty} \mu(O_i \setminus E_i) \leq \epsilon/2$. 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 $\mu(O \cap E) < \epsilon$. $\exists \tilde{F} \subseteq O$ with $F$ being a compact $G_{\delta}$ and $F \subseteq E$ such that $\mu(E \cap F) < \epsilon$. Then, $\tilde{O}$ is a closed set, $\tilde{O} \subseteq \tilde{F}$, and $\mu(\tilde{F} \setminus \tilde{O}) = \mu(O \setminus E) < \epsilon$. 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 $\epsilon > 0$ and $\mu(\tilde{F} \setminus \tilde{O}) = \mu(\tilde{F} \setminus (\tilde{O} \cap \bigcup_{i=1}^{\infty} K_i)) = \mu(\tilde{F} \setminus (\bigcup_{i=1}^{\infty} (\tilde{O} \cap K_i))) = \mu(\bigcap_{i=1}^{\infty} (\tilde{E} \cap (\tilde{O} \cap K_i))) = \mu(\cap_{i=1}^{\infty} (\tilde{E} \cap (\tilde{O} \cap K_i))) = \lim_{i \to \infty} \mu(\tilde{E} \setminus (\tilde{O} \cap K_i))$.
where the last equality follows from Proposition 11.5. Then, \( \exists n \in \mathbb{N} \) such that \( \mu(E \setminus (\overline{O} \cap K_n)) < \epsilon \). 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, \( \overline{O} \cap K_n \subseteq E \).

Hence, \( E \) satisfies (ii). \( F \in \mathcal{O} \) and is \( \sigma \)-compact by Claim 11.198.2. Note that \( E \subseteq F \) and \( \mu(F \setminus E) = \mu(E \setminus F) < \epsilon \). Then, \( E \) satisfies (i). Hence, \( \tilde{E} \in \mathcal{R} \). \( \tilde{E} \) is closed under set complements. Clearly, \( \emptyset \in \mathcal{R} \). This proves that \( \mathcal{R} \) is a \( \sigma \)-algebra.

For every \( \forall E \in \mathcal{O} \) with \( E \) being a compact \( G_\delta \), \( \forall \epsilon \in (0, +\infty) \subset \mathbb{R} \), then \( E \) satisfies (ii) trivially. By Lemma 11.197, there exists a continuous function \( \phi : X \to [0, 1] \subset \mathbb{R} \) 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, \( \mu(O_n \setminus E) = \mu(O_n) - \mu(E) < \epsilon \). 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 \( \mathcal{E} := \{ E \in \mathcal{B} X \mid E \) is a compact \( G_\delta \} \subseteq \mathcal{R} \). By Proposition 11.196, \( \mathcal{R} = \mathcal{B} X \). Then, \( X \) is a finite metric measure space.

Thus, we have \( M_f(X, \mathcal{B}(X)) \subseteq M_{f_1}(X) \). Clearly, \( M_f(X, \mathcal{B}(X)) \supseteq M_{f_1}(X) \). Hence, \( M_f(X, \mathcal{B}(X)) = M_{f_1}(X) \).

For every \( \forall \tilde{\mu} \in M_f(X, \mathcal{B}(X), \gamma) \) then \( \mathcal{P} \circ \tilde{\mu} \in M_f(X, \mathcal{B}(X)) = M_{f_1}(X) \). This implies that \( \tilde{\mu} \in M_f(X, \gamma) \). Hence, \( M_f(X, \mathcal{B}(X), \gamma) \subseteq M_{f_1}(X, \gamma) \). Clearly, \( M_f(X, \mathcal{B}(X), \gamma) \supseteq M_{f_1}(X, \gamma) \). Then, \( M_f(X, \mathcal{B}(X), \gamma) = M_{f_1}(X, \gamma) \).

This completes the proof of the theorem.

\[ \Box \]

**Definition 11.199** Let \( X := (X, \mathcal{O}) \) be a topological space, \( \gamma \) be a normed linear space, \( f : X \to \gamma \). The support of \( f \) is the set \( \text{supp}(f) := \{ x \in X \mid f(x) \neq \gamma \} \).

**Lemma 11.200** Let \( X := (X, \mathcal{O}) \) be compact Hausdorff topological space, \( \mathcal{Z} := C(X, \mathbb{R}) \), and \( f \in \mathcal{Z}^* \). \( f \in \mathcal{Z}^* \) is said to be a positive linear functional if \( f(z) \geq 0, \forall z \in P := \{ h \in \mathcal{Z} \mid h : X \to [0, \infty) \subset \mathbb{R} \} \). Then, \( f = f_+ - f_- \), where \( f_+, f_- \in \mathcal{Z}^* \) are positive linear functionals.

**Proof** For every \( \forall z \in P \), define \( f_+(z) := \sup_{\phi \in \mathcal{Z}, 0 \leq \phi(x) \leq z(x), \forall x \in X} f(\phi) \). Then, \( 0 \leq f_+(z) \leq \| f \| \| z \| < \infty \) and \( f_+(z) \geq f(z), \forall z \in P \). Clearly,

(i) \( f_+(\alpha z) = \alpha f_+(z), \forall z \in P \) and \( \forall \alpha \in [0, \infty) \subset \mathbb{R} \).

(ii) \( f_+(z_1 + z_2) \leq f_+(z_1) + f_+(z_2) \). On the other hand, \( \forall z \in \mathcal{Z} \) with \( 0 \leq \phi(x) \leq z_1(x), \forall x \in X \), we have \( f(\phi_1) + f(\phi_2) = f(\phi_1 + \phi_2) \leq f_+(z_1 + z_2) \). Then, \( f_+(z_1) + f_+(z_2) \leq f_+(z_1 + z_2) \). Therefore,

(ii) \( f_+(z_1) + f_+(z_2) = f_+(z_1 + z_2), \forall z_1, z_2 \in P \).
\( \forall z \in \mathbb{Z}, \) define \( z_+ := z \vee 0 \in P \) and \( z_- := (-z) \vee 0 \in P \). Clearly, \( z = z_+ - z_- \). Define \( f_+(z) := f_+(z_+) - f_+(z_-) \in \mathbb{R} \). Then, \( f_+ : \mathbb{Z} \to \mathbb{R} \) is well-defined. We will show that \( f_+ \in \mathbb{Z}^* \).

\[ \forall z \in \mathbb{Z}, \quad |f_+(z)| = |f_+(z_+) - f_+(z_-)| \leq f_+(z_+) + f_+(z_-) = f_+(z_+ + z_-) \leq \|f\| \|z_+ + z_-\| = \|f\| \|z\|. \]

\[ \forall z_1, z_2 \in \mathbb{Z}, \text{ let } z := z_1 + z_2. \text{ Then, } f_+(z_1) + f_+(z_2) = f_+(z_1+z_2) - f_+(z_1) - f_+(z_2) = f_+(z_1) + f_+(z_2) - f_+(z_1) - f_+(z_2) = f_+(z_1) + \alpha z_1 \text{ and } f_+(z_2) - f_+(z_1) + f_+(z_2) - f_+(z_2) = f_+(z) + f_+(z_1 + z_2) - f_+(z_1 + z_2) - f_+(z_1 - z_2) = f_+(z_1) + f_+(z_2) - f_+(z_1 - z_2), \]

where the first equality follows from the definition of \( f_+ \); the second equality follows from (ii); the third equality follows from (ii); and the last equality follows from the definition of \( f_+ \).

Note that, \( \forall x \in X, \)

\[
(z_1 + z_2 - z)(x) = z_1(x) \vee 0 + z_2(x) \vee 0 - (z_1(x) + z_2(x)) \vee 0
\]

\[
= \begin{cases} 
0 & z_1(x) \geq 0 \text{ and } z_2(x) \geq 0 \\
- z_2(x) & z_1(x) \geq 0 > z_2(x) \geq - z_1(x) \\
- z_1(x) & z_2(x) \geq 0 > z_1(x) \geq - z_2(x) \\
z_2(x) & z_2(x) \geq 0 > z_2(x) \geq - z_1(x) \\
0 & z_1(x) < 0 \text{ and } z_2(x) < 0
\end{cases}
\]

Hence, (iii) \( f_+(z_1) + f_+(z_2) = f_+(z) = f_+(z_1 + z_2), \forall z_1, z_2 \in \mathbb{Z}. \)

\[ \forall z \in \mathbb{Z}, \forall \alpha \in \mathbb{R}. \text{ If } \alpha = 0, \text{ then } f_+(\alpha z) = 0 = \alpha f_+(z). \text{ If } \alpha > 0, \text{ then } f_+(\alpha z) = f_+(\alpha z_+) - f_+(\alpha z_-) = \alpha f_+(z_+) - \alpha f_+(z_-) = \alpha f_+(z), \text{ where the first equality follows from the definition of } f_+; \text{ the second equality follows from (i); and the third equality follows from the definition of } f_+. \text{ If } \alpha < 0, \text{ then } f_+(\alpha z) = f_+(-\alpha z) = - f_+(-\alpha z_+) = - \alpha f_+(z_-) = \alpha f_+(z), \text{ where the first equality follows from the definition of } f_+; \text{ the second equality follows from (i); and the third equality follows from the definition of } f_+. \]

Hence, we have \( \forall z \in \mathbb{Z}: f_+(\alpha z) = \alpha f_+(z), \forall z \in \mathbb{Z} \text{ and } \forall \alpha \in \mathbb{R}. \)

Hence, \( f_+ \in \mathbb{Z}^* \) and is a positive linear functional.

Let \( f_- := f_+ - f \in \mathbb{Z}^*. \) 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 \( \mathcal{X} := (X, \mathcal{O}) \) be a compact Hausdorff topological space, \( \mathcal{Y} \) be a normed linear space over \( \mathbb{K} \), \( \mathcal{Z} := \mathcal{C}(\mathcal{X}, \mathcal{Y}) \), and \( Z := \{ z \in \mathbb{Z} \mid z = hy, h \in \mathcal{C}(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y} \} \). Assume that span \( \mathcal{Z} \rangle = \mathbb{Z}. \) Then, \( \forall f \in \mathbb{Z}^*, \exists ! \mu \in \mathcal{M}_f(\mathcal{X}, \mathcal{Y}^*) \) such that \( f(z) = \int_X \langle \mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle, \forall z \in \mathbb{Z}. \) Furthermore, \( \mathbb{Z}^* = \mathcal{M}_f(\mathcal{X}, \mathcal{Y}^*) \) isometrically isomorphically.
Proof. By Proposition 5.14, \( X \) is normal. By Example 7.31, \( Z \) is a normed linear space over \( K \). Fix any \( f \in Z^* \). Define \( \nu_0 : O \to [0, \| f \|] \subset \mathbb{R} \) by \( \nu_0(O) := \sup_{z \in Z, \| z \| \leq 1} \supp(z) \| f(z) \|, \forall O \in \mathcal{O} \). 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 \( O_1 \subseteq O_2 \).

\( \forall (O_i)_{i=1}^\infty \subseteq \mathcal{O} \), let \( O := \bigcup_{i=1}^\infty O_i \in \mathcal{O} \). \( \forall z \in Z \) with \( K := \supp(z) \subseteq O \) 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, \( \exists (\varphi_n)_n \subseteq \mathcal{C}(X, \mathbb{R}) \) such that \( \varphi_n : X \to [0, 1] \subset \mathbb{R} \), \( \supp(\varphi_n) \subseteq O_n, i = 1, \ldots, m, \sum_{i=1}^m \varphi_n(x) = 1, \forall x \in K, \) and \( \| z \varphi_n \| \leq 1, \forall i \in \{1, \ldots, m\}, z \varphi_n \in Z \). Then, \( z = \sum_{i=1}^m z_i \varphi_n, \forall i \in \{1, \ldots, m\}, z \varphi_n \in Z \). By Proposition 11.193, there exists a finite topological measure space \( \bar{X} := (X, \mathcal{B}(X), \nu) \) such that \( \nu(O) = \nu_0(O), \forall O \in \mathcal{O} \) and \( \nu(E) = \inf_{O \in \mathcal{O}, E \subseteq O} \nu_0(O) \), \( \forall E \in \mathcal{B}(X) \). Furthermore, \( \nu \) is unique among measures \( \tilde{\nu} \) on \( (X, \mathcal{B}(X)) \) with \( \tilde{\nu}_0 = \nu_0 \). Clearly, \( \nu \in \mathcal{M}_f(X) \).

Fix any \( \bar{F} \in \mathcal{O} \). Let \( \mathcal{A}_\rho := \{(U, V, h) \in \mathcal{O} \times \mathcal{O} \times \mathcal{C}(X, \mathbb{R}) \mid F \subseteq U \subseteq \overline{U} \subseteq V, h : X \to [0, 1] \subset \mathbb{R} \} \). Define a relation \( \prec \) on \( \mathcal{A}_\rho \) 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, \( \mathcal{A}_\rho := (\mathcal{A}_\rho, \prec) \) is a directed system. \( \forall (U, V, h) \in \mathcal{A}_\rho \), the function \( f_h : Y \to X \), defined by \( f_h(y) = f(hy), \forall y \in Y \), is a bounded linear functional since \( f \in Z^* \). Then, \( f_h \in \mathcal{Y}^* \). This defines a net \( (f_h)_{(U, V, h) \in \mathcal{A}_\rho} \subseteq \mathcal{Y}^* \).
11.10. DUAL OF $C(\mathcal{X}, \mathcal{Y})$

Claim 11.201.1 The net $(f_h)_{(U,V,h) \in \mathcal{A}_F} \subseteq \mathcal{Y}^*$ is Cauchy.

Proof of claim: \( \forall \epsilon \in (0, +\infty) \subseteq \mathbb{R} \), \( \exists V \in \mathcal{O} \) with \( F \subseteq V \) such that \( \nu(V \setminus F) < \epsilon \). By Proposition 3.35, \( \exists U, \bar{U} \in \mathcal{O} \) such that \( F \subseteq U \subseteq \bar{U} \subseteq \bar{U} \subseteq V \). By Urysohn's Lemma 3.55, \( \exists h \in C(\mathcal{X}, \mathbb{R}) \) with \( h : \mathcal{X} \to [0,1] \subseteq \mathbb{R} \) such that \( h|_U = 1 \) and \( h|_{\bar{U}} = 0 \). Then, \( \text{supp}(h) \subseteq \bar{U} \subseteq V \) and \( (U, V, h) \in \mathcal{A}_F \). \( \forall (U_1, V_1, h_1) \in \mathcal{A}_F \) with \( (U, V, h) \prec (U_1, V_1, h_1) \), we have \( V_1 \subseteq V \) and \( \| f_h - f_{h_1} \| = \sup_{y \in V, \| y \| \leq 1} |\langle f_h, y \rangle - \langle f_{h_1}, y \rangle | = \sup_{y \in V, \| y \| \leq 1} |f(hy) - f(h_1 y)| = \sup_{y \in V, \| y \| \leq 1} |f((h - h_1)y)| \). Note that, \( \forall y \in \mathcal{Y} \) with \( \| y \| \leq 1 \), \( (h - h_1)y \in \mathbb{Z} \), \( \| (h - h_1)y \| \leq 1 \), and

\[
\text{supp}(h - h_1) \subseteq \text{supp}(h - h_1)
\]

\[
\subseteq (\text{supp}(h) \cup \text{supp}(h_1)) \setminus \{ x \in X \mid h(x) = 1 = h_1(x) \}
\]

\[
\subseteq (\text{supp}(h) \cup \text{supp}(h_1)) \setminus (U \cap U_1)
\]

\[
= \text{supp}(h) \setminus (U \cap U_1) \cup \text{supp}(h_1) \setminus (U \cap U_1)
\]

\[
\subseteq \text{supp}(h) \setminus (U \cap U_1) \cup \text{supp}(h_1) \setminus (U \cap U_1)
\]

\[
= (V \setminus F) \cup (V \setminus F) \subseteq V \setminus F \subseteq \mathcal{O}
\]

where the third containment follows from the fact that \( h|_U = 1 \) and \( h|_{\bar{U}} = 1 \); 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.

By Propositions 4.44 and 7.72, (vii) \( \lim_{(U,V,h) \in \mathcal{A}_F} f_h =: \hat{\mu}(F) \in \mathcal{Y}^* \), \( \forall F \in \mathcal{O} \).

Thus, we have defined a function \( \hat{\mu} \) that assigns a vector \( \hat{\mu}(F) \in \mathcal{Y}^* \) to each closed subset \( F \subseteq \mathcal{X} \). We will next show that \( \hat{\mu} \) 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 \( \epsilon \in (0, \infty) \subseteq \mathbb{R} \). 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 \( \nu(V_i \setminus F_i) < \epsilon/4 \), \( i = 1, 2 \). Fix any \( i \in \{ 1, 2 \} \). By \( \lim_{(U,V,h) \in \mathcal{A}_{F_i}} f_h = \hat{\mu}(F_i) \), \( \exists (U_i, V_i, \hat{h}_i) \in \mathcal{A}_{F_i} \) with \( \hat{h}_i \subseteq V_i \) such that \( \| \hat{\mu}(F_i) - f_{\hat{h}_i} \| < \epsilon/6 \). By \( \lim_{(U,V,h) \in \mathcal{A}_{F_1 \cup F_2}} f_h = \hat{\mu}(F_1 \cup F_2) \), \( \exists (U, V, \hat{h}) \in \mathcal{A}_{F_1 \cup F_2} \) with \( \hat{h} \subseteq V_1 \cup V_2 \) such that \( \| \hat{\mu}(F_1 \cup F_2) - f_{\hat{h}} \| < \epsilon/6 \). This implies that \( \| \hat{\mu}(F_1 \cup F_2) - \hat{\mu}(F_1) - \hat{\mu}(F_2) \| < \epsilon/2 + \sup_{y \in \mathcal{Y}, \| y \| \leq 1} |\langle f_{\hat{h}}, y \rangle - \langle f_{\hat{h}_1}, y \rangle - \langle f_{\hat{h}_2}, y \rangle | = \epsilon/2 + \sup_{y \in \mathcal{Y}, \| y \| \leq 1} |f(hy) - f(h_1 y) - f(h_2 y)| = \epsilon/2 + \sup_{y \in \mathcal{Y}, \| y \| \leq 1} |f((\hat{h} - \hat{h}_1 - \hat{h}_2)y)| \). Note that, \( \forall y \in \mathcal{Y} \) with \( \| y \| \leq 1 \), \( \hat{z} := (\hat{h} - \hat{h}_1 - \hat{h}_2)y \in \mathbb{Z} \), \( \| \hat{z} \| \leq 1 \), since \( \text{supp}(\hat{h}_1) \cap \text{supp}(\hat{h}_2) \subseteq \hat{V}_1 \cap \hat{V}_2 = \emptyset \), and \( \text{supp}(\hat{z}) \subseteq \hat{V}_1 \cap \hat{V}_2 = \emptyset \).
We have
\[\bigcup \in \{ \bigcup \in \mathbb{R} \} = \bigcup \in \mathbb{R} \}

This leads to \(\| A \| \leq \epsilon/4\). By the arbitrarity of \(\epsilon\), we have
(viii) \(\tilde{\mu}(F_1) + \tilde{\mu}(F_2) = \tilde{\mu}(F_1 \cup F_2), \forall F_1, F_2 \in O \) with \(F_1 \cap F_2 = \emptyset\).

Fix any \(\tilde{F}_1, \tilde{F}_2, O \in O \) with \(F_1 \triangle F_2 \subseteq O\). \(\forall \epsilon \in (0, +\infty) \subseteq \mathbb{R}\), \(\forall i \in \{1, 2\}\), \(\exists V_i \in O \) with \(F_i \subseteq V_i \) such that \(\nu(V_i \setminus F_i) < \epsilon/4\). By \(\lim_{l(U, V, h) \in \mathcal{A}_F} f_h = \tilde{\mu}(F_1), \exists \tilde{U}_i, \tilde{V}_i, \tilde{h}_i \in \mathcal{A}_F\), such that \(\| \tilde{\mu}(F_1) - f_h \| < \epsilon/4\). This leads to \(\| \tilde{\mu}(F_1) - f_{h_1} \| + \| f_{h_1} - f_{h_2} \| + \| f_{h_2} - \tilde{\mu}(F_2) \| < \epsilon/2 + \sup_{y \in \mathcal{Y}, \| y \| \leq 1} \| \tilde{f}(h_1 - h_2) y \|\). Note that, \(\forall y \in \mathcal{Y}\) with \(\| y \| \leq 1\), \(\| h_1 - h_2 \| \leq 1, \) and \(\tilde{\mu}(\tilde{F}_1) \leq \tilde{\mu}(\tilde{F}_2) \leq \tilde{\mu}(O), \forall \tilde{F}_1, \tilde{F}_2, O \in O \) with \(F_1 \triangle \tilde{F}_2 \subseteq O\).

By Proposition 11.194, \(\exists \mu \in \mathcal{M}_f(\mathcal{X}, \mathcal{Y}^*)\) such that \(\mu(F) = \tilde{\mu}(F), \forall \tilde{F} \in O\). Furthermore, \(\mathcal{P} \circ \mu \leq \nu\). Then, \(\| \mu \| = \mathcal{P} \circ \mu(\mathcal{X}) \leq \nu(\mathcal{X}) \leq \| f \| \leq \infty\). This defines a mapping \(\Phi_\mathcal{Z}: \mathcal{Z}^* \to \mathcal{M}_f(\mathcal{X}, \mathcal{Y}^*)\) by \(\Phi_\mathcal{Z}(f) = \mu, \forall f \in \mathcal{Z}^*\). The above can be expressed as
(x) \(\| \Phi_\mathcal{Z}(f) \| \leq \| f \|, \forall f \in \mathcal{Z}^*\).
11.10. DUAL OF $\mathcal{C}(X, Y)$

$\|z\| P \circ \mu(X) = \|z\| \mu < \infty$. Hence, $z$ is absolutely integrable over $X := (X, B_0(X), \mu)$. By Proposition 5.7, $z(X) \subseteq Y$ is compact. Then, by Propositions 11.190, 7.35, and 7.17, $H := \text{span} \{z(X)\} \subseteq Y \subseteq Y^\ast$ is a separable subspace. By Proposition 11.132, $(\langle \mu, z \rangle) = \int_X (\langle d\mu(x), z(x) \rangle) \in K$ and $|\langle \mu, z \rangle| = \int_X (\|d\mu(x), z(x)\|) \leq \|z\| \|\mu\| < \infty$. Hence, we have (xi) $\forall \mu \in \mathcal{M}_f(X, Y^\ast)$, $\forall z \in Z$, we have $z$ is absolutely integrable and integrable over $(X, B_0(X), \mu)$, and $|\langle \mu, z \rangle| \leq \|z\| \|\mu\| < \infty$.

We will show that $f(z) = \int_X (\langle d\mu(x), z(x) \rangle) = \langle \mu, z \rangle$, $\forall z \in Z$, $\forall f \in Z^\ast$, where $\mu = \Phi(z)(f)$, in four steps. In the first step, we consider the special case: $Y = R$ and $f$ is a positive linear functional. $\forall F \in O$, $\forall (U, V, h) \in \tilde{A}_F$, $f_h = f(h) \subseteq [0, \|f\|] \subseteq R$, since $h \in P := \{z \in \mathcal{C}(X, R) \mid z : X \to [0, \infty) \subseteq R\} \subseteq \mathcal{C}(X, R) = Z$. Then, $\tilde{\mu}(F) = \lim_{(U, V, h) \in \tilde{A}_F} f_h \subseteq [0, \|f\|] \subseteq R$. Hence, by Proposition 11.194 and its proof, $\mu : B_0(X) \to [0, \|f\|] \subseteq R$ and $\mu \in \mathcal{M}_f(X)$. Fix any $z \in Z$. We will distinguish three exhaustive cases: Case 1: $z = \vartheta_z$; Case 2: $z \neq \vartheta_z$ and $z \in P$; Case 3: $z \in Z$. Case 1: $z = \vartheta_z$. Then, $f(z) = 0 = \int_X z(x) d\mu(x) = \langle \mu, z \rangle$.

Case 2: $z \neq \vartheta_z$ and $z \in P$. Let $\bar{z} := z/\|z\| \in P$. Then, $\bar{z} : X \to [0, 1] \subseteq R$. Fix any $n \in N$ and fix any $i \in \{1, \ldots, n\}$. Define $\phi_i := \{(n\bar{z} - i + 1) \vee 0 \wedge 1 \subseteq P$, $O_i := \{x \in X \mid n\bar{z}(x) > i - 1\}$, and $F_i := \{x \in X \mid n\bar{z}(x) \geq i - 1\}$. Let $O_0 := \{x \in X \mid n\bar{z}(x) > 1\} = X$, $O_{i+1} := \{x \in X \mid n\bar{z}(x) > n\} = \emptyset$, $F_{i+1} := \{x \in X \mid n\bar{z}(x) \geq n\}$, and $F_{i+2} := \{x \in X \mid n\bar{z}(x) \geq n + 1\} = \emptyset$. Clearly, $F_{i+2} \subseteq O_{i+1} \subseteq O_i \subseteq O_i \subseteq F_i \subseteq F_{i+1} \subseteq O_{i-1}$, $\phi_i : X \to [0, 1] \subseteq R$, $\phi_i|_{F_{i+1}} = 1$, and supp$(\phi_i) \subseteq O_{i-1} \subseteq F_i \subseteq O_i \subseteq O_i$. Clearly $O_i \in O$ and $F_{i+1} \in O$, $i = 0, \ldots, n + 1$, and $\bar{z} = \frac{1}{n} \sum_{i=1}^n \phi_i$. Fix any $i \in \{1, \ldots, n\}$. By the definition of $\tilde{\mu}$ and the fact that $f$ is a positive linear functional, $\forall (U, V, h) \in \tilde{A}_{F_i+}$, with $V \subseteq O_{i+1}$, we have $h(x) \leq \chi_{O_{i+1}, X}(x) \leq \chi_{F_{i+1}, X}(x) \leq \phi(x)$, $\forall x \in X$, and $f_h = f(h) \leq f(\phi_i)$. This implies that $\mu(F_{i+2}) \leq f(\phi_i)$. $\forall (U, V, h) \in \tilde{A}_F$, we have $\phi(x) \leq \chi_{F_i, X}(x) \leq h(x)$, $\forall x \in X$, and $f_h = f(h) \leq f(\phi_i)$. This leads to $\langle \mu(z) \rangle \leq \mu(F_i)$. By Propositions 11.75 and 11.83, we have $\mu(F_{i+2}) \leq \mu(F_{i+1}) = \int_X \chi_{F_{i+1}, X} d\mu \leq \int_X \phi_i d\mu \leq \int_X \chi_{F_i, X} d\mu \leq \mu(F_i)$. Then, $f(\phi_i) - \int_X \phi_i d\mu \leq \mu(F_i) - \mu(F_{i+2})$. Summing both side from 1 to $n$, we have $f(\bar{z}) = \int_X \bar{z} d\mu = \frac{1}{n} \sum_{i=1}^n (f(\phi_i) - \int_X \phi_i d\mu) \leq \frac{1}{n} \sum_{i=1}^n |f(\phi_i) - \int_X \phi_i d\mu| \leq \frac{1}{n} \sum_{i=1}^n |f(\phi_i) + \mu(F_i) - 2\mu(X)/n = 2\|\mu\|/n \leq 2\|f\|/n$, where the first equality 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, $f(z) = f(\|z\| \bar{z}) = \|z\| f(\bar{z}) = \|z\| \int_X \bar{z} d\mu = \int_X z d\mu = \langle \mu, z \rangle$.

Case 3: $z \in Z$. Let $z_+ := z \vee 0 \in P$ and $z_- := (-z) \wedge 0 \in P$. Clearly, $z = z_+ - z_-$. By Cases 1 and 2, $f(z) = f(z_+) - f(z_-) = \int_X z_+ d\mu - \int_X z_- d\mu = \int_X z d\mu = \langle \mu, z \rangle$, where the third equality follows from (xi) and Proposition 11.92.

Hence, $f(z) = \int_X z d\mu = \langle \mu, z \rangle$, $\forall z \in Z$. This completes the first step.
526

CHAPTER 11. GENERAL MEASURE AND INTEGRATION

In the second step, we consider the special case: Y = IR. By
Lemma 11.200, f = f+ − f− , where f+ , f− ∈ Z∗ are positive linear
functionals.
R Let µ+ := ΦZ (f+ )Rand µ− := ΦZ (f− ). By the first step,
f+ (z) = X z dµ+ and f− (z) = X z dµ− , ∀z ∈ Z. Then, by linearity of
ΦZ , µ = ΦZ (f ) = ΦZ (f+ − f− ) = ΦZ (f+ ) − ΦZ (f− ) R= µ+ − µ− .RBy PropoRsitions 11.145 and 11.146, f (z) = f+ (z) − f− (z) = X z dµ+ − X z dµ− =
X z dµ = hh µ, z ii. This completes the second step.

In the third step, we consider the special case when Y is a real
normed linear space. ∀y ∈ Y, define µy : BB ( X ) → IR by µy (E) =
hh µ(E), y ii, ∀E ∈ BB ( X ). Then, µy ∈ Mf t (X , IR) by Proposition 11.136.
Define fy : C(X , IR) → IR by fy (h) = f (hy), ∀h ∈ C(X , IR). By
f ∈ Z∗ , we have fy ∈ (C(X , IR))∗ . ∀Fe ∈ O, µy (F ) = hh µ(F ), y ii =
lim(U,V,h)∈ĀF fh , y = lim(U,V,h)∈ĀF hh fh , y ii = lim(U,V,h)∈ĀF f (hy) =
lim(U,V,h)∈ĀF fy (h) = ΦC(X ,IR) (fy )(F ). By Proposition 11.194, we have
µ
R y = ΦC(XR,IR) (fy ). By the second step, we have f (hy) = fy (h) =
h dµy = X hh dµ(x), h(x)y ii = hh µ, hy ii, ∀h ∈ C(X , IR), where the third
X
equality follows from Proposition 11.146. ∀z ∈ Z, ∀ǫ ∈ (0, +∞)
P ⊂ IR,
∃z̄ ∈ span ( Z ) such that k z − z̄ k < ǫ/(2 k f k + 1). Then, z̄ = ni=1 hi yi ,
where n ∈ Z+ , h1 , . . . , hn ∈ C(X , IR), and y1 , . . . , yn ∈ Y. Then, | f (z) −
hh µ, z iiP
| ≤ | f (z)−f (z̄)R|+| f (z̄)−hhP
µ, z̄ ii |+| hh µ, z̄ ii−hh
R µ, z ii | ≤ k f k k z −
n
n
hh dµ(x), z̄(x) ii −
hh
dµ(x),
h
(x)y
ii
+
z̄ k +
f
(h
y
)
−
i
i i
i=1
X
X
R
R
Pn i=1 i
hh
dµ(x),
hi (x)y ii) +
hh
dµ(x),
z(x)
ii
≤
ǫ/2
+
(f
(h
y)
−
iR
i=1
X
RX
hh dµ(x), z̄(x) − z(x) ii ≤ ǫ/2 + 0 + X P ◦ (z̄ − z)(x) dP ◦ µ(x) ≤
X
ǫ/2 + k z − z̄ k P ◦ µ(X) ≤ ǫ/2 + 2kfǫk+1 k µ k ≤ ǫ/2 + 2kfǫk+1 k f k < ǫ, where
the second inequality follows from Proposition 7.72; the third inequality
follows from (xi) and 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
R follows from (x). By the arbitrariness of
ǫ, we have f (z) = hh µ, z ii = X hh dµ(x), z(x) ii. This completes the third
step.

In the fourth step, we consider the special case when Y is a complex normed linear space. Let g := Re ◦f . By Lemmas 7.40 and
7.81, g is a bounded linear functional of the real normed linear space
ZIR := (C(X , YIR ), IR, k · kC(X ,YIR ) ) and f (z) = g(z) − ig(iz), ∀z ∈ ZIR . Let
ḡ ∈ Z∗IR be defined by ḡ(z) = g(iz), ∀z ∈ ZIR , µgr := ΦZIR (g) ∈ Mf t (X , Y∗IR ),
and µgi := ΦZIR (ḡ) ∈ Mf t (X , Y∗IR ). ∀Fe ∈ O, µ(F ) = lim(U,V,h)∈ĀF fh =
lim(U,V,h)∈ĀF (gh − iḡh ) = lim(U,V,h)∈ĀF gh − i lim(U,V,h)∈ĀF ḡh =
ΦZIR (g)(F ) − iΦZIR (ḡ)(F ) = µgr (F ) − iµgi (F ) = (µgr − iµgi )(F ). By
the definition of ḡ, we have µgi (E)(y) = µgr (E)(iy) ∈ C, ∀E ∈ BB ( X ),
∀y ∈ YIR . Then, by Lemma 7.81, µgr (E) − iµgi (E) ∈ Y∗ , ∀E ∈ BB ( X ).
By Proposition 11.136, µgr − iµgi ∈ Mf t (X , Y∗ ). By RProposition 11.194,
we have µ =R µgr − iµgi . By the third step, g(z) = X hh dµgr (x), z(x) ii
∀z ∈ Z, f (z) =
and ḡ(z) = X hh dµgi (x), z(x)R ii, ∀z ∈ ZIR . This implies,
R
g(z) − ig(iz) = g(z) − iḡ(z) = X hh dµgr (x), z(x) ii − i X hh dµgi (x), z(x) ii =


\[ \int_X \langle \langle d(\mu_{gr} - i\mu_{gi})(x), z(x) \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle, \] where the fourth equality follows from Propositions 11.145 and 11.146. This completes the fourth step.

Thus, we have the representation
\[(\text{xii}) \quad f(z) = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle = \langle \langle \mu, z \rangle \rangle, \quad \forall z \in Z, \forall f \in Z^*, \text{ where } \mu = \Phi_{\text{z}}(f) \in \mathcal{M}_{fI}(X, Y^*).\]

Fix any \( \mu \in \mathcal{M}_{fI}(X, Y^*). \) Define \( f : Z \to K \) by \( f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle, \) \( \forall z \in Z. \) We will show that \( f \in Z^*, \| f \| \leq \| \mu \|, \) and \( \Phi_{\text{z}}(f) = \mu. \) \( \forall z \in Z. \) By Proposition 5.7, \( z(X) \subseteq Y \) is compact. Then, by Propositions 11.190, 7.35, and 7.17, \( H := \text{span}(z(X)) \subseteq Y \) is a separable subspace. By (xi) and Proposition 11.132, \( f(z) = \langle \langle \mu, z \rangle \rangle = \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \in K \) is well-defined. By Proposition 11.132, \( f \in Z^* \) and \( \| f \| = \sup_{z \in Z, \| z \| \leq 1} |\langle \langle \mu, z \rangle \rangle| = \sup_{z \in Z, \| z \| \leq 1} \int_X \langle \langle d\mu(x), z(x) \rangle \rangle \leq \sup_{z \in Z, \| z \| \leq 1} \| \mu \|. \)

Let \( \tilde{\mu} := \Phi_{\text{z}}(\langle \langle \mu \rangle \rangle) \in \mathcal{M}_{fI}(X, Y^*). \) \( \forall F \in O, \forall y \in Y, \langle \langle \tilde{\mu}(F), y \rangle \rangle = \langle \langle \lim_{(U,V,h) \in \mathcal{A}_F} f(h) \rangle \rangle = \lim_{(U,V,h) \in \mathcal{A}_F} \langle \langle f(h) \rangle \rangle = \lim_{(U,V,h) \in \mathcal{A}_F} \int_X \langle \langle d\mu(x), h(x) y \rangle \rangle. \) \( \forall \epsilon \in (0, +\infty) \in \mathbb{R}, \exists O \in O \) with \( F \subseteq O \) such that \( \mathcal{P} \circ \mu(O \setminus F) < \epsilon/(1 + \| y \|). \) \( \forall (U,V,h) \in \mathcal{A}_F \) with \( V \subseteq O, \) \( \int_X \langle \langle d\mu(x), h(x) y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle = \int_{\mathcal{P} \circ h} \langle \langle d\mu(x), h(x) y \rangle \rangle - \langle \langle \mu(F), y \rangle \rangle = \langle \langle f(\mu(F)), y \rangle \rangle + \int_{\mathcal{P} \circ h} \langle \langle d\mu(x), h(x) y \rangle \rangle \leq \epsilon. \) By the arbitrariness of \( \epsilon \) and \( y, \) we have \( \tilde{\mu}(F) = \mu(F). \) By Proposition 11.194, \( \tilde{\mu} = \mu. \) This shows that \( \Phi_{\text{z}} \) is surjective.

\( \forall f \in Z^*, \) let \( \mu := \Phi_{\text{z}}(f) \in \mathcal{M}_{fI}(X, Y^*). \) Then, by the above and (x), we have \( \| \mu \| = \| \Phi_{\text{z}}(f) \| \leq \| f \| \leq \| \mu \|. \) Hence, \( \| \Phi_{\text{z}}(f) \| = \| f \|. \) Then, \( \Phi_{\text{z}} \) is injective. This shows that \( \Phi_{\text{z}} \) is bijective, continuous, linear, and norm preserving. Hence, \( \Phi_{\text{z}} : Z^* \to \mathcal{M}_{fI}(X, Y^*) \) is an isometrical isomorphism.

This completes the proof of the theorem.

**Definition 11.202** Let \( m \in \mathbb{Z}_+, \) \( I := [0, 1] \subset \mathbb{R}, \) \( Y \) be a normed linear space, \( f : I^m \to Y, \) and \( n \in \mathbb{N}. \) The nth Bernstein function for \( f, \) \( B_n : I^m \to Y, \) 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 \( m \in \mathbb{Z}_+ \), \( I := [0,1] \subset \mathbb{R} \), \( Y \) 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, \( \lim_{n \to \infty} \| f - B_n \|_{C(I^m,Y)} = 0 \).

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

\[
y = y \cdot 1 = y \left( \sum_{k=0}^{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right)
\]

\[
= \sum_{k=0}^{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k} = \sum_{k=1}^{n} \binom{n-1}{k-1} y^k (1 - y)^{n-k}
\]

\[
= \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} = \sum_{k=0}^{n-1} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k} \quad (11.7)
\]

where the second equality follows from (11.6). Similarly, \( \forall n - 1 \in \mathbb{N} \), \( \forall y \in I \),

\[
y^2 = y \left( \sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^k (1 - y)^{n-1-k} \right)
\]

\[
= \sum_{k=0}^{n-1} \frac{k}{n-1} \binom{n-1}{k} y^{k+1} (1 - y)^{n-1-k}
\]

\[
= \sum_{k=1}^{n} \frac{k-1}{n-1} \binom{n}{k-1} y^k (1 - y)^{n-k} = \sum_{k=1}^{n} \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k}
\]

\[
= \sum_{k=0}^{n-1} \frac{k(k-1)}{n(n-1)} \binom{n}{k} y^k (1 - y)^{n-k}
\]

where the first equality follows from (11.7). This implies that

\[
(1 - \frac{1}{n})y^2 = \sum_{k=0}^{n} \frac{k^2 - k}{n^2} \binom{n}{k} y^k (1 - y)^{n-k}
\]

\[
= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k}
\]

\[
= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} - \frac{1}{n} y
\]
11.10. DUAL OF $C(X, Y)$

where the last equality follows from (11.7). Rearranging terms in the above yields, $\forall y \in I$, $\forall n - 1 \in \mathbb{N},$

$$\frac{1}{n} y (1 - y) = \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} + y^2 - 2y^2$$

$$= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} y^k (1 - y)^{n-k} + y^2 \sum_{k=0}^{n} \binom{n}{k} y^k (1 - y)^{n-k}$$

$$= -2y \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} y^k (1 - y)^{n-k}$$

$$= \sum_{k=0}^{n} (y^2 - 2y \frac{k}{n} + \frac{k^2}{n^2} \binom{n}{k}) y^k (1 - y)^{n-k}$$

$$= \sum_{k=0}^{n} \left(y - \frac{k}{n}\right)^2 \binom{n}{k} y^k (1 - y)^{n-k} \tag{11.8}$$

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

$$1 = \prod_{i=1}^{m} (x_i + 1 - x_i)^n = \prod_{i=1}^{m} \left(\sum_{k=0}^{n} \binom{n}{k} x_i^k (1 - x_i)^{n-k}\right)$$

$$= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \prod_{i=1}^{m} \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i} \tag{11.9}$$

Then, $\forall x := (x_1, \ldots, x_m) \in I^m,$ $\forall n - 1 \in \mathbb{N},$ we have

$$\frac{1}{n} \sum_{i=1}^{m} x_i (1 - x_i) = \sum_{i=1}^{m} \sum_{k_i=0}^{n} \left(x_i - \frac{k_i}{n}\right)^2 \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i}$$

$$= \sum_{i=1}^{m} \left(\sum_{k_i=0}^{n} \left(x_i - \frac{k_i}{n}\right)^2 \binom{n}{k_i} x_i^{k_i} (1 - x_i)^{n-k_i}\right)$$

$$\cdot \prod_{j=1}^{m} \left(\sum_{k_j=0}^{n} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j}\right)$$

$$= \sum_{i=1}^{m} \sum_{k_i=0}^{n} \left(\sum_{k_i=0}^{n} \left(x_i - \frac{k_i}{n}\right)^2 \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j}\right)$$

$$= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \left(\prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j}\right) \left(\sum_{i=1}^{m} \left(x_i - \frac{k_i}{n}\right)^2\right)$$

$$= \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \left|x - \left(\frac{k_1}{n}, \ldots, \frac{k_m}{n}\right)\right|^2 \prod_{j=1}^{m} \binom{n}{k_j} x_j^{k_j} (1 - x_j)^{n-k_j} \tag{11.10}$$
where the first equality follows from (11.8); 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. $\forall \varepsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x, \bar{x} \in I^m$ with $|x - \bar{x}| < \delta$, we have $\|f(x) - f(\bar{x})\| < \varepsilon/2$. Let $n_0 := \left\lceil \max\{\delta^{-4}, m^2M^2/\varepsilon^2, 2\} \right\rceil \in \mathbb{N}$. $\forall n \in \mathbb{N}$ with $n_0 \leq n$, $\forall x := (x_1, \ldots, x_m) \in I^m$, we have

$$
\|f(x) - B_n(x)\| = \left\| f(x) \sum_{k_1=0}^{4} \cdots \sum_{k_m=0}^{n} \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \right.
$$

$$
- \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right) \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \left. \right\|
$$

$$
= \left\| \sum_{k_1=0}^{n} \cdots \sum_{k_m=0}^{n} \|f(x) - f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \right\|
$$

$$
\leq \sum_{(k_1, \ldots, k_m) \in J_x} \|f(x) - f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i}
$$

$$
+ \sum_{(k_1, \ldots, k_m) \in J^c_x} \|f(x) - f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i}
$$

Note that, $\forall (k_1, \ldots, k_m) \in J_x$, $\|x(1, \ldots, x_m) - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})\| < n^{-1/4} \leq \delta$, then $\|f(x) - f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| < \varepsilon/2$. This implies that

$$
\sum_{(k_1, \ldots, k_m) \in J_x} \|f(x) - f\left( \frac{k_1}{n}, \ldots, \frac{k_m}{n} \right)\| \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i}
$$

$$
< \frac{\varepsilon}{2} \sum_{(k_1, \ldots, k_m) \in J_x} \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i}
$$

$$
\leq \frac{\varepsilon}{2} \sum_{(k_1, \ldots, k_m) \in J_x} \prod_{i=1}^{m} \left( \frac{n}{k_i} \right) x_i^{k_i} (1 - x_i)^{n-k_i} = \frac{\varepsilon}{2}
$$
where the equality follows from (11.9). Note also that, \( \forall (k_1, \ldots, k_m) \in J \setminus J_x, \| f(x) - f(\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \leq 2M = 2M \frac{|x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})|^2}{|x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})|} \leq 2M \sqrt{n} |x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n})|^2. \) Then,

\[
\sum_{(k_1, \ldots, k_m) \in J \setminus J_x} \| f(x) - f(\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \prod_{i=1}^m \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \leq 2M \sqrt{n} \sum_{(k_1, \ldots, k_m) \in J} \| x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \prod_{i=1}^m \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} \leq 2M \sqrt{n} \sum_{(k_1, \ldots, k_m) \in J} \| x - (\frac{k_1}{n}, \ldots, \frac{k_m}{n}) \| \prod_{i=1}^m \left( \begin{array}{c} n \\ k_i \end{array} \right) x_i^{k_i} (1 - x_i)^{n-k_i} = 2M \sqrt{n} \frac{1}{n} \sum_{i=1}^m x_i (1 - x_i) \leq \frac{Mm}{2\sqrt{n}} \leq \epsilon/2
\]

Hence, we have \( \| f(x) - B_n(x) \| < \epsilon. \) By the arbitrariness of \( x \in I^m, \) we have \( \| f - B_n \|_{C(I^m, y)} < \epsilon. \) Therefore, \( \lim_{n \to \infty} \| f - B_n \|_{C(I^m, y)} = 0. \)

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 \( \mathcal{O}, \mathcal{X} := (X, \mathcal{O}), \mathcal{Y} \) be a normed linear space, and \( Z := C(\mathcal{X}, \mathcal{Y}). \) Then, \( Z^* = \mathcal{M}_f(X, \mathcal{Y}^*) = \mathcal{M}_f(X, \mathcal{B}_B(\mathcal{X}), \mathcal{Y}^*). \)

**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 \( Z := \{ z \in Z \mid z = hy, h \in C(\mathcal{X}, \mathbb{R}), y \in \mathcal{Y} \}. \) We will show that \( Z = \text{span}(\mathbb{Z}) \) by distinguishing two exhaustive and mutually exclusive cases: Case 1: \( \exists i_0 \in \{1, \ldots, m\} \) such that \( I_{i_0} = \emptyset; \) Case 2: \( I_1, \ldots, I_m \) are nonempty. Case 1: \( \exists i_0 \in \{1, \ldots, m\} \) such that \( I_{i_0} = \emptyset. \) Then, \( X = \emptyset \) and \( Z \) is the trivial Banach space with a singleton element. Clearly, \( Z = \mathbb{Z}. \) Hence, \( Z = \text{span}(\mathbb{Z}). \)

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 \( a_i \leq b_i, i = 1, \ldots, m, \) and \( \exists \bar{m} \in \mathbb{Z}_+ \) with \( \bar{m} \leq m \) such that \( a_i < b_i, i = 1, \ldots, \bar{m}, \) and \( a_i = b_i, i = \bar{m} + 1, \ldots, m. \) Let \( I := [0, 1] \subset \mathbb{R}. \) Then, we may define a homeomorphism \( \psi : \mathcal{X} \to I^m \) by \( \psi(x) = (\frac{x_1 - a_1}{b_1 - a_1}, \ldots, \frac{x_{\bar{m}} - a_{\bar{m}}}{b_{\bar{m}} - a_{\bar{m}}}), \forall x := (x_1, \ldots, x_{\bar{m}}) \in \mathcal{X}. \) \( \forall f \in Z, \) let \( \tilde{f} : I^m \to \mathcal{Y} \) be defined by \( \tilde{f} := f \circ \psi \). By Proposition 3.12, \( \tilde{f} \in C(I^m, \mathcal{Y}). \) By Bernstein Approximation Theorem 11.203, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n \in \mathbb{N} \) such that \( \| \tilde{f} - B_n \|_{C(I^m, \mathcal{Y})} < \epsilon, \) where \( B_n : I^m \to \mathcal{Y} \) is the \( n \)th Bernstein function.
Representation Theorem 11.204 and its proof, we have $Z$. Let $f = (\bar{f} - B_n) \circ \psi \in \mathbb{B}(Z)$. By the arbitrariness of $\epsilon$, we have $f \in \mathbb{B}(Z)$.

By the arbitrariness of $f$, we have $Z = \mathbb{B}(Z)$.

Hence, in both cases, we have $Z = \mathbb{B}(Z)$. By Riesz Representation Theorem 11.201, we have $Z^* = \mathcal{M}_f(X, Y^*)$. By Theorem 11.198, $Z^* = \mathcal{M}_f(X, Y^*) = \mathcal{M}_f(X, B_B(X), Y^*)$. 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 $Z := \mathcal{C}(X, Y)$. Then, $Z^* = \mathcal{M}_f(X, Y^*)$.

Proof Let $Z := \{ z \in Z \mid z = hy, h \in \mathcal{C}(X, \mathbb{R}), y \in Y \}$. We will show that $Z = \mathbb{B}(Z)$. Then, the theorem is a direct consequence of Riesz Representation Theorem 11.201.

Let the dimension of $Y$ be $m \in \mathbb{Z}_+$ 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: $K = \mathbb{R}$; Case 2: $K = \mathbb{C}$.

Case 1: $K = \mathbb{R}$. $\forall f \in Z$, let $g_i := \pi_i \circ \phi \circ f \in \mathcal{C}(X, \mathbb{R}), i = 1, \ldots, m$. 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, $f \in \mathbb{B}(Z)$. This case is proved.

Case 2: $K = \mathbb{C}$. $\forall f \in Z$, let $g_i := \text{Re} \circ \pi_i \circ \phi \circ f \in \mathcal{C}(X, \mathbb{R})$ and $h_i := \text{Im} \circ \pi_i \circ \phi \circ f \in \mathcal{C}(X, \mathbb{R}), i = 1, \ldots, m$. Then, $f(x) = \sum_{i=1}^m (\pi_i \circ \phi(f(x)))y_i = \sum_{i=1}^m (\text{Re} \circ \pi_i \circ \phi \circ f(x) + i \text{Im} \circ \pi_i \circ \phi \circ f(x))y_i = \sum_{i=1}^m (g_i(x)y_i + h_i(x)|y_i|), \forall x \in X$. Hence, $f \in \mathbb{B}(Z)$. This case is also proved.

This completes the proof of the theorem.

Proposition 11.206 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 $\mathcal{O}$, $X := (X, \mathcal{O})$, and $Y$ be a separable normed linear space. Then, $Z := \mathcal{C}(X, Y)$ is separable.

Proof Let $Z := \{ z \in Z \mid z = hy, h \in \mathcal{C}(X, \mathbb{R}), y \in Y \}$. By Riesz Representation Theorem 11.204 and its proof, we have $Z = \mathbb{B}(Z)$. By Proposition 11.190, $\mathcal{C}(X, \mathbb{R})$ is separable. Let $D_1 \subset \mathcal{C}(X, \mathbb{R})$ be a countable dense subset. Let $D_2 \subset Y$ be a countable dense subset. Then, $Z_D := \{ z \in Z \mid z = hy_1, h_1 \in D_1, y_1 \in D_2 \}$ is a countable set. Clearly, $\mathbb{B}(Z_D) \supset Z$. Then, $Z$ is separable. By Proposition 7.35, $Z$ is separable. This completes the proof of the proposition.

□
12.1 Carathéodory Extension Theorem

Definition 12.1 Let $X$ be a set, $A \subseteq X^2$ be an algebra on $X$, $Y$ be a normed linear space over $K$, and $\nu : A \to [0, \infty) \subseteq \mathbb{R}_+$ be a measure on the algebra $A$. A $Y$-valued measure on the algebra $A$ is a function $\mu$ from $A$ to $Y$ that satisfies

(i) $\mu(\emptyset) = \emptyset_Y$;

(ii) $\forall A \in A$ with $\nu(A) = \infty$, $\mu(A)$ is undefined;

(iii) $\forall A \in A$ with $\nu(A) < \infty$, $\mu(A) \in Y$ and $\forall$ pairwise disjoint $(A_i)_{i=1}^\infty \subseteq A$ 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) $\forall A \in A$ with $\nu(A) < \infty$, we have

$$\nu(A) = \sup_{n \in \mathbb{Z}_+, \,(A_i)_{i=1}^n \subseteq A, \,A = \bigcup_{i=1}^n A_i, \,A_i \cap A_j = \emptyset, \,\forall 1 \leq i < j \leq n} \sum_{i=1}^n \|\mu(A_i)\|$$

Then, $\nu$ is said to be the total variation of $\mu$ and denoted by $P \circ \mu$. $\mu$ 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, $A \subseteq X^2$ be an algebra on $X$, $Y$ be a normed linear space over $K$, $\mu$ be a $Y$-valued measure on the algebra $A$, and $E \in A$. Then, $A_E := \{C \in A \mid C \subseteq E\}$ is an algebra on $E$, $\mu_E := \mu|_{A_E}$ is a $Y$-valued measure on the algebra $A_E$, and $P \circ \mu_E = (P \circ \mu)|_{A_E}$.
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 $\nu_E := \left(\mathcal{P} \circ \mu\right)|_{\mathcal{A}_E}$. Then, $\nu_E(\emptyset) = \mathcal{P} \circ \mu(\emptyset) = 0$. For every 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 $\mathcal{A}_E$.

$$\mu_E(\emptyset) = \mu(\emptyset) = \lambda_E. \forall A \in \mathcal{A}_E$$

with $\nu_E(A) = \infty$, we have $\mathcal{P} \circ \mu(A) = \infty, \mu(A)$ is undefined, and therefore, $\mu_E(A)$ is undefined. For all $A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\mathcal{P} \circ \mu(A) < \infty$ and $\mu_E(A) = \mu(A) \in \mathcal{Y}$. For every 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 $\sum_{i=1}^{\infty} \mu_E(A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \mu(A) = \mu_E(A)$. For all $A \in \mathcal{A}_E$ with $\nu_E(A) < \infty$, we have $\nu_E(A) = \mathcal{P} \circ \mu(A) = \sup_{n \in \mathbb{Z}^+, (A_i)_{i=1}^{n} \subseteq \mathcal{A}_E, A = \bigcup_{i=1}^{n} A_i, A \cup A_i = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \|\mu(A_i)\|$. Hence, $\mu_E$ is a $\mathcal{Y}$-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 $\rho(E_1, E_2) = \mu(E_1 \Delta E_2), \forall E_1, E_2 \in \mathcal{B}$. Then, $\rho$ defines a pseudo-metric on $\mathcal{B}$.

Proof. For all $E_1, E_2, E_3 \in \mathcal{B}$, $\rho(E_1, E_2) = \mu(E_1 \Delta E_2) \in [0, +\infty) \subset \mathbb{R}$ since $\mu$ is a finite measure on $(X, \mathcal{B})$. Clearly, $\rho(E_1, E_3) = \mu(E_1 \Delta E_3) = \rho(\emptyset) = 0$ and $\rho(E_1, E_2) = \mu(E_1 \Delta E_2) = \rho(E_2, E_1)$. Note that $\rho(E_1, E_2) + \rho(E_2, E_3) = \mu(E_1 \Delta E_2) + \mu(E_2 \Delta E_3) = \mu(E_1 \setminus E_2) + \mu(E_2 \setminus E_1) + \mu(E_2 \setminus E_3) + \mu(E_3 \setminus E_2) \geq \mu((E_1 \setminus E_2) \cup (E_2 \setminus E_3)) + \mu((E_2 \setminus E_1) \cup (E_1 \setminus E_3))$. Clearly, we have $(E_1 \setminus E_2) \cup (E_2 \setminus E_3) \subseteq E_1 \setminus E_3$ and $(E_2 \setminus E_1) \cup (E_3 \setminus E_2) \subseteq E_2 \setminus E_3$. Then, $\rho(E_1, E_2) + \rho(E_2, E_3) \geq \mu(E_1 \setminus E_3) + \mu(E_3 \setminus E_1) = \mu(E_1 \Delta E_3) = \rho(E_1, E_3)$. 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 $X$, $\mathcal{Y}$ be a Banach space over $\mathbb{K}$, $\mu$ be a $\sigma$-finite $\mathcal{Y}$-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 $\mathcal{Y}$-valued measure $\tilde{\mu}$ on $(X, \mathcal{B})$ such that $\exists \mathcal{Y} = \mu$ and $\left(\mathcal{P} \circ \tilde{\mu}\right)|_{\mathcal{A}} = \mathcal{P} \circ \mu$. Furthermore, $\tilde{\mu}$ is $\sigma$-finite.

Proof. We first consider the special case where $\mu$ is finite. Then, $\mu : A \to \mathcal{Y}$ and $\mathcal{P} \circ \mu : A \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$. 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 $\nu(E) = \inf\left(\mathcal{A}_{\mathcal{Y}}, \mathcal{A}_{\mathcal{Y}} \cup A \mathcal{B}, \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i), \forall E \in \mathcal{B}\right)$ for $E \in \mathcal{B}$. Define $\rho : \mathcal{B} \times \mathcal{B} \to [0, \mathcal{P} \circ \mu(X)] \subset \mathbb{R}$ by $\rho(E_1, E_2) = \nu(E_1 \setminus E_2), \forall E_1, E_2 \in \mathcal{B}$. 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 $\rho(E_1, E_2) = \nu(E_1 \setminus E_2) = 0$. 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 \( \tilde{\rho}([E_1], [E_2]) = \rho(E_1, E_2), \forall [E_1], [E_2] \in \mathcal{B}/ \equiv \) 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 \( E \in \mathcal{B} \), \( \exists \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{A} \) 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 \( \sum_{i=n_0+1}^{\infty} \mathcal{P} \circ \mu(A_i) < \epsilon/2 \). Let \( A := \bigcup_{i=n_0+1}^{\infty} A_i \). 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 \( 0 \leq \nu(E \setminus A) \leq \nu\left(\bigcup_{i=n_0+1}^{\infty} A_i \setminus A\right) \leq \sum_{i=n_0+1}^{\infty} \nu(A_i) = \sum_{i=n_0+1}^{\infty} \mathcal{P} \circ \mu(A_i) < \epsilon/2 \). Note also that \( A \setminus E \subseteq \bigcup_{i=1}^{\infty} A_i \setminus E \). This further implies that \( 0 \leq \nu(A \setminus E) \leq \nu\left(\bigcup_{i=1}^{\infty} A_i \setminus E\right) = \nu\left(\bigcup_{i=1}^{\infty} A_i \setminus A\right) - \nu(E) \leq \sum_{i=1}^{\infty} \nu(A_i) - \nu(E) = \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(A_i) - \nu(E) < \epsilon/2 \). Hence, \( \tilde{\rho}\left([E], [A]\right) = \rho(E, A) = \nu(E \setminus A) + \nu(A \setminus E) < \epsilon \). Hence, \( \overline{\mathcal{A}_\equiv} = \mathcal{B}/ \equiv \). This completes the proof of the claim. \( \square \)

Define \( g : \mathcal{A}_\equiv \rightarrow \mathcal{Y} \) 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 \( \rho(A_1, A_2) = \nu(A_1 \Delta A_2) = \mathcal{P} \circ \mu(A_1 \Delta A_2) = 0 \). Then, \( \mathcal{P} \circ \mu(A_1 \setminus A_2) = 0 = \mathcal{P} \circ \mu(A_2 \setminus A_1) \). This implies that \( \mu(A_1 \setminus A_2) = \vartheta_y = \mu(A_2 \setminus A_1) \) and \( \mu(A_1) = \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) = \mu(A_1 \cap A_2) + \mu(A_2 \setminus A_1) = \mu(A_2) \). 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 \( \tilde{\rho}([A_1], [A_2]) < \delta \), we have \( \| g([A_1]) - g([A_2]) \| = \| \mu(A_1) - \mu(A_2) \| = \| \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) - \mu(A_2 \cap A_1) - \mu(A_2 \setminus A_1) \| = \| \mu(A_1 \setminus A_2) - \mu(A_2 \setminus A_1) \| \leq \| \mu(A_1 \setminus A_2) \| + \| \mu(A_2 \setminus A_1) \| \leq \mathcal{P} \circ \mu(A_1 \setminus A_2) + \mathcal{P} \circ \mu(A_2 \setminus A_1) = \mathcal{P} \circ \mu(A_1 \Delta A_2) = \nu(A_1 \Delta A_2) = \rho(A_1, A_2) = \tilde{\rho}([A_1], [A_2]) < \delta = \epsilon \). Hence, \( g \) is uniformly continuous. This completes the proof of the claim. \( \square \)

By Proposition 4.46 and the completeness of \( \mathcal{Y} \), there exists a unique continuous function \( \tilde{g} : \mathcal{B}/ \equiv \rightarrow \mathcal{Y} \) such that \( \tilde{g}|_{\mathcal{A}_\equiv} = g \), and \( \tilde{g} \) is uniformly continuous. Define \( \tilde{\mu} : \mathcal{B} \rightarrow \mathcal{Y} \) by \( \tilde{\mu}(E) = \tilde{g}\left([E]\right), \forall E \in \mathcal{B} \). We will show that \( \tilde{\mu} \) is the \( \mathcal{Y} \)-valued measure we seek. Clearly, \( \tilde{\mu}(A) = \tilde{g}\left([A]\right) = g([A]) = \mu(A), \forall A \in \mathcal{A} \). Then, \( \tilde{\mu}(\emptyset) = \mu(\emptyset) = \vartheta_y \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), by the uniform continuity of \( \tilde{g} \), \( \exists \delta(\epsilon) \in (0, \epsilon) \subset \mathbb{R} \) such that \( \forall [E_1], [E_2] \in \mathcal{B}/ \equiv \) with \( E_1, E_2 \in \mathcal{B} \) and \( \tilde{\rho}([E_1], [E_2]) = \rho(E_1, E_2) = \nu(E_1 \Delta E_2) < \delta(\epsilon) \), we have \( \| \tilde{\mu}(E_1) - \tilde{\mu}(E_2) \| = \| \tilde{g}\left([E_1]\right) - \tilde{g}\left([E_2]\right) \| < \epsilon \).

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 \( \tilde{\rho}([A_1], [E_i]) = \rho(A_i, E_i) = \nu(A_i \setminus E_i) < \delta(\epsilon/4)/2, i = 1, 2 \). Then, \( \| \tilde{\mu}(E_i) - \tilde{\mu}(A_i) \| < \epsilon/4, i = 1, 2 \). Note that
(E_1 \cup E_2) \setminus (A_1 \cup A_2) = (E_1 \setminus A_2) \cap (\tilde{A}_1 \cap A_2) = (E_1 \cap A_1 \cap \tilde{A}_2) \cup (E_2 \cap \tilde{A}_1 \cap A_2) \subseteq (E_1 \setminus A_1) \cup (E_2 \setminus A_2) and \ (A_1 \cup A_2) \setminus (E_1 \cup E_2) \subseteq (A_1 \setminus E_1) \cup (A_2 \setminus E_2). This leads to (E_1 \cup E_2) \triangle (A_1 \cup A_2) \subseteq (E_1 \setminus A_1) \cup (E_2 \setminus A_2) \cup (A_1 \setminus E_1) \cup (A_2 \setminus E_2) = (E_1 \triangle A_1) \cup (E_2 \triangle A_2). This implies that \( \tilde{\mu}[(E_1 \cup E_2), [A_1 \cup A_2]] = \nu((E_1 \cup E_2) \triangle (A_1 \cup A_2)) \leq \nu(E_1 \triangle A_1) + \nu(E_2 \triangle A_2) < \delta(\epsilon/4). Hence, \( \| \tilde{\mu}(E_1 \cup E_2) - \tilde{\mu}(A_1 \cup A_2) \| < \epsilon/4. Note also that (A_1 \cup A_2) \triangle (E_1 \cup E_2) = ((A_1 \cup A_2) \setminus (E_1 \cup E_2)) \cup ((E_1 \cup E_2) \setminus (A_1 \cup A_2)) = (A_1 \cap A_2 \cap \tilde{E}_1) \cup (A_1 \cap A_2 \cap \tilde{E}_2) \cup (E_1 \cap E_2 \cap A_1) \cup (E_1 \cap E_2 \cap A_2) \subseteq (E_1 \setminus A_1) \cup (E_2 \setminus A_2) \cup (E_1 \setminus A_1) \cup (E_2 \setminus A_2) = (E_1 \triangle A_1) \cup (E_2 \triangle A_2). This leads to \( \tilde{\mu}[(E_1 \cap E_2), [A_1 \cap A_2]] = \tilde{\mu}(\emptyset, [A_1 \cap A_2]) = \nu((E_1 \triangle A_1) \cup (E_2 \triangle A_2)) < \nu(E_1 \triangle A_1) + \nu(E_2 \triangle A_2) < \delta(\epsilon/4). This implies that \( \| \tilde{\mu}(A_1 \cup A_2) - \tilde{\mu}(A_2) \| = \| \mu(A_1 \cup A_2) - \mu(A_1) - \mu(A_2) \| = \| \mu(A_1) + \mu(A_2) - \mu(A_1) - \mu(A_2) \| = \| \mu(A_1 \cap A_2) - \mu(\emptyset) \| = \| \tilde{\mu}(A_1 \cap A_2) - \tilde{\mu}(\emptyset) \| < \epsilon/4. Therefore, we have \( \| \tilde{\mu}(E_1 \cup E_2) - \tilde{\mu}(E_1) - \tilde{\mu}(E_2) \| < \| \tilde{\mu}(E_1 \cup E_2) - \tilde{\mu}(A_1 \cup A_2) \| + \| \tilde{\mu}(A_1 \cup A_2) - \tilde{\mu}(A_1) - \tilde{\mu}(A_2) \| + \| \tilde{\mu}(A_1) - \tilde{\mu}(E_1) \| + \| \tilde{\mu}(A_2) - \tilde{\mu}(E_2) \| < \epsilon. By the arbitrariness of \( \epsilon, we have \| \tilde{\mu}(E_1 \cup E_2) - \tilde{\mu}(E_1) \| < \epsilon. Hence, \( \tilde{\mu} \) is finitely additive. This completes the proof of the claim. □

Claim 12.4.4 ∀E ∈ B, \( \| \tilde{\mu}(E) \| \leq \nu(E) \).

Proof of claim: ∀E ∈ B, ∀E \subseteq (0, \infty) \subseteq \mathbb{R}, by Claim 12.4.1, \( \exists A \in A \) such that \( \tilde{\mu}[(E), A] = \nu(E \cap A) < \delta(\epsilon/2) \leq \epsilon/2. Then, \( \| \tilde{\mu}(E) - \mu(A) \| < \epsilon/2. Note that \( \| \tilde{\mu}(A) \| = \| \mu(A) \| \leq \mathcal{P}(\mu(A) = \nu(A)). Then, \( \| \tilde{\mu}(E) \| \leq \| \tilde{\mu}(E) - \mu(A) \| + \| \mu(A) \| < \epsilon/2 + \nu(A \cap E) \leq \epsilon/2 + \nu(A \setminus E) \leq \epsilon/2 + \nu(E \setminus A) < \nu(E) + \epsilon. By the arbitrariness of \( \epsilon, we have \( \| \tilde{\mu}(E) \| \leq \nu(E). This completes the proof of the claim. □

∀ pairwise disjoint \( (E_n)_{n=1}^{\infty} \subseteq B, \) let \( E := \bigcup_{n=1}^{\infty} E_n \in B. \) By Claim 12.4.4, we have \( \sum_{n=1}^{\infty} \| \tilde{\mu}(E_n) \| \leq \sum_{n=1}^{\infty} \nu(E_n) = \nu(E) < +\infty. \) Let \( \tilde{E}_n := \bigcup_{n=1}^{\infty} E_n \in B, \forall n \in \mathbb{N}. \) By Claim 12.4.3, \( \tilde{\mu}(\tilde{E}_n) = \sum_{i=1}^{n} \tilde{\mu}(E_i). \) We will show that \( \lim_{n \to \infty} \tilde{\mu}(\tilde{E}_n) = \tilde{\mu}(E), \) which then implies that \( \mu \) is countably additive. \( \forall E \subseteq (0, \infty) \subseteq \mathbb{R}, \exists n_0 \in \mathbb{N} \) such that \( \forall n \in \mathbb{N} \) with \( n \geq n_0, \) we have \( 0 \leq \sum_{n=n_0+1}^{\infty} \nu(E_n) < \delta(\epsilon). \) Then, \( \tilde{\mu}[(\tilde{E}_n), [E]] = \rho(\tilde{E}_n, E) = \nu(\tilde{E}_n \setminus E) = \nu(E \setminus \tilde{E}_n) = \sum_{i=n+1}^{\infty} \nu(E_i) < \delta(\epsilon). \) This implies that \( \| \tilde{\mu}(E) - \tilde{\mu}(\tilde{E}_n) \| < \epsilon. \) Hence, \( \tilde{\mu}(E) = \lim_{n \to \infty} \tilde{\mu}(\tilde{E}_n) = \sum_{n=1}^{\infty} \tilde{\mu}(E_n). \) Therefore, \( \tilde{\mu} \) is countably additive. \( \forall E \subseteq B, \forall t \in \mathbb{Z}, \forall \text{ pairwise disjoint} \ (E_i)_{i=1}^{t} \subseteq B \) with \( E = \bigcup_{i=1}^{t} E_i, \) by Claim 12.4.4, we have \( \sum_{i=1}^{t} \| \tilde{\mu}(E_i) \| \leq \sum_{i=1}^{t} \nu(E_i) \leq \nu(E). \) Therefore, \( \mathcal{P}(\mu(E) \leq \nu(E). \) By Proposition 11.100, \( \mathcal{P} \circ \tilde{\mu} \) is a measure on \( (X, \mathcal{B}). \) By the above argument, \( \mathcal{P}(\mu(E) \leq \nu(E) \leq \mathcal{P}(\mu(X) < +\infty. \) Hence, \( \mathcal{P} \circ \tilde{\mu} \) is a finite measure on \( (X, \mathcal{B}). \) \( \forall A \subseteq A \), by Definitions 11.108 and 12.1, \( \nu(A) = \mathcal{P}(\mu(A) \leq \mathcal{P}(\tilde{\mu}(A) \leq \nu(A). \) Then, \( \mathcal{P} \circ \tilde{\mu}(A) = \nu(A) = \mathcal{P} \circ \mu(A), \forall A \subseteq A \). By Carathéodory Extension Theorem 11.19, \( \mathcal{P} \circ \tilde{\mu} = \nu. \)

To show the uniqueness of \( \tilde{\mu}, \) let \( \tilde{\mu} \) be another \( \tilde{\mu} \)-valued measure on \( (X, \mathcal{B}) \) such that \( \tilde{\mu}|_A = \mu \) and \( \mathcal{P} \circ \tilde{\mu}|_A = \mathcal{P} \circ \mu. \) By Carathéodory Extension Theorem 11.19, \( \mathcal{P} \circ \tilde{\mu} = \nu. \) Hence, \( \tilde{\mu} \) is a finite \( \tilde{\mu} \)-valued measure.
and \( \bar{\mu} : B \rightarrow \gamma \). Define \( \bar{g} : B/ \equiv \rightarrow \gamma \) by \( \bar{g}(E) = \tilde{\mu}(E), \ \forall [E] \in B/ \equiv \) with \( E \in B \). We need to show that \( \bar{g} \) is well defined. \( \forall E_1, E_2 \in B \) with \( E_1 \equiv E_2 \), we have \( \nu(E_1, E_2) = \nu(E_1 \Delta E_2) = 0 \). This leads to \( \nu(E_1 \setminus E_2) = 0 = \nu(E_2 \setminus E_1) \) and \( \bar{\mu}(E_1 \setminus E_2) = \bar{\mu}(E_2 \setminus E_1) \). Then, \( \bar{\mu}(E_1) \) and \( \bar{\mu}(E_2) \) is well defined. Clearly, \( \bar{g}([A]) = \tilde{\mu}(A) = \mu(A) = g([A]) \), \( \forall [A] \in \mathcal{A} \) with \( A \in \mathcal{A} \). Hence, \( \bar{g}|\mathcal{A} = g. \ \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \) let \( \delta = \epsilon. \) \( \forall [E_1], [E_2] \in B/ \equiv \) with \( E_1, E_2 \in B \) and \( \tilde{\mu}([E_1], [E_2]) = \nu(E_1 \Delta E_2) < \delta \), we have \( \| \bar{g}(E_1) - \tilde{\mu}(E_1) \| = \| \bar{\mu}(E_1) - \bar{\mu}(E_2) \| = \| \bar{\mu}(E_1 \setminus E_2) + \bar{\mu}(E_2 \setminus E_1) \| \leq \| \bar{\mu}(E_1 \setminus E_2) \| + \| \bar{\mu}(E_2 \setminus E_1) \| \leq \| \bar{\mu}(E_1) \| + \| \bar{\mu}(E_2) \| = \nu(E_1 \Delta E_2) < \delta \). Hence, \( \bar{g} \) is uniformly continuous. By Proposition 4.46, \( \bar{g} = \tilde{\mu} \). Next, consider the general case where \( \mu \) is \( \sigma \)-finite. Then, \( \exists (X_n)_{n=1}^{\infty} \subseteq \mathcal{A} \) 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 \( \mathcal{A}_n := \{ E \in \mathcal{A} \mid E \subseteq X_n \} \) and \( \mu_n := \mu|\mathcal{A}_n \). By Proposition 12.2, \( \mu_n \) is a \( \gamma \)-valued measure on the algebra \( \mathcal{A}_n \) and \( \mathcal{P} \circ \mu_n = (\mathcal{P} \circ \mu)|\mathcal{A}_n \). By the special case, there exists a unique \( \gamma \)-valued measure \( \mu_n \) on \( (X_n, \mathcal{B}_n) \), where \( \mathcal{B}_n \) is the \( \sigma \)-algebra generated by \( \mathcal{A}_n \), such that \( \mu_n(A) = \mu_n(A) \) and \( \mathcal{P} \circ \mu_n(A) = \mathcal{P} \circ \mu_n(A), \forall A \in \mathcal{A}_n \). 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 \( \gamma \)-valued measure space \( \mathcal{X} := (X, \mathcal{B}, \bar{\mu}) \) on \( X \), where \( \mathcal{B} := \{ E \subseteq X \mid E \cap X_n \subseteq \mathcal{B}_n, \forall n \in \mathbb{N} \} \), \( \forall E \in \mathcal{B}, \mathcal{P} \circ \bar{\mu}(E) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(E \cap X_n), \bar{\mu}(E) \) is undefined if \( \mathcal{P} \circ \bar{\mu}(E) = \infty \), \( \bar{\mu}(E) = \sum_{n=1}^{\infty} \mu_n(E \cap X_n) \in \gamma \) if \( \mathcal{P} \circ \bar{\mu}(E) < \infty \), and \( X_n \) is the finite \( \gamma \)-valued measure subspace of \( \mathcal{X} \), \( \forall n \in \mathbb{N} \).

Claim 12.4.5 \( B = \mathcal{B} \). 

Proof of claim: \( \forall A \in \mathcal{A}, \forall n \in \mathbb{N}, A \cap X_n \subseteq \mathcal{A}_n \subseteq \mathcal{B}_n \). This implies that \( A \in \mathcal{B} \). Then, \( A \subseteq B \). Since \( B \) is a \( \sigma \)-algebra, then \( B \subseteq \mathcal{B} \). On the other hand, \( \forall E \in \mathcal{B}, \forall n \in \mathbb{N} \), we have \( E \cap X_n \subseteq \mathcal{B}_n \subseteq B \). Then, \( E = \bigcup_{n=1}^{\infty} (E \cap X_n) \) in \( B \), and hence, \( B \subseteq \mathcal{B} \). This shows that \( B = B \) and completes the proof of the claim. 

\( \forall A \in \mathcal{A}, \forall n \in \mathbb{N} \), we have \( A \cap X_n \subseteq \mathcal{A}_n \subseteq \mathcal{B}_n \) and \( \mathcal{P} \circ \bar{\mu}(A) = \mathcal{P} \circ \mu_n(A) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \mathcal{P} \circ \mu(A \cap X_n) = \mathcal{P} \circ \mu(A). \) Hence, \( \mathcal{P} \circ \bar{\mu} \) is defined and \( \mathcal{P} \circ \mu(A) = \infty \), which implies that \( \mu(A) \) is undefined. \( \forall A \in \mathcal{A} \) with \( \mathcal{P} \circ \mu(A) < \infty \), then \( \mathcal{P} \circ \mu(A) < \infty \) and \( \bar{\mu}(A) = \sum_{n=1}^{\infty} \mu_n(A \cap X_n) = \sum_{n=1}^{\infty} \mu(A \cap X_n) = \mu(A) \in \gamma \). Hence, \( \mu = \bar{\mu} \). This shows that \( \bar{\mu} \) is the \( \sigma \)-finite \( \gamma \)-valued measure on \( (X, \mathcal{B}) \) that we seek.

This completes the proof of the theorem.
Definition 12.5 Let \( X \) be a set, \( A \subseteq X^2 \) be an algebra on \( X \), \( Y \) be a normed linear space over \( K \), and \( \mu : A \rightarrow Y \). \( \mu \) is said to be a \( Y \)-valued pre-measure on the algebra \( A \) if

(i) \( \mu(\emptyset) = 0 \);

(ii) \( \forall \) pairwise disjoint \( (A_i)^n_{i=1} \subseteq A \) with \( A := \bigcup_{i=1}^\infty A_i \in A \), we have
\[
\sum_{i=1}^\infty \| \mu(A_i) \| < \infty \text{ and } \mu(A) = \sum_{i=1}^\infty \mu(A_i) \in Y.
\]

Define \( P \circ \mu : A \rightarrow [0, \infty] \subset \mathbb{R}_{\infty} \) by, \( \forall A \in A \), \( P \circ \mu(A) := \sup_{n \in \mathbb{Z}_+} (A_i)^n_{i=1} \subseteq A \), \( A = \bigcup_{i=1}^\infty A_i \), \( A \cap A_i = \emptyset \), \( \forall 1 \leq i < j \leq n \sum_{i=1}^n \| \mu(A_i) \| \). \( P \circ \mu \) 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, \( A \subseteq X^2 \) be an algebra on \( X \), \( Y \) be a normed linear space over \( K \), and \( \mu : A \rightarrow Y \) be a \( Y \)-valued pre-measure on the algebra \( A \). Then, \( P \circ \mu : A \rightarrow [0, \infty] \subset \mathbb{R}_{\infty} \) defines a measure on the algebra \( A \).

Proof Clearly, \( P \circ \mu(\emptyset) = 0 \). \( \forall \) pairwise disjoint \( (A_i)^n_{i=1} \subseteq A \) with \( A := \bigcup_{i=1}^\infty A_i \in A \), \( \forall i \in \mathbb{N} \) with \( P \circ \mu(A_i) > 0 \), \( \exists \) \( t_i < P \circ \mu(A_i) \), \( \exists n_i \in \mathbb{Z}_+ \) and \( \exists \) pairwise disjoint \( (A_{i,j})^{n_i}_{j=1} \subseteq A \) with \( A_i = \bigcup_{j=1}^{n_i} A_{i,j} \) such that \( t_i < \sum_{j=1}^{n_i} \| \mu(A_{i,j}) \| \leq P \circ \mu(A_i) \). \( \forall i \in \mathbb{N} \) with \( P \circ \mu(A_i) = 0 \), 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 \( t_i \geq 0 \), \( \forall i \in \mathbb{N} \). 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 \( t - \epsilon < \sum_{i=1}^N \sum_{j=1}^{n_i} \| \mu(A_{i,j}) \| \leq \sum_{i=1}^N \sum_{j=1}^{n_i} \| \mu(A_{i,j}) \| + \| \mu(A \setminus \left( \bigcup_{i=1}^N \bigcup_{j=1}^{n_i} A_{i,j} \right) ) \| \leq P \circ \mu(A) \). 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
\[
M \leq \sum_{i=1}^N \sum_{j=1}^{n_i} \| \mu(A_{i,j}) \| \\
\leq \sum_{i=1}^N \sum_{j=1}^{n_i} \| \mu(A_{i,j}) \| + \| \mu(A \setminus \left( \bigcup_{i=1}^N \bigcup_{j=1}^{n_i} A_{i,j} \right) ) \| \leq P \circ \mu(A)
\]
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 \( (E_j)_{j=1}^{n} \subseteq A \) with \( A = \bigcup_{j=1}^{n} E_j \), we have

\[
\sum_{j=1}^{n} \| \mu(E_j) \| = \sum_{j=1}^{n} \| \mu \left( \bigcup_{i=1}^{\infty} (A_i \cap E_j) \right) \| = \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(A_i \cap E_j) \| \\
\leq \sum_{j=1}^{n} \sum_{i=1}^{\infty} \| \mu(A_i \cap E_j) \| = \sum_{i=1}^{\infty} \sum_{j=1}^{n} \| \mu(A_i \cap E_j) \| \leq \sum_{i=1}^{\infty} P \circ \mu(A_i)
\]

By the arbitrariness of \( n \) and \( (E_j)_{j=1}^{n} \), 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, \( \mathcal{X} := (\bar{X}, \mathcal{B}, \nu) \) be a \( \mathcal{Y} \)-valued measure space, \( \mathcal{Y} \) be a normed linear space, and \( g : X \to \bar{X} \). \( g \) is said to be an isomeasure between \( \mathcal{X} \) and \( \bar{X} \) 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 \( \nu(E) = \mu(g_{\text{inv}}(E)) \).

We will say \( \mathcal{X} \) and \( \bar{X} \) are isomeasuric if there is an isomeasure between \( \mathcal{X} \) and \( \bar{X} \).

Two \( \mathcal{Y} \)-valued measure spaces are isomeasuric, 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 \( \mathcal{X} := (X, \mathcal{B}, \mu) \) be a \( \mathcal{Y} \)-valued topological measure space, \( \bar{X} := (\bar{X}, \mathcal{B}, \nu) \) be a \( \mathcal{Y} \)-valued topological measure space, \( \mathcal{Y} \) be a normed linear space, and \( g : X \to \bar{X} \). \( g \) is said to be a homeomorphical isomeasure between \( \mathcal{X} \) and \( \bar{X} \) if \( g : \mathcal{X} \to \bar{X} \) is a homeomorphism and \( g \) is an isomeasure between \( \mathcal{X} \) and \( \bar{X} \). We will say \( \mathcal{X} \) and \( \bar{X} \) are homeomorphically isomeasuric if there is an homeomorphical isomeasure between \( \mathcal{X} \) and \( \bar{X} \).

Two \( \mathcal{Y} \)-valued topological measure spaces are homeomorphically isomeasuric, then they are equivalent up to relabeling of elements.
Proposition 12.9 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $(\bar{X}, \bar{B})$ 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, $\forall E \in \mathcal{B}$, $g(E) \in \mathcal{B}$. Then, we may define the induced measure $g_\mu : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+$ by $g_\mu(E) = \mu(g_{\text{inv}}(E)), \forall E \in \mathcal{B}$. Then, $\mathcal{X} := (\bar{X}, \bar{B}, g_\mu)$ is a measure space and $g$ is an isomasure between $\mathcal{X}$ and $\bar{X}$.

**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, $(\bar{X}, \bar{B})$ 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, $\forall E \in \mathcal{B}$, $g(E) \in \mathcal{B}$. Then, we may define the induced measure $g_\mu : \mathcal{B} \to [0, \infty] \subset \mathbb{R}_+$ by $g_\mu(E) = \mu(g_{\text{inv}}(E)), \forall E \in \mathcal{B}$. Then, $\mathcal{X} := (\bar{X}, \bar{B}, g_\mu)$ is a measure space and $g_\mu$ is an isomasure between $\mathcal{X}$ and $\bar{X}$.

**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, $(\bar{X}, \bar{B}, g_\mu(\mathcal{P} \circ \mu))$ is a measure space and $g_\mu$ is an isomasure between $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ and $(\bar{X}, \bar{B}, g_\mu(\mathcal{P} \circ \mu))$.

(i) $g_\mu(\emptyset) = \mu(g_{\text{inv}}(\emptyset)) = \mu(\emptyset) = 0_{\mathcal{Y}} \in \mathcal{Y}$.

(ii) $\forall E \in \mathcal{B}$ with $g_\mu(\mathcal{P} \circ \mu)(E) = +\infty$, we have $\mathcal{P} \circ \mu(g_{\text{inv}}(E)) = +\infty$ and $g_{\text{inv}}(E) \in \mathcal{B} \setminus \text{dom } (\mu)$. This implies that $g_\mu(E)$ is undefined.

(iii) $\forall E \in \mathcal{B}$ with $g_\mu(\mathcal{P} \circ \mu)(E) < \infty$, 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 $\left(E_i\right)_{i=1}^{\infty} \subseteq \mathcal{B}$ with $E = \bigcup_{i=1}^{\infty} E_i$, $\forall i \in \mathbb{N}$, $E_i \subseteq \bar{E}$, $g_{\text{inv}}(E_i) \subseteq g_{\text{inv}}(\bar{E})$, $\mathcal{P} \circ \mu(g_{\text{inv}}(E_i)) \leq \mathcal{P} \circ \mu(g_{\text{inv}}(E))$, and $g_{\text{inv}}(E_i) \in \text{dom } (\mu)$. Then, $\sum_{i=1}^{\infty} \Vert g_\mu(E_i) \Vert = \sum_{i=1}^{\infty} \Vert \mu(g_{\text{inv}}(E_i)) \Vert \leq \sum_{i=1}^{\infty} \mathcal{P} \circ \mu(g_{\text{inv}}(E_i)) = \mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} g_{\text{inv}}(E_i)) = \mathcal{P} \circ \mu(g_{\text{inv}}(\bigcup_{i=1}^{\infty} E_i)) = \mathcal{P} \circ \mu(g_{\text{inv}}(E)) < \infty$, 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, $\sum_{i=1}^{\infty} g_\mu(E_i) = \sum_{i=1}^{\infty} \mu(g_{\text{inv}}(E_i)) = \mu(\bigcup_{i=1}^{\infty} g_{\text{inv}}(E_i)) = \mu(g_{\text{inv}}(\bigcup_{i=1}^{\infty} E_i)) = \mu(g_{\text{inv}}(E))$, $\mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} E_i) \subseteq \bigcup_{i=1}^{\infty} \mathcal{P} \circ \mu(E_i)$, where the second equality follows from Definition 11.108; and the third equality follows from Proposition 2.5.

(iv) $\forall E \in \mathcal{B}$ with $g_\mu(\mathcal{P} \circ \mu)(E) < \infty$, by the previous paragraph, we have $\sup_{n \in \mathbb{N}} g_\mu(\mathcal{P} \circ \mu)(E) = \mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} E_i) = \mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} E_i) = \mathcal{P} \circ \mu(\bigcup_{i=1}^{\infty} E_i)$, $\forall E \subseteq \bigcup_{i=1}^{\infty} E_i$, $E_i \subseteq E$, $\forall i \in \mathbb{N}$, $\forall \epsilon > 0$, $\exists \delta > 0$, $\exists (E_i)_{i=1}^{n} \subseteq \mathcal{B}$ pairwise disjoint $\left(E_i\right)_{i=1}^{n} \subseteq \mathcal{B}$ with $g_{\text{inv}}(E) = \bigcup_{i=1}^{n} E_i$, such that $\mathcal{P} \circ \mu(g_{\text{inv}}(E)) - \epsilon < \sum_{i=1}^{n} \Vert \mu(E_i) \Vert \leq \mathcal{P} \circ \mu(g_{\text{inv}}(E))$.

Then, $s_{\mathcal{E}} := \left(g_{\text{inv}}(E_i)\right)_{i=1}^{n} \subseteq \mathcal{B}$ pairwise disjoint, $\bigcup_{i=1}^{n} E_i = g_{\text{inv}}(E)$, and $\sum_{i=1}^{n} \Vert g_\mu(E_i) \Vert = \sum_{i=1}^{n} \Vert \mu(g_{\text{inv}}(E_i)) \Vert = \sum_{i=1}^{n} \Vert g(E_i) \Vert > g_\mu(\mathcal{P} \circ \mu)(E) - \epsilon$. Hence, $s_{\mathcal{E}} \geq g_\mu(\mathcal{P} \circ \mu)(E)$ and $s_{\mathcal{E}} = g_\mu(\mathcal{P} \circ \mu)(E)$.\]
Therefore, by Definition 11.108, \((X, \mathcal{B}, g_* \mu)\) is a \(Y\)-valued measure space with \(P \circ (g_* \mu) = g_* (P \circ \mu)\). It is straightforward to show that \(g\) is an isomeasure between \(\mathcal{X}\) and \(\tilde{\mathcal{X}}\). This completes the proof of the proposition.

\[\square\]

**Proposition 12.11** Let \(\mathcal{X} := (X, \mathcal{O})\) and \(\tilde{\mathcal{X}} := (\tilde{X}, \hat{\mathcal{O}})\) be topological spaces, \(g: \mathcal{X} \rightarrow \tilde{\mathcal{X}}\) be a homeomorphism, \((Y, \mathcal{B}, \mu)\) be a \(Y\)-valued) topological measure space. Then, \(X := (\tilde{X}, \mathcal{B}_\mu(\tilde{X}), g_* \mu)\) is a \(Y\)-valued) topological measure space and \(g\) is an isomeasure between \(\mathcal{X}\) and \(\tilde{\mathcal{X}}\).

**Proof** Clearly, \(g\) is bijective and \(g\) and \(g_{\text{inv}}\) are continuous. By Proposition 11.34, \(\forall E \in \mathcal{B}_\mu(\mathcal{X})\), we have \(g(E) \in \mathcal{B}_\mu(\tilde{\mathcal{X}})\); and \(\forall E \in \mathcal{B}(\mathcal{X})\), we have \(g_{\text{inv}}(E) \in \mathcal{B}(\tilde{\mathcal{X}})\). Then, by Propositions 12.9 and 12.10, the \(Y\)-valued) induced measure \(g_* \mu\) is well defined, \((X, \mathcal{B}_\mu(\mathcal{X}), g_* \mu)\) is a \(Y\)-valued) measure space, and \(g\) is an isomeasure between \(\mathcal{X}\) and \(\tilde{\mathcal{X}}\).

We will show that \(\tilde{\mathcal{X}}\) is a \(Y\)-valued topological measure space. \(\forall E \in \mathcal{B}_\mu(\mathcal{X})\), \(\forall \varepsilon \in (0, \infty) \subset \mathbb{R}\), we have \(g_{\text{inv}}(E) \in \mathcal{B}_\mu(\mathcal{X})\). Since \(\mathcal{X}\) is a \(Y\)-valued) topological measure space, then \(\exists U \in \mathcal{O}\) with \(g_{\text{inv}}(E) \subseteq U\) such that \(\mu(U \setminus g_{\text{inv}}(E)) < \varepsilon\). Let \(\tilde{U} := g(U) \in \hat{\mathcal{O}}\). Then, \(E \subseteq g(U) = \tilde{U}\) and \(g_* \mu(U \setminus E) = \mu(g_{\text{inv}}(U \setminus \tilde{E})) = \mu(U \setminus g_{\text{inv}}(E)) = \mu(U \setminus g_{\text{inv}}(E)) < \varepsilon\) (in the case when \(\mu\) is a \(Y\)-valued measure, we have \(P \circ (g_* \mu)(U \setminus \tilde{E}) = P \circ \mu(g_{\text{inv}}(U \setminus \tilde{E})) = P \circ \mu(g_{\text{inv}}(U \setminus g_{\text{inv}}(E))) = P \circ \mu(U \setminus g_{\text{inv}}(E)) < \varepsilon\), where the first equality follows from Proposition 12.10). Hence, \(\tilde{\mathcal{X}}\) is a \(Y\)-valued topological measure space. Then, \(g\) is a homeomorphical isomeasure between \(\mathcal{X}\) and \(\tilde{\mathcal{X}}\).

\[\square\]

**Proposition 12.12** Let \(\mathcal{X} := (X, \mathcal{B}, \mu)\) be a measure space, \(\tilde{\mathcal{X}} := (\tilde{X}, \mathcal{B}, \nu)\) be a measure space, \((Y, \mathcal{B}, \mu)\) be a \(\mathcal{B}\)-measurable, and \(g: X \rightarrow Y\) be an isomeasure between \(\mathcal{X}\) and \(\tilde{\mathcal{X}}\). Then, \(\int_X (f \circ g) \, d\mu = \int_{\tilde{X}} f \, d\nu\), 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: \(\mu(X) = +\infty\). \(\nu(\tilde{X}) = \mu(g_{\text{inv}}(\tilde{X})) = \mu(X) = +\infty\). Let \(\mathcal{I}(Y)\) be the integration system on \(Y\). \(\forall R = \{ (y_\alpha, U_\alpha) \mid \alpha \in \Lambda \} \in \mathcal{I}(Y)\), let \(F_R := \sum_{\alpha \in \Lambda} y_\alpha \nu(f_{\text{inv}}(U_\alpha)) \in \mathcal{Y}\) and \(F_R := \sum_{\alpha \in \Lambda} y_\alpha \mu((f \circ g)_{\text{inv}}(U_\alpha)) \in \mathcal{Y}\). Note that \(F_R = \sum_{\alpha \in \Lambda} y_\alpha \mu(g_{\text{inv}}(f_{\text{inv}}(U_\alpha))) = F_R\). Then, by Definition 11.70, we have \(\int_{\tilde{X}} f \, d\nu = \lim_{R \in \mathcal{I}(Y)} F_R = \lim_{R \in \mathcal{I}(Y)} F_R = \int_X (f \circ g) \, d\mu\), whenever one of the integrals exists.

Case 2: \(\mu(X) = +\infty\). \(\nu(\tilde{X}) = \mu(g_{\text{inv}}(\tilde{X})) = \mu(X) = +\infty\). Let \((F_A)_{A \in \mathcal{M}(\tilde{X})}\) be the net for \(\int_{\tilde{X}} f \, d\nu\) and \((F_A)_{A \in \mathcal{M}(X)}\) be the net for \(\int_X (f \circ g) \, d\mu\) as defined in Definition 11.71. Without loss of generality,
assume that $\int_X f \, d\nu$ exists. Then,

$$\int_X f \, d\nu = \lim_{A \in \mathcal{M}(\bar{X})} \bar{F}_A = \lim_{A \in \mathcal{M}(X)} \int_A f \, d\nu$$

$$= \lim_{A \in \mathcal{M}(\bar{X})} \int_{\mathcal{g}\text{inv}(\bar{A})} (f \circ g) \, d\mu_{\mathcal{g}\text{inv}(\bar{A})},$$

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 $Y = \mathbb{R}$ or $U \subseteq Y$ if $Y \neq \mathbb{R}$) with $\int_X f \, d\nu \in U$, $\exists A_0 \in \mathcal{M}(\bar{X})$, $\forall \bar{A} \in \mathcal{M}(\bar{X})$ with $\bar{A}_0 \subseteq \bar{A}$, $F_{\bar{A}} = \int_{\mathcal{g}\text{inv}(\bar{A})} (f \circ g) \, d\mu_{\mathcal{g}\text{inv}(\bar{A})} \in U$. $\forall \bar{A} \in \mathcal{M}(\bar{X})$ with $\mathcal{g}\text{inv}(A_0) \subseteq \bar{A}$, we have $A \in \mathcal{B}$ and $\mu(A) < +\infty$. Then, $g(A) \in \mathcal{B}$ and $\mu(A) = \nu(g(A)) < +\infty$. Hence, $g(A) \in \mathcal{M}(\bar{X})$ and $g(A) \subseteq g(\mathcal{g}\text{inv}(A_0)) = \bar{A}_0$. Then, $F_A = \int_{\bar{A}} (f \circ g) \, d\mu_{\bar{A}} = \int_{\mathcal{g}\text{inv}(\bar{A})} (f \circ g) \, d\mu_{\mathcal{g}\text{inv}(\bar{A})} \subseteq \mathcal{g}\text{inv}(\bar{A}) \in U$. Hence, $\int_X (f \circ g) \, d\mu = \lim_{A \in \mathcal{M}(X)} F_A = \int_X f \, d\nu$. This completes the proof of the proposition.

\[\square\]

**Proposition 12.13** Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a measure space, $\bar{\mathcal{X}} := (\bar{X}, \mathcal{B}, \nu)$ be a measure space, $\mathcal{Y}$ be a Banach space, $\mathcal{W}$ be a separable subspace of $\mathcal{Y}$, $f : \bar{X} \to \mathcal{W}$ be absolutely integrable over $\bar{X}$, and $g : X \to \bar{X}$ be an isomorphism between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $f \circ g$ is absolutely integrable over $\mathcal{X}$ and $\int_X (f \circ g) \, d\mu = \int_X f \, d\nu \in \mathbb{Y}$.

**Proof** Clearly, $f \circ g$ is $\mathcal{B}$-measurable. By the assumption, $0 \leq \int_X (f \circ g) \, d\mu < +\infty$. By Proposition 12.12, $\int_X (f \circ g) \, d\mu = \int_X f \, d\nu \in [0, \infty) \subseteq \mathbb{R}$. Hence, $f \circ g$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, $f$ is integrable over $\bar{X}$. By Proposition 12.12, $\int_X (f \circ g) \, d\mu = \int_X f \, d\nu \in \mathbb{Y}$. This completes the proof of the proposition.

\[\square\]

**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 $\mathcal{K}$, $\bar{\mathcal{X}} := (\bar{X}, \mathcal{B}, \nu)$ be a $\mathcal{Y}$-valued measure space, $\mathcal{Z}$ be a normed linear space over $\mathcal{K}$, $\mathcal{W} := \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \to \mathcal{W}$ be $\mathcal{B}$-measurable, and $g : X \to \bar{X}$ is an isomorphism between $\mathcal{X}$ and $\bar{\mathcal{X}}$. Then, $\int_X f \, d\gamma = \int_X (f \circ g) \, d\mu$, 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: $\mathcal{P} \circ \mu(X) = +\infty$. $\mathcal{P} \circ \nu(\bar{X}) = \mathcal{P} \circ \mu(\mathcal{g}\text{inv}(\bar{X})) = \mathcal{P} \circ \mu(X) = +\infty$. Let $\mathcal{J} \subseteq \mathcal{W}$ be the integration system on $\mathcal{W}$. $\forall \mathcal{R} = \{(w, U) : w \in \mathbb{R} \} \subseteq \mathcal{J} \subseteq \mathcal{W}$, let $F_R := \sum_{\alpha \in \Lambda} w_\alpha \nu(\mathcal{g}\text{inv}(U)) \subseteq \mathcal{Z}$ and $F_R := \sum_{\alpha \in \Lambda} w_\alpha \mu((f \circ g)\text{inv}(U)) \subseteq \mathcal{Z}$. Then, $F_R = \sum_{\alpha \in \Lambda} w_\alpha \mu(\mathcal{g}\text{inv}(f\text{inv}(U))) = F_R$. Then, by Definition 11.119, we have $\int_X f \, d\nu = \lim_{\mathcal{R} \in \mathcal{J} \subseteq \mathcal{W}} F_R = \lim_{\mathcal{R} \in \mathcal{W}} F_R = \int_X (f \circ g) \, d\mu$, whenever one of the integrals exists.

Case 2: $\mathcal{P} \circ \mu(X) = +\infty$. $\mathcal{P} \circ \nu(X) = \mathcal{P} \circ \mu(\mathcal{g}\text{inv}(X)) = \mathcal{P} \circ \mu(X) = +\infty$. Let $\{F_{\bar{A}}\}_{\bar{A} \in \mathcal{M}(\bar{X})}$ be the net for $\int_X f \, d\nu$ and $\{F_A\}_{A \in \mathcal{M}(X)}$ be the net for
is some measure between $I_K$ be a separable subspace of $W$ space, where

Theorem 12.16 (Change of Variable)

Let $f$ is absolutely integrable over $X$. Hence, $\int_X (f \circ g) \, d\mu = \lim_{A \to \mathcal{M}(X)} \int_{g_{inv}(A)} (f \circ g) \, d\mu_{g_{inv}(A)}$

where the first two equalities follow from Definition 11.120; and the third equality follows from Case 1. Fix any open set $U (U \subseteq \mathbb{R}$ if $\mathcal{Z} = \mathbb{R}$ or $U \subseteq \mathcal{Z}$ if $\mathcal{Z} \neq \mathbb{R})$ with $\int_X f \, d\nu = \int_X f \, d\nu$ and $g_{inv}(A) \subseteq A$, we have $A \in \mathcal{B}$ and $\mathcal{P} \circ \mu(A) = \mathcal{P} \circ \mu(A) < +\infty$. Hence, $g(A) \in \mathcal{M}(\mathcal{X})$ and $g(A) \geq ginv(A) \in \mathcal{A}_0$. Then, $F = \int_A (f \circ g) |A \, d\mu_A = \int_{g_{inv}(g(A))} (f \circ g) |g_{inv}(g(A)) \, d\mu_{g_{inv}(g(A))}$ is absolutely integrable over $X$. Hence, $\int_X (f \circ g) \, d\mu = \lim_{A \to \mathcal{M}(X)} F = \int_X f \, d\nu$.

This completes the proof of the proposition.

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 $\mathcal{K}$, $\bar{X} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a $\mathcal{Y}$-valued measure space, $\mathcal{Z}$ be a Banach space over $\mathcal{K}$, $W$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \to W$ be absolutely integrable over $\bar{X}$, and $g : X \to \bar{X}$ be an isomorphism between $\mathcal{X}$ and $\bar{X}$. Then, $f \circ g$ is absolutely integrable over $\mathcal{X}$, and $\int_X f \, d\nu = \int_X (f \circ g) \, d\mu \in \mathcal{Z}$.

Proof Clearly, $f \circ g$ is $\mathcal{B}$-measurable. By Definition 12.7, $g$ is an isomorphism between $(X, \mathcal{B}, \mathcal{P} \circ \mu)$ and $(\bar{X}, \bar{\mathcal{B}}, \mathcal{P} \circ \nu)$. By Proposition 12.12, $\int_X (f \circ g) \, d\nu = \int_X (f \circ g) \, d\nu \in \mathcal{K}$, where the last step follows from the assumption that $f$ is absolutely integrable over $\bar{X}$. Hence, $f \circ g$ is absolutely integrable over $\mathcal{X}$. By Proposition 12.14, $\int_X (f \circ g) \, d\mu = \int_X f \, d\nu \in \mathcal{Z}$. □

Theorem 12.16 (Change of Variable) Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite $\mathcal{K}$-valued measure space, $\bar{X} := (\bar{X}, \bar{\mathcal{B}}, \nu)$ be a $\sigma$-finite $\mathcal{Y}$-valued measure space, where $\mathcal{Y}$ is a separable Banach space over $\mathcal{K}$, $g : \bar{X} \to X$ be an isomorphism between $\mathcal{X}$ and $\bar{X} := (X, \mathcal{B}, g_{new})$, $\mathcal{Z}$ be a Banach space over $\mathcal{K}$, $W$ be a separable subspace of $\mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $f : \bar{X} \to W$ be $\mathcal{B}$-measurable, and $\frac{d(g_{new})}{d\mu} = h : X \to \mathcal{Y}$. Then, $f$ is absolutely integrable over $\bar{X}$ if, and only if, $(P \circ f \circ g_{new})(P \circ h)$ is integrable over $(X, \mathcal{B}, P \circ \mu)$. In which case, $(f \circ g_{new})h$ is absolutely integrable over $\bar{X}$ and $\int_X f \, d\nu = \int_X (f \circ g_{new})h \, d\mu \in \mathcal{Z}$.

Proof By Proposition 12.15, $f$ is absolutely integrable over $\bar{X}$ if, and only if, $f \circ g_{inv}$ is absolutely integrable over $\bar{X}$. By (ii) of Proposition 11.168, $f \circ g_{inv}$ is absolutely integrable over $\bar{X}$ if, and only if, $(P \circ f \circ g_{inv})(P \circ h)$ is integrable over $(X, \mathcal{B}, P \circ \mu)$. Hence, $f$ is absolutely integrable over $\bar{X}$ if,
and only if, \((P \circ f \circ g_{inv})(P \circ h)\) is integrable over \((X, B, P \circ \mu)\). In this case, 
\((f \circ g_{inv})h\) is absolutely integrable over \(X\) and, by Proposition 12.15 and (ii) 
of Proposition 11.168, 
\[
\int_X f \, d\nu = \int_X (f \circ g_{inv}) \, d(g_*\nu) = \int_X (f \circ g_{inv})h \, d\mu \in \mathbb{Z}.
\]
This completes the proof of the theorem. \(\square\)

12.3 Product Measure

**Definition 12.17** Let \(X\) be a set. \(\mathcal{A} \subseteq \mathcal{P}(X)^2\) is said to be a \(\pi\)-system on \(X\) if: (i) \(\emptyset, X \in \mathcal{A}\); and (ii) \(\forall A, B \in \mathcal{A}\), we have \(A \cap B \in \mathcal{A}\). \(\mathcal{D} \subseteq \mathcal{P}(X)^2\) is said to be a monotone class on \(X\) if

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

(ii) \(\forall A, B \in \mathcal{D}\) 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 \(\mathcal{A} \subseteq \mathcal{P}(X)^2\). 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. \(\mathcal{B} \subseteq \mathcal{P}(X)^2\) 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 \(X \setminus A = \overline{A} \in \mathcal{B}\), since \(\mathcal{B}\) is a monotone class on \(X\). \(\forall A_1, A_2 \in \mathcal{B}\), we have \(A_1 \cup A_2 = \overline{A_1 \cap A_2}\). 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 \(\bigcup_{n=1}^\infty A_n = \bigcup_{i=1}^\infty E_i \in \mathcal{B}\). This shows that \(\mathcal{B}\) is a \(\sigma\)-algebra.

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

**Lemma 12.19 (Monotone Class Lemma)** Let \(X\) be a set, \(\mathcal{A} \subseteq \mathcal{P}(X)^2\) 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 \(\mathcal{D} = \mathcal{D}_1\).
Clearly, $\emptyset \in A \subseteq D$ and $X \in A \subseteq D$. For all $E_1, E_2 \in D$ with $E_1 \subseteq E_2$, we have $E_1, E_2 \in D$ and $E_2 \setminus E_1 \in D$. For all $A \subseteq D$, 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 $E_2 \setminus E_1 \in D$. For all $i \in \mathbb{N}$, we have $\bigcup_{i=1}^{\infty} E_i \in D$ since $D$ is a monotone class on $X$. For all $A \subseteq X$, all $i \in \mathbb{N}$, $E_i \in D$, implies that $A \cap E_i \in D$. Clearly, $A \cap E_i \subseteq A \cap E_{i+1}$, all $i \in \mathbb{N}$. 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 $\bigcup_{i=1}^{\infty} E_i \in D$. This shows that $D$ is a monotone class on $X$.

Define $D_2 := \{ E \in D \mid E \cap B \in D, \forall B \in D \}$. 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 $X \in A \subseteq D_2$. For all $E_1, E_2 \in D_2$ with $E_1 \subseteq E_2$, we have $E_2 \setminus E_1 \in D$ since $D$ is a monotone class on $X$. For all $B \in D$, we have $E_1 \cap B \in D$ and $E_2 \cap B \in D$ since $E_1, E_2 \in D_2$. Clearly, $E_1 \cap B \subseteq E_2 \cap B$. Then, $D \ni (E_2 \cap B) \setminus (E_1 \cap B) = (E_2 \setminus E_1) \cap B$, since $D$ is a monotone class on $X$. By the arbitrariness of $B$, we have $E_2 \setminus E_1 \in D_2$. For all $i \in \mathbb{N}$, we have $\bigcup_{i=1}^{\infty} E_i \in D$ since $D$ is a monotone class on $X$. For all $B \in D$, all $i \in \mathbb{N}$, $E_i \in D_2$ implies that $B \cap E_i \in D$. Clearly, $B \cap E_i \subseteq B \cap E_{i+1}$, all $i \in \mathbb{N}$. 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 $\bigcup_{i=1}^{\infty} E_i \in D_2$. 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.

**Proposition 12.20** Let $X$ be a set, $\mathcal{C}$ be a semialgebra on $X$, $\mathcal{Y}$ be a normed linear space, $\mathcal{A}$ be the algebra on $X$ generated by $\mathcal{C}$, $\nu : \mathcal{C} \rightarrow [0, \infty] \subset \mathbb{R}_+$, and $\mu$ be a function from $\mathcal{C}$ to $\mathcal{Y}$. Assume that

(i) $\mu(\emptyset) = \emptyset y$ and $\nu(\emptyset) = 0$;

(ii) For all $C \in \mathcal{C}$, all $n \in \mathbb{Z}_+$, all 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) For all $C \in \mathcal{C}$, all 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) For all $C \in \mathcal{C}$ with $\nu(C) = \infty$, $\mu(C)$ is undefined;

(v) For all $C \in \mathcal{C}$ with $\nu(C) < \infty$, $\mu(C) \in \mathcal{Y}$ and all $n \in \mathbb{Z}_+$, all 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) For all $C \in \mathcal{C}$ with $\nu(C) < \infty$, we have

$$\nu(C) = \sup_{n \in \mathbb{Z}_+, (C_i)_{i=1}^{n} \subseteq \mathcal{C}, C = \bigcup_{i=1}^{n} C_i, C_i \cap C_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \| \mu(C_i) \|$$
Then, \( \nu \) admits a unique extension to a measure \( \bar{\nu} \) on the algebra \( \mathcal{A} \); and \( \mu \) admits a unique extension to a \( \mathbb{Y} \)-valued measure \( \bar{\mu} \) on the algebra \( \mathcal{A} \) with \( (\mathcal{P} \circ \bar{\mu})|_\mathcal{C} = \nu \). Furthermore, \( \mathcal{P} \circ \bar{\mu} = \bar{\nu} \).

**Proof**  
By Proposition 11.31, \( \mathcal{A} = \{ A \subseteq X \mid A \) is the finite disjoint union of sets in \( \mathcal{C} \} \). Define \( \bar{\nu} : \mathcal{A} \to [0, \infty) \subseteq \mathbb{R}_e \) 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 \( \mathbb{Y} \) by: \( \bar{\mu}(A) \) is undefined if \( \bar{\nu}(A) = \infty \); and \( \bar{\mu}(A) = \sum_{i=1}^{n} \mu(C_i) \) 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, \( \bar{\mu}|_\mathcal{C} = \mu \) and hence \( \bar{\mu} \) is an extension of \( \mu \).

We will show that \( \bar{\mu} \) is a \( \mathbb{Y} \)-valued measure on the algebra \( \mathcal{A} \) with \( \mathcal{P} \circ \bar{\mu} = \bar{\nu} \). Then, \( \bar{\mu} \) is the \( \mathbb{Y} \)-valued measure on the algebra \( \mathcal{A} \) that we seek. By (i), \( \bar{\mu}(\emptyset) = \emptyset_y \). \( \forall A \in \mathcal{A} \) with \( \bar{\nu}(A) = \infty \), \( \bar{\mu}(A) \) is undefined. \( \forall A \in \mathcal{A} \) with \( \bar{\nu}(A) < \infty \), \( \forall \) pairwise disjoint \( (A_i)_{i=1}^{n} \subseteq \mathcal{A} \) with \( A = \bigcup_{i=1}^{n} A_i \), we have \( \bar{\mu}(A) \in \mathbb{Y} \), \( \bar{\nu}(A_i) < \infty \), and \( \bar{\mu}(A_i) \in \mathbb{Y} \), \( \forall i \in \mathbb{N} \). 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 \( \sum_{k=1}^{\infty} ||\bar{\mu}(A_i)|| = \sum_{k=1}^{\infty} \left( \sum_{j=1}^{n_i} \mu(C_{i,j}) \right) \leq \sum_{k=1}^{\infty} \sum_{j=1}^{n_i} \mu(C_{i,j}) \leq \sum_{k=1}^{\infty} \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{k=1}^{\infty} \sum_{j=1}^{n_i} \bar{\nu}(C_{i,j}) = \bar{\nu}(\bigcup_{k=1}^{\infty} \bigcup_{j=1}^{n_k} C_{i,j}) = \bar{\nu}(A) < \infty \), where the first equality follows from the definition of \( \bar{\mu} \); the second inequality follows from (vi); the second equality follows from (vi); the seventh equality follows from the fact that \( \bar{\nu} \) is a measure on the algebra \( \mathcal{A} \). Hence, \( \bar{\mu}(A) = \sum_{i=1}^{\infty} \bar{\mu}(A_i) \).
∀A ∈ A with \( \tilde{\nu}(A) < \infty \), then \( A = \bigcup_{i=1}^{n_0} C_i \), where \( n_0 \in \mathbb{Z}_+ \) and \( (C_i)_{i=1}^{n_0} \subseteq \mathcal{C} \) is pairwise disjoint. ∀\( n \in \mathbb{Z}_+ \), ∀ pairwise disjoint \( (A_i)_{i=1}^{n} \subseteq A \) with \( A = \bigcup_{i=1}^{n} A_i \), ∀i \in \{1, \ldots, n\}, \( A_i \in A \) 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.

Then, \( \sum_{i=1}^{n} \| \tilde{\mu}(A_i) \| = \sum_{i=1}^{n} \| \sum_{j=1}^{n_i} \mu(C_{i,j}) \| \leq \sum_{i=1}^{n} \sum_{j=1}^{n_i} \| \mu(C_{i,j}) \| \leq \sum_{i=1}^{n} \sum_{j=1}^{n_i} \nu(C_{i,j}) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \tilde{\nu}(C_{i,j}) = \sum_{i=1}^{n} \tilde{\nu}(A_i) = \tilde{\nu}(A) \), where the first equality follows from the definition of \( \tilde{\mu} \); the second inequality follows from (vi); the second equality follows from the fact that \( \tilde{\nu} \) is a measure on the algebra \( \mathcal{A} \). Hence, \( s_A := \sup \{ \tilde{\nu}(A) : A = \bigcup_{i=1}^{n} A_i, A \in \mathcal{A}, \bigcap_{i=1}^{n} A_i = \emptyset, \forall i \} \), \( \tilde{\nu} \) is a measure on the algebra \( \mathcal{A} \) that is an extension of \( \tilde{\nu} \), and \( \tilde{\nu} = \mathcal{P} \circ \tilde{\mu} \) by Proposition 11.32. ∀\( A \in \mathcal{A} \) with \( \tilde{\nu}(A) = \infty \), then \( \tilde{\nu} \) and \( \tilde{\mu}(A) \) are both undefined. ∀\( A \in \mathcal{A} \) with \( \tilde{\nu}(A) < \infty \), then \( A = \bigcup_{i=1}^{n} C_i \), where \( n \in \mathbb{Z}_+ \) and \( (C_i)_{i=1}^{n} \subseteq \mathcal{C} \) is pairwise disjoint. Then, \( \tilde{\mu}(A) = \sum_{i=1}^{n} \mu(C_i) = \sum_{i=1}^{n} \tilde{\mu}(C_i) = \tilde{\mu}(A) \), where the first equality follows from the fact that \( \tilde{\mu} \) is a \( \mathcal{C} \)-valued measure on the algebra \( \mathcal{A} \); the second equality follows from the definition of \( \tilde{\mu} \). Hence, \( \tilde{\mu} = \mu \). Therefore, \( \tilde{\mu} \) 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 \( \mathbb{K} \)-valued measure space, ∀\( j \in \{1, \ldots, m\} \), \( \mathcal{C}_m := \{ \prod_{j=1}^{m} E_j \subseteq \prod_{j=1}^{m} X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \ldots, m\}\} \), 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 \( \mu(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mu_j(E_j) \), ∀\( \prod_{j=1}^{m} E_j \in \mathcal{C}_m \), and \( \nu : \mathcal{C}_m \to [0, \infty) \subset \mathbb{R} \) be defined by \( \nu(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mathcal{P} \circ \mu_j(E_j) \), ∀\( \prod_{j=1}^{m} E_j \in \mathcal{C}_m \). Then, \( \mathcal{C}_m \) is a semialgebra on \( \prod_{j=1}^{m} X_j \), \( \mu \) and \( \nu \) satisfy the assumptions of Proposition 12.20. Therefore, there exists a unique \( \mathbb{K} \)-valued measure \( \lambda := \prod_{j=1}^{m} \mu_j \) on the measurable space \( (\prod_{j=1}^{m} X_j, \mathcal{B}_m, \mathcal{C}_m) \) such that \( \lambda(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \lambda_j(E_j) \) and \( \mathcal{P} \circ \lambda(\prod_{j=1}^{m} E_j) = \prod_{j=1}^{m} \mathcal{P} \circ \mu_j(E_j) \), ∀\( \prod_{j=1}^{m} E_j \in \mathcal{C}_m \). Furthermore, \( \lambda \) is finite. The \( \mathbb{K} \)-valued measure space \( \mathcal{X} := (\prod_{j=1}^{m} X_j, \mathcal{B}_m, \lambda) \) is said to be the product \( \mathbb{K} \)-valued measure space of \( X_1, \ldots, X_m \), and denoted by \( \prod_{j=1}^{m} \mathcal{X}_j \). \( \lambda \) is said to be the product \( \mathbb{K} \)-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 \( \prod_{j=1}^1 \mu_j = \mu_1 \).

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.

\[ \forall \prod_{j=1}^k E_{j,1}, \prod_{j=1}^k E_{j,2} \in \mathcal{C}_k, \text{ we have } (\prod_{j=1}^k E_{j,1}) \cap (\prod_{j=1}^k E_{j,2}) = \prod_{j=1}^k (E_{j,1} \cap E_{j,2}) \in \mathcal{C}_m. \]

Note that \( (\prod_{j=1}^k E_{j,1}) \sim (\prod_{j=1}^{k-1} E_{j,1}) \times E_{k,1} \cup (\prod_{j=1}^{k-1} E_{j,1}) \times E_{k,1} ^\sim \cup (\prod_{j=1}^{k-1} E_{j,1}) \times E_{k,1} \). By inductive assumption, \( \prod_{j=1}^{k-1} E_{j,1} \in \mathcal{C}_{k-1} \) and \( \mathcal{C}_{k-1} \) is a semialgebra on \( \prod_{j=1}^{k-1} X_j \), then \( (\prod_{j=1}^{k-1} E_{j,1}) \sim \) is a finite disjoint union of sets in \( \mathcal{C}_{k-1} \). Hence, \( (\prod_{j=1}^{k-1} E_{j,1}) \sim \) is a finite disjoint union of sets in \( \mathcal{C}_k \). Clearly, \( \emptyset = \prod_{j=1}^k \emptyset \in \mathcal{C}_k \neq \emptyset \). Hence, \( \mathcal{C}_k \) is a semialgebra on \( X_1 \times X_2 \).

Next, we will show that \( \mu \) and \( \nu \) satisfy (i) – (vi) of Proposition 12.20.

\[ \mu(\emptyset) = \mu(\prod_{j=1}^k \emptyset) = \prod_{j=1}^k \mu_j(\emptyset) = 0 \text{ and } \nu(\emptyset) = \nu(\prod_{j=1}^k \emptyset) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(\emptyset) = 0. \]

Hence, (i) of Proposition 12.20 holds.

(iv) of Proposition 12.20 holds trivially since \( \nu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) \leq \prod_{j=1}^k \mathcal{P} \circ \mu_j(X_j) < \infty, \forall \prod_{j=1}^k E_j \in \mathcal{C}_k. \)

\[ \forall \prod_{j=1}^k E_j \in \mathcal{C}_k, \forall \text{ pairwise disjoint } \left( \prod_{j=1}^k E_{j,i} \right)_{i=1}^\infty \subseteq \mathcal{C}_k \text{ with } \prod_{j=1}^k E_j = \bigcup_{i=1}^\infty \left( \prod_{j=1}^k E_{j,i} \right). \]

Fix any \( x_k \in E_k. \) \( \forall (x_1, \ldots, x_{k-1}) \in \prod_{j=1}^{k-1} E_j, \exists i \in \mathbb{N} \text{ such that } (x_1, \ldots, x_k) \in \prod_{j=1}^k E_{j,i}. \)

Then, \( \prod_{j=1}^k E_{j,i} \) and the sets in the union are pairwise disjoint.

Then, \( \sum_{i \in \mathbb{N}} \chi_{E_{k,i} \cdot X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i \in \mathbb{N}, \prod_{j=1}^k E_{j,i}} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i \in \mathbb{N}, \prod_{j=1}^k E_{j,i}} \mathcal{P} \circ \mu_j(E_{j,i}) = \chi_{E_k \cdot X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) \in [0, \infty) \subset \mathbb{R}, \) where the last four equalities follow from the inductive assumption. \( \forall x_k \in X_k \setminus E_k, \) we have \( \chi_{E_k \cdot X_k}(x_k) \cdot \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) = 0 = \sum_{i \in \mathbb{N}} \chi_{E_{k,i} \cdot X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}). \)

Then, \( \forall x_k \in X_k, \chi_{E_k \cdot X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_j) = \sum_{i \in \mathbb{N}} \chi_{E_{k,i} \cdot X_k}(x_k) \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}). \) By Monotone Convergence Theorem 11.81 and Proposition 11.75, \( \nu(\prod_{j=1}^k E_j) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) = \int_{X_k} \chi_{E_k \cdot X_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) \, d\mathcal{P} \circ \mu_k = \int_{X_k} \sum_{i \in \mathbb{N}} \chi_{E_{k,i} \cdot X_k} \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) \, d\mathcal{P} \circ \mu_k = \sum_{i=1}^\infty (\prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i})) \mathcal{P} \circ \mu_k(E_{k,i}) = \sum_{i=1}^\infty \prod_{j=1}^{k-1} \mathcal{P} \circ \mu_j(E_{j,i}) = \sum_{i=1}^\infty \nu(\prod_{j=1}^k E_{j,i}). \) 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, \( \prod_{j=1}^{k-1} E_j = \bigcup_{i \in N, x_k \in E_k} \prod_{j=1}^{k-1} E_{j,i} \) and the sets in the union are pairwise disjoint. Then, \( \sum_{i \in N} \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \lambda_j(E_{j,i}) = \sum_{i \in N, x_k \in E_k} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \)

\[ \sum_{i \in N, x_k \in E_k} \lambda \left( \prod_{j=1}^{k-1} E_{j,i} \right) = \lambda \left( \bigcup_{i \in N} \prod_{j=1}^{k-1} E_{j,i} \right) = \lambda \left( \prod_{j=1}^{k-1} E_j \right) = \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \in \mathbb{K} \], where the last four equalities follow from the inductive assumption. \( \forall x_k \in X_k \setminus E_k \), we have \( \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = 0 = \sum_{x_k \in X_k \setminus E_k} \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \). Then, \( \forall x_1 \in X_1 \), \( \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \prod_{j=1}^{k-1} \mu_j(E_{j,i}) = \sum_{i \in N} \chi_{E_k,i}(x_k) \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \). By Propositions 11.125 and 11.132, we have \( \mu(\prod_{j=1}^{k} E_j) = \prod_{j=1}^{k} \mu_j(E_{j,i}) = \int_{X_k} \chi_{E_k,i} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \)

\[ \int_{X_k} \left( \sum_{i \in N} \chi_{E_k,i} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \right) d\mu_k = \sum_{i \in N} \prod_{j=1}^{k} \mu_j(E_{j,i}) = \sum_{i \in N, x_k \in E_k} \prod_{j=1}^{k-1} \mu_j(E_{j,i}) \]. 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 \( \sum_{i=1}^{n} \mu(\prod_{j=1}^{k} E_{j,i}) = \)

\[ \sum_{i=1}^{n} \prod_{j=1}^{k} \mu_j(E_{j,i}) = \sum_{i=1}^{n} \prod_{j=1}^{k} \mu_j(E_{j,i}) \] and the last equality follows from Definition 11.108; and the last equality follows from (ii). Then, \( \nu(\prod_{j=1}^{k} E_j) \geq s_{\prod_{j=1}^{k} E_j} := \sup_{n \in \mathbb{Z}_+, \prod_{j=1}^{k} E_{j,i} \subseteq \mathcal{E}_k} \sum_{i=1}^{n} \mu(\prod_{j=1}^{k} E_{j,i}) \) on the other hand, \( \forall \epsilon \in (0,1) \subseteq \mathbb{R} \), \( \forall j \in \{1, \ldots, k\}, \exists n_{j} \in \mathbb{Z}_+ \), \( \exists \prod_{j=1}^{k} E_{j,i} \subset \mathcal{E}_k \), \( \forall \epsilon \in (0,1) \subseteq \mathbb{R} \), \( \nu(\prod_{j=1}^{k} E_{j,i}) \) and the sets in the union are pairwise disjoint and in \( \mathcal{E}_k \). Then, \( \sum_{i=1}^{n_{j}} \prod_{j=1}^{k} \mu_j(E_{j,i}) = \sum_{i=1}^{n_{j}} \prod_{j=1}^{k} \mu_j(E_{j,i}) \geq (1 - \frac{1 + \nu(\prod_{j=1}^{k} E_{j,i}))}{1 + \nu(\prod_{j=1}^{k} E_{j,i})}) \prod_{j=1}^{k} \mu_j(E_{j,i}) \) 

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 \( \tilde{\nu} \) on the algebra \( A_k \); and \( \mu \) admits a unique extension to a \( \mathbb{K} \)-valued measure \( \tilde{\mu} \) on the
algebra \( \mathcal{A}_k \) with \( (\mathcal{P} \circ \mu)|_{\mathcal{C}_k} = \nu \). Furthermore, \( \mathcal{P} \circ \mu = \nu \). Clearly, \( \mu \) 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 \( \lambda|_{\mathcal{A}} = \mu \) and \( (\mathcal{P} \circ \lambda)|_{\mathcal{A}} = \mathcal{P} \circ \mu = \nu \). 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, \( \forall j \in \{1, \ldots, k\} \), \( \mathcal{C} := \{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_j \mid E_j \in \mathcal{B}_j, \forall j \in \{1, \ldots, k\} \} \) 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 \( \mathcal{X} := (\prod_{j=1}^k X_j, \mathcal{B}, \lambda) \) 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 \( X_j, \forall n \in \mathbb{N} \). Fix any \( n_1, \ldots, n_k \in \mathbb{N} \). Let \( \mathcal{C}_{n_1, \ldots, n_k} := \{ \prod_{j=1}^k E_j \subseteq \prod_{j=1}^k X_{j,n_j} \mid E_j \in \mathcal{B}_{j,n_j}, j = 1, \ldots, k \} \) and \( \mathcal{B}_{n_1, \ldots, n_k} \) be the \( \sigma \)-algebra on \( \prod_{j=1}^k X_{j,n_j} \) 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 \( \mathcal{X} := (\prod_{j=1}^k X_j, \mathcal{B}, \lambda) \) on \( \prod_{j=1}^k X_j \) such that \( \lambda_{n_1, \ldots, n_k} \) is the finite \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X}, \forall n_1, \ldots, n_k \in \mathbb{N} \), where \( \mathcal{B} := \{ E \subseteq \prod_{j=1}^k X_j \mid E \cap (\prod_{j=1}^k X_{j,n_j}) \in \mathcal{B}_{n_1, \ldots, n_k}, \forall n_j \in \mathbb{N}, \forall j \in \{1, \ldots, k\} \}, \mathcal{P} \circ \lambda(E) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \ldots, n_k}(E \cap (\prod_{j=1}^k X_{j,n_j})), \forall E \in \mathcal{B} \), and \( \lambda(E) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \ldots, n_k}(E \cap (\prod_{j=1}^k X_{j,n_j})), \forall E \in \mathcal{B} \) with \( \mathcal{P} \circ \lambda(E) < \infty \).
Claim 12.22.1 $\mathcal{B} = \overline{\mathcal{B}}$.

Proof of claim: $\forall j$, $\forall n_1, \ldots, n_k \in \mathbb{N}$, $E_j \cap X_{j,n_j} \in \mathcal{B}_{j,n_j}$, $\forall j \in \{1, \ldots, k\}$, by Proposition 11.115. Then, $\left(\prod_{j=1}^k E_j\right) \cap \left(\prod_{j=1}^k X_{j,n_j}\right) \in \mathcal{C}_{n_1, \ldots, n_k} \subseteq \overline{\mathcal{B}}_{n_1, \ldots, n_k}$. This implies that $\prod_{j=1}^k E_j \in \overline{\mathcal{B}}$. Then, $\mathcal{C} \subseteq \overline{\mathcal{B}}$. Since $\overline{\mathcal{B}}$ is a $\sigma$-algebra, then $\mathcal{B} \subseteq \overline{\mathcal{B}}$.

On the other hand, $\forall n_1, \ldots, n_k \in \mathbb{N}$, $\prod_{j=1}^k E_j \in \mathcal{C}_{n_1, \ldots, n_k}$, we have $E_j \in \overline{\mathcal{B}}_{j,n_j} \subseteq \overline{\mathcal{B}}_j$, $\forall j \in \{1, \ldots, k\}$. Thus, $\prod_{j=1}^k E_j \in \mathcal{C} \subseteq \overline{\mathcal{B}}$. Then, $\mathcal{C}_{n_1, \ldots, n_k} \subseteq \mathcal{B}$ and $\overline{\mathcal{B}}_{n_1, \ldots, n_k} \subseteq \mathcal{B}$. This further implies that $\mathcal{B} \subseteq \overline{\mathcal{B}}$. Hence, $\mathcal{B} = \overline{\mathcal{B}}$. This completes the proof of the claim. □

$\forall j \in \{1, \ldots, k\}$, $\forall E_j \in \text{dom}(\mu_j)$, we have $\infty > \mathcal{P} \circ \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_j(E_j \cap X_{j,n_j}) = \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_j,n_j(E_j \cap X_{j,n_j})$, $\mathbb{K} \ni \mu_j(E_j) = \sum_{n_j=1}^{\infty} \mu_j(E_j \cap X_{j,n_j})$. Then,$$
\mathcal{P} \circ \lambda \left( \prod_{j=1}^k E_j \right) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^k E_j \right) \cap \left( \prod_{j=1}^k X_{j,n_j} \right) \\
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \mathcal{P} \circ \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^k (E_j \cap X_{j,n_j}) \right) \\
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^k \mathcal{P} \circ \mu_j,n_j(E_j \cap X_{j,n_j}) \\
= \prod_{j=1}^k \left( \sum_{n_j=1}^{\infty} \mathcal{P} \circ \mu_j,n_j(E_j \cap X_{j,n_j}) \right) = \prod_{j=1}^k \mathcal{P} \circ \mu_j(E_j) < \infty$$

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. We also have

$$\lambda \left( \prod_{j=1}^k E_j \right) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^k E_j \right) \cap \left( \prod_{j=1}^k X_{j,n_j} \right) \\
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \lambda_{n_1, \ldots, n_k} \left( \prod_{j=1}^k (E_j \cap X_{j,n_j}) \right) \\
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_k=1}^{\infty} \prod_{j=1}^k \lambda_{n_j} \left( E_j \cap X_{j,n_j} \right) \\
= \prod_{j=1}^k \left( \sum_{n_j=1}^{\infty} \lambda_{n_j} \left( E_j \cap X_{j,n_j} \right) \right) = \prod_{j=1}^k \lambda_j(E_j) \in \mathbb{K}$$

where the first equality follows from Proposition 11.118; and the third equality follows from Proposition 12.21. Hence, $\lambda$ is the $\sigma$-finite $\mathbb{K}$-valued...
measure space that we seek.

This completes the proof of the proposition. \qed

**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, \( X_i \in \mathcal{B}_i \), \( \mathcal{X}_i := (\bar{X}_i, \bar{\mathcal{B}}_i, \bar{\mu}_i) \) be the \( \sigma \)-finite \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X}_i \), \( i = 1, \ldots, m \), and \( \mathcal{X} := \left( \prod_{i=1}^m X_i, \mathcal{B}, \lambda \right) := \prod_{i=1}^m \mathcal{X}_i \) be the \( \sigma \)-product \( \mathbb{K} \)-valued measure space of \( X_1, \ldots, X_m \). Then, \( \mathcal{X} := \left( \prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}}, \bar{\lambda} \right) \) is the \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \) if, and only if, \( \bar{X} \subseteq \bar{X} \).

**Proof** Let \( \bar{X} := (\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}}, \bar{\lambda}) \) be the \( \sigma \)-finite \( \mathbb{K} \)-valued measure subspace of \( \mathcal{X} \) and \( \bar{X} := (\prod_{i=1}^m \bar{X}_i, \bar{\mathcal{B}}, \bar{\lambda}) \) be the \( \sigma \)-finite product \( \mathbb{K} \)-valued measure space of \( \bar{X}_1, \ldots, \bar{X}_m \). All we need to show is that \( \bar{X} = \bar{X} \).

Let \( \bar{X}_{i,1} := \bar{X}_i \) and \( \bar{X}_{i,2} := X_i \setminus \bar{X}_i \), \( i = 1, \ldots, m \), \( \mathcal{C}_{n_1, \ldots, n_m} := \{ \prod_{i=1}^m B_i \subseteq \prod_{i=1}^m \bar{X}_{i,n_i} \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \ldots, m\} \} \) be the set of measurable rectangles on \( \prod_{i=1}^m \bar{X}_{i,n_i}, \bar{\mathcal{B}}_{n_1, \ldots, n_m} \) be the \( \sigma \)-algebra generated by \( \mathcal{C}_{n_1, \ldots, n_m} \) on \( \prod_{i=1}^m \bar{X}_{i,n_i}, \forall n_1, \ldots, n_m \in \{1, 2\} \), and \( \mathcal{C} := \left( \prod_{i=1}^m B_i \subseteq \prod_{i=1}^m \bar{X}_i \mid B_i \in \mathcal{B}_i, \forall i \in \{1, \ldots, m\} \right) \) be the set of measurable rectangles on \( \prod_{i=1}^m \bar{X}_i \). Then, by Proposition 12.22, \( \mathcal{B} \) is the \( \sigma \)-algebra generated by \( \mathcal{C} \) on \( \prod_{i=1}^m X_i \).

**Claim 12.23.1** \( \mathcal{B} = \left\{ E \subseteq \prod_{i=1}^m X_i \mid E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \bar{\mathcal{B}}_{n_1, \ldots, n_m}, \forall n_i \in \{1, 2\}, \forall i \in \{1, \ldots, m\} \right\} =: \hat{\mathcal{B}}. \)

**Proof of claim:** \( \forall \prod_{i=1}^m B_i \in \mathcal{C} \), we have \( B_i \subseteq \mathcal{B}_i \), \( i = 1, \ldots, m \), which implies that \( \mathcal{B}_i \supseteq B_i \cap X_{i,n_i} \subseteq \bar{X}_{i,n_i}, n_i = 1, 2, i = 1, \ldots, m \). Hence, \( \prod_{i=1}^m B_i \subseteq \hat{\mathcal{B}}. \) By the arbitrariness of \( \prod_{i=1}^m B_i \), we have \( \mathcal{C} \subseteq \hat{\mathcal{B}}. \)

Now, we will show that \( \hat{\mathcal{B}} \) is a \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \). Clearly, \( \emptyset = \prod_{i=1}^m \emptyset \in \mathcal{C} \subseteq \hat{\mathcal{B}} \) and \( \prod_{i=1}^m X_i \subseteq \mathcal{C} \subseteq \hat{\mathcal{B}} \). \( \forall (E_n)_{n=1}^\infty \subseteq \hat{\mathcal{B}}, \forall n_i = 1, 2, \forall i = 1, \ldots, m, \forall n \in \mathbb{N} \), by the definition of \( \hat{\mathcal{B}}, E_n \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \bar{\mathcal{B}}_{n_1, \ldots, n_m} \), which implies \( \left( \bigcup_{n=1}^\infty E_n \right) \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) = \bigcup_{n=1}^\infty (E_n \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \bar{\mathcal{B}}_{n_1, \ldots, n_m} \), since \( \hat{\mathcal{B}}_{n_1, \ldots, n_m} \) is a \( \sigma \)-algebra. Hence, \( \bigcup_{n=1}^\infty E_n \in \hat{\mathcal{B}} \). \( \forall \mathcal{E} \subseteq \hat{\mathcal{B}}, \forall n_i = 1, 2, \forall i = 1, \ldots, m, E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \bar{\mathcal{B}}_{n_1, \ldots, n_m} \), then, \( (\prod_{i=1}^m X_i) \setminus E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) = (\prod_{i=1}^m \bar{X}_{i,n_i}) \setminus E = (\prod_{i=1}^m \bar{X}_{i,n_i}) \setminus (E \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \hat{\mathcal{B}}_{n_1, \ldots, n_m} \), since \( \hat{\mathcal{B}}_{n_1, \ldots, n_m} \) is a \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \). This implies that \( (\prod_{i=1}^m X_i) \setminus E \subseteq \hat{\mathcal{B}}. \) Hence, \( \hat{\mathcal{B}} \) is a \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \).

Since \( \mathcal{B} \) is the \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \) generated by \( \mathcal{C} \) and \( \mathcal{C} \subseteq \hat{\mathcal{B}} \), then \( \mathcal{B} \subseteq \hat{\mathcal{B}}. \)

On the other hand, \( \mathcal{C}_{n_1, \ldots, n_m} \subseteq \mathcal{C} \), \( n_i = 1, 2, i = 1, \ldots, m \), implies that \( \hat{\mathcal{B}}_{n_1, \ldots, n_m} \subseteq \mathcal{B} \). \( \forall E \in \hat{\mathcal{B}}, \) we have \( E \cap (\prod_{i=1}^m \bar{X}_{i,n_i}) \in \hat{\mathcal{B}}_{n_1, \ldots, n_m} \subseteq \mathcal{B}, \) \( n_i = 1, 2, i = 1, \ldots, m \), which implies that \( E = \bigcup_{n_1=1}^2 \cdots \bigcup_{n_m=1}^2 (E \cap (\prod_{i=1}^m \bar{X}_{i,n_i})) \in \mathcal{B}. \) Hence, \( \hat{\mathcal{B}} \subseteq \mathcal{B} \). Therefore, \( \mathcal{B} = \hat{\mathcal{B}}. \) This completes the proof of the claim. \qed
Note that \( \tilde{\mathcal{B}} = \tilde{\mathcal{B}}_{1, \ldots, 1} = \{ E \subseteq \prod_{i=1}^{m} \tilde{X}_i \mid E \in \mathcal{B} \} = \{ E \subseteq \prod_{i=1}^{m} \tilde{X}_i \mid E \in \mathcal{B} \}, \) where the first equality follows from Proposition 12.22; the second equality follows from the definition of \( \tilde{\mathcal{B}}; \) the third equality follows from Claim 12.23.1; and the last equality follows from Proposition 11.115.

Finally, we will show that \( \tilde{\lambda} = \hat{\lambda} \) by Propositions 12.22 and 11.115. By Proposition 11.115 and Claim 12.23.1, \( \hat{\lambda} \) is a \( \mathbb{K} \)-valued measure on \( \prod_{i=1}^{m} \tilde{X}_i, \mathcal{B} \), \( \hat{\lambda} = \lambda_{|\mathcal{B}} = \lambda_{|\mathcal{B}} \), and \( \mathcal{P} \circ \hat{\lambda} = (\mathcal{P} \circ \lambda)|_{\mathcal{B}}. \) By Proposition 12.22, \( \hat{\lambda} \) is the unique \( \mathbb{K} \)-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, \( \mu_i = \mu_i|_{\mathcal{B}} \) and \( \mathcal{P} \circ \mu_i = (\mathcal{P} \circ \mu_i)|_{\mathcal{B}}. \) Hence, \( E_i \in \text{dom}(\mu_i). \) This implies that \( \hat{\lambda}(\prod_{i=1}^{m} E_i) = \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mu_i(E_i) \) and \( \mathcal{P} \circ \hat{\lambda}(\prod_{i=1}^{m} E_i) = \mathcal{P} \circ \lambda(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \mu_i(E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \mu_i(E_i) \).

By Proposition 12.22, \( \hat{\lambda} \) is the unique \( \mathbb{K} \)-valued measure on \( (\prod_{i=1}^{m} X_i, \mathcal{B}) \) such that \( \hat{\lambda}(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \tilde{\mu}_i(E_i) \) and \( \mathcal{P} \circ \hat{\lambda}(\prod_{i=1}^{m} E_i) = \prod_{i=1}^{m} \mathcal{P} \circ \tilde{\mu}_i(E_i), \forall E_i \in \text{dom}(\tilde{\mu}_i), \forall i \in \{1, \ldots, m\}. \) Hence, we have \( \lambda = \hat{\lambda}. \)

Hence, \( \tilde{X} = \hat{X}. \) This completes the proof of the proposition. \( \square \)

**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 \( \mathbb{K} \)-valued measure space, \( \forall \alpha \in \Lambda, \ X_{\beta} := \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} := (\prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}, \mathcal{B}_{\beta}, \mu_{\beta}) \) be the \( \sigma \)-finite product \( \mathbb{K} \)-valued measure space, \( \forall \beta \in \Gamma, \ X := \prod_{\beta \in \Gamma} X_{\beta} := (\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}, \mathcal{B}, \mu) \) be the \( \sigma \)-finite product \( \mathbb{K} \)-valued measure space, and \( \tilde{X} := \prod_{\alpha \in \Lambda} X_{\alpha} := (\prod_{\alpha \in \Lambda} X_{\alpha}, \tilde{\mathcal{B}}, \tilde{\mu}) \) be the \( \sigma \)-finite product \( \mathbb{K} \)-valued measure space. Then, \( \hat{X} \) and \( \tilde{X} \) are isomeric.

**Proof** Define \( \Psi : \hat{X} \to \tilde{X} \) by, \( \forall x \in \hat{X}, \forall \alpha \in \Lambda, \ \exists ! \beta_{\alpha} \in \Gamma \) \( \exists \cdot \alpha \in \Lambda_{\beta_{\alpha}}, \pi_{\alpha}(\tilde{\Psi}(x)) = \pi_{\alpha}^{(\alpha)}(\Psi^{(\beta_{\alpha})(x)}). \) By the proof of Proposition 3.30, \( \Psi \) is bijective with inverse \( \Psi_{\text{inv}} : \tilde{X} \to \hat{X}. \)

Let \( \hat{\mathcal{E}}_{\beta} := \{ \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha} \subseteq \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \mid E_{\alpha} \in \mathcal{B}_{\alpha}, \forall \alpha \in \Lambda_{\beta} \}, \) \( \forall \beta \in \Gamma, \) \( \tilde{\mathcal{E}} := \{ \prod_{\beta \in \Gamma} E_{\beta} \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \mid E_{\beta} \in \tilde{\mathcal{B}}_{\beta}, \forall \beta \in \Gamma \}, \) \( \hat{\mathcal{E}} := \{ \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha} \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha} \mid E_{\alpha} \in \mathcal{B}_{\alpha}, \forall \alpha \in \Lambda \}, \) and \( \tilde{\mathcal{E}} := \{ \prod_{\alpha \in \Lambda} E_{\alpha} \subseteq \prod_{\alpha \in \Lambda} X_{\alpha} \mid E_{\alpha} \in \mathcal{B}_{\alpha}, \forall \alpha \in \Lambda \}. \) By Proposition 12.22, \( \hat{\mathcal{B}}_{\beta} \) is the \( \sigma \)-algebra generated by \( \hat{\mathcal{E}}_{\beta} \) on \( \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}, \forall \beta \in \Gamma, \) \( \hat{\mathcal{B}} \) is the \( \sigma \)-algebra generated by \( \hat{\mathcal{E}} \) on \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} X_{\alpha}, \) and \( \tilde{\mathcal{B}} \) is the \( \sigma \)-algebra generated by \( \tilde{\mathcal{E}} \) on \( \prod_{\alpha \in \Lambda} X_{\alpha}. \) Clearly, \( \hat{\mathcal{E}} \subseteq \hat{\mathcal{E}}. \)

Fix any \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_{\beta}} E_{\alpha} \subseteq \hat{\mathcal{E}}, \) 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 \( \pi_{\alpha}(\hat{\Psi}(x)) = \pi_{\alpha}^{(\alpha)}(\hat{\Psi}(x)). \)
\[
\pi_{\alpha}(x) \in E_\alpha. \quad \text{Hence, } \Psi(x) \in \prod_{\alpha \in \Lambda} E_\alpha. \quad \text{This implies that } \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) \subseteq \prod_{\alpha \in \Lambda} E_\alpha. \quad \text{On the other hand, } \forall y \in \prod_{\alpha \in \Lambda} E_\alpha, \forall \beta \in \Gamma, \forall \alpha \in \Lambda_\beta, \text{ we have } \pi_{\alpha}(\pi_{\beta}(\Psi(\Psi_{\text{inv}}(y)))) = \pi_{\alpha}(y) \in E_\alpha. \quad \text{Then, } \Psi_{\text{inv}}(y) \in \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha. \quad \text{This leads to } \Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_\alpha) \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha. \quad \text{This is equivalent to } \prod_{\alpha \in \Lambda} E_\alpha \subseteq \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha). \quad \text{Therefore,} \quad \Psi(\prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha) = \prod_{\alpha \in \Lambda} E_\alpha \in \bar{\mathcal{E}}. \quad \text{Conversely, } \forall \prod_{\alpha \in \Lambda} E_\alpha \in \bar{\mathcal{E}}, \text{ we have } \Psi_{\text{inv}}(\prod_{\alpha \in \Lambda} E_\alpha) = \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} E_\alpha \in \bar{\mathcal{E}}.
\]

Let \( \tilde{B} \) be the \( \sigma \)-algebra on \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \) generated by \( \bar{\mathcal{E}} \). We will show that \( \tilde{B} = \{ E \subseteq \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \mid \Psi(E) \in \bar{\mathcal{B}} \} =: \hat{B} \). Since \( \bar{\mathcal{E}} \subseteq \bar{\mathcal{B}} \), by the previous paragraph, we have \( \bar{\mathcal{E}} \subseteq \hat{B} \). 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 \( \bar{\mathcal{B}} \) is a \( \sigma \)-algebra. Then, \( \tilde{B} \subseteq \hat{B} \). On the other hand, let \( \check{B} := \{ E \subseteq \prod_{\alpha \in \Lambda} X_\alpha \mid \Psi_{\text{inv}}(E) \in \check{B} \} \). Since \( \check{B} \) is a \( \sigma \)-algebra on \( \prod_{\beta \in \Gamma} \prod_{\alpha \in \Lambda_\beta} X_\alpha \), then \( \check{B} \) is a \( \sigma \)-algebra on \( \prod_{\alpha \in \Lambda} X_\alpha \). Since \( \bar{\mathcal{E}} \subseteq \hat{B} \), by the previous paragraph, then \( \bar{\mathcal{E}} \subseteq \check{B} \). By the definition of \( \check{B} \), we have \( \check{B} \subseteq \hat{B} \). Since \( \tilde{B} \subseteq \check{B} \), then, by the definitions of \( \check{B} \) and \( \hat{B} \), \( \tilde{B} \subseteq \check{B} \). Thus, we have \( \tilde{B} = \hat{B} = \check{B} \). We can now prove the following result.

**Claim 12.24.1** \( \hat{E} \subseteq \check{E} \).

**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:

- \( \forall i \in \{1, \ldots, n+1\}, \forall E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{E}_{\beta_j}, j = i, \ldots, n, \forall E_{\beta_j} \in \hat{B}_{\beta_j}, j = 1, \ldots, i-1, \) we have \( \prod_{j=1}^{n} E_{\beta_j} \in \hat{B} \).

by mathematical induction on \( i \).

1° \( i = 1 \). \( \forall E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{E}_{\beta_j}, j = 1, \ldots, n, \) we have \( \prod_{j=1}^{n} E_{\beta_j} = \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{E} \). This case is proved.

2° Assume that the statement holds for \( i = k \in \{1, \ldots, n\} \).

3° Consider the case \( i = k + 1 \). Fix any \( E_{\beta_j} := \prod_{\alpha \in \Lambda_{\beta_j}} E_\alpha \in \hat{E}_{\beta_j}, j = k + 1, \ldots, n, \) and any \( E_{\beta_k} \in \hat{B}_{\beta_k}, j = 1, \ldots, k - 1 \). Define \( \mathcal{F} := \{ E_{\beta_k} \in \hat{B}_{\beta_k} \mid \prod_{j=1}^{k-1} E_{\beta_j} \in \hat{B} \} \). By the inductive assumption, we have \( \mathcal{E}_{\beta_k} \subseteq \mathcal{F} \). Then, \( \emptyset = \prod_{\alpha \in \Lambda_{\beta_k}} \emptyset \in \mathcal{E}_{\beta_k} \subseteq \mathcal{F} \) and \( X_{\beta_k} := \prod_{\alpha \in \Lambda_{\beta_k}} X_\alpha \in \hat{E}_{\beta_k} \subseteq \mathcal{F} \). \( \forall E_{\beta_k} \in \mathcal{F}, \) we have \( E_{\beta_k} \in \mathcal{E}_{\beta_k} \) and \( \prod_{j=1}^{n} E_{\beta_j} \in \hat{B} \). By \( X_{\beta_k} \in \hat{E}_{\beta_k} \subseteq \mathcal{F} \), 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 \( (\prod_{j=1}^{k-1} E_{\beta_j}) \times (\prod_{j=k+1}^{n} E_{\beta_j}) \subseteq \hat{B} \). Then, \( X_{\beta_k} \setminus E_{\beta_k} = \ldots \)
12.3. PRODUCT MEASURE

Let \( \gamma \in (1,2) \), then \( \pi_1 ^{\gamma} X \) is a \( \sigma \)-finite measure space, \( i = 1,2 \), and \( \lambda \) is the product measure space.

Then, \( \forall E \in B, \forall x_1 \in X_1, E_{x_1, x_2} := \{ x_2 \in X_2 \mid (x_1, x_2) \in E \} \in B_2 \); and, \( \forall x_2 \in X_2, E(x_2) := \{ x_1 \in X_1 \mid (x_1, x_2) \in E \} \in B_1 \).
Proof Let $\mathcal{E} := \{ E \in \mathcal{B} \mid E_{x_1} \in \mathcal{B}_2, \forall x_1 \in X_1 \} \subseteq \mathcal{B}$. Then, $\mathcal{E} \subseteq \mathcal{C}$. We will show that $\mathcal{E}$ is a $\sigma$-algebra on $X_1 \times X_2$. Therefore, 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 $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. For every $E \in \mathcal{E}$, $\forall x_1 \in X_1$, we have $(X_1 \times X_2) \setminus E_{x_1} = X_2 \setminus E_{x_1} \in \mathcal{B}_2$ since $\mathcal{B}_2$ is a $\sigma$-algebra. By the arbitrariness of $x_1$, we have $(X_1 \times X_2) \setminus E \in \mathcal{E}$. Further, by Proposition 12.25, $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E}$, let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. Then, $E_{x_1} \in X_1$, we have $\bigcup_{i=1}^{\infty} E_{ix_1} \in \mathcal{B}_2$ 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 $\forall E \in \mathcal{B} = \mathcal{E}$, $\forall x_1 \in X_1$, $E_{x_1} \in \mathcal{B}_2$. 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 $\mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times \mathcal{X}_2$ be the product measure space. Then, $\forall E \in \mathcal{B}$, the function $g_E : X_1 \to [0, \infty] \subseteq \mathbb{R}_e$, defined by $g_E(x_1) = \mu_2(E_{x_1})$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$-measurable, where $E_{x_1}$ is as defined in Proposition 12.25. By symmetry, $\forall E \in \mathcal{B}$, the function $h_E : X_2 \to [0, \infty] \subseteq \mathbb{R}_e$, defined by $h_E(x_2) = \mu_1(E_{x_2})$, $\forall x_2 \in X_2$, 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, $g_E(x_1) \in [0, \infty] \subseteq \mathbb{R}_e$. Hence, the function $g_E$ is well-defined.

First, consider the special case when $X_1$ and $X_2$ are finite. Let $\mathcal{E} := \{ E \in \mathcal{B} \mid g_E \text{ is } \mathcal{B}_1\text{-measurable } \}$, and $\mathcal{C}$ be the set of measurable rectangles in $X_1 \times X_2$, $\forall E := E_1 \times E_2 \in \mathcal{C}$, $g_E(x_1) = \mu_2(E_{x_1})$, $\forall x_1 \in X_1$. Clearly, $g_E$ is $\mathcal{B}_1$-measurable. Then, $E \in \mathcal{E}$ and $\mathcal{C} \subseteq \mathcal{E}$.

Clearly, $\mathcal{C}$ is a $\pi$-system 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 $X_1 \times X_2 \in \mathcal{C} \subseteq \mathcal{E}$. $\forall E_1, E_2 \in \mathcal{C}$ with $E_1 \subseteq E_2$, then $g_{E_1} : X_1 \to [0, \mu_2(X_2)] \subseteq \mathbb{R}$ is $\mathcal{B}_1$-measurable, $i = 1, 2$. Let $E := E_2 \setminus E_1$, $\forall x_1 \in X_1$, we have $g_E(x_1) = \mu_2(E_{2x_1}) = \mu_2(E_{2x_1}) - \mu_2(E_{1x_1}) = g_{E_2}(x_1) - g_{E_1}(x_1) \in [0, \mu_2(X_2)] \subseteq \mathbb{R}$. By Propositions 7.23, 11.39, and 11.38, $g_E$ is $\mathcal{B}_1$-measurable. Then, $E = E_2 \setminus E_1 \in \mathcal{E}$. $\forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E}$ with $E_i \subseteq E_{i+1}, \forall i \in \mathbb{N}$, then $g_{E_i} : X_1 \to [0, \mu_2(X_2)] \subseteq \mathbb{R}$ is $\mathcal{B}_1$-measurable, $\forall i \in \mathbb{N}$. Let $E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}$. $\forall x_1 \in X_1$, $g_E(x_1) = \mu_2(\bigcup_{i=1}^{\infty} E_{ix_1}) = \lim_{i \in \mathbb{N}} \mu_2(E_{ix_1}) = \lim_{i \in \mathbb{N}} g_{E_i}(x_1)$, where the first equality follows from the definition of $g_E$; the second equality follows from Proposition 11.7; and the third equality follows from the definition of $g_{E_i}$’s. Hence, by Proposition 11.48, $g_E$ is $\mathcal{B}_1$-measurable. Then, $E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{E}$. This shows that $\mathcal{E}$ is a monotone class on $X_1 \times X_2$.

Then, $\mathcal{E} = \mathcal{B}$ and $\forall E \in \mathcal{B}$, $g_E$ is $\mathcal{B}_1$-measurable. By symmetry, $\forall E \in \mathcal{B}$, $h_E$ is $\mathcal{B}_2$-measurable. The proposition holds 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 $\exists (X_{1,n})_{n=1}^{\infty} \subseteq B_1$ such that $X_1 = \bigcup_{n=1}^{\infty} X_{1,n}$ and
12.3. PRODUCT MEASURE

\( \mu_1(X_{1,n}) < \infty, \forall n \in \mathbb{N} \). Without loss of generality, we may assume that \( X_{1,n} \subseteq X_{1,n+1}, \forall n \in \mathbb{N} \). Since \( X_2 \) is \( \sigma \)-finite, then \( \exists (X_{2,m})_{m=1}^{\infty} \subseteq B_2 \) such that \( X_2 = \bigcup_{m=1}^{\infty} X_{2,m} \) and \( \mu_2(X_{2,m}) < \infty, \forall m \in \mathbb{N} \). Without loss of generality, we may assume that \( X_{2,m} \subseteq X_{2,m+1}, \forall m \in \mathbb{N} \). 

Let \( X_{1,n} := (X_{1,n}, B_{1,n}, \mu_{1,n}) \) be the finite measure subspace of \( X_1 \), \( X_{2,m} := (X_{2,m}, B_{2,m}, \mu_{2,m}) \) be the finite measure subspace of \( X_2 \), \( (X_{1,n} \times X_{2,m}, B_{n,m}, \lambda_{n,m}) := \mathcal{X}_{1,n} \times \mathcal{X}_{2,m} \), which is the finite measure subspace of \( \mathcal{X} \) by Proposition 12.22 and its proof. \( \forall E \in \mathcal{B}, \forall n, m \in \mathbb{N} \), let \( E_{n,m} := E \cap (X_{1,n} \times X_{2,m}) \in B_{n,m} \). By the special case, the function \( g_{n,m} : X_{1,n} \rightarrow [0, \infty) \subseteq \mathbb{R} \), 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 \( B_{1,n} \)-measurable. Define \( g_{n,m} : X_1 \rightarrow [0, \infty) \subseteq \mathbb{R} \) by \( \bar{g}_{n,m}(x_1) = \mu_2((E_{n,m})_{x_1}), \forall x_1 \in X_1 \). Then, \( \bar{g}_{n,m}(x_1) = \begin{cases} g_{n,m}(x_1) & \forall x_1 \in X_{1,n} \\ 0 & \forall x_1 \in X_1 \setminus X_{1,n} \end{cases} \), \( \forall x_1 \in X_1 \). By Proposition 11.41, \( \bar{g}_{n,m} \) is \( B_1 \)-measurable. Note that \( g_E(x_1) = \mu_2(E_{x_1}) = \mu_2((\bigcup_{n=1}^{\infty} (E_{n,n})_{x_1}) = \lim_{n \in \mathbb{N}} \mu_2((E_{n,n})_{x_1}) = \lim_{n \in \mathbb{N}} \bar{g}_{n,m}(x_1), \forall x_1 \in X_1 \), 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.

\[ \square \]

**Lemma 12.27** Let \( \mathcal{X}_i := (X_i, B_i, \mu_i) \) be a \( \sigma \)-finite measure space, \( i = 1, 2 \), and \( \mathcal{X} := (X_1 \times X_2, B, \mu) := \mathcal{X}_1 \times \mathcal{X}_2 \) be the product measure space. Then, \( \forall E \in \mathcal{B}, \) we have \( \mu(E) = \int_{X_1} g_E \, d\mu_1 = \int_{X_2} h_E \, d\mu_2 \), where the function \( g_E : X_1 \rightarrow [0, \infty) \subseteq \mathbb{R}_+ \), as defined in Proposition 12.26, is \( B_1 \)-measurable; and the function \( h_E : X_2 \rightarrow [0, \infty) \subseteq \mathbb{R}_+ \), 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 \( g_E : X_1 \rightarrow [0, \mu_2(X_2)] \subseteq \mathbb{R} \) is \( B_1 \)-measurable. Let \( \mathcal{E} := \{ E \in \mathcal{B} \mid \mu(E) = \int_{X_1} g_E \, d\mu_1 \} \subseteq \mathcal{B} \). 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 \( \mathcal{E} \subseteq \mathcal{E} \).

We will shown that \( \mathcal{E} \) is a monotone class on \( X_1 \times X_2 \). By Monotone Class Lemma 12.19, we have \( \mathcal{E} = \mathcal{B} \).

Clearly, \( \emptyset \in \mathcal{E} \subseteq \mathcal{E} \) 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 \( \mathcal{E} \); and 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 \( E = E_2 \setminus E_1 \in \mathcal{E} \).
Chapter 12. Differentiation and Integration

\[ \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E} \text{ with } E_i \subseteq E_{i+1}, \forall i \in \mathbb{N}, \text{ let } E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B}. \forall i \in \mathbb{N}, \text{ by Proposition 12.26, } g_E \text{ is } \mathcal{B}_1\text{-measurable. } \forall x_1 \in X_1, \text{ we have } g_{E_i}(x_1) \leq g_{E_{i+1}}(x_1) \text{ since } E_i \subseteq E_{i+1}. \text{ By Proposition 12.26, } g_E \text{ is } \mathcal{B}_1\text{-measurable.}

As shown in the proof of Proposition 12.26, } g_E(x_1) = \lim_{n \in \mathbb{N}} g_{E_i}(x_1), \forall x_1 \in X_1. \text{ This implies that } \mu(E) = \lim_{n \in \mathbb{N}} \mu(E_i) = \lim_{n \in \mathbb{N}} \int_{X_1} g_{E_i} \, d\mu_1 = \int_{X_1} g_E \, d\mu_1, \text{ where the first equality follows from Proposition 11.7; the second equality follows from the definition of } \mathcal{E}; \text{ 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} \text{ is a monotone class on } X_1 \times X_2 \text{ and } \mathcal{E} = \mathcal{B}. \text{ Therefore, } 

\mu(E) = \int_{X_1} g_E \, d\mu_1, \forall E \in \mathcal{B}. \text{ By symmetry, } \mu(E) = \int_{X_2} h_E \, d\mu_2, \forall E \in \mathcal{B}. \text{ This completes the proof for the special case.}

Next, consider the general case when } X_1 \text{ and } X_2 \text{ are } \sigma\text{-finite. Since } X_1 \text{ is } \sigma\text{-finite, then } \exists (X_1)_1^{\infty} \subseteq \mathcal{B}_1 \text{ such that } X_1 = \bigcup_{n=1}^{\infty} X_{1,n} \text{ and } \mu_1(X_{1,n}) < \infty, \forall n \in \mathbb{N}. \text{ Without loss of generality, we may assume that } X_{1,n} \subseteq X_{1,n+1}, \forall n \in \mathbb{N}. \text{ Since } X_2 \text{ is } \sigma\text{-finite, then } \exists (X_2)_m^{\infty} \subseteq \mathcal{B}_2 \text{ such that } X_2 = \bigcup_{m=1}^{\infty} X_{2,m} \text{ and } \mu_2(X_{2,m}) < \infty, \forall m \in \mathbb{N}. \text{ Without loss of generality, we may assume that } X_{2,m} \subseteq X_{2,m+1}, \forall m \in \mathbb{N}. \forall n,m \in \mathbb{N}, \text{ let } X_{1,n} := (X_{1,n}, \mu_{1,n}(n)) \text{ be the finite measure subspace of } X_1, X_{2,m} := (X_{2,m}, \mu_{2,m}) \text{ be the finite measure subspace of } X_2, (X_1 \times X_2, \mathcal{B}_{1,m}, \lambda_{n,m}) := X_{1,n} \times X_{2,m}, \text{ which is the finite measure subspace of } \mathcal{X} \text{ by Proposition 12.22. } \forall E \in \mathcal{B}, \forall n,m \in \mathbb{N}, \text{ let } E_{n,m} := E \cap (X_{1,n} \times X_{2,m}) \in \mathcal{B}_{1,m}. \text{ By the special case, the function } g_{n,m} : X_{1,n} \to [0, \infty) \subseteq \mathbb{R}, \text{ defined by } g_{n,m}(x_1) = \mu_{2,m}((E_{n,m})_{x_1} \cap X_{2,m}), \forall x_1 \in X_{1,n}, \text{ is } \mathcal{B}_{1,m}\text{-measurable and } \mu(E_{n,m}) = \lambda_{n,m}(E_{n,m}) = \int_{X_{1,n}} g_{n,m} \, d\mu_{1,n}. \text{ Define } g_{n,m} : X_{1,n} \to [0, \infty) \subseteq \mathbb{R} \text{ by } \tilde{g}_{n,m}(x_1) = \mu_{2}((E_{n,m})_{x_1}), \forall x_1 \in X_{1,n}.

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,n} \setminus X_{1,n} \end{cases}, \forall x_1 \in X_{1,n}. \text{ By Proposition 11.41, } \tilde{g}_{n,m} \text{ is } \mathcal{B}_{1,m}\text{-measurable. By Definition 11.79, } \mu(E_{n,m}) = \int_{X_{1,n}} \tilde{g}_{n,m} \, d\mu_{1,n}. \text{ By Proposition 12.26 and its proof, we have } g_{E_{n,m}}(x_1) = \lim_{n \in \mathbb{N}} g_{n,m}(x_1), \forall x_1 \in X_{1,n}, \text{ and } g_E \text{ is } \mathcal{B}_1\text{-measurable. Clearly, } \tilde{g}_{n,m}(x_1) \leq \tilde{g}_{n+1,m+1}(x_1), \forall x_1 \in X_1, \forall n \in \mathbb{N}, \text{ since } E_{n,m} \subseteq E_{n+1,m+1}. \text{ 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,n} = \int_{X_1} g_E \, d\mu_1, \text{ where the first equality follows from Proposition 11.7; 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}. \text{ This completes the proof of the proposition.} \square

Proposition 12.28 Let } X_i := (X_i, \mathcal{B}_i, \mu_i) \subseteq \text{ a } \sigma\text{-finite measure space, } i = 1, 2, \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2 \text{ be the product measure space, } \mathcal{Y} := (Y, \mathcal{O}) \text{ be a topological space, and } f : X_1 \times X_2 \to \mathcal{Y} \text{ be } \mathcal{B}\text{-measurable. Then,}

(i) } \forall x_1 \in X_1, f_{x_1} : X_2 \to \mathcal{Y}, \text{ defined by } f_{x_1}(x_2) = f(x_1, x_2), \forall x_2 \in X_2, \text{ is } \mathcal{B}_2\text{-measurable;}

(ii) } \forall x_2 \in X_2, f_{x_2} : X_1 \to \mathcal{Y}, \text{ defined by } f_{x_2}(x_1) = f(x_1, x_2), \forall x_1 \in X_1, \text{ is } \mathcal{B}_1\text{-measurable.}
12.3. PRODUCT MEASURE

**Proof**  (i) ∀ open set \( O \subseteq \mathcal{O} \), \( f_{\text{inv}}(O) \in \mathcal{B} \) 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. \( \Box \)

**Theorem 12.29 (Tonelli)** Let \( X_1 := (X_1, \mathcal{B}_1, \mu_1) \) be a \( \sigma \)-finite measure space, \( i = 1, 2, \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2 \) be the product measure space, and \( f : X_1 \times X_2 \to [0, \infty] \subset \mathbb{R}_+ \) be \( \mathcal{B} \)-measurable. Define \( f_{x_1} : X_2 \to \mathbb{R} \) and \( f_{(x_1)} : X_1 \to \mathbb{R} \) 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 \( p : X_1 \to [0, \infty] \subset \mathbb{R}_+ \), 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 \( q : X_2 \to [0, \infty] \subset \mathbb{R}_+ \), 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 \( (\varphi_n)_{n=1}^{\infty} \), \( \varphi_n : \mathcal{X} \to [0, \infty] \subset \mathbb{R}, \forall n \in \mathbb{N} \), such that \( 0 \leq \varphi_n(x_1, x_2) \leq \varphi_{n+1}(x_1, x_2) \leq f(x_1, x_2), \forall (x_1, x_2) \in X_1 \times X_2, \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} \varphi_n(x_1, x_2) = f(x_1, x_2), \forall (x_1, x_2) \in X_1 \times X_2 \). By Proposition 12.28 and Definition 11.79, \( p \) is well-defined. Define \( p_n : X_1 \to [0, \infty] \subset \mathbb{R}_+ \) 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, \( p(x_1) = \lim_{n \to \infty} p_n(x_1) \), \( \forall x_1 \in X_1 \).

Fix any \( n \in \mathbb{N} \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m_n} a_i \chi_{E_i} \times X_1, \) 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 \( g_i : X_1 \to [0, \infty] \subset \mathbb{R}_+ \) by \( g_i(x_1) = \mu_2(E_{i_1}, x_1), \forall x_1 \in X_1, \) where \( (E_{i_1})_{x_1} := \{ x_2 \in X_2 \mid (x_1, x_2) \in E_i \} \). 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, \( \int_{X_1 \times X_2} \varphi_n \, d\mu = \sum_{i=1}^{m_n} a_i \mu(E_i) = \sum_{i=1}^{m_n} a_i \int_{X_1} \chi_{E_i} \times X_1 \, d\mu_1 = \int_{X_1} (\sum_{i=1}^{m_n} a_i g_i) \, d\mu_1 =: \int_{X_1} p_n \, d\mu_1 \), where the third equality follows from Proposition 11.83. By Propositions 3.89, 11.38, and 11.39, \( p_n : X_1 \to [0, \infty] \subset \mathbb{R}_+ \) is \( \mathcal{B}_1 \)-measurable. Note that \( p_n(x_1) = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \int_{X_2} (\sum_{i=1}^{m_n} a_i \chi_{E_i} \times X_1, x_2) \, d\mu_2(x_2) = \sum_{i=1}^{m_n} a_i \int_{X_2} \chi_{E_i}(x_1, x_2) \, d\mu_2 = \sum_{i=1}^{m_n} a_i \mu_2(E_{i_1})(x_2) = \sum_{i=1}^{m_n} a_i g_i(x_1) = p_n(x_1), \forall x_1 \in X_1 \), 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 $B_1$-measurable. By Proposition 11.48 and the fact that $[0,\infty] \subset \mathbb{R}_+$ is metrizable, $p$ is $B_1$-measurable. Then, $\int_{X_1 \times X_2} \varphi_n \, d\mu = \int_{X_1} p_n \, d\mu_1 = \int_{X_1} p_n \, d\mu_1 \in [0,\infty] \subset \mathbb{R}$. 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.

\section*{Theorem 12.30 (Fubini)} Let $X_i := (X_i, B_i, \mu_i)$ be a $\sigma$-finite measure space, $i = 1, 2$, $X := (X_1 \times X_2, B, \mu) := X_1 \times X_2$ 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 $\bar{U} \in B_1$ 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 $\bar{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y}$; $\forall x_1 \in X_1 \setminus \bar{U}$, $\bar{p}(x_1) = \bar{v}_y$; $\bar{p}$ is absolutely integrable over $X_1$; and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} \bar{p} \, d\mu_1 \in \mathcal{Y}$.

(ii) There exists $\bar{q} : X_2 \to \mathcal{Y}$ and exists $\bar{V} \in B_2$ 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 $\bar{q}(x_2) = \int_{X_1} f_{(x_2)} \, d\mu_1 \in \mathcal{Y}$; $\forall x_2 \in X_2 \setminus \bar{V}$, $\bar{q}(x_2) = \bar{v}_y$; $\bar{q}$ is absolutely integrable over $X_2$; and $\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} \bar{q} \, d\mu_2 \in \mathcal{Y}$.

(iii) $\forall p : X_1 \to \mathcal{Y}$ satisfying: $p$ is $B_1$-measurable; and $\exists U \in B_1$ 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 $\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \in \mathcal{Y}$.

(iv) $\forall q : X_2 \to \mathcal{Y}$ satisfying: $q$ is $B_2$-measurable; and $\exists V \in B_2$ 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 $\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2 \in \mathcal{Y}$.

\section*{Proof} (i) Since $f$ is absolutely integrable over $X$, then $P \circ f$ is $B$-measurable and $\int_{X_1 \times X_2} \varphi \, d\mu < \infty$. 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 $B_1$-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 $\tilde{g} : X_1 \to [0,\infty] \subset \mathbb{R}$, defined by $\tilde{g}(x_1) = \{ g(x_1) \, g(x_1) < \infty \} \{ 0 \, g(x_1) = \infty \}$, $\forall x_1 \in X_1$, is $B_1$-measurable, $g = \tilde{g}$ a.e. in $X_1$, and $\int_{X_1 \times X_2} P \circ f \, d\mu = \int_{X_1} \tilde{g} \, d\mu_1 < \infty$. Let $U := \{ x_1 \in X_1 \mid g(x_1) < \infty \}$. Then, $U \in B_1$ and $\mu_1(X_1 \setminus U) = 0$. $\forall x_1 \in U$, we have $g(x_1) = \int_{X_2} P \circ f(x_1, x_2) \, d\mu_2(x_2) = \int_{X_2} P \circ f_{x_1} \, d\mu_2 < \infty$. Then, $f_{x_1}$ is absolutely integrable over $X_2$. By
Propositions 12.28 and 11.92, \( \int_{X_2} f_{x_1} \, d\mu_2 \in \mathcal{Y}, \forall x_1 \in U \). Define \( \tilde{p} : X_1 \to \mathcal{Y} \) by \( \tilde{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \), \( x_1 \in U \), \( \forall x_1 \in X_1 \).

By Proposition 11.66, there exists a sequence of simple functions \( (\varphi_n)_{n=1}^{\infty} \), \( \varphi_n : X \to \mathcal{Y}, \forall n \in \mathbb{N} \), 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 \( \lim_{n \to \infty} \varphi_n = f \) a.e. in \( X \). 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 \( A_{x_1} := \{ \bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in A \} \) as defined in Proposition 12.25, and \( \hat{g}_A : X_1 \to [0, \infty) \subset \mathbb{R} \) by \( \hat{g}_A(x_1) = \left\{ \begin{array}{ll} g_A(x_1) & g_A(x_1) < \infty \\ 0 & g_A(x_1) = \infty \end{array} \right., \forall x_1 \in X_1 \).

Then, by Lemma 12.27 and Proposition 11.82, \( g_A \) and \( \hat{g}_A \) are well-defined and \( \mathcal{B}_1 \)-measurable, \( g_A = \hat{g}_A \) a.e. in \( X_1 \), and \( \mu(A) = \int_{X_1} g_A \, d\mu_1 = \int_{X_1} \hat{g}_A \, d\mu_1 = 0 \). Since \( \hat{g}_A(x_1) \geq 0, \forall x_1 \in X_1 \), then, by Proposition 11.96, we have \( \hat{g}_A = 0 \) a.e. in \( X_1 \). Then, by Lemma 11.44 and the fact that \( [0, \infty) \subset \mathbb{R} \) is second countable and metricizable, \( g_A = 0 \) a.e. in \( X_1 \). Let \( A_1 := \{ x_1 \in X_1 \mid g_A(x_1) \neq 0 \} \). Then, \( A_1 \in \mathcal{B}_1 \) and \( \mu_1(A_1) = 0 \). Fix any \( x_1 \in U \setminus A_1 \). \( g_A(x_1) = 0 \) implies that \( \mu_2(A_{x_1}) = 0 \), which further implies that \( \lim_{n \to \infty} \varphi_n(x_1, x_2) = f(x_1, x_2) \) a.e. \( x_2 \in X_2 \). Then, by Lebesgue Dominated Convergence Theorem 11.91, \( \tilde{p}(x_1) = \lim_{n \to \infty} \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \in \mathcal{Y} \), \( \forall x_1 \in U \setminus A_1 \).

Fix any \( n \in \mathbb{N} \). Let \( \varphi_n \) admit the canonical representation \( \varphi_n = \sum_{i=1}^{m} y_i \chi_{E_{i,x_1} \times X_2} \), where \( m \in \mathbb{Z}_+, y_1, \ldots, y_m \in \mathcal{Y} \) are distinct and nonequal to \( \vartheta_{y_i}, E_{i,x_1} \times X_2 \subseteq \mathcal{B} \) 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 \( E_{i,x_1} := \{ \bar{x}_2 \in X_2 \mid (x_1, \bar{x}_2) \in E_i \} \). 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 \( \hat{g}_i(x_1) = \left\{ \begin{array}{ll} g_i(x_1) & g_i(x_1) < \infty \\ 0 & g_i(x_1) = \infty \end{array} \right., \forall x_1 \in X_1 \).

By Proposition 11.82, \( \hat{g}_i \) is \( \mathcal{B}_1 \)-measurable, \( g_i = \hat{g}_i \) a.e. in \( X_1 \), and \( \mu(E_i) = \int_{X_1 \times X_2} \hat{g}_i \, d\mu = \sum_{i=1}^{m} y_i \mu(E_i) = \sum_{i=1}^{m} y_i \int_{X_1} \hat{g}_i \, d\mu_1 \). Then, \( \int_{X_1} \chi_{E_{i,x_1}} \, d\mu_2 = \sum_{i=1}^{m} y_i \chi_{E_{i,x_1}}(x_2) \, d\mu_2 = \int_{X_1} \sum_{i=1}^{m} y_i \chi_{E_{i,x_1}} \, d\mu_1 =: \int_{X_1} \hat{g}_i \, d\mu_1 \), 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, \( A_n \in \mathcal{B}_1 \) and \( \mu_1(A_n) = 0 \). \( \forall x_1 \in X_1 \setminus A_n \), \( \tilde{p}_n(x_1) = \sum_{i=1}^{m} y_i \hat{g}_i(x_1) = \sum_{i=1}^{m} y_i g_i(x_1) = \sum_{i=1}^{m} y_i \mu_2(E_{i,x_1}) = \sum_{i=1}^{m} y_i \int_{X_2} \chi_{E_{i,x_1} \times X_2} \, d\mu_2 = \int_{X_2} \sum_{i=1}^{m} y_i \chi_{E_{i,x_1} \times X_2} \, d\mu_2 = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \in \mathcal{Y}, \) 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 \( g_i \)'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, \( A_2 \in \mathcal{B}_1 \) and \( \mu_1(A_2) = 0 \). \( \forall x_1 \in U \setminus (A_1 \cup A_2) := U \), we have \( \tilde{p}(x_1) = \lim_{n \to \infty} \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \lim_{n \to \infty} \tilde{p}_n(x_1) \in \mathcal{Y} \).
Clearly, \( \bar{U} \in B_1 \) and \( 0 \leq \mu_1(X_1 \setminus \bar{U}) = \mu_1((X_1 \setminus U) \cup A_1 \cup A_2) \leq \mu_1(X_1 \setminus U) + \mu_1(A_1) + \mu_1(A_2) = 0 \). Fix any \( n \in \mathbb{N} \). Define \( \hat{p}_n : X_1 \to \mathbb{Y} \) by
\[
\hat{p}_n(x_1) = \begin{cases} 
\hat{p}_n(x_1) & x_1 \in \bar{U} \\
\bar{p}(x_1) & x_1 \in X_1 \setminus \bar{U}
\end{cases}, \forall x_1 \in X_1. \]
By Proposition 11.41, \( \hat{p}_n \) is \( B_1 \)-measurable. By Lemma 11.43, \( \hat{p}_n = \hat{p}_n \) a.e. in \( X_1 \). Define \( \bar{p} : X_1 \to \mathbb{Y} \) by
\[
\bar{p}(x_1) = \begin{cases} 
\hat{p}(x_1) & x_1 \in \bar{U} \\
\bar{p}(x_1) & x_1 \in X_1 \setminus \bar{U}
\end{cases}, \forall x_1 \in X_1. \]
Then, \( \bar{p}(x_1) = \lim_{n \to \infty} \hat{p}_n(x_1) \), \( \forall x_1 \in X_1 \). By Proposition 11.48, \( \bar{p} \) is \( B_1 \)-measurable. \( \forall x_1 \in \bar{U} \),
\[
\bar{p}(x_1) = \bar{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \text{ and } f_{x_1} \text{ is absolutely integrable over } X_2.
\]
Note that by symmetry, (ii) also holds.

(iii) Fix any \( \gamma \in X_1 \) satisfying: \( \gamma \) is \( B_1 \)-measurable; and \( \exists \mu \in B_1 \) 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 \( U_0 := U \cap U_1 \in B_1 \) and \( \mu_1(X_1 \setminus U_0) = \mu_1((X_1 \setminus U) \cup (X_1 \setminus U)) = 0 \). \( \forall x_1 \in U_0 \), \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 = \bar{p}(x_1) \). By Lemma 11.43, \( \hat{p} \) is \( p \) a.e. in \( X_1 \). By Proposition 11.83, we have \( p \) is absolutely integrable over \( X_1 \) since \( \bar{p} \) is so. By Proposition 11.92,
\[
\int_{X_1} p \, d\mu_1 = \int_{X_1} \bar{p} \, d\mu_1 = \int_{X_1 \times X_2} f \, d\mu \in \mathbb{Y}.
\]
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 \( U \) of the previous theorem takes a simpler form. This is established in the following theorem.

Theorem 12.31 (Fubini) Let \( X_i := (X_i, B_i, \mu_i) \) be a \( \sigma \)-finite measure space, \( i = 1, 2 \), \( X := (X_1 \times X_2, B, \mu) := X_1 \times X_2 \) be the product measure space, \( \mathbb{Y} \) be a separable Banach space, \( U \subseteq \mathbb{Y} \) be a \( \sigma \)-compact conic segment, \( f : X_1 \times X_2 \to U \) be absolutely integrable over \( X \). Define \( f_{x_1} : X_2 \to \mathbb{Y} \) and \( \int_{x_2} : X_1 \to \mathbb{Y} \) as in Proposition 12.28. Then, the following statements hold.

(i) Define \( U := \{ x_1 \in X_1 \mid f_{x_1} \text{ is absolutely integrable over } X_2 \} \), and \( p : X_1 \to \mathbb{Y} \) by \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \mid \begin{array}{ll} x_1 \in U \\
\bar{p}(x_1) & x_1 \in X_1 \setminus U \end{array}, \forall x_1 \in X_1 \).

Then, \( U \in B_1 \) with \( \mu_1(X_1 \setminus U) = 0 \), \( p \) is \( B_1 \)-measurable and is absolutely integrable over \( X_1 \), and \( \int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \in \mathbb{Y} \).
(ii) Define $V := \{ x_2 \in X_2 \mid f(x_2) \text{ is absolutely integrable over } X_1 \}$, and $q : X_2 \to \mathbb{Y}$ by $q(x_2) = \int_{X_1} f(x_2) \, \mathrm{d}\mu_1 \, x_2 \in V$, $\forall x_2 \in X_2$.

Then, $V \in \mathcal{B}_2$ with $\mu_2(V \setminus V) = 0$, $q$ is $\mathcal{B}_2$-measurable and is absolutely integrable over $X_2$, and $\int_{X_1 \times X_2} f \, \mathrm{d}\mu = \int_{X_2} q \, \mathrm{d}\mu_2 \in \mathbb{Y}$.

**Proof**

(i) Since $f$ is absolutely integrable over $\mathcal{X}$, then $f$ and $\mathcal{P} \circ f$ are $\mathcal{B}$-measurable and $\int_{X_1 \times X_2} \mathcal{P} \circ f \, \mathrm{d}\mu < \infty$. By Tonelli Theorem 12.29, $g : X_1 \to [0, \infty] \subset \mathbb{R}$, defined by $g(x_1) = \int_{X_1} \mathcal{P} \circ f(x_1, x_2) \, \mathrm{d}\mu_2(x_2)$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$-measurable and $\int_{X_1 \times X_2} \mathcal{P} \circ f \, \mathrm{d}\mu = \int_{X_1} g \, \mathrm{d}\mu_1$.

By Proposition 11.82, the function $\tilde{g} : X_1 \to [0, \infty] \subset \mathbb{R}$, defined by $\tilde{g}(x_1) = \begin{cases} g(x_1) & g(x_1) < \infty, \\ 0 & g(x_1) = \infty \end{cases}$, $\forall x_1 \in X_1$, is $\mathcal{B}_1$-measurable, $g = \tilde{g}$ a.e. in $X_1$, and $\int_{X_1 \times X_2} \mathcal{P} \circ f \, \mathrm{d}\mu = \int_{X_1} \tilde{g} \, \mathrm{d}\mu_1$. Let $\tilde{U} := \{ x_1 \in X_1 \mid g(x_1) < \infty \}$.

Then, $\tilde{U} \in \mathcal{B}_1$, and $\mu_1(X_1 \setminus \tilde{U}) = 0$, $\forall x_1 \in \tilde{U}$, we have $g(x_1) = \int_{X_1} \mathcal{P} \circ f(x_1, x_2) \, \mathrm{d}\mu_2(x_2) = \int_{X_1} \mathcal{P} \circ f_{x_1} \, \mathrm{d}\mu_2 < \infty$. Then, $f_{x_1}$ is absolutely integrable over $X_2$, $\forall x_1 \in \tilde{U}$. By Propositions 12.28 and 11.92, $\int_{X_2} f_{x_1} \, \mathrm{d}\mu_2 \in \mathbb{Y}$, $\forall x_1 \in \tilde{U}$, $\forall x_1 \in X_1 \setminus \tilde{U}$, $\int_{X_2} \mathcal{P} \circ f_{x_1} \, \mathrm{d}\mu_2 = \infty$ and $f_{x_1}$ is not absolutely integrable over $X_2$. Hence, $U = \tilde{U}$.

By Lemma 11.68, $\exists$ a sequence of simple functions $(\varphi_n)_{n=1}^{\infty}, \varphi_n : \mathcal{X} \to U$, $\forall n \in \mathbb{N}$, 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 $\lim_{n \to \infty} \varphi_n(x_1, x_2) = f(x_1, x_2)$, $\forall (x_1, x_2) \in X_1 \times X_2$.

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, $p(x_1) = \lim_{n \to \infty} \int_{X_2} \varphi_n(x_1, x_2) \, \mathrm{d}\mu_2(x_2), \forall x_1 \in X_1$.

Fix any $n \in \mathbb{N}$. Let $\varphi_n$ admit the canonical representation $\varphi_n = \sum_{i=1}^{m} y_i \chi_{E_{i}, X_1 \times X_2}$, where $m \in \mathbb{Z}_+$, $y_1, \ldots, y_m \in U$ are distinct and none equals to $\emptyset$, $E_1, \ldots, E_m \in \mathcal{B}$ 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 $g_i(x_1) = \mu_2(E_{i,x_1}), \forall x_1 \in X_1, \text{ where } E_{i,x_1} := \{ x_2 \in X_2 \mid (x_1, x_2) \in E_i \}$.

By Lemma 12.27, $g_i$ is $\mathcal{B}_1$-measurable and $\mu(E_i) = \int_{X_1} g_i \, \mathrm{d}\mu_1$.

Define $\tilde{g}_i : X_1 \to [0, \infty] \subset \mathbb{R}$ by $\tilde{g}_i(x_1) = \begin{cases} g_i(x_1) & g_i(x_1) < \infty, \\ 0 & g_i(x_1) = \infty \end{cases}$, $\forall x_1 \in X_1$. 

By Proposition 11.82, $\tilde{g}_i$ is $\mathcal{B}_1$-measurable, $\tilde{g}_i = \tilde{g}_i$ a.e. in $X_1$, and $\mu(E_i) = \int_{X_1} \tilde{g}_i \, \mathrm{d}\mu_1$. Then, $\int_{X_1 \times X_2} \varphi_n \, \mathrm{d}\mu = \sum_{i=1}^{m} y_i \mu(E_i) = \sum_{i=1}^{m} y_i \int_{X_1} \tilde{g}_i \, \mathrm{d}\mu_1 = \int_{X_1} \sum_{i=1}^{m} y_i \tilde{g}_i \, \mathrm{d}\mu_1 =: \int_{X_1} \tilde{p}_n \, \mathrm{d}\mu_1$, 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, $\tilde{p}_n : X_1 \to \mathbb{Y}$ is $\mathcal{B}_1$-measurable and, by Proposition 11.92, $\tilde{p}_n$ is absolutely integrable over $X_1$. $\forall x_1 \in U$, $g(x_1) < \infty$, then $f_{x_1}$ is absolutely integrable over $X_2$. Since $\| \varphi_n(x_1, x_2) \| \leq \| f(x_1, x_2) \|$, $\forall (x_1, x_2) \in X_1 \times X_2$, $\forall n \in \mathbb{N}$, then $\forall x_1 \in U$, $\| \varphi_n(x_1, x_2) \|$, as a function of $x_2$, is absolutely integrable over $X_2$. Then, $\tilde{g}_i(x_1) = g_i(x_1), i = 1, \ldots, m, \forall x_1 \in U$. Hence, $\tilde{p}_n(x_1) = \sum_{i=1}^{m} y_i \tilde{g}_i(x_1) = \sum_{i=1}^{m} y_i g_i(x_1) = \sum_{i=1}^{m} y_i \mu_2(E_{i,x_1}) = \sum_{i=1}^{m} y_i \int_{X_2} \chi_{E_{i,x_1}} \, \mathrm{d}\mu_2 = \int_{X_2} \sum_{i=1}^{m} y_i \chi_{E_{i,x_1}} \, \mathrm{d}\mu_2$.
\[ = \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \in \mathcal{Y}, \forall x_1 \in U, \] where the first equality follows from the definition of \( \hat{p}_n \); 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 11.92; and the fifth equality follows from Proposition 11.92. Define \( \hat{p}_n : X_1 \to \mathcal{Y} \) by \( \hat{p}_n(x_1) = \begin{cases} \hat{p}_n(x_1) & x_1 \in U \\ \vartheta & x_1 \in X_1 \setminus U \end{cases}, \forall x_1 \in X_1. \) By Proposition 11.41, \( \hat{p}_n \) is \( B_1 \)-measurable. Then, \( \hat{p}_n = \hat{p}_n \) a.e. in \( X_1 \).

Note that, \( \forall x_1 \in U, \) we have
\[
\| \hat{p}_n(x_1) \| = \| \hat{p}_n(x_1) \| = \left\| \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) \right\|
\leq \int_{X_2} \mathcal{P} \circ \varphi_n(x_1, x_2) \, d\mu_2(x_2) \leq \int_{X_2} \mathcal{P} \circ f(x_1, x_2) \, d\mu_2(x_2) = \tilde{g}(x_1) < \infty
\]
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 \( \| \hat{p}_n(x_1) \| = 0 = \tilde{g}(x_1). \) Then, \( \| \hat{p}_n(x_1) \| \leq \tilde{g}(x_1), \forall x_1 \in X_1. \) Hence, \( \hat{p}_n \) is absolutely integrable over \( X_1. \)

By Proposition 11.92, \( \int_{X_1 \times X_2} \varphi_n \, d\mu = \int_{X_1} \hat{p}_n \, d\mu_1 = \int_{X_1} \hat{p}_n \, d\mu_1, \) where the first equality follows from the definition of \( \hat{p}_n. \)

By the result of the third preceding paragraph, \( p(x_1) = \lim_{n \in \mathbb{N}} \int_{X_2} \varphi_n(x_1, x_2) \, d\mu_2(x_2) = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1) = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in U, \) where the second equality follows from the result of the second preceding paragraph; the last equality follows from the definition of \( \hat{p}_n. \) Note that \( p(x_1) = \vartheta = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in X_1 \setminus U. \) Then, \( p(x_1) = \lim_{n \in \mathbb{N}} \hat{p}_n(x_1), \forall x_1 \in X_1. \) This implies that, by Proposition 11.51, \( p \) is \( B_1 \)-measurable. Then, we have
\[
\int_{X_1 \times X_2} f \, d\mu = \lim_{n \in \mathbb{N}} \int_{X_1 \times X_2} \varphi_n \, d\mu = \lim_{n \in \mathbb{N}} \int_{X_1} \hat{p}_n \, d\mu_1 = \int_{X_1} p \, d\mu_1, \]
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.

\[ \]

**Proposition 12.32** Let \( m \in \mathbb{N}, X_i := (X_i, \mathcal{O}_i) \) be a second countable topological space, \( X_i := (X_i, B_i, \mu_i) \) 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 \( (\prod_{i=1}^m X_i, \mathcal{B}, \mu) := \prod_{i=1}^m (X_i, B_i, \mu_i) \) 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 \( \mathcal{X}, \)
Let \( U \) being a topological measure space, we have \( \mathcal{B}_i = \mathcal{B}_\mathcal{B}(\mathcal{X}_i), \ i = 1, \ldots, m \). Let \( \mathcal{X}_1 := \mathcal{X}_1 \) and \( \mathcal{X}_i := \mathcal{X}_{i-1} \times \mathcal{X}_i := ((\cdots (X_1 \times X_2) \times \cdots) \times X_i, \mathcal{O}_i), \ i = 2, \ldots, m \), \((X_1, \mathcal{B}_1, \mu_1) := (X_1, \mathcal{B}_1, \mu_1)\) and \((\cdots (X_1 \times X_2) \times \cdots) \times X_i, \mathcal{B}_i, \mu_i) := ((\cdots (X_1 \times X_2) \times \cdots) \times X_i, \mathcal{B}_i, \mu_i), \ i = 2, \ldots, m \). By Proposition 3.28, \( \mathcal{X}_i, \ i = 1, \ldots, m \), are second countable topological spaces, and by Proposition 11.24, \( \mathcal{B}_i = \mathcal{B}_\mathcal{B}(\mathcal{X}_i), \ i = 1, \ldots, m \). By Propositions 3.30 and 12.24, \( \mathcal{X}_m \) is homeomorphic to \( \mathcal{X} \) and \(((\cdots (X_1 \times X_2) \times \cdots) \times X_m, \mathcal{B}_m, \mu_m)\) is isomearic 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 a second countable finite topological measure space.

First, consider the special case when \( \mathcal{X}_i, \ i = 1, \ldots, m \), are finite. Let \( \mathcal{E} := \left\{ E \in \mathcal{B} \mid \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \ (i) \ \forall U \in \mathcal{O} \text{ with } E \subseteq U \text{ such that } \mu(U \setminus E) < \epsilon; \ (ii) \ \exists F \in \mathcal{O} \text{ with } F \subseteq E \text{ such that } \mu(E \setminus F) < \epsilon \right\} \). We will show that \( \mathcal{E} \) is a \( \sigma \)-algebra on \( \prod_{i=1}^m X_i \) and contains \( \mathcal{E} \). Then, \( \mathcal{E} = \mathcal{B} \) and \( \mathcal{X} \) is a second countable finite topological measure space.

\[
\forall E := \prod_{i=1}^m B_i \in \mathcal{E}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists O_i \in \mathcal{O}_i \text{ with } B_i \subseteq O_i \text{ such that } \mu(O_i \setminus B_i) < 1 + m \prod_{j=1}^m \frac{\epsilon}{1 + \mu_j(B_j)}.
\]

Let \( U := \prod_{i=1}^m O_i \in \mathcal{O} \) and \( F := \prod_{i=1}^m F_i \). Clearly, \( \prod_{i=1}^m ((\prod_{j=1}^{i-1} X_j) \times (X_i \setminus F_i) \times (\prod_{j=i+1}^m X_j)) \in \mathcal{O} \). Then, \( F \subseteq E \subseteq U \),

\[
\mu(U \setminus E) = \mu\left(\prod_{i=1}^m O_i \setminus \left(\prod_{i=1}^m B_i\right)\right) = \mu\left(\bigcup_{i=1}^m \left(\prod_{j=1}^{i-1} O_j \times (O_i \setminus B_i) \times \prod_{j=i+1}^m O_j\right)\right) \leq \sum_{i=1}^m \mu\left(\prod_{j=1}^{i-1} O_j \times (O_i \setminus B_i) \times \prod_{j=i+1}^m O_j\right) = \sum_{i=1}^m \mu_i(O_i \setminus B_i) \prod_{j=1,j \neq i}^m \mu_j(O_j) \leq \sum_{i=1}^m \frac{\epsilon}{1 + m \prod_{j=1,j \neq i}^m (1 + \mu_j(B_j))} \prod_{j=1,j \neq i}^m (1 + \mu_j(B_j)) < \epsilon
\]

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, \( E \in \mathcal{E} \). This shows that \( \mathcal{E} \subseteq \mathcal{E} \).

\( \forall E \in \mathcal{E}, \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \tilde{U} \in \mathcal{O} \) with \( E \subseteq \tilde{U} \) such that \( \mu(U \setminus E) < \epsilon \), \( \exists \tilde{F} \in \mathcal{O} \) with \( F \subseteq E \) such that \( \mu(E \setminus F) < \epsilon \). Then, \( \mu(\tilde{E} \setminus \tilde{U}) = \mu(U \setminus E) < \epsilon \) and \( \mu(F \setminus \tilde{E}) = \mu(F \setminus F) < \epsilon \). Hence, \( \tilde{E} \in \mathcal{E} \).

\( \forall (E_i)_{i=1}^{\infty} \subseteq \mathcal{E}, \) let \( E := \bigcup_{i=1}^{\infty} E_i \in \mathcal{B} \). \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R}, \forall i \in \mathbb{N}, \exists U_i \in \mathcal{O} \) with \( E_i \subseteq U_i \) such that \( \mu(U_i \setminus E_i) < 2^{-i} \epsilon \); \( \exists \tilde{F}_i \in \mathcal{O} \) with \( F_i \subseteq U_i \) such that \( \mu(E_i \setminus F_i) < 2^{i-1} \epsilon \). Let \( U := \bigcup_{i=1}^{\infty} U_i \in \mathcal{O} \). Then, \( E = \bigcup_{i=1}^{\infty} E_i \subseteq \bigcup_{i=1}^{\infty} U_i = U \) and \( \mu(U \setminus E) = \sum_{i=1}^{\infty} \mu(U_i \setminus E_i) \leq \sum_{i=1}^{\infty} \mu(U_i \setminus E_i) < \epsilon \), where the first inequality follows from Proposition 11.6; and the second inequality follows from Proposition 11.4. Hence, \( E \) satisfies (i). By Proposition 11.7, \( \mu(E) = \mu(\bigcup_{i=1}^{\infty} E_i) = \lim_{n \to \infty} \mu(\bigcup_{i=1}^{n} E_i) \). Then, by \( \mu(E) \leq \mu(\prod_{i=1}^{m} X_i) = \prod_{i=1}^{m} \mu_i(X_i) < \infty \), \( \exists \epsilon \in \mathbb{R} \) such that \( \mu(\prod_{i=1}^{m} E_i) \leq \mu(E) < \mu(\prod_{i=1}^{m} E_i) + \epsilon/2 \). Let \( F := \bigcap_{i=1}^{m} F_i \). Clearly, \( F = \bigcap_{i=1}^{m} \tilde{F}_i \in \mathcal{O} \). Then, \( F \subseteq E \) and \( \mu(E \setminus F) = \mu(E \setminus (\prod_{i=1}^{m} E_i)) + \mu(\prod_{i=1}^{m} E_i \setminus F) = \mu(E) - \mu(\prod_{i=1}^{m} E_i) + \mu(\prod_{i=1}^{m} E_i \setminus F) < \epsilon/2 + \sum_{i=1}^{m} \mu(E_i \setminus F_i) < \epsilon \), where the first and the second equalities follow from Fact 11.3; the first inequality follows from Proposition 11.6; and the second inequality follows from Proposition 11.4. Hence, \( E \) satisfies (ii). Then, \( E = \bigcup_{i=1}^{\infty} E_i \in \mathcal{E} \).

Clearly, \( \emptyset = \prod_{i=1}^{m} \emptyset \in \mathcal{E} \subseteq \mathcal{E} \) and \( \prod_{i=1}^{m} X_i \in \mathcal{E} \subseteq \mathcal{E} \). This shows that \( \mathcal{E} \) is a \( \sigma \)-algebra on \( \prod_{i=1}^{m} X_i \) and contains \( \mathcal{C} \). This completes the proof of the special case.

Next, consider the general case when \( X_i, i = 1, \ldots, m, \) are \( \sigma \)-finite. Fix any \( i = 1, \ldots, m \). By \( X_i \) being \( \sigma \)-finite, \( \exists (X_{i,n})_{n=1}^{\infty} \subseteq \mathcal{B}_i \) such that \( X_i = \bigcup_{n=1}^{\infty} X_{i,n} \) and \( \mu_i(X_{i,n}) < \infty, \forall n \in \mathbb{N} \). 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 \( X_{i,n} \in \mathcal{O}_i, \forall n \in \mathbb{N} \). \( \forall n \in \mathbb{N} \), 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 subspace of \( X_i, \forall n \in \mathbb{N}, \) by the special case, let \( \tilde{X}_n := \prod_{i=1}^{m} X_{i,n} := (\prod_{i=1}^{m} X_{i,n}, \mathcal{O}_n), \mathcal{B}_n, \mu_n) \) be the second countable finite product topological measure space. By Proposition 12.23, \( (\prod_{i=1}^{m} X_{i,n}, \mathcal{B}_n, \mu_n) \) is the finite...
12.3. PRODUCT MEASURE

measure subspace of \((\prod_{i=1}^{m} X_i, B, \mu)\). For all \(E \in B\), \(\forall \epsilon \in (0, \infty) \subset \mathbb{R}\), \(\forall n \in \mathbb{N}\), let \(E_n := E \cap (\prod_{i=1}^{m} X_i, \mathcal{B}_n) \in \mathcal{B}_n\). Since \(\mathcal{X}_n\) is a topological measure, \(\exists U_n \in \mathcal{O}_n\), with \(E_n \subseteq U_n\) such that \(\mu_n(U_n \setminus E_n) < 2^{-n}\epsilon\). 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, \(E = \bigcup_{n=1}^{\infty} E_n \leq \bigcup_{n=1}^{\infty} U_n = U\) and \(\mu(U \setminus E) = \mu(U) - \mu(U \setminus E) = \bigcup_{n=1}^{\infty} \mu(U_n \setminus E_n) = \sum_{n=1}^{\infty} \mu(U_n \setminus E_n) < \epsilon\), where the first inequality follows from Proposition 12.22; the second inequality follows from Proposition 11.14; and the second equality follows from Proposition 11.13. By the arbitrariness of \(E\), we have \(\mathcal{X}\) is a \(\sigma\)-finite topological measure space. It is second countable since \(\mathcal{X}\) 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, \(\mathcal{X} := (X_i, \mathcal{B}_i, \mu_i)\) be a second countable \(\sigma\)-finite \(K\)-valued topological measure space, \(i = 1, \ldots, m\), \(\mathcal{X} := (\prod_{i=1}^{m} X_i, \mathcal{O}) := \prod_{i=1}^{m} \mathcal{X}_i\) be the product topological space, and \((\prod_{i=1}^{m} X_i, B, \mu) := \prod_{i=1}^{m}(X_i, B, \mu_i)\) be the product \(K\)-valued measure space. Then, \(\mathcal{X} := (\mathcal{X}, B, \mu)\) 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 \(\mathcal{X} = \prod_{i=1}^{m} X_i\).

**Proof** Let \(\mathcal{X} := (X_i, B_i, \mathcal{P} \circ \mu_i =: \nu_i), i = 1, \ldots, m\), and \(\mathcal{X} := (\prod_{i=1}^{m} X_i, B, \mathcal{P} \circ \mu =: \nu)\). By Proposition 12.22, \(\mathcal{X}_i, i = 1, \ldots, m\), and \(\mathcal{X}\) are \(\sigma\)-finite measure spaces and \(\mathcal{X} = \prod_{i=1}^{m} \mathcal{X}_i\). Since \(\mathcal{X}_i\) is a second countable \(\sigma\)-finite \(K\)-valued topological measure space, then, by Definition 11.191, \(\mathcal{X}_i := (X_i, B_i, \nu_i)\) is a second countable \(\sigma\)-finite topological measure space, \(i = 1, \ldots, m\). By Proposition 12.32, \(\mathcal{X} := (\mathcal{X}, B, \nu)\) is a second countable \(\sigma\)-finite topological measure space. Then, \(\mathcal{X}\) is a second countable \(\sigma\)-finite \(K\)-valued topological measure space. This completes the proof of the proposition.

**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 \(\frac{d\nu_i}{d\mu_i} = f_i : X_i \rightarrow K\) a.e. in \(X_i, i = 1, \ldots, m\), and \(f : \prod_{i=1}^{m} X_i \rightarrow K\) 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 \(\mathcal{X}\), where \(\nu := \prod_{i=1}^{m} \nu_i, \mu := \prod_{i=1}^{m} \mu_i, \text{and } \mathcal{X} := (\prod_{i=1}^{m} X_i, B, \mu) := \prod_{i=1}^{m} \mathcal{X}_i\).

**Proof** By Proposition 12.22, \(\mathcal{X}\) 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 \(\nu\) on \((\prod_{i=1}^{m} X_i, B)\) with respect to \(\mu\). By Proposition 11.167, \(\frac{d\nu}{d\mu} = f\) a.e. in \(\mathcal{X}\).

First, consider the special case when \(\nu_i, i = 1, \ldots, m, \) are finite \(K\)-valued measures. Let \(\mathcal{E} := \{ E \in B \mid \nu(E) = \bar{\nu}(E)\}\) and \(\mathcal{F} := \{ \prod_{i=1}^{m} B_i \subseteq \)
\[ \prod_{i=1}^{m} X_i \mid B_i \in \mathcal{B}, \forall i \in \{1, \ldots, m\} \}, \] 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

\[
\int_{\prod_{i=1}^{m} X_i} \mathcal{P} \circ f \, d\mu = \left( \int_{\prod_{i=1}^{m-1} X_i} m-1 \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d\left( \prod_{i=1}^{m-1} \mu_i \right) \right) \left( \int_{X_m} \mathcal{P} \circ f_m \, d\mu_m \right)
\]

\[
= \left( \int_{\prod_{i=1}^{m-1} X_i} m-1 \prod_{i=1}^{m-1} \mathcal{P} \circ f_i \, d\left( \prod_{i=1}^{m-1} \mu_i \right) \right) \mathcal{P} \circ \nu_m(X_m) = \cdots = \prod_{i=1}^{m} \mathcal{P} \circ \nu_i(X_i)
\]

< \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 11.166. By Proposition 11.83, \( f \mathcal{P} \circ \nu_i(B_i) \mathcal{P} \circ \nu_i(X_i) \), is absolutely integrable over \( \mathcal{X} \). By repeated application of Fubini Theorem 12.30,

\[
\tilde{\nu}(E) = \int_{\prod_{i=1}^{m} B_i} f \, d\mu = \int_{\prod_{i=1}^{m} X_i} f \mathcal{P} \circ \nu_m = \prod_{i=1}^{m} \mathcal{P} \circ \nu_i \mathcal{P} \circ \nu_i \mathcal{P} \circ \nu_m(B_m)
\]

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 \( \frac{d\tilde{\nu}}{d\nu} = f \) a.e. in \( \mathcal{X} \), which completes the proof of the special case.

Note that \( \emptyset = \prod_{i=1}^{m} \emptyset \subseteq \mathcal{C} \subseteq \mathcal{E} \) and \( \prod_{i=1}^{m} X_i \Subset \mathcal{C} \subseteq \mathcal{E} \). \( \forall E_1, E_2 \in \mathcal{E} \) 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, \( \nu(E) = \lim_{i \in \mathbb{N}} \nu(E_i) = \lim_{i \in \mathbb{N}} \tilde{\nu}(E_i) = \tilde{\nu}(E) \), 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 \( \tilde{\nu} \) 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 \( \tilde{X}_{i,n} := (X_{i,n}, \mathcal{B}_{i,n}, \mu_{i,n}) \) be the \( \sigma \)-finite measure subspace of \( X_i \), \( \tilde{\tilde{X}}_{i,n} := (X_{i,n}, B_{i,n}, \nu_{i,n}) \) be the finite \( \mathbb{K} \)-valued measure subspace of \( \tilde{X}_i := (X_i, \mathcal{B}_i, \nu_i) \). Fix any \( n_1, \ldots, n_m \in \mathbb{N} \). Let \( \tilde{X}_{n_1,\ldots,n_m} := (\prod_{i=1}^{m} X_{i,n_i}, \tilde{B}_{n_1,\ldots,n_m}, \tilde{\mu}_{n_1,\ldots,n_m}) := \prod_{i=1}^{m} \tilde{X}_{i,n_i} \) be the \( \sigma \)-finite product measure space, and \( \tilde{\tilde{X}}_{n_1,\ldots,n_m} := (\prod_{i=1}^{m} X_{i,n_i}, \tilde{B}_{n_1,\ldots,n_m}, \tilde{\nu}_{n_1,\ldots,n_m}) := \prod_{i=1}^{m} \tilde{X}_{i,n_i} \) be the finite product \( \mathbb{K} \)-valued measure space. By Proposition 12.23, \( \tilde{\tilde{X}}_{n_1,\ldots,n_m} \) is the measure subspace of \( \tilde{X} \) and \( \tilde{X}_{n_1,\ldots,n_m} \) is the finite \( \mathbb{K} \)-valued measure subspace of \( \tilde{\tilde{X}} := (\prod_{i=1}^{m} X_i, B, \nu) \). By Definition 11.166, we have

\[
\frac{d\nu_i}{d\tilde{\nu}_{1,n_i}} = f|_{X_{i,n_i}} \text{ a.e. in } \tilde{X}_{i,n_i}, i = 1, \ldots, m.
\]

By the special case, we have

\[
\frac{d\tilde{\nu}_{1,n_i}}{d\tilde{\mu}_{n_1,\ldots,n_m}} = f|_{\prod_{i=1}^{m} X_{i,n_i}} \text{ a.e. in } \tilde{\tilde{X}}_{n_1,\ldots,n_m}, \forall E \in \tilde{B} \text{ with } \mathcal{P} \circ \nu(E) < \infty,
\]

let \( E_{n_1,\ldots,n_m} := E \cap (\prod_{i=1}^{m} X_{i,n_i}) \in \tilde{B}_{n_1,\ldots,n_m} \). Then, we have

\[
\mathcal{P} \circ \nu(E) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \mathcal{P} \circ \nu(E_{n_1,\ldots,n_m})
\]

\[
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \mathcal{P} \circ \tilde{\nu}_{n_1,\ldots,n_m}(E_{n_1,\ldots,n_m})
\]

\[
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1,\ldots,n_m}} \mathcal{P} \circ (f|_{\prod_{i=1}^{m} X_{i,n_i}}) d\tilde{\mu}_{n_1,\ldots,n_m}
\]

\[
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1,\ldots,n_m}} \mathcal{P} \circ f d\mu = \int_{E} \mathcal{P} \circ f d\mu < \infty
\]

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 \( \frac{d\tilde{\nu}_{1,n_i}}{d\tilde{\mu}_{n_1,\ldots,n_m}} = f|_{\prod_{i=1}^{m} X_{i,n_i}} \text{ a.e. in } \tilde{\tilde{X}}_{n_1,\ldots,n_m} \); and the last two equalities follow from Proposition 11.83. We also have

\[
\nu(E) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \nu(E_{n_1,\ldots,n_m}) = \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \tilde{\nu}_{n_1,\ldots,n_m}(E_{n_1,\ldots,n_m})
\]

\[
= \sum_{n_1=1}^{\infty} \cdots \sum_{n_m=1}^{\infty} \int_{E_{n_1,\ldots,n_m}} f|_{\prod_{i=1}^{m} X_{i,n_i}} d\tilde{\mu}_{n_1,\ldots,n_m}
\]
where the first equality follows from Definition 11.108; the second equality follows from Proposition 11.115; the third equality follows from the fact that \(\frac{\partial}{\partial n_1 \cdots n_m} = f\big|_{\prod_{i=1}^m X_i \times n_1} \) a.e. in \(X_{n_1} \times \cdots \times X_{n_m}\); the fourth equality follows from Proposition 11.92; the fifth equality follows from Proposition 11.92 and the fact that \(\int k \circ f \, d\mu < \infty\); and the last equality follows from the fact that \(\frac{d\nu}{d\mu} = f\) a.e. in \(X\). Therefore, by Proposition 11.137, we have \(\nu = \nu\) and \(\frac{d\nu}{d\mu} = f\) a.e. in \(X\).

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

**Theorem 12.35 (Fubini)** Let \(X_i := (X_i, \mathcal{B}_i, \mu_i)\) be a \(\sigma\)-finite \(K\)-valued measure space, \(i = 1, 2, \mathcal{X} := (X_1 \times X_2, \mathcal{B}, \mu) := X_1 \times X_2\) be the product \(K\)-valued measure space, \(\gamma\) be a separable Banach space over \(K\), and \(f : X_1 \times X_2 \to \gamma\) be absolutely integrable over \(X\). Define \(f_{x_1} : X_2 \to \gamma\) and \(f(x_2) : X_1 \to \gamma\) as in Proposition 12.28. Then, the following statements hold.

(i) There exists \(\bar{p} : X_1 \to \gamma\) and exists \(\bar{U} \in \mathcal{B}_1\) such that \(\mathcal{P} \circ \mu_1(X_1 \backslash \bar{U}) = 0\); \(\forall x_1 \in \bar{U}\), \(f_{x_1}\) is absolutely integrable over \(X_2\) and \(\bar{p}(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \gamma\); \(\gamma(x_1) = \bar{\gamma}(\gamma)\); \(\bar{p}\) is absolutely integrable over \(X_1\); and \(\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} \bar{p} \, d\mu_1 \in \gamma\).

(ii) There exists \(\bar{q} : X_2 \to \gamma\) and exists \(\bar{V} \in \mathcal{B}_2\) such that \(\mathcal{P} \circ \mu_2(X_2 \backslash \bar{V}) = 0\); \(\forall x_2 \in \bar{V}\), \(f(x_2)\) is absolutely integrable over \(X_1\) and \(\bar{q}(x_2) = \int_{X_1} f(x_2) \, d\mu_1 \in \gamma\); \(\bar{q}\) is absolutely integrable over \(X_2\); and \(\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} \bar{q} \, d\mu_2 \in \gamma\).

(iii) \(\forall p : X_1 \to \gamma\) satisfying: \(p\) is \(\mathcal{B}_1\)-measurable; and \(\exists U \in \mathcal{B}_1\) with \(\mathcal{P} \circ \mu_1(X_1 \backslash U) = 0\) such that, \(\forall x_1 \in U\), \(f_{x_1}\) is absolutely integrable over \(X_2\) and \(p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \gamma\), we have \(p\) is absolutely integrable over \(X_1\) and \(\int_{X_1 \times X_2} f \, d\mu = \int_{X_1} p \, d\mu_1 \in \gamma\).

(iv) \(\forall q : X_2 \to \gamma\) satisfying: \(q\) is \(\mathcal{B}_2\)-measurable; and \(\exists V \in \mathcal{B}_2\) with \(\mathcal{P} \circ \mu_2(X_2 \backslash V) = 0\) such that, \(\forall x_2 \in V\), \(f(x_2)\) is absolutely integrable over \(X_1\) and \(q(x_2) = \int_{X_1} f(x_2) \, d\mu_1 \in \gamma\), we have \(q\) is absolutely integrable over \(X_2\) and \(\int_{X_1 \times X_2} f \, d\mu = \int_{X_2} q \, d\mu_2 \in \gamma\).

**Proof**

(i) Let \(X : = (X_i, \mathcal{B}_i, \mathcal{P} \circ \mu_i =: \nu_i), i = 1, 2,\) and \(X : = (X_1 \times X_2, \mathcal{B}, \mathcal{P} \circ \mu =: \nu)\). By Proposition 12.22, \(X_1, X_2,\) and \(X\) are \(\sigma\)-finite measure spaces and \(X = X_1 \times X_2\). Fix any \(i = 1, 2\). By Radon-Nikodym Theorem 11.171, we may define functions \(\frac{d\nu}{d\nu_i} =: g_i : X_i \to K\). \(\forall E \in \mathcal{B}_i, \nu_i(E) = \mathcal{P} \circ \mu_i(E) = \int_E \mathcal{P} \circ g_i \, d\nu_i\). By Proposition 11.167, we have \(\mathcal{P} \circ g_i = 1\) a.e. in \(X_i\). By Proposition 12.34, we have \(\frac{d\nu}{d\mu} = g_1 g_2\) a.e. in \(X\).
12.4 Functions of Bounded Variation

Since \( f \) is absolutely integrable over \( X \), then, by Proposition 11.168,
\[
\int_{X_1 \times X_2} f \, d\mu = \int_{X_1 \times X_2} (f \circ g_1) \, d\nu \text{ and } (P \circ f)(P \circ (g_1 \circ g_2)) = (P \circ f)(P \circ g_1)(P \circ g_2)
\]
is integrable over \( X \). Define \( \bar{\rho} := f \circ g_1 \circ g_2 : X_1 \times X_2 \to \mathbb{Y} \). By Fubini Theorem 12.30, there exists \( \rho : X_1 \to \mathbb{Y} \) and exists \( \bar{U} \in B_1 \) such that \( \nu_1(X_1 \setminus \bar{U}) = 0 \); \( \forall x_1 \in \bar{U}, f_{x_1} \) is absolutely integrable over \( X_2 \) and \( \bar{\rho}(x_1) = \int_{X_2} f_{x_1} \, d\nu_2; \forall x_1 \in X_1 \setminus \bar{U}, \bar{\rho}(x_1) = \vartheta(y) \). \( \bar{\rho} \) is absolutely integrable over \( X_1 \); and \( \int_{X_1 \times X_2} \bar{f} \, d\nu = \int_{X_1} \bar{\rho} \, d\nu_1 \). Let \( A_1 := \{ x_1 \in X_1 \mid g_1(x_1) = 0 \} \). Clearly, \( A_1 \in B_1 \) and \( \nu_1(A_1) = 0 \). Let \( U := U \setminus A_1 \in B_1 \). Then, \( P \circ \mu_1(X_1 \setminus \bar{U}) = \nu_1(X_1 \setminus \bar{U}) = \nu_1((X_1 \setminus \bar{U}) \cup A_1) = 0 \). \( \forall x_1 \in \bar{U}, f_{x_1} \) is absolutely integrable over \( X_2 \) implies that \( f_{x_1} \circ g_1(x_1) \) is absolutely integrable over \( X_2 \), which further implies that \( (P \circ f_{x_1})(P \circ g_2) \) is integrable over \( X_2 \). By Proposition 11.168, \( f_{x_1} \) is absolutely integrable over \( X_2 \) and \( \int_{X_1} f_{x_1} \, d\mu_2 = \int_{X_1} (f_{x_1} \circ g_2) \, d\nu_2 \). Then, \( \bar{\rho}(x_1) = g_1(x_1) \int_{X_2} f_{x_1} \, d\mu_2 \).

Define \( \bar{\rho} : X_1 \to \mathbb{Y} \) by \( \bar{\rho}(x_1) = \left\{ \begin{array}{ll} \int_{X_2} f_{x_1} \, d\mu_2 & x_1 \in \bar{U} \\ \vartheta(y) & x_1 \in X_1 \setminus \bar{U} \end{array} \right. \). By Propositions 11.41, 11.38, and 11.39, \( \bar{\rho} \) is \( B_1 \)-measurable. Since \( \bar{\rho} \) and \( g_1 \bar{\rho} \) are \( B_1 \)-measurable, then, by Lemma 11.43, \( \bar{\rho} = g_1 \bar{\rho} \) a.e. in \( X_1 \). Since \( \bar{\rho} \) is absolutely integrable over \( X_2 \), then, \( P \circ (g_1 \bar{\rho}) = (P \circ \bar{\rho})(P \circ g_1) \) is integrable over \( X_1 \). By Proposition 11.168, \( \bar{\rho} \) is absolutely integrable over \( X_1 \) and \( \int_{X_1} (\bar{\rho} \circ g_1) \, d\nu_1 = \int_{X_1} \bar{\rho} \, d\nu_1 \). By Proposition 11.92, we have \( \int_{X_1} f \, d\mu = \int_{X_1} \bar{\rho} \, d\nu_1 = \int_{X_1} \bar{\rho} \, d\mu_1 \). Hence, (i) holds.

By symmetry, (ii) holds.

(iii) Fix any \( p : X_1 \to \mathbb{Y} \) satisfying: \( p \) is \( B_1 \)-measurable; and \( \exists U \in B_1 \) with \( P \circ \mu_1(X_1 \setminus U) = 0 \) such that, \( \forall x_1 \in U, f_{x_1} \) is absolutely integrable over \( X_2 \) and \( p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 \in \mathbb{Y} \). By (i), we have \( U := U \cap U \in B_1 \) and \( P \circ \mu_1(X_1 \setminus U_0) = P \circ \mu_1((X_1 \setminus U) \cup (X_1 \setminus U)) = 0 \). \( \forall x_1 \in U_0, p(x_1) = \int_{X_2} f_{x_1} \, d\mu_2 = \bar{\rho}(x_1) \in \mathbb{Y} \). By Lemma 11.43, \( p = \bar{\rho} \) a.e. in \( X_1 \). By Proposition 11.83, we have \( p \) is absolutely integrable over \( X_1 \) since \( \bar{\rho} \) is so. By Proposition 11.92, \( \int_{X_1} p \, d\mu_1 = \int_{X_1} \bar{\rho} \, d\mu_1 = \int_{X_1} \bar{\rho} \, d\nu_1 = \int_{X_1} \bar{\rho} \, d\mu_1 \). Hence, (iii) holds.

By symmetry, (iv) holds. This completes the proof of the theorem. \( \square \)

12.4 Functions of Bounded Variation

Definition 12.36 Let \( m \in \mathbb{N}, \mathbb{R}^m \) be endowed with the usual positive cone, \( \mathbb{Y} \) be a normed linear space over \( \mathbb{K} \), \( D \subseteq \mathbb{R}^m \), and \( F : D \to \mathbb{Y} \). \( F \) is said to be continuous on the right at \( x_0 \in D \) if \( \forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta \in (0, \infty) \subseteq \mathbb{R}, \forall x \in B_{\mathbb{R}^m}(x_0, \delta) \cap D \) with \( x_0 \leq x \), 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 \( x_\) for \( \forall x \in D \). \( F \) is said to be continuous on the right on \( E \subseteq D \) if it is continuous on the
right at $x$, $\forall x \in E$.

**Proposition 12.37** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone $P$, $Y$ be a normed linear space over $K$, $x_0 \in \Omega \subseteq \mathbb{R}^m$, and $F : \Omega \to Y$. 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), $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$ with $x_0 \preceq x$, we have $\|F(x) - F(x_0)\| < \epsilon$. Then, $\forall (h_n)_{n=1}^\infty \subseteq (P \cap (\Omega - x_0)) \setminus \{0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, $\exists N \in \mathbb{N}$ such that $\forall n \in N$ with $N \leq n$, $h_n \in \mathcal{B}_{\mathbb{R}^m}(0_m, \delta)$. This implies that $x_0 + h_n \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$, $x_0 + h_n \supsetneq x_0$, and $\|F(x_0 + h_n) - F(x_0)\| < \epsilon$. Hence, $\lim_{h \in \mathbb{R}} F(x_0 + h_n) = F(x_0)$.

(iii) $\Rightarrow$ (ii). Let $D := P \cap (\Omega - x_0)$. Define $\bar{F} : D \to Y$ by $\bar{F}(h) = F(x_0 + h)$, $\forall h \in D$. By (iii), we have $\forall (h_n)_{n=1}^\infty \subseteq D \setminus \{0_m\}$ with $\lim_{n \in \mathbb{N}} h_n = 0_m$, we have $\lim_{h \in \mathbb{R}} \bar{F}(h_n) = F(x_0)$. By Proposition 4.16, we have $\lim_{h \cdot a_m} \bar{F}(h) = F(x_0)$. Hence, $\lim_{h \cdot a_m} F(x_0 + h) = F(x_0)$.

(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 $\mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap ((\Omega \cap (x_0 + P)) \setminus \{x_0\}) = \emptyset$ and $\mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega \cap (x_0 + P) = \{x_0\}$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega$ with $x_0 \preceq x$, 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), $\lim_{h \cdot a_m} F(x_0 + h) = F(x_0)$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall h \in \mathcal{B}_{\mathbb{R}^m}(0_m, \delta) \cap P \cap (\Omega - x_0) \setminus \{0_m\}$, we have $\|F(x_0 + h) - F(x_0)\| < \epsilon$. Then, $\forall x \in \mathcal{B}_{\mathbb{R}^m}(x_0, \delta) \cap \Omega \cap (x_0 + P)$, we have $x - x_0 \in \mathcal{B}_{\mathbb{R}^m}(0_m, \delta) \cap (\Omega - x_0) \cap P$ 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. $\forall x_1, x_2 \in \mathbb{R}^m$ with $x_1 \preceq x_2$, the semi-open rectangle with corners $x_1$ and $x_2$ is $R_{x_1, x_2} := \{x \in \mathbb{R}^m \mid x_1 < x \preceq x_2\}$. The closed rectangle
with corners $x_1$ and $x_2$ is $r_{x_1,x_2} := \{ x \in \mathbb{R}^m \mid x_1 \leq x \leq x_2 \}$. The open rectangle with corners $x_1$ and $x_2$ is $r_{x_1,x_2}^0 := \{ x \in \mathbb{R}^m \mid x_1 < x < x_2 \}$.

$\Omega \subseteq \mathbb{R}^m$ is said to be a region if $\exists N \subseteq \mathbb{N}$ and $(\bar{x}_i)_{i \in N}, (\hat{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \leq \hat{x}_i$ and $r_{\bar{x}_i,\hat{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} r_{\bar{x}_i,\hat{x}_i}$, $(r_{\bar{x}_i,\hat{x}_i})_{i \in N}$ is pairwise disjoint, $P(\Omega) = \bigcup_{i \in N} r_{\bar{x}_i,\hat{x}_i}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1,x_2} \subseteq \Omega$, we have $r_{x_1,x_2} \subseteq P(\Omega)$. $P(\Omega)$ 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 $\Omega \in \mathcal{B}(\mathbb{R}^m)$ and $P(\Omega) \in \mathcal{B}(\mathbb{R}^m)$.

**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, $\Omega \in \mathcal{B}(\mathbb{R}^m)$, and $P(\Omega) \in \mathcal{B}(\mathbb{R}^m)$.

(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 $\Omega \in \mathcal{B}(\mathbb{R}^m)$ be a rectangle, that is $\exists$ intervals $I_1, \ldots, I_m$ such that $\Omega = \prod_{i=1}^m I_i$, $\forall i \in \{1, \ldots, m\} =: J$, if $I_i = [a_i, b_i] \subset \mathbb{R}$ with $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, let $\bar{I}_i := (a_i, b_i) \subset \mathbb{R}$; if $I_i = [a_i, \infty) \subset \mathbb{R}$ with $a_i \in \mathbb{R}$, let $\bar{I}_i := [a_i, \infty) \subset \mathbb{R}$; if $I_i = [\infty, b_i) \subset \mathbb{R}$ with $b_i \in \mathbb{R}$, let $\bar{I}_i := [\infty, b_i) \subset \mathbb{R}$; in all other cases, let $\bar{I}_i := I_i$. Then, $\Omega$ is a region and $P(\Omega) = \prod_{i=1}^m \bar{I}_i$ 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 $r_{x_1,x_2}$ are regions with $P(\mathbf{r}_{x_1,x_2}) = P(r_{x_1,x_2}) = r_{x_1,x_2}$.

(vii) Let $M \subseteq \mathbb{N}$ and $(\bar{x}_j)_{j \in M}, (\hat{x}_j)_{j \in M} \subseteq \mathbb{R}^m$ with $\bar{x}_j \leq \hat{x}_j$, $\forall j \in M$. Then, $\Omega := \bigcup_{j \in M} r_{\bar{x}_j,\hat{x}_j} \subseteq \mathbb{R}^m$ 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 $(\bar{x}_{j,i})_{i \in N_j}, (\hat{x}_{j,i})_{i \in N_j} \subseteq \Omega$ with $\bar{x}_{j,i} \leq \hat{x}_{j,i}$ and $r_{\bar{x}_{j,i},\hat{x}_{j,i}} \subseteq \Omega$, $\forall i \in N_j$, such that $\Omega = \bigcup_{i \in N_j} r_{\bar{x}_{j,i},\hat{x}_{j,i}}$, $(r_{\bar{x}_{j,i},\hat{x}_{j,i}})_{i \in N_j}$ is pairwise disjoint, $\Omega_j := \bigcup_{i \in N_j} r_{\bar{x}_{j,i},\hat{x}_{j,i}}$, and $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1,x_2} \subseteq \Omega$, we have $r_{x_1,x_2} \subseteq \Omega_j$. We will show that $\Omega_1 = \Omega_2$. Then, the result follows.
∀i ∈ N₁, we have \( \bar{x}_{i,i} \leq \bar{x}_{1,i} \) and \( \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \subseteq \Omega \). This implies that \( \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \subseteq \Omega_2 \). Hence, \( \Omega_1 = \bigcup_{i \in N} \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \subseteq \Omega_2 \). By symmetry, we have \( \Omega_2 \subseteq \Omega_1 \).

(ii) By \( \Omega_1 \) being a region, \( \exists N_1 \subseteq \mathbb{N} \) and \((\bar{x}_{i,i})_{i \in N_1},(\bar{x}_{1,i})_{i \in N_1} \subseteq \Omega_1 \subseteq \Omega_2 \) with \( \bar{x}_{i,i} \leq \bar{x}_{1,i} \) and \( \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \subseteq \Omega_1 \subseteq \Omega_2 \), ∀i ∈ N₁, such that \( \Omega_1 = \bigcup_{i \in N_1} \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \text{ and } \mathbb{P}(\Omega_1) = \bigcup_{i \in N_1} \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \). By \( \Omega_2 \) being a region, we have \( \bar{r}_{\bar{x}_{i,i},\bar{x}_{1,i}} \subseteq \mathbb{P}(\Omega_2), \forall i \in N_1 \). Hence, \( \mathbb{P}(\Omega_1) \subseteq \mathbb{P}(\Omega_2) \).

(iii) By (ii), \( \mathbb{P}(\Omega_1 \cup \Omega_2) \supseteq \mathbb{P}(\Omega_i), \ i = 1, 2 \). Then, the result follows.

(iv) By (ii), \( \mathbb{P}(\Omega_1 \cap \Omega_2) \subseteq \mathbb{P}(\Omega_i), \ i = 1, 2 \). 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 \((\bar{x}_{j,i})_{j \in M} =: (\Omega_j)_{j \in M} \) is pairwise disjoint. ∀j ∈ M, by (vi), \( \Omega_j \) is a region and \( \mathbb{P}(\Omega_j) = \Omega_j \). Then, \( \exists N_j \subseteq \mathbb{N} \) and \((\bar{x}_{j,i})_{i \in N_j},(\bar{x}_{j,i})_{i \in N_j} \subseteq \Omega_j \) with \( \bar{x}_{j,i} \leq x_{j,i} \), ∀i ∈ N_j, such that \( \Omega_j = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} \) is pairwise disjoint, \( \mathbb{P}(\Omega_j) = \Omega_j = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} \), and \( \forall x_1, x_2 \in \Omega_j \) with \( x_1 \leq x_2 \) and \( \bar{r}_{x_1,x_2} \subseteq \Omega_j \), we have \( r_{x_1,x_2} \subseteq \mathbb{P}(\Omega_j) = \Omega_j \). For \( \Omega := \bigcup_{j \in M} \Omega_j \), consider \((\bar{x}_{j,i})_{j \in M},i \in N_j, (\bar{x}_{j,i})_{j \in M},i \in N_j \subseteq \Omega \). We have \( \bar{x}_{j,i} \equiv \bar{x}_{j,i} \), ∀j ∈ M, \( \forall i \in N_j \), \((\bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}})_{j \in M},i \in N_j \subseteq \Omega \). We have \( \bar{x}_{j,i} \equiv \bar{x}_{j,i} \), ∀j ∈ M, \( \forall i \in N_j \), \((\bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}})_{j \in M},i \in N_j \subseteq \Omega \). We have \( \mathbb{P}(\Omega) = \bigcup_{j \in M} \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} = \bigcup_{j \in M} \Omega_j = \Omega_j \), and \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \bar{r}_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq \mathbb{P}(\Omega) = \Omega_j \). Hence, \( \Omega \) is a region with \( \mathbb{P}(\Omega) = \Omega \).

Next, we prove the general case when \((\Omega_j)_{j \in M} \) is not necessarily pairwise disjoint. Without loss of generality, take \( M = \mathbb{N} \). Then, \( \Omega = \bigcup_{n=1}^{\infty} \left( \Omega_n \setminus \left( \bigcup_{k=1}^{n-1} \Omega_k \right) \right) \). Since for open semi-rectangles \( A \) and \( B \), \( A \setminus B \) 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 \( \bar{x}_n, \bar{x}_n \in \mathbb{R}^m \) with \( \bar{x}_n \equiv \bar{x}_n \), ∀n ∈ N, and the sets in the union are pairwise disjoint. By the special case, we have \( \Omega \) is a region and \( \mathbb{P}(\Omega) = \Omega \).

(viii) By \( \Omega_j \) being a region, \( \exists N_j \subseteq \mathbb{N} \) 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 \( \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} \subseteq \Omega_j \), ∀i ∈ N_j, such that \( \Omega_j = \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} \) is pairwise disjoint, \( \mathbb{P}(\Omega_j) := \bigcup_{i \in N_j} \bar{r}_{\bar{x}_{j,i},\bar{x}_{j,i}} \), and \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \bar{r}_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq \mathbb{P}(\Omega_j) \), \( j = 1, 2 \).

Consider \((\bar{x}_{1,i},\bar{x}_{2,j})_{i \in N_1,j \in N_2},(\bar{x}_{1,i},\bar{x}_{2,j})_{i \in N_1,j \in N_2} \subseteq \Omega_1 \times \Omega_2 \). We have \((\bar{x}_{1,i},\bar{x}_{2,j}) \leq (\bar{x}_{1,i},\bar{x}_{2,j}) \) and \( r_{(\bar{x}_{1,i},\bar{x}_{2,j}),((\bar{x}_{1,i},\bar{x}_{2,j}))_{i \in N_1,j \in N_2}} \subseteq \Omega_1 \times \Omega_2 \), ∀i ∈ N₁, ∀j ∈ N₂. Furthermore, \( \Omega_1 \times \Omega_2 = \bigcup_{i \in N_1} \bigcup_{j \in N_2} \bar{r}_{(\bar{x}_{1,i},\bar{x}_{2,j}),((\bar{x}_{1,i},\bar{x}_{2,j}))_{i \in N_1,j \in N_2}} \). Therefore, \( \mathbb{P}(\Omega_1) \times \mathbb{P}(\Omega_2) = \bigcup_{i \in N_1} \bigcup_{j \in N_2} r_{(\bar{x}_{1,i},\bar{x}_{2,j}),((\bar{x}_{1,i},\bar{x}_{2,j}))_{i \in N_1,j \in N_2}} \). ∀(\bar{x}, \bar{y}),
\( (\bar{x}, \bar{y}) \in \Omega_1 \times \Omega_2 \) with \( (\bar{x}, \bar{y}) \leq (\bar{x}, \bar{y}) \) and \( \bar{r}(\bar{x}, \bar{y}), (\bar{x}, \bar{y}) \subset \Omega_1 \times \Omega_2 \), we have \( \bar{x} \leq \bar{x} \) and \( \bar{y} \leq \bar{y} \). \( \bar{r}(\bar{x}, \bar{y}), (\bar{x}, \bar{y}) = \bar{r}(\bar{x}, \bar{y}) \times \bar{r}(\bar{y}, \bar{y}) \subset \Omega_1 \times \Omega_2 \), \( \bar{r}(\bar{x}, \bar{z}) \subset \Omega_1 \), \( \bar{r}(\bar{y}, \bar{z}) \subset \Omega_2 \), which further implies that \( \bar{r}(\bar{x}, \bar{z}) \subset P(\Omega_1) \), \( \bar{r}(\bar{y}, \bar{z}) \subset P(\Omega_2) \), and \( \bar{r}(\bar{x}, \bar{z}), \bar{r}(\bar{y}, \bar{z}) \subset P(\Omega_1) \times P(\Omega_2) \). Hence, \( \Omega_1 \times \Omega_2 \) is a region with \( P(\Omega_1) \times P(\Omega_2) = P(\Omega_1) \times P(\Omega_2) \).

This completes the proof of the proposition.

\[ \square \]

**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 \( D \subset \Omega \) be a countable dense subset and define \( B := \{ r^{\delta}_{x-\delta 1_m, x+\delta 1_m} : x \in D, \delta \in \mathbb{Q}, \delta > 0, \bar{r}(x-\delta 1_m, x+\delta 1_m) \subset \Omega \} \). It is easy to show, by Definition 3.17, that \( B \) is a countable basis for \( \mathcal{O} \). Then, \( \Omega = \bigcup_{B \in \mathcal{B}} B =: \bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} \bar{r}(x-\delta 1_m, x+\delta 1_m) \subset \bigcup_{n \in \mathbb{N}} \bar{r}(x-\delta 1_m, x+\delta 1_m) \), where \( n \in \mathbb{N} \). By (vii) of Proposition 12.39, \( \Omega \) is a region with \( P(\Omega) = \Omega \). This completes the proof of the proposition.

\[ \square \]

**Definition 12.41** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \mathcal{O} \) be a normed linear space over \( \mathbb{K} \), \( \Omega \in \mathcal{O}_{\mathbb{R}^m} \) be a region, and \( F : \Omega \to \mathbb{Y} \). For \( x_1, x_2 \in \mathbb{R}^m \) with \( x_1 \leq x_2 \), \( \forall J : J = \{1, \ldots, m\} \), the vertex 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 \( x_J = \bigwedge_{J \subseteq \{1, \ldots, m\}} \pi_J(x_1) \). We define the set of vertexes of \( r_{x_1, x_2} \) with \( i \) coordinates equal to that of \( x_1 \) to be \( \text{VRect}_{i, x_1, x_2} := \{ x_J \in \mathbb{R}^m \mid J \subset J, \text{card}(J) = i \} \), if \( x_1 \neq x_2 \). When \( r_{x_1, x_2} \subset \Omega \), the increment of \( F \) on \( r_{x_1, x_2} \) is \( \Delta F(r_{x_1, x_2}) := \sum_{J \subseteq \{1, \ldots, m\}} (-1)^{\text{card}(J)} F(x_J) \). The total variation of \( F \) on the semi-open rectangle \( r_{x_1, x_2} \) is

\[ T_F (r_{x_1, x_2}) := \sup_{n \in \mathbb{N}^+} \sum_{i=1}^n \| \Delta F(r_{\bar{x}_i, \bar{x}_i}) \| \]

where \( T_F (r_{x_1, x_2}) \in [0, \infty) \subset \mathbb{R}_+ \). \( F \) 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 \( r_{x_1, x_2} \subset \Omega \), we have \( T_F (r_{x_1, x_2}) < \infty; \)
(iii) \( \forall x_1, x_2 \in \Omega \) with \( x_1 < x_2 \) and \( \tau_{x_1, x_2} \subseteq \Omega \), \( \tilde{G} : \tau_{x_1, x_2} \to [0, \infty) \subset \mathbb{R} \) defined by \( \tilde{G}(x) = T_F(\tau_{x_1, x_2}) \), \( \forall x \in \tau_{x_1, x_2} \), is continuous on the right.

The total variation of \( F \) is defined by \( T_F := \sum_{i \in \mathbb{N}} T_F(\tau_{\tilde{x}_i, \tilde{x}_i}) \) 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, \( \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) be a region with the subset topology, \( \mu \) be a \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on \( (\mathbb{P}(\Omega), \mathcal{B}_B(\mathbb{P}(\Omega))) \), 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 \( \tau_{x_1, x_2} \subseteq \Omega \) and \( x_1 \leq x_2 \), we have \( \mu(\tau_{x_1, x_2}) = \Delta_F(\tau_{x_1, x_2}) \). \( \mu \) is said to be locally finite if \( \mathcal{P} \circ \mu(K) < \infty \), \( \forall \) compact \( K \subseteq \mathbb{P}(\Omega) \).

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 \( F(x) = \mu(\mathcal{P}(\Omega) \cap \{ \tilde{x} \in \mathbb{R}^m \mid \tilde{x} \leq x \}) \), \( \forall x \in \Omega \), 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 \( \mathcal{M} := \{ \prod_{i=1}^m B_i \subseteq X \mid \prod_{i=1}^m B_i \in \mathcal{B}_i, i = 1, \ldots, m \} \). 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 \( C_m \subseteq \mathcal{F} \). Clearly, \( \emptyset \in \mathcal{C}_m \subseteq \mathcal{F} \). Then, by \( \mathcal{C}_m \) being a semi-algebra on \( X_m \), we have \( X_m = X_m \setminus \emptyset = \bigcup_{n=1}^\infty C_{m,i} \), where \( n \in \mathbb{Z}_+ \) and \( (C_{m,i})_{i=1}^n \subseteq \mathcal{C}_m \) is pairwise disjoint. Then, \( (\prod_{i=1}^{m-1} C_i) \times X_m = \bigcup_{n=1}^\infty (\prod_{i=1}^{m-1} C_i) \times C_{m,i} \in \mathcal{B} \). This implies that \( X_m \in \mathcal{F} \). \( \forall E \in \mathcal{F} \), we have \( (\prod_{i=1}^{m-1} C_i) \times E \in \mathcal{B} \). Then, by \( \mathcal{B} \) being a \( \sigma \)-algebra and \( (\prod_{i=1}^{m-1} C_i) \times X_m \in \mathcal{B} \), we have \( (\prod_{i=1}^{m-1} C_i) \times (X_m \setminus E) = (\prod_{i=1}^{m-1} C_i) \times (X_m \setminus X_m) \subseteq (\prod_{i=1}^{m-1} C_i) \times E \) \( \subseteq \mathcal{B} \). Thus, \( X_m \setminus E \in \mathcal{F} \). \( \forall (E_j)_{j=1}^\infty \in \mathcal{F} \), we have \( (\prod_{i=1}^{m-1} C_i) \times (\bigcup_{j=1}^\infty E_j) \in \mathcal{B} \). Then, by \( \mathcal{B} \) being a \( \sigma \)-algebra, we have \( (\prod_{i=1}^{m-1} C_i) \times (\bigcup_{j=1}^\infty E_j) = \bigcup_{j=1}^\infty ((\prod_{i=1}^{m-1} C_i) \times E_j) \in \mathcal{B} \). Thus, \( \bigcup_{j=1}^\infty E_j \in \mathcal{F} \). This shows that \( \mathcal{F} \) is a \( \sigma \)-algebra. Hence, \( \mathcal{B}_m \subseteq \mathcal{F} \) since \( \mathcal{B}_m \) is the \( \sigma \)-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

\[ X_i = X_i \setminus \emptyset = \bigcup_{j=1}^{n_i} C_{i,j}, \text{ where } n_i \in \mathbb{Z}^+ \text{ and } (C_{i,j})_{j=1}^{n_i} \subseteq \mathcal{C}_i \text{ is pairwise disjoint. By Claim 12.43.1, } \left( \prod_{i=1}^{m-1} C_{i,j} \right) \times B_m \in \mathcal{B}, j = 1, \ldots, n_i, i = 1, \ldots, m - 1. \text{ Then, } \left( \prod_{i=1}^{m-1} X_i \right) \times B_m = \left( \prod_{i=1}^{m-1} \left( \prod_{j=1}^{n_i} C_{i,j} \right) \times B_m \right) \in \mathcal{B}. \text{ By symmetry, we have } \left( \prod_{i=1}^{k-1} X_i \right) \times B_k \times \left( \prod_{i=k+1}^{m} X_i \right) \in \mathcal{B}, k = 1, \ldots, m. \text{ Then, } \prod_{i=1}^{m} B_i = \bigcap_{k=1}^{m} \left( \prod_{i=1}^{k-1} X_i \right) \times B_k \times \left( \prod_{i=k+1}^{m} X_i \right) \in \mathcal{B}. \text{ By the arbitrariness of } B_i \text{'s, we have } \mathcal{C} \subseteq \mathcal{M} \subseteq \mathcal{B}. \text{ Since } \mathcal{B} \text{ is the } \sigma\text{-algebra on } X \text{ generated by } \mathcal{C}, \text{ then } \mathcal{B} \text{ is the } \sigma\text{-algebra on } X \text{ generated by } \mathcal{M}. \]

Proposition 12.44 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \bar{x}, \tilde{x} \in \mathbb{R}^m \) with \( \bar{x} \leq \tilde{x} \), \( X := r_{\bar{x}, \tilde{x}} \subseteq \mathbb{R}^m \) with subset topology \( \mathcal{O} \), \( \mathcal{X} := (X, \mathcal{O}) \) be the topological space, \( \mathcal{C} := \{ r_{x_1, x_2} \mid \bar{x} \leq x_1 \leq x_2 \leq \tilde{x} \} \), and \( \mathcal{B} \) be the \( \sigma\text{-algebra on } X \) generated by \( \mathcal{C} \). 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 \( X \), \( \forall i \in \mathbb{N} \), and \( \mathcal{B} := \{ E \subseteq X \mid E \cap X_i \in \mathcal{B}_{\mathcal{B}}(X_i), \forall i \in \mathbb{N} \} \). 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, \( \forall \bar{x}, \tilde{x} \in \Omega \) with \( \bar{x} \leq \tilde{x} \) and \( r_{\bar{x}, \tilde{x}} \subseteq \Omega \), \( \forall n \in \mathbb{Z}^+, \forall (\bar{x}_i)_{i=1}^{n} \subseteq \Omega \) with \( \bar{x}_i \leq \tilde{x}_i \), \( r_{\bar{x}_i, \tilde{x}_i} \subseteq \Omega \), \( \forall i \in \{ 1, \ldots, n \} \), \( r_{\bar{x}, \tilde{x}} = \bigcup_{i=1}^{n} r_{\bar{x}_i, \tilde{x}_i} \), and the sets in the union being pairwise disjoint, we have \( \Delta_F(r_{\bar{x}, \tilde{x}}) = \sum_{i=1}^{n} \Delta_F(r_{\bar{x}_i, \tilde{x}_i}) \).

Proof We will prove the result by mathematical induction on \( n \). Clearly, \( r_{\bar{x}, \tilde{x}} = \emptyset \) and \( \bar{x} \neq \tilde{x} \). Then, \( \Delta_F(r_{\bar{x}, \tilde{x}}) = \emptyset \gamma = \sum_{i=1}^{0} \Delta_F(r_{\bar{x}_i, \tilde{x}_i}) \). 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 \( \bar{x}_i < \bar{x}_i \), \( i = 1, 2 \). Otherwise, the result follows from \( n = 1 \) or \( n = 0 \) cases, since \( \Delta^n_F(\tau_{x_1, x_2}) = \emptyset \) if \( x_1 \leq x_2 \) and \( x_1 \neq x_2 \). Then, we have \( \left( \bar{x}_i \right)_{i=1}^n, \left( \bar{x}_i \right)_{i=1}^n \subseteq \tau_{\bar{x}, \bar{x}} \). Without loss of generality, assume \( \bar{x} \in \tau_{\bar{x}_2, \bar{x}_2} \). Then, \( \bar{x} = \bar{x}_2 \) and \( \bar{x} \neq \bar{x}_2 = \bar{x} \).

Clearly, \( \bar{x} \neq \bar{x}_2 \) since \( \tau_{\bar{x}, \bar{x}} = \bigcup_{i=1}^2 \tau_{\bar{x}_i, \bar{x}_i} \) and the sets in the union are nonempty and pairwise disjoint. Then, \( \exists j_0 \in \{1, \ldots, m\} \) such that \( \pi_{j_0}(\bar{x}) > \pi_{j_0}(\bar{x}_2) \). Without loss of generality, assume that \( j_0 = m \).

Define \( x_C \in \tau_{\bar{x}, \bar{x}} \) be such that \( \pi_j(x_C) = \begin{cases} \pi_j(x) & j = 1, \ldots, m - 1 \\ \pi_m(x_2) & j = m \end{cases} \) and \( x_D \in \tau_{\bar{x}_1, \bar{x}_1} \) be such that \( \pi_j(x_D) = \begin{cases} \pi_j(x) & j = 1, \ldots, m - 1 \\ \pi_m(x_2) & j = m \end{cases} \). Clearly, we have \( x_D \in \tau_{\bar{x}_1, \bar{x}_1} \setminus \tau_{\bar{x}_2, \bar{x}_2} \). 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

\[
\sum_{i=1}^2 \Delta^n_F(\tau_{\bar{x}_i, \bar{x}_i}) = \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_i, x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_i, x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_1, x_D}} \sum_{\pi_m(s) = \pi_m(x)} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_1, x_C}} (-1)^i F(s)
\]

\[
+ \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_2, x_D}} \sum_{\pi_m(s) = \pi_m(x_2)} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_2, x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_1, x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_2, x_C}} (-1)^i F(s)
\]

\[
+ \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_1, x_D}} (-1)^i F(s) + \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_2, x_C}} (-1)^i F(s)
\]

\[
= \sum_{i=0}^m \sum_{s \in \text{VRect}_{\bar{x}_1, x_D}} (-1)^i F(s) = \Delta^n_F(\tau_{\bar{x}, \bar{x}})
\]

where the first two equalities follow from Definition 12.41; the third equality follows from the fact that, \( \forall i \in \{0, \ldots, m\}, \forall s \in \text{VRect}_{\bar{x}, x_{i+1}} \) with \( \pi_m(s) = \pi_m(x_D) \) if, and only if, \( s \in \text{VRect}_{\bar{x}_1, x_{i+1}} \) with \( \pi_m(s) = \pi_m(x_C) \), \( \forall s \in \text{VRect}_{\bar{x}_1, x_{i+1}} \) with \( \pi_m(s) = \pi_m(\bar{x}) \) if, and only if, \( s \in \text{VRect}_{\bar{x}_1, x_{i+1}} \) with \( \pi_m(s) = \pi_m(\bar{x}) \), and \( \forall s \in \text{VRect}_{\bar{x}_2, x_{i+1}} \) with \( \pi_m(s) = \pi_m(\bar{x}) \) if, and only if, \( s \in \text{VRect}_{\bar{x}_2, x_{i+1}} \) with \( \pi_m(s) = \pi_m(\bar{x}) \); and the fourth equality follows from the fact that the second and fourth summations cancel each other out. This case is proved.

Assume that the result holds for \( n \leq k - 1 \), where \( k \in \mathbb{N} \) and \( k \geq 3 \).  

\[\text{Case } n = k \leq k - 1, \text{ where } k \in \mathbb{N} \text{ and } k \geq 3.\]
3° Consider the case when \( n = k \in \{3, 4, \ldots\} \). Without loss of generality, assume that \( \hat{x} \not\preceq \hat{x} \) and \( \hat{x} \not\preceq \hat{x}_i, \ i = 1, \ldots, n \). Otherwise, the result follows from 1° or 2°, since \( \Delta F(x_1, x_2) = \partial y \) if \( x_1 \equiv x_2 \) and \( x_1 \not\preceq x_2 \). Then, we have \((\hat{x}_i)_{i=1}^n \subseteq T_{\hat{x}, \hat{x}}\). Without loss of generality, assume \( \hat{x} \in r_{\hat{x}, \hat{x}} \). Then, \( \hat{x} = \hat{x}_n \) and \( \hat{x} \not\preceq \hat{x}_n < \hat{x}_n = \hat{x} \).

Clearly, \( \hat{x} \not\preceq \hat{x}_n \) since \( r_{\hat{x}, \hat{x}} = \bigcup_{i=1}^n r_{\hat{x}, \hat{x}_i} \) 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 \( x_C \in T_{\hat{x}, \hat{x}} \) 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 \( x_D \in T_{\hat{x}, \hat{x}} \) 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, \( r_{\hat{x}, \hat{x}} = r_{\hat{x}, x_D} \cup r_{x_C, \hat{x}} = (\bigcup_{i=1}^n (r_{\hat{x}, x_D} \cap r_{\hat{x}, \hat{x}_i})) \cup (\bigcup_{i=1}^n (r_{x_C, \hat{x}} \cap r_{\hat{x}, \hat{x}_i})). \)

Clearly, \( r_{x_C, \hat{x}} \cap r_{\hat{x}, \hat{x}_i} = \emptyset \). Note that \( x_D \in r_{\hat{x}, \hat{x}} \setminus r_{\hat{x}, \hat{x}_n} \). Without loss of generality, assume that \( x_D \not\preceq \hat{x}_n \). Then, \( \hat{x} \not\preceq \hat{x}_n - 1 < x_D \not\preceq \hat{x}_n - 1 \not\preceq \hat{x} \).

Then, we must have \( x_D = \hat{x}_n - 1 \) since \( r_{\hat{x}_n - 1, \hat{x}_n - 1} \cap r_{\hat{x}, \hat{x}_n} = \emptyset \). Thus, \( r_{x_C, \hat{x}} \cap r_{\hat{x}, \hat{x}_n - 1} = \emptyset \). Then, \( r_{\hat{x}, x_D} = \bigcup_{i=1}^n (r_{\hat{x}, x_D} \cap r_{\hat{x}, \hat{x}_i}) \) and \( r_{x_C, \hat{x}} = \bigcup_{i=1}^n (r_{x_C, \hat{x}} \cap r_{\hat{x}, \hat{x}_i}). \)

By the inductive assumption, we have \( \Delta F(r_{\hat{x}, \hat{x}}) = \Delta F(r_{\hat{x}, x_D}) = \sum_{i=1}^{n-1} \Delta F(r_{\hat{x}, x_D} \cap r_{\hat{x}, \hat{x}_i}) \). Hence, \( \Delta F(r_{\hat{x}, \hat{x}}) = \sum_{i=1}^{n-1} \Delta F(r_{\hat{x}, x_D} \cap r_{\hat{x}, \hat{x}_i}) + \sum_{i=1}^{n} \Delta F(r_{x_C, \hat{x}} \cap r_{\hat{x}, \hat{x}_i}) \). This case is proved.

Therefore, the result holds for any \( n \in \mathbb{N} \). This completes the proof of the proposition.

Proposition 12.47 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \mathbb{Y} \) be a normed linear space, \( \Omega \subseteq \mathbb{R}^m \), and \( F : \Omega \to \mathbb{Y} \). Assume that \( F \) satisfies (i) and (ii) of Definition 12.41. Then, \( \forall \hat{x}, \hat{x} \in \Omega \) with \( \hat{x} \not\preceq \hat{x} \) and \( r_{\hat{x}, \hat{x}} \subseteq \Omega \), \( \forall i \in \mathbb{Z}_+ \), \( \forall (\hat{x}_i)_{i=1}^{n} \subseteq T_{\hat{x}, \hat{x}} \) with \( \hat{x}_i \not\preceq \hat{x}_i \), \( i = 1, \ldots, n \), \( r_{\hat{x}, \hat{x}_i} = \bigcup_{j=1}^{n} r_{\hat{x}, \hat{x}_j} \), and \( \{r_{\hat{x}, \hat{x}_j} \} \) being pairwise disjoint, we have \( T_F(r_{\hat{x}, \hat{x}}) = \sum_{i=1}^{n} T_F(r_{\hat{x}, \hat{x}_i}) \).

Proof \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \forall i \in \{1, \ldots, n\} \), \( \exists n_i \in \mathbb{Z}_+ \), \( \exists (\hat{x}_{i,j})_{j=1}^{n_i} \subseteq T_{\hat{x}_i, \hat{x}_i} \) with \( \hat{x}_{i,j} \not\preceq \hat{x}_{i,j} \), \( j = 1, \ldots, n_i \), \( r_{\hat{x}, \hat{x}_i} = \bigcup_{j=1}^{n_i} r_{\hat{x}_i, \hat{x}_{i,j}} \), and \( \{r_{\hat{x}_i, \hat{x}_{i,j}} \} \) being pairwise disjoint, such that \( \sum_{j=1}^{n_i} \Delta F(r_{\hat{x}_i, \hat{x}_{i,j}}) > T_F(r_{\hat{x}_i, \hat{x}_i}) - \epsilon/(n + 1) \). Then, \( \sum_{i=1}^{n} T_F(r_{\hat{x}, \hat{x}_i}) - \epsilon < \sum_{i=1}^{n} \sum_{j=1}^{n_i} \Delta F(r_{\hat{x}_i, \hat{x}_{i,j}}) \leq \sum_{i=1}^{n} \Delta F(r_{\hat{x}, \hat{x}_i}) \). By the arbitrariness of \( \epsilon \), we have \( \sum_{i=1}^{n} \Delta F(r_{\hat{x}, \hat{x}_i}) \leq T_F(r_{\hat{x}, \hat{x}}) \).

On the other hand, \( \forall \epsilon \in (0, \infty) \subseteq \mathbb{R} \), \( \exists n_i \in \mathbb{Z}_+ \), \( \exists (\hat{x}_i)_{i=1}^{n} \subseteq T_{\hat{x}, \hat{x}} \) with \( \hat{x}_i \not\preceq \hat{x}_i \), \( i = 1, \ldots, n \), \( r_{\hat{x}, \hat{x}_i} = \bigcup_{j=1}^{n_i} r_{\hat{x}_i, \hat{x}_j} \), and \( \{r_{\hat{x}_i, \hat{x}_j} \} \) being pairwise disjoint, such that \( \sum_{j=1}^{n_i} \Delta F(r_{\hat{x}_i, \hat{x}_j}) > T_F(r_{\hat{x}_i, \hat{x}_i}) - \epsilon \). Then, \( T_F(r_{\hat{x}, \hat{x}}) - \epsilon < \sum_{i=1}^{n} \Delta F(r_{\hat{x}_i, \hat{x}_i}) \leq \sum_{i=1}^{n} T_F(r_{\hat{x}_i, \hat{x}_i}) \).
∀ \sum_{i=1}^{n} \sum_{j=1}^{n} \| \Delta F(r_{x,i} \cap r_{x,j}) \| \leq \sum_{i=1}^{n} T_F(r_{x,i}), \text{ where the second equality follows from Proposition 12.46. By the arbitrariness of } \epsilon, \text{ we have } T_F(r_{x,i}) \leq \sum_{i=1}^{n} T_F(r_{x,i}). \text{ This completes the proof of the proposition.} \square

**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 \( \Delta F(r_{x_i} \cap r_{x_j}) \geq 0 \), \( \forall x_1, x_2 \in \Omega \) with \( r_{x_1} \subseteq \Omega \) and \( x_1 \leq x_2 \). Then, there exists a unique \( \sigma \)-finite measure \( \mu \) on the measurable space \((P(\Omega), \mathcal{B}(P(\Omega)))\) such that \( F \) is a cumulative distribution function of \( \mu \). Furthermore, \( T_F(r_{x_1} \cap r_{x_2}) = \Delta F(r_{x_1}, r_{x_2}) = \mu(r_{x_1}, r_{x_2}), \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1}, r_{x_2} \subseteq \Omega \), and \( T_F = \mu(P(\Omega)) \in [0, \infty) \subseteq \mathbb{R} \).

**Proof:** Since \( \Omega \) is a region, then \( \exists N \subseteq \mathbb{N}, \exists (\bar{x}_i)_{i \in N}, (\bar{x}_i)_{i \in N} \subseteq \Omega \) with \( \bar{x}_i \leq \bar{x}_i \) and \( r_{\bar{x}_i} \subseteq \Omega, \forall i \in N \), such that \( \Omega = \bigcup_{i \in N} r_{\bar{x}_i} \). By Proposition 11.32, \( (i) \) of Proposition 11.32 is satisfied. We will prove the claim by Proposition 11.32. By (i) of Definition 12.41, \( \forall r \in (0, \infty) \subseteq \mathbb{R}, \exists y \in r_{\bar{x}} \) such that \( \Delta F(r_{\bar{x}}, r_{\bar{y}}) \geq \Delta F(r_{\bar{x}}, r_{\bar{x}}) - \epsilon/2 \). Since \( r_{\bar{x}} \subseteq r_{\bar{x}} \), by the assumption and (i) of Proposition 11.32, we have \( \Delta F(r_{\bar{x}}, r_{\bar{y}}) \leq \Delta F(r_{\bar{x}}, r_{\bar{x}}) \).

Fix any \( j \in \mathbb{N} \), \( C_j = r_{\bar{x}_j} \subseteq \mathcal{C} \subseteq \bar{\mathcal{C}} \subseteq \Omega \). By (i) of Definition 12.41, \( \exists y \in \mathbb{R}^m \) with \( \bar{x}_j < \bar{y}_j \) such that \( \Delta F(r_{\bar{x}_j}, r_{\bar{y}_j}) \leq \Delta F(r_{\bar{x}_j}, r_{\bar{x}_j}, r_{\bar{x}_j}) + 2^{-j-1} \epsilon \). Since \( C_j \subseteq r_{\bar{x}_j} \cap r_{\bar{y}_j} \cap r_{\bar{x}_j}, \) by the assumption and (i) of Proposition 11.32, we have \( \Delta F(r_{\bar{x}_j}, r_{\bar{x}_j}) \leq \Delta F(r_{\bar{x}_j}, r_{\bar{x}_j}, r_{\bar{x}_j}) \).

Obviously, \( r_{\bar{y}_j} \subseteq r_{\bar{x}_j} = C = \bigcup_{j=1}^{\infty} C_j = \bigcup_{j=1}^{\infty} r_{\bar{x}_j} \subseteq \bigcup_{j=1}^{\infty} (r_{\bar{x}_j} \cap r_{\bar{x}_j}) \). Since \( r_{\bar{y}_j} \) is compact, then \( \exists N \subseteq \mathbb{N} \) such that \( r_{\bar{y}_j} \subseteq r_{\bar{x}_j} \subseteq \bigcup_{j=1}^{N} (r_{\bar{x}_j} \cap r_{\bar{x}_j}) \subseteq \bigcup_{j=1}^{N} (r_{\bar{x}_j} \cap r_{\bar{x}_j}) \). By (i) of Proposition 11.32, \( \mu(C) \leq \Delta F(r_{\bar{x}_j}, r_{\bar{x}_j}) - \epsilon/2 \leq \sum_{j=1}^{N} \Delta F(r_{\bar{x}_j} \cap r_{\bar{x}_j}) + \epsilon/2 < \sum_{j=1}^{N} (\Delta F(r_{\bar{x}_j}, r_{\bar{x}_j}) + 2^{-j-1} \epsilon) + \epsilon/2 < \sum_{j=1}^{N} \mu(C_j) + \epsilon \leq \sum_{j=1}^{\infty} \mu(C_j) + \epsilon. \)
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 \((r_{\bar{x}_i}, \bar{x}_i, B_i)\) that is an extension of \( \hat{\mu} \), where \( B_i \) is the \( \sigma \)-algebra on \( r_{\bar{x}_i}, \bar{x}_i \)
generated by \( \mathcal{C}_i := \{ \bar{x}_i, x_2 \mid \bar{x}_i \leq x_1 \leq x_2 \leq \bar{x}_i \} \). By Proposition 12.44, \( B_i = B_B(r_{\bar{x}_i}, \bar{x}_i) \). Let \( \mathcal{X}_i := (r_{\bar{x}_i}, \bar{x}_i, \hat{\mu}_i) \) 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 \( \mathcal{X} := (P(\Omega), B, \hat{\mu}) \) such that \( \mathcal{X}_i \) is the finite measure subspace of \( \mathcal{X} \), \( \forall i \in N \).

By Proposition 12.45, \( B = B_B(P(\Omega)) \). Then, \( \mathcal{X} = (P(\Omega), B_B(P(\Omega)), \hat{\mu}) \).

Finally, we will show that \( T_F = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_r \). Note that \( T_F = \sum_{i \in N} T_F(r_{\bar{x}_i}, \bar{x}_i) = \sum_{i \in N} \Delta_F(r_{\bar{x}_i}, \bar{x}_i) = \sum_{i \in N} \hat{\mu}(r_{\bar{x}_i}, \bar{x}_i) = \hat{\mu}(P(\Omega)) \in [0, \infty] \subset \mathbb{R}_r \), where the first equality follows from Definition 12.41; the second equality follows from the assumption of the proposition and Definition 12.41; the third equality follows from the fact that \( T_F \) is a cumulative distribution function of \( \hat{\mu} \); and 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 \((\bar{x}_i)_{i \in N}\), 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, \( \Omega \in B_B(\mathbb{R}^m) \) 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 \( r_{x_1, x_2} \subseteq \Omega \), we have \( \Delta_F(r_{x_1, x_2}) \geq 0 \).

Then,

1. \( F \) is of locally bounded variation;

2. there exists a unique \( \sigma \)-finite measure \( \hat{\mu} \) on the measurable space \((P(\Omega), B_B(P(\Omega))) \) 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 \( \tau_{x,x} \subseteq \Omega \), \( \forall n \in \mathbb{Z}_+ \), \( \forall \bar{x}_1, \ldots, \bar{x}_n \in \tau_{x,x} \) with \( \bar{x}_i \leq \bar{x}_i \), \( \forall i \in \{1, \ldots, n\} \), \( r_{x,x} = \bigcup_{i=1}^n r_{x_i,x_i} \), and the sets in the union being pairwise disjoint, we have \( \sum_{i=1}^n |\Delta_F(r_{x_i,x_i})| = \sum_{i=1}^n \Delta_F(r_{x_i,x_i}) = \Delta_F(r_{x,x}) \in [0,\infty) \subset \mathbb{R} \), where the first equality follows from (ii); and the second equality follows from Proposition 12.46. Then, \( T_F(r_{x,x}) = \Delta_F(r_{x,x}) \in [0,\infty) \subset \mathbb{R} \). 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 \( G : \tau_{x_1,x_2} \rightarrow [0,\infty) \subset \mathbb{R} \) by \( G(x) = T_F(r_{x_1,x_2}) \), \( \forall x \in \tau_{x_1,x_2} \). By the previous paragraph, \( G(x) = \Delta_F(r_{x_1,x_2}) \), \( \forall x \in \tau_{x_1,x_2} \). Then, \( \forall x \in \tau_{x_1,x_2} \) with \( x \neq x_2 \), \( G(x) = \Delta_F(r_{x_1,x_2}) = \lim_{h \downarrow 0} \Delta_F(r_{x_1,x_2} + h) = \lim_{h \downarrow 0} G(x + h) \), 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. \( \square \)

**Theorem 12.50**  
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 \( \mathcal{O} \), \( \mathcal{Y} \) be a Banach space, \( F : \Omega \rightarrow \mathcal{Y} \) 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 \( \Omega \in \mathcal{B}_B(\mathbb{R}^m) \) is a region, then \( \exists N \subseteq \mathbb{N} \) and \( (\hat{x}_i)_{i \in N} \), \( (\hat{x}_i)_{i \in N} \subseteq \Omega \) with \( \hat{x}_i \leq \hat{x}_i \) and \( \tau_{x_i,x_i} \subseteq \Omega \), \( \forall i \in \mathbb{N} \), such that \( \Omega = \bigcup_{i \in N} \tau_{x_i,x_i} \). \( (\hat{x}_i, \hat{x}_j)_{i \in N} \) is pairwise disjoint, \( P(\Omega) := \bigcup_{i \in N} \tau_{\hat{x}_i,\hat{x}_i} \), and \( \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( \tau_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq P(\Omega) \). We need the following result.

**Claim 12.50.1**  
Fix any \( \bar{x}, \bar{x} \in \Omega \) with \( \bar{x} \leq \bar{x} \) and \( \tau_{x,x} \subseteq \Omega \), let \( \hat{\Omega} := r_{x,x} \subseteq \Omega \) be endowed with subset topology, then, there exists a unique finite \( \mathcal{Y} \)-valued measure \( \hat{\mu} \) on the measurable space \( (\hat{\Omega}, \mathcal{B}(\hat{\Omega})) \) such that \( \hat{\mu}(r_{x_1,x_2}) = \Delta_F(r_{x_1,x_2}) \) and \( \mathcal{P} \circ \hat{\mu}(r_{x_1,x_2}) = T_F(r_{x_1,x_2}) \), \( \forall x_1 \leq x_2 \leq \hat{x} \). Furthermore, \( \hat{\mu} \in \mathcal{M}_f(\hat{\Omega}, \mathcal{Y}) \).

**Proof of claim:**  
Let \( \hat{\Omega} := \tau_{x,x} \), \( \hat{F} := F|_{\hat{\Omega}} : \hat{\Omega} \rightarrow \mathcal{Y} \), and \( \hat{G} : \hat{\Omega} \rightarrow [0,\infty) \subset \mathbb{R} \) be defined by \( \hat{G}(x) = T_F(r_{x,x}) \), \( \forall x \in \hat{\Omega} \). 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, $\Delta_G(x_1, x_2) = T_P(x_1, x_2) \geq 0$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. By Proposition 12.49, $\tilde{G}$ is of locally bounded variation, there exists a unique $\sigma$-finite measure $\hat{\nu}$ on the measurable space $(\Omega, \mathcal{B}_\Omega)$ such that $\tilde{G}$ is a cumulative distribution function of $\hat{\nu}$, and $\hat{\nu}(\Omega) = T_G(\Omega) < \infty$. Hence, $\hat{\nu}$ is finite.

Let $\mathcal{C} := \left\{ x_{x_1, x_2} \mid \tilde{x} \leq x_1 \leq x_2 \leq \tilde{x} \right\}$ be the collection of semi-open rectangles contained in $\Omega$. Clearly, $\mathcal{C}$ is a semialgebra on $\Omega$. Define $\nu : \mathcal{C} \rightarrow [0, \infty) \subset \mathbb{R}$ by $\nu := \hat{\mu}|_{\mathcal{C}}$ and $\mu : \mathcal{C} \rightarrow \mathbb{Y}$ by $\mu(x_1, x_2) = \Delta_F(x_1, x_2) \in \mathbb{Y}$, $\forall x_1, x_2 \in \mathcal{C}$. It is easy to show that (i) – (v) of Proposition 12.20 are satisfied by $\mu$ and $\nu$. $\forall x_1, x_2 \in \mathcal{C}$, we have

$$
\nu(x_1, x_2) = \hat{\nu}(x_1, x_2) = T_F(x_1, x_2)
$$

$$
= \sup_{n \in \mathbb{Z}_+, (C_i)_{i=1}^n \subseteq \mathcal{C}, x_{x_1, x_2} = \bigcup_{i=1}^n C_i, C_i \cap C_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^n \| \Delta_F(x_{x_i, x_j}) \|
$$

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 $\Omega$, $\mathcal{A}$, 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 $\mathbb{Y}$-valued measure $\hat{\mu}$ on $(\Omega, \mathbb{B})$ such that $\hat{\mu}|_{\mathcal{A}} = \hat{\mu}$ and $(\mathcal{P} \circ \hat{\mu})|_{\mathcal{A}} = \nu$, where $\mathbb{B}$ is the $\sigma$-algebra on $\Omega$ generated by $\mathcal{A}$. Then, $\mathbb{B}$ equals to the $\sigma$-algebra on $\Omega$ generated by $\mathcal{C}$. By Proposition 12.44, $\mathbb{B} = \mathbb{B}_\Omega(\Omega)$. By Carathéodory Extension Theorem 11.19, we have $\mathcal{P} \circ \hat{\mu} = \nu$. Then, $\hat{\mu}$ is a finite $\mathbb{Y}$-valued measure on $(\Omega, \mathbb{B}_\Omega(\Omega))$. By Theorem 11.198, $\hat{\mu} \in M_{\mathbb{Y}}(\Omega, \mathbb{Y})$. This completes the proof of the claim. $\Box$

$\forall i \in N$, let $X_i := (x_{i, \tilde{x}}, B_{\Omega}(x_{i, \tilde{x}}), \hat{\mu}_i)$ be the unique finite $\mathbb{Y}$-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 $\mathbb{Y}$-valued measure space $X := (\mathcal{P}(\Omega), B, \hat{\mu})$ such that $X_i$ is the finite $\mathbb{Y}$-valued measure subspace of $X$, $\forall i \in N$. By Proposition 12.45, we have $B = B_{\Omega}(\mathcal{P}(\Omega))$.

Next, we will show that $F$ is a cumulative distribution function of $\hat{\mu}$. Fix any $x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $T_{x_1, x_2} \subseteq \Omega$. Let $(r_{x_1, x_2}, B_{\Omega}(r_{x_1, x_2}), \hat{\mu})$ be the finite $\mathbb{Y}$-valued metric measure space as prescribed in Claim 12.50.1. Note that $\mathcal{P} \circ \hat{\mu}(r_{x_1, x_2}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1, x_2} \cap r_{x_i, \tilde{x}_i}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1, x_2} \cap r_{x_i, \tilde{x}_i}) = \sum_{i \in N} T_F(r_{x_1, x_2} \cap r_{x_i, \tilde{x}_i}) = \sum_{i \in N} \mathcal{P} \circ \hat{\mu}(r_{x_1, x_2} \cap r_{x_i, \tilde{x}_i}) = \mathcal{P} \circ \hat{\mu}(r_{x_1, x_2}) = T_F(r_{x_1, x_2}) < \infty$, 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 \( \hat{\mu}(r_{x_1},x_2) = \sum_{i \in N} \hat{\mu}(r_{x_1},x_2 \cap r_{\bar{x}_i},\bar{x}_i) = \sum_{i \in N} \Delta_F(r_{x_1} \cap r_{\bar{x}_i},\bar{x}_i) = \sum_{i \in N} \mu(r_{x_1} \cap r_{\bar{x}_i},\bar{x}_i) = \mu(r_{x_1}) = \Delta_F(r_{x_1},x_2) \in \gamma \), where the first equality follows from the fact that \( \hat{\mu} \) is a \( \gamma \)-valued 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 \( \hat{\mu} \) is a finite \( \gamma \)-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, \( T_F = \sum_{i \in N} T_F(r_{\bar{x}_i},\bar{x}_i) = \sum_{i \in N} P \circ \hat{\mu}(r_{\bar{x}_i},\bar{x}_i) = P \circ \hat{\mu}(P(\Omega)), \) 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 \( (\bar{x}_i)_{i \in N} \) and \( (\bar{x}_i)_{i \in N} \), by the uniqueness of \( \hat{\mu} \).

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

**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 \( O \), \( \gamma \) be a normed linear space over \( K \), and \( \mu \in M_f(\mathcal{P}(\Omega),\mathcal{B}(\mathcal{P}(\Omega)),\gamma) \). Define \( F : \Omega \to \gamma \) and \( G : \Omega \to [0,\infty) \subset \mathbb{R} \) by \( F(x) = \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \}) \) and \( G(x) = P \circ \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \}), \forall x \in \Omega \). 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 \( T_F \leq P \circ \mu(\mathcal{P}(\Omega)) \) 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, \( \gamma \) 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 \( T_F = P \circ \mu(\mathcal{P}(\Omega)) \). Therefore, \( F \) is of bounded variation.

**Proof.** Since \( \Omega \) 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 \( r_{\bar{x}_i,\bar{x}_i} \subseteq \Omega \), \( \forall i \in N \), such that \( \Omega = \bigcup_{i \in N} r_{\bar{x}_i,\bar{x}_i} = \bigcup_{i \in N} r_{\bar{x}_i,\bar{x}_i} \), \( (\bar{x}_i)_{i \in N} \) is pairwise disjoint, \( P(\Omega) := \bigcup_{i \in N} r_{\bar{x}_i,\bar{x}_i}, \forall x_1, x_2 \in \Omega \) with \( x_1 \leq x_2 \) and \( r_{x_1,x_2} \subseteq \Omega \), we have \( r_{x_1,x_2} \subseteq \mathcal{P}(\Omega) \).

\[
\forall x_1, x_2 \in \Omega \text{ with } x_1 \leq x_2 \text{ and } r_{x_1,x_2} \subseteq \Omega, \text{ we have }
\]

\[
\Delta_F(r_{x_1},x_2) = \left\{ \sum_{i=0}^{m} \sum_{x \in \mathcal{V} r_{x_1,x_2}} (-1)^i F(\bar{x}) \mid x_1 \leq x_2, \bar{x} \neq x, \bar{x} \neq x \right\} = \mu(P(\Omega) \cap r_{x_1,x_2}) = \mu(r_{x_1},x_2)
\]
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, $\forall x \in \Omega$, $G(x) = P \circ \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \}) = P \circ \mu(\bigcap_{i=1}^{\infty} \{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + 1/m/i \}) = \lim_{i \to \infty} P \circ \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + 1/m/i \})$. Let $\delta := 1/n_0 \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$ such that $P \circ \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + 1/m/n_0 \}) < G(x) + \epsilon$. Hence, $G$ is continuous on the right. By Proposition 12.49, $G$ is of locally bounded variation. Then, $T_G = \sum_{i \in \mathbb{N}} T_G(r_{x_i, x_i}) = \sum_{i \in \mathbb{N}} \Delta_G(r_{x_i, x_i}) = \sum_{i \in \mathbb{N}} P \circ \mu(r_{x_i, x_i}) = P \circ \mu(\bar{P}(\Omega)) < \infty$, where the first equality follows from Definition 12.41; the second equality follows from the fact that $\Delta_G(r_{x_i, x_i}) = P \circ \mu(r_{x_i, x_i}) \geq 0$, $\forall x_1, x_2 \in \Omega$ 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 $(\bar{x}_i)_{i \in \mathbb{N}}$ and $(\bar{x}_i)_{i \in \mathbb{N}}$. Hence, $G$ is of bounded variation. Thus, (ii) holds.

$\forall x \in \Omega$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall 0 < h < \delta$, $\exists \bar{h}$ with $x + h, \bar{h} \in \Omega$, we have $G(x) \leq G(x + h) < G(x) + \epsilon$. Then, $\|F(x + h) - F(x)\| = \|\mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + h \}) - \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \})\| \leq \|\mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + h \}) - \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \})\| = \|P \circ \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x + h \}) - \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \})\| = G(x + h) - G(x) < \epsilon$, where the second equality and the first inequality follow from the fact that $\mu$ is a finite $\mathfrak{m}$-valued measure; and the third equality follows from Fact 11.3. Hence, $F$ is continuous on the right. Then, $F$ satisfies (i) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $r_{x_1, x_2} \subseteq \Omega$, we have

$$T_F(r_{x_1, x_2}) = \sup_{n \in \mathbb{N}} \sum_{i=1}^{n} \| \Delta_F(r_{\bar{x}_i, \bar{x}_i}) \| \sum_{i=1}^{n} \| \Delta_F(r_{\bar{x}_i, \bar{x}_i}) \|$$

$$= \sup_{n \in \mathbb{N}} \sum_{i=1}^{n} \| \mu(r_{\bar{x}_i, \bar{x}_i}) \| \sum_{i=1}^{n} \| \mu(r_{\bar{x}_i, \bar{x}_i}) \|$$
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 inequality follows from the fact that $P \circ \mu$ is a measure; the second inequality follows from Proposition 11.4; and the last inequality follows from the fact that $\mu$ is finite. Then, $F$ satisfies (ii) of Definition 12.41.

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\mathbf{r}_{x_1, x_2} \subseteq \Omega$, define $\tilde{G} : \mathbf{r}_{x_1, x_2} \rightarrow [0, \infty) \subseteq \mathbb{R}$ by $\tilde{G}(x) = T_F(r_{x_1, x})$, $\forall x \in \mathbf{r}_{x_1, x_2}$. $\forall x \in \mathbf{r}_{x_1, x_2}$, by $\tilde{G}$ being continuous on the right, $\forall \varepsilon \in (0, \infty) \subseteq \mathbb{R}, \exists \delta \in (0, \infty) \subseteq \mathbb{R}, \forall \delta \mathbf{0}_m \leq h < \delta \mathbf{1}_m$ with $x + h \in \Omega$, we have $G(x) \leq G(x + h) < G(x) + \varepsilon$. Then, $\forall \delta \mathbf{0}_m \leq h < \delta \mathbf{1}_m$ with $x + h \in \mathbf{r}_{x_1, x_2}$, $0 \leq \tilde{G}(x + h) - \tilde{G}(x) = T_F(r_{x_1, x + h} - T_F(r_{x_1, x}) = \sum_{(\tilde{x}, \tilde{\pi}) \in M} T_F(r_{x, \tilde{x}}) \leq \sum_{(\tilde{x}, \tilde{\pi}) \in M} P \circ \mu(r_{\tilde{x}, \tilde{\pi}}) = P \circ \mu(r_{x_1, x + h} - r_{x_1, x}) \leq P \circ \mu(P(\Omega) \mid \tilde{x} = x + h) - P \circ \mu(P(\Omega) \mid \tilde{x} = x) = G(x + h) - G(x) < \varepsilon$, where the first inequality follows from the monotonicity of $G$; $M := \{(\tilde{x}, \tilde{\pi}) \in \mathbf{r}_{x_1, x} \times \mathbf{r}_{x_1, x + h} \mid i \in \{0, \ldots, m - 1\}, J \subseteq J_m := \{1, \ldots, m\}, \text{card}(J) = i, \pi_k(\tilde{x}) = \pi_k(x_1), \forall k \in J, \pi_k(\tilde{x}) = \pi_k(x), \forall k \in J_m \setminus J, \pi_k(\tilde{x}) = \pi_k(x), \forall k \in J_m \setminus J\}$; the second equality follows from Proposition 12.47: the second inequality follows from the last paragraph; the third equality follows from the fact that $P \circ \mu$ is a measure; the third inequality follows from Proposition 11.4; and the fourth equality follows from the fact that $P \circ \mu$ is a finite measure. Hence, $G$ is continuous on the right. Thus, $F$ satisfies (iii) of Definition 12.41.

Therefore, $F$ is of locally bounded variation. When $T_F$ is well-defined, we have $T_F = \sum_{i \in N} T_F(r_{\tilde{x}_i, \tilde{\pi}_i}) \leq \sum_{i \in N} P \circ \mu(r_{\tilde{x}_i, \tilde{\pi}_i}) = P \circ \mu(P(\Omega)) < \infty$, where the first equality follows from Definitions 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 $P \circ \mu$ is a finite measure. Thus, (i) holds.

(iii) Let $\gamma$ be a Banach space. By Theorem 12.50 and (i), $\mu$ is the unique $\sigma$-finite $\gamma$-valued measure on the measurable space $(P(\Omega), B_B(P(\Omega)))$ such that $F$ is a cumulative distribution function of $\mu$. Theorem 12.50, coupled with (ii), further concludes that $T_F(r_{x_1, x_2}) = P \circ \mu(r_{x_1, x_2}) = \Delta_G(r_{x_1, x_2})$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $\mathbf{r}_{x_1, x_2} \subseteq \Omega$, and $T_F = P \circ \mu(P(\Omega)) < \infty$. Since $T_F$ is well-defined and is finite, then $F$ is of bounded variation.

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

**Proposition 12.52** Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone $P$, $\Omega \in B_B(\mathbb{R}^m)$ be a rectangle with the subset topology $\mathcal{O}$, $\gamma$
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 $\pi_i(x) = \pi_i(x_0)$ for some $i_0 \in J := \{1, \ldots, m\}$; and $\forall x \in \Omega$ with $\pi_i(x) \neq \pi_i(x_0)$, $\forall i \in J$, $F(x)$ is such that $\Delta F(x, \tilde{x}) = \mu(x, \tilde{x})$, where $\tilde{x} := x \wedge x_0$ 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 $G : \Omega \to \mathcal{R}$ 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 $T_F(x_1, x_2) = \mathcal{P} \circ \mu(x_1, x_2) = \Delta G(x_1, x_2)$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$ and $T(x_1, x_2) \subseteq \Omega$, and $T_F = \mathcal{P} \circ \mu(\mathcal{P}(\Omega))$.

Proof

(i) By the assumption on $\mu$, we have $\mathcal{P} \circ \mu(x_1, x_2) < \infty$, $\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. Then, $T(x_1, x_2) \in \text{dom}(\mu)$. This implies that $F(x) \in \mathcal{Y}$ and $G(x) \in \mathcal{R}$, $\forall x \in \Omega$. 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:

$\forall x_1, x_2 \in \Omega$ with $x_1 \leq x_2$. If $x_1 \neq x_2$, then by Definition 12.41, we have $\mu(x_1, x_2) = \mu(\emptyset) = \emptyset y = \Delta F(x_1, x_2)$. On the other hand, if $x_1 \leq x_2$, then let $n := \text{card}(\{ i \in J \mid \pi_i(x_1) < \pi_i(x_0) \}) = \{ i \in J \mid \pi_i(x_1) < \pi_i(x_0) \}$, by the definition of $F$ and Definition 12.41, we have $\Delta F(x_1, x_2) = \mu(x_1, x_2) = (1)^n F(x)$. On the other hand, if $\pi_i(x) = \pi_i(x_0)$ for some $i_0 \in J$, then, by the definition of $F$ and Definition 12.41, $\Delta F(x_1, x_2) = \emptyset y = \mu(\emptyset) = \mu(x_1, x_2) = F(x) = (1)^n F(x)$. Hence, we have $\mu(x_1, x_2) = (1)^n F(x)$, $\forall x \in \text{VRect}_{j, x_1, x_2}$, $\forall j \in \{0, \ldots, m\}$.

Then, we have

$$
\mu(x_1, x_2) = \mu\left( \left\{ x \in \mathcal{P}(\Omega) \mid x_1 \leq x \leq x_2 \right\} \right)
$$

$$
= \mu(\left\{ x \in \mathcal{P}(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2) \leq \pi_i(x_0), \forall i \in J; \pi_i(x_0) \leq \pi_i(x_1) < \pi_i(x_2) \leq \pi_i(x_0), \forall i \in J \right\})
$$

$$
= \mu(\left\{ x \in \mathcal{P}(\Omega) \mid \pi_i(x_1) < \pi_i(x) < \pi_i(x_0), \forall i \in J \right\})
$$
\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in \tilde{J} \setminus J \}

\[-\mu \left( \bigcup_{i_0=1}^m \{ x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \setminus \{i_0\}; \right. \]

\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in (\tilde{J} \setminus J) \setminus \{i_0\}; \]

\[ \pi_i(x_2) < \pi_i(x) \leq \pi_i(x_0), \forall i \in J \cap \{i_0\}; \]

\[ \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_1), \forall i \in (\tilde{J} \setminus J) \cap \{i_0\} \}

\[ \mu(r_{x_1,x_2}) = \sum_{j=0}^m (-1)^j \sum_{J \subseteq J, \text{card}(J)=j} \mu(r_{\tilde{x},\tilde{x}}) \]

\[ = \sum_{j=0}^m (-1)^j \sum_{J \subseteq J, \text{card}(J)=j} (-1)^{\text{card}\,(J \triangle \tilde{J})} F(x_j) \]

\[ = \sum_{j=0}^m \sum_{\tilde{J} \subseteq \tilde{J}, \text{card}(\tilde{J})=j} (-1)^{\text{card}(J \setminus \tilde{J})} F(x_j) \]

\[ = \sum_{i=0}^m J_{\tilde{x},x_1,x_2} (-1)^i F(\tilde{x}) = \Delta_F(r_{x_1,x_2}) \]

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 \( n = k \in J. \) Since \( n > 0, \) there exists \( i_0 \in J \) 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 \( \lim_{h \to 0} \cdot \) that, \( \forall i \in J \}\) = \( \mu(\{x \in P(\Omega) \mid \pi_i(x_0) < \pi_i(x) \leq \pi_i(x_2), \forall i \in J \}) = \mu(\{x \in P(\Omega) \mid \pi_i(x_1) < \pi_i(x) \leq \pi_i(x_2), \forall i \in J \}) = \mu(r_{x_1,x_0} + \mu(r_{x_2,x_2})), \)

where \( x_3, x_4 \in \Omega \) are defined by \( \pi_i(x_3) = \{ \pi_i(x_0) \text{ if } i = 1 \}

\pi_i(x_4) = \{ \pi_i(x_1) \text{ if } i \in J \} \}; \text{ and all equalities follow from the assumption on } \mu. \text{ By the inductive assumption, we have } \mu(r_{x_1,x_2}) = \mu(r_{x_1,x_3}) + \mu(r_{x_2,x_2}) = \Delta_F(r_{x_1,x_1}) + \Delta_F(r_{x_2,x_2}) = \Delta_F(r_{x_1,x_2}), \text{ where the last equality follows from Proposition 12.46. This completes the induction process.}

Therefore, \( F \) is a cumulative distribution function of \( \mu. \) \( \square \)

By Claim 12.52.1, \( F \) and \( G \) are cumulative distribution functions of \( \mu \) and \( \mathcal{P} \circ \mu, \) respectively. We need the following intermediate result.

Claim 12.52.2 \( \forall x_1, x_2 \in \Omega \text{ with } x_1 \leq x_2, \text{ if } x_2 \text{ is an accumulation point of } \Omega \cap (x_2 + P), \text{ then } \lim_{h \to 0} \mathcal{P} \circ \mu(r_{x_1,x_2+h} \setminus r_{x_1,x_2}) = 0. \forall x_1, x_2 \in \Omega \text{ with } x_1 \leq x_2 \text{ and } x_1 \neq x_2, \text{ then } \lim_{h \to 0, h \geq 0, h \neq 0, h \neq m} \mathcal{P} \circ \mu(r_{x_1,x_2+h} \setminus r_{x_1,x_2}) = 0. \)

Proof of claim: \( \forall (h_n)_{n=1}^{\infty} \subseteq P \cap (\Omega - x_2) \setminus \{0_m\} \) with \( \lim_{n \to N} h_n = 0_m, \) we have \( \lim_{n \to N} h_n = \lim_{n \to N} \bigvee_{i=n}^{\infty} h_i = 0_m, \) \( 0_m \leq h_{n+1} = h_n \neq 0_m, \) and \( h_n \in P \cap (\Omega - x_2) \setminus \{0_m\}, \forall n \in N \). Then, \( 0 \leq \lim_{n \to N} \mathcal{P} \circ \mu(r_{x_1,x_2+h} \setminus r_{x_1,x_2}) \leq \lim_{n \to N} \mathcal{P} \circ \mu \big( \bigcap_{n=1}^{\infty} (r_{x_1,x_2+h} \setminus r_{x_1,x_2}) \big) = \mathcal{P} \circ \mu(\emptyset) = 0, \) where the first inequality follows from the fact that \( \mathcal{P} \circ \mu \) is a measure; the second inequality follows from Proposition 11.4; and the first equality follows from Proposition 11.5. Then, \( \lim_{n \to N} \mathcal{P} \circ \mu(r_{x_1,x_2+h} \setminus r_{x_1,x_2}) = 0. \) By Proposition 4.16, we have \( \lim_{h \to 0, h \geq 0, h \neq 0} \mathcal{P} \circ \mu(r_{x_1,x_2+h} \setminus r_{x_1,x_2}) = 0. \) This proves the first statement.

\( \forall (h_n)_{n=1}^{\infty} \subseteq \{ h \in R^m \mid x_1 \leq x_1 + h \leq x_2, h \neq 0_m \} \) with \( \lim_{n \to N} h_n = 0_m, \) we have \( \lim_{n \to N} h_n \subseteq \{ h \in R^m \mid x_1 \leq x_1 + h \leq x_2, h \neq 0_m \}, \forall n \in N \). Then, \( 0 \leq \lim_{n \to N} \mathcal{P} \circ \mu(r_{x_1,x_2} \setminus r_{x_1+h_n,x_2}) \leq \lim_{n \to N} \mathcal{P} \circ \mu(r_{x_1,x_2} \setminus r_{x_1+h_n,x_2}) = \mathcal{P} \circ \mu(\bigcap_{n=1}^{\infty} (r_{x_1,x_2} \setminus r_{x_1+h_n,x_2})) = \mathcal{P} \circ \mu(\emptyset) = 0, \) where the first inequality follows from the fact that \( \mathcal{P} \circ \mu \) is a measure; the second inequality follows from Proposition 11.4; and the first equality follows from Proposition 11.5. Then, \( \lim_{n \to N} \mathcal{P} \circ \mu(r_{x_1,x_2} \setminus r_{x_1+h_n,x_2}) = 0. \) By Proposition 4.16, we have \( \lim_{h \to 0, h \geq 0, h \neq 0} \mathcal{P} \circ \mu(r_{x_1,x_2} \setminus r_{x_1+h_n,x_2}) = 0. \) This proves the second statement.

\( \forall x \in \Omega, \) let \( J := \{ i \in J \mid \pi_i(x) < \pi_i(x_0) \} \) and \( n := \text{card}(J). \) Let \( x \) be an accumulation point of \( \Omega \cap (x + P). \) Then, \( \exists \delta \in (0, \infty) \subseteq R \) such that, \( \forall h \in P \cap (\Omega - x) \) with \( h \neq 0_m \) and \( h < \delta 1_m, \) we have \( x + h \in \Omega \)
and \( \{ i \in J \mid \pi_i(x + h) < \pi_i(x_h) \} = J \). Then, \( 0 \leq |G(x + h) - G(x)| = \left| (-1)^n (\mathcal{P} \circ \mu)(r_{x,h}, \hat{x}) - \mathcal{P} \circ \mu(r_{x,h}) \right| \leq \left| \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) - \mathcal{P} \circ \mu(r_{x,h}) \right| + \left| \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) - \mathcal{P} \circ \mu(r_{x,h}) \right| \leq \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) - \mathcal{P} \circ \mu(r_{x,h}) \leq \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) + \mathcal{P} \circ \mu(r_{x,h}) \), 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 \( \hat{x} \leq \hat{x}_h \); the second equality follows from Fact 11.3; and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have \( 0 \leq \lim_{h \to 0_m} |G(x + h) - G(x)| \leq \lim_{h \to 0_m} \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) + \mathcal{P} \circ \mu(r_{x,h}) = 0 \). Hence, by Proposition 12.37, \( G \) is continuous on the right. By Proposition 12.49, \( G \) is of locally bounded variation.

For all \( x \in \Omega \), let \( x \) be an accumulation point of \( \Omega \cap (x + P) \), and let \( J, n, \delta \), and \( x_k \) be as defined in the previous paragraph. For all \( h \in P \cap (\Omega - x) \) with \( h \neq 0_m \) and \( h < \delta \), we have \( 0 \leq \| F(x + h) - F(x) \| = \left( -1 \right)^n \| \mu(r_{x,h}, \hat{x}) - \mu(r_{x,h}) \| = \| \mu(r_{x,h}, \hat{x}) \| + \| \mu(r_{x,h}, \hat{x}) \| \leq \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) + \mathcal{P} \circ \mu(r_{x,h}) \leq \mathcal{P} \circ \mu(x, \hat{x}) + \mathcal{P} \circ \mu(x, \hat{x}) \), where the first equality follows from the definition of \( F \); the second inequality follows from the fact that \( x \leq x_h \) and \( \hat{x} \leq \hat{x}_h \); the second equality and the third inequality follow from the fact that \( \mu \) is a locally finite \( \mathcal{F} \)-valued measure; and the last inequality follows from Proposition 11.4. By Claim 12.52.2, we have \( 0 \leq \lim_{h \to 0_m} \| F(x + h) - F(x) \| \leq \lim_{h \to 0_m} \mathcal{P} \circ \mu(r_{x,h}, \hat{x}) + \lim_{h \to 0_m} \mathcal{P} \circ \mu(r_{x,h}) = 0 \). Hence, by Proposition 12.37, \( F \) is continuous on the right and \( F \) satisfies (i) of Definition 12.41.

For all \( x, y \in \Omega \) with \( x \sim y \), we have

\[
T_F(r_{x, y}) = \sup_{n \geq 2, x, \hat{x} \in \mathcal{F}_{x, y} \cap \{ x, \hat{x} \} \in \{ 1, \ldots, n \}} \sum_{i=1}^{n} \| \Delta F(r_{x, i}, \hat{x}) \|
\]

\[
= \sup_{n \geq 2, x, \hat{x} \in \mathcal{F}_{x, y} \cap \{ x, \hat{x} \} \in \{ 1, \ldots, n \}} \sum_{i=1}^{n} \| \mu(r_{x, i}, \hat{x}) \|
\]

\[
\leq \sup_{n \geq 2, x, \hat{x} \in \mathcal{F}_{x, y} \cap \{ x, \hat{x} \} \in \{ 1, \ldots, n \}} \sum_{i=1}^{n} \mathcal{P} \circ \mu(r_{x, i}, \hat{x})
\]

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.
∀x₁, x₂ ∈ Ω with x₁ < x₂, define \( \tilde{G} : \mathbb{T}_{x₁,x₂} \to [0, \infty) \subset \mathbb{R} \) by \( \tilde{G}(x) = T_{F} (r_{x₁,x}) \), \( \forall x \in \mathbb{T}_{x₁,x₂} \). \( \forall x \in \mathbb{T}_{x₁,x₂} \) with \( x \neq x₂ \), we have

\[
0 \leq \lim_{\alpha \to \alpha_0} \frac{\tilde{G}(x+h) - \tilde{G}(x)}{h} = \lim_{\alpha \to \alpha_0} \frac{T_{F} (r_{x₁,x+h}) - T_{F} (r_{x₁,x})}{h}.
\]

\[
= \lim_{\alpha \to \alpha_0} \sum_{\bar{x} \in M} \frac{T_{F} (r_{\bar{x},\bar{x}})}{h} \leq \lim_{\alpha \to \alpha_0} \sum_{\bar{x} \in M} P \circ \mu(r_{\bar{x},\bar{x}}).
\]

where the first inequality follows from the monotonicity of \( \tilde{G} \); the first equality follows from the definition of \( \tilde{G} \); the second equality follows from Proposition 12.47; the second inequality follows from Proposition 11.35; the third equality follows from Proposition 12.47; the second inequality follows from the last paragraph; the third equality follows from the fact that \( P \circ \mu \) is a measure; and the last equality follows from Claim 12.52. 2.

Hence, by Proposition 12.37, \( \tilde{G} \) is continuous on the right. Thus, \( F \) satisfies (iii) of Definition 12.41.

Therefore, \( F \) is of locally bounded variation. This proves (i).

(ii) Let \( \mathcal{Y} \) be a Banach space. By Theorem 12.50 and (i), \( \mu \) is the unique \( \sigma \)-finite \( \mathcal{Y} \)-valued measure on the measurable space \((P(\Omega), \mathcal{B}_B (P(\Omega)))\) such that \( F \) is a cumulative distribution function of \( \mu \). Theorem 12.50, coupled with (i), further concludes that \( T_{F} (r_{x₁,x₂}) = P \circ \mu(r_{x₁,x₂}) = \Delta_{G}(r_{x₁,x₂}), \forall x₁, x₂ \in \Omega \) with \( x₁ \leq x₂ \) and \( r_{x₁,x₂} \subseteq \Omega \), and \( T_{F} = P \circ \mu(P(\Omega)) \).

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

**Lemma 12.53** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone \( \mathcal{P} \), \( \Omega \in \mathcal{B}_B (\mathbb{R}^m) \) be a region with the subset topology \( \mathcal{O} \), and \( F : \Omega \to \mathbb{R} \). Assume that the following holds.

(i) \( F \) is continuous on the right.

(ii) \( F \) is monotonically nondecreasing, i.e., \( F(x₁) \leq F(x₂), \forall x₁, x₂ \in \Omega \) with \( x₁ \leq x₂ \).

Then, \( F \) is \( \mathcal{B}_B (\mathbb{R}^m) \)-measurable.

**Proof** Consider the special case when \( \Omega \) is a rectangle. \( \forall \alpha \in \mathbb{R} \), let \( B := \{ \bar{x} \in \Omega \mid F(\bar{x}) < \alpha \} \). We will show that \( B \in \mathcal{O} \subseteq \mathcal{B}_B (\Omega) \subseteq \mathcal{B}_B (\mathbb{R}^m) \). Then, by Proposition 11.35, \( F \) is \( \mathcal{B}_B (\mathbb{R}^m) \)-measurable.

\( \forall x \in B \), \( F(x) \in (0, \infty) \). Since \( F \) is continuous on the right, then \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( F(\bar{x}) < \alpha, \forall \bar{x} \in \Omega \cap (x + P) \) with \( \bar{x} < x + \delta 1_m \).

\( \forall \bar{x} \in \{ \bar{x} \in \Omega \mid \bar{x} < x + \delta 1_m \} =: O \in \mathcal{O} \), since \( \bar{x}, x \in \Omega \) and \( \Omega \) is a rectangle, we have \( \exists x \in \Omega \cap (x + P) \) with \( \exists \bar{x} < x + \delta 1_m \). Then, \( F(\bar{x}) \leq F(x \vee \bar{x}) < \alpha. \)
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, $\exists N \subseteq \mathbb{N}$, $\exists (\bar{x}_i)_{i \in N}, (\tilde{x}_i)_{i \in N} \subseteq \Omega$ with $\bar{x}_i \bar{=} \tilde{x}_i$ and $\mathcal{T}_{\tilde{x}_i, \bar{x}_i} \subseteq \Omega$, $\forall i \in N$, such that $\Omega = \bigcup_{i \in N} \mathcal{T}_{\bar{x}_i, \tilde{x}_i}$. Let $B := \{ \bar{x} \in \Omega \mid F(\bar{x}) < \alpha \} = \bigcup_{i \in N} (B \cap \mathcal{T}_{\bar{x}_i, \tilde{x}_i})$, By the special case, $B_i \in \mathcal{B}(\mathbb{R}^m)$, $\forall i \in N$. 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.

**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 $\mathfrak{K}$ with $\mathcal{Y}^*$ being separable, $\mu \in \mathcal{M}_f(\mathbb{P}(\Omega), \mathcal{B}(\mathbb{P}(\Omega)), \mathcal{Y})$, and $F : \Omega \to \mathcal{Y}$ be the cumulative distribution function of $\mu$ as defined by $F(x) = \mu(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \})$, $\forall x \in \Omega$, 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 $(\mathbb{P}(\Omega), \mathcal{B}(\mathbb{P}(\Omega)))$ such that $\mu = \mu_+ - \mu_-$. Furthermore, $\mathcal{P} \circ \mu = \mu_+ + \mu_-$. Let $F_+ : \Omega \to [0, \infty) \subseteq \mathbb{R}$ and $F_- : \Omega \to [0, \infty) \subseteq \mathbb{R}$ be the cumulative distribution functions of $\mu_+$ and $\mu_-$, respectively, as defined by $F_+(x) = \mu_+(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \})$ and $F_-(x) = \mu_-(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \})$, $\forall x \in \Omega$. 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 $x_1 \bar{=} x_2$. Hence, $F_+$ and $F_-$ are monotonic nondecreasing. By Lemma 12.53, $F_+$ and $F_-$ are $\mathcal{B}(\mathbb{R}^m)$-measurable. Clearly, $F(x) = \mu_+(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \}) - \mu_-(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \}) = F_+(x) - F_-(x)$, $\forall x \in \Omega$. 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 $(\mathbb{P}(\Omega), \mathcal{B}(\mathbb{P}(\Omega)))$. Let $F_r : \Omega \to \mathbb{R}$ and $F_i : \Omega \to \mathbb{R}$ be defined by $F_r(x) = \mu_r(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \})$ and $F_i(x) = \mu_i(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \})$, $\forall x \in \Omega$. 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, $F(x) = \mu_+(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \}) + i\mu_i(\{ \bar{x} \in \mathbb{P}(\Omega) \mid \bar{x} \bar{=} x \}) = F_r(x) + iF_i(x)$, $\forall x \in \Omega$. 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 \( Y \) is a separable normed linear space over \( K \) with \( Y^* \) being separable. Fix any \( y_0 \in Y^* \). Define \( \mu_{y_0} : B_{K} (P(\Omega)) \to K \) by \( \mu_{y_0}(E) = \langle \langle y_0, \mu(E) \rangle \rangle, \forall E \in B_{K}(P(\Omega)) \). By Proposition 11.136, \( \mu_{y_0} \) is a finite \( K \)-valued measure on \( (P(\Omega), B_{K}(P(\Omega))) \) and \( \mathcal{P} \circ \mu_{y_0} \leq \| y_0 \| \mathcal{P} \circ \mu \). Define \( F_{y_0} : \Omega \to K \) by \( F_{y_0}(x) = \langle \langle y_0, F(x) \rangle \rangle, \forall x \in \Omega \). Clearly, \( F_{y_0}(x) = \langle \langle y_0, F(x) \rangle \rangle = \langle \langle y_0, \mu(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \}) \rangle \rangle = \mu_{y_0}(\{ \bar{x} \in P(\Omega) \mid \bar{x} \leq x \}), \forall x \in \Omega \). 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 \( B_{K}(R^{m}) \)-measurable. By Proposition 11.170, \( F \) is \( B_{K}(R^{m}) \)-measurable. This completes the last step, and hence, the proof of the proposition. \( \square \)

**Proposition 12.55** Let \( m \in \mathbb{N} \), \( R^{m} \) be endowed with the usual positive cone, \( \Omega \in B_{K}(R^{m}) \) be a rectangle with the subset topology \( \mathcal{O} \), \( Y \) be a separable normed linear space over \( K \) with \( Y^* \) being separable, \( \mu \) be a \( \sigma \)-finite \( Y \)-valued measure on \( (P(\Omega), B_{K}(P(\Omega))) \) such that \( \mathcal{P} \circ \mu(x, x_{2}) < \infty \), \( \forall x_{1}, x_{2} \in \Omega \) with \( x_{1} \leq x_{2} \), and \( F : \Omega \to Y \) be the cumulative distribution function of \( \mu \) with \( \mu_{x_{0}} \) \( E_{0} \) \( \Omega \neq \emptyset \). Then, \( F \) is \( B_{K}(R^{m}) \)-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, \( P(\Omega) \) is also a rectangle and \( \Omega \) is a region. Then, \( \exists (\bar{x}_{i})_{i=1}^{\infty}, (\check{x}_{i})_{i=1}^{\infty} \subseteq \Omega \) such that \( \check{x}_{i} \leq x_{0} \leq \bar{x}_{i}, \forall i \in \mathbb{N} \), \( \bigcup_{i=1}^{\infty} \bar{r}_{\bar{x}_{i}, \check{x}_{i}} = \Omega \), and \( P(\Omega) = \bigcup_{i=1}^{\infty} \bar{r}_{\bar{x}_{i}, \check{x}_{i}} \).

Fix any \( i \in \mathbb{N} \), let \( E_{i} := \bar{r}_{\bar{x}_{i}, \check{x}_{i}} \in B_{K}(R^{m}) \) and \( E_{i} := r_{\bar{x}_{i}, \check{x}_{i}} \in B_{K}(R^{m}) \). Note that \( \mathcal{P} \circ \mu(E_{i}) < \infty \), by the assumption on \( \mu \). Let \( (E_{i}, B_{E_i}, \mu_{E_i}) \) be the finite \( Y \)-valued measure subspace of \( (P(\Omega), B_{K}(P(\Omega))) \), \( \mathcal{O}_{E_i} \) be the subset topology on \( E_{i} \subseteq \Omega \), and \( \check{E}_{i} := (E_{i}, \mathcal{O}_{E_i}) \). Define \( \check{F} : \check{E}_{i} \to Y \) by \( \check{F}(x) = \mu_{E_i}(\{ \bar{x} \in E_{i} \mid \bar{x} \leq x \}), \forall x \in \check{E}_{i} \). By Proposition 12.51, \( \check{F} \) is a cumulative distribution function of \( \mu_{E_i} \), and is of locally bounded variation. By Proposition 12.54, \( \check{F} \) is \( B_{K}(R^{m}) \)-measurable. \( \forall \bar{x} \in \check{E}_{i} \), let \( \check{x} := x \land x_{0} \in \check{E}_{i} \) and \( \check{x} : = x \lor x_{0} \in \check{E}_{i} \). Then, \( \Delta_{F}(\check{x}, \check{x}) = \mu(\check{x}, \check{x}) = \mu_{E_i}(\check{x}, \check{x}) = \Delta_{F}(\check{x}, \check{x}) \), where the first equality follows from the fact that \( F \) is a cumulative distribution function of \( \mu_{E_i} \); the second equality follows from Proposition 11.115; and the third equality follows from the fact that \( \check{F} \) is a cumulative distribution function of \( \mu_{E_i} \). If \( \exists k_{0} \in \check{J} := \{ 1, \ldots, m \} \) such that \( \pi_{k_{0}}(x) = \pi_{k_{0}}(x_{0}) \), then \( F(x) = \check{y} = \Delta_{F}(\check{x}, \check{x}) \). Otherwise, we have \( (-1)^{n(x)}\check{F}(x) = \Delta_{F}(\check{x}, \check{x}) \), where \( n(x) = \text{card}(\{ k \in \check{J} \mid \pi_{k}(x) < \pi_{k}(x_{0}) \}) \). Hence, we have \( F(x) = (-1)^{n(x)}\Delta_{F}(\check{x}, \check{x}) = (-1)^{n(x)}\Delta_{F}(\check{x}, \check{x}) = (-1)^{n(x)}\sum_{J \subseteq \check{J}}(-1)^{\text{card}(J)}\check{F}(\check{x}, \check{x}) \), where \( \check{x}, \check{x} \in \check{E}_{i} \) is defined by \( \pi_{k}(\check{x}) = \left\{ \begin{array}{ll} \pi_{k}(\check{x}) & \forall k \in \check{J} \\ \pi_{k}(\check{x}) & \forall k \in \check{J} \setminus \check{J} \end{array} \right. \).

\( \forall j \in \{ 0, \ldots, 2^{m} - 1 \} \), let \( j_{1} \ldots j_{m} \) be the representation of \( j \) in base 2, \( J_{j} := \{ k \in \check{J} \mid j_{k} = 0 \} \), \( \eta_{j} := \text{card}(J_{j}) \). Define \( O_{j} := \{ x \in \check{E}_{i} \mid \eta_{j} \} \).
Clearly, \( \mathcal{E}_i \mid \pi_k(x) > \pi_k(x_0), \forall k \in J \setminus J, \pi_k(x) < \pi_k(x_0), \forall k \in J \} \in \mathcal{O}_{\mathcal{E}_i} \). Let \( D := \{ x \in \mathcal{E}_i \mid \pi_{i_0}(x) = \pi_{i_0}(x_0) \text{ for some } i_0 \in \bar{J} \} \in \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \).

Clearly, \( \mathcal{E}_i = (\bigcup_{j=0}^{2^m-1} O_j) \cup D \). Then, \( F_{|D}(x) = \psi_y, \forall x \in D \). Then, \( F_{|D} \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. Fix any \( j \in \{0, \ldots, 2^m - 1\} \). Denote \( J_j := J \) and \( n_j := n \). Then, \( F_{|O_j}(x) = (-1)^n \sum_{j \subseteq J} (-1)^{\text{card}(j)} \tilde{F}(\tilde{x}_j), \forall x \in O_j \). Note that \( \pi_2(\tilde{x}_j) = \begin{cases} \pi_2(x) & j \in J \setminus J \setminus j \setminus \tilde{J} \\ \pi_2(x_0) & j \in J \triangle \tilde{J} \end{cases}, \forall k \in J, \forall j \subseteq J. \) We may define function \( \tilde{F}_j : O_j \to Y \) by \( \tilde{F}_j(x) = \tilde{F}(\tilde{x}_j), \forall x \in O_j \). Since \( \mathcal{E}_i \) is a closed rectangle, it can be viewed as product of a closed set \( E_{i,\delta} \) in the \( J \setminus \bar{J} \) collection of coordinates and another closed set \( E_{i,\sigma} \) in the \( J \setminus (J \setminus \bar{J}) \) collection of coordinates. \( \forall x \in \mathcal{E}_i, \pi_2(x) \in E_{i,\delta} \) gives the one coordinate and \( \pi_1(x) \in E_{i,\sigma} \) gives the other coordinate. Then, \( (\pi_2(x), \pi_2(x)) \). For any open set \( O_Y \in O_Y, \tilde{F}_{\text{inv}}(O_Y) \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{E}_i) \subseteq \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \), since \( \tilde{F} \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. Define \( A_1 := \tilde{F}_{\text{inv}}(O_Y)_{s_1(x_0)} := \{ s \in E_{i,\sigma} \mid (\pi_2(x_0), s) \in \tilde{F}_{\text{inv}}(O_Y) \} \). By Proposition 12.25, \( A_1 \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{E}_i) \). By Proposition 11.24, \( E_{i,\sigma} \times A_1 \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{E}_i) \). Then, \( \tilde{F}_{\text{inv}}(O_Y) = O_j \cap (E_{i,\delta} \times A_1) \subseteq \mathcal{B}_{\mathcal{B}}(\mathcal{E}_i) \subseteq \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \). Hence, \( \tilde{F}_j \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. Clearly, \( F_{|O_j} = (-1)^n \sum_{j \subseteq J} (-1)^{\text{card}(j)} \tilde{F}_j \). By Propositions 7.23, 11.39, and 11.38, \( F_{|O_j} \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. By the arbitrariness of \( j \) and Proposition 11.41, \( F_{|E} \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. By the arbitrariness of \( i \) and Proposition 11.41, \( F \) is \( \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m) \)-measurable. This completes the proof of the proposition.

**Example 12.56** Let \( m \in \mathbb{N} \), \( \mathbb{R} \) 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 \( \mathbb{R}^m = (\mathbb{R}^m, \mathbb{R}, | \cdot |), \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m), \mu_{BM} \) 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 \( \mu_{Lm0} : \mathbb{R}^m \to [0, \infty] \subset \mathbb{R} \), which is defined by \( \mu_{Lm0}(E) = \inf_{(A_i)_{i=1}^{\infty} \subseteq \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m), E \subseteq \bigcup_{i=1}^{\infty} A_i, \sum_{i=1}^{\infty} \mu_{BM}(A_i), \forall E \subseteq \mathbb{R}^m}, B_{Lm} := \{ E \subseteq \mathbb{R}^m \mid E \text{ is measurable with respect to } \mu_{Lm0} \} \) is the collection of \( m \)-dimensional Lebesgue measurable sets, and \( \mu_{Lm} := \mu_{Lm0}|_{B_{Lm}} \) is the \( m \)-dimensional Lebesgue measure. By Carathéodory Extension Theorem 11.19, \( (\mathbb{R}^m, B_{Lm}, \mu_{Lm}) \) is the completion of the \( m \)-dimensional Borel measure space \( (\mathbb{R}^m, \mathcal{B}_{\mathcal{B}}(\mathbb{R}^m), \mu_{BM}) \), and is said to be the \( m \)-dimensional Lebesgue measure space. When \( m = 1 \), \( \mu_{BM}, \mu_{Lm0}, \mu_{Lm}, \) and \( B_{Lm} \) are also denoted by \( \mu_B, \mu_{L0}, \mu_L, \) and \( B_L \), respectively.

\( \forall E \in \mathbb{R}_{\geq}^m \), we will show that

\[
\mu_{Lm0}(E) = \inf_{(O_i)_{i=1}^{\infty} \subseteq \mathbb{R}_{\geq}^m, E \subseteq \bigcup_{i=1}^{\infty} O_i} \sum_{i=1}^{\infty} \mu_{BM}(O_i) =: \mu_o(E)
\]
12.5. Absolute and Lipschitz Continuity

Since $\mathcal{O}_{\mathbb{R}^m} \subseteq \mathcal{B}(\mathbb{R}^m)$, then $\mu_{Lm}(E) \leq \mu_o(E)$. On the other hand, if $\mu_{Lm}(E) = \infty$, then $\infty = \mu_{Lm}(E) \leq \mu_o(E) \leq \infty$ and $\mu_{Lm}(E) = \mu_o(E)$. If $\mu_{Lm}(E) < \infty$, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists (A_i)_{i=1}^{\infty} \subseteq \mathcal{B}(\mathbb{R}^m)$ such that $E \subseteq \bigcup_{i=1}^{\infty} A_i$ and $\mu_{Lm}(E) \leq \sum_{i=1}^{\infty} \mu_{Bm}(A_i) < \mu_{Lm}(E) + \epsilon / 2$. $\forall i \in \mathbb{N}$, since $\mathbb{R}^m$ is a topological measure space, then $\exists \Omega_i \in \mathcal{O}_{\mathbb{R}^m}$ with $A_i \subseteq \Omega_i$ such that $\mu_{Bm}(\Omega_i \setminus A_i) < 2^{-i-1}\epsilon$. Then, $E \subseteq \bigcup_{i=1}^{\infty} \Omega_i$ and $\mu_o(E) \leq \sum_{i=1}^{\infty} \mu_{Bm}(\Omega_i) = \sum_{i=1}^{\infty} (\mu_{Bm}(A_i) + \mu_{Bm}(\Omega_i \setminus A_i)) < \mu_{Lm}(E) + \epsilon / 2 + \epsilon / 2 = \mu_{Lm}(E) + \epsilon$. By the arbitrariness of $\epsilon$, we have $\mu_o(E) \leq \mu_{Lm}(E)$. Hence, $\mu_{Lm}(E) = \mu_o(E)$.

By the translational invariance of $\mathbb{R}$, we have $\mathbb{R}^m$ is translational invariant. Then, by the above, $\mu_{Lm}$ 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, $\gamma$ be a Banach space over $\mathbb{K}$, $F : \Omega \to \gamma$ be of locally bounded variation, $(P(\Omega), \mathcal{B}(P(\Omega)), \mu)$ be the unique $\sigma$-finite $\gamma$-valued measure space that admits $F$ as a cumulative distribution function as prescribed in Theorem 12.50, $\mathbb{Z}$ be a normed linear space over $\mathbb{K}$, $U \in \mathcal{B}(P(\Omega))$, and $g : U \to \mathcal{B}(\mathbb{Y}, \mathbb{Z})$ be $\mathcal{B}(P(\Omega))$-measurable (or $\mathcal{B}_{\mathbb{R}^m}$)-measurable. We will denote the integral $\int_U g(x) \, d\mu(x)$ by $\int_U g \, dF$ 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 $\gamma$ be a normed linear space. A function $f : I \to \gamma$ is said to be absolutely continuous at $x_0 \in I$ if $\exists \alpha \in (0, \infty) \subseteq \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\exists \delta(\epsilon) \in (0, a] \subseteq \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - \alpha \leq x_1 \leq \bar{x}_1 \leq x_2 \leq \bar{x}_2 \leq \cdots \leq x_n \leq \bar{x}_n \leq x_0 + \alpha$ with $x_i, \bar{x}_i \in I$, $i = 1, \ldots, n$, such that

$\sum_{i=1}^{n} \mu_{\gamma}(x_i, \bar{x}_i) = \sum_{i=1}^{n} \|\gamma \| \|f(x_i) - f(x_1)\| < \epsilon$.

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 $\gamma$ be a normed linear space over $\mathbb{K}$. $F : \Omega \to \gamma$ is said to be absolutely continuous at $x_0 \in \Omega$ if the following conditions holds.

(i) When $m > 1$, $\forall i_0 \in J := \{1, \ldots, m\}$, let $D_{i_0} := \{ x \in \mathbb{R}^{m-1} : \exists \bar{x} \in \Omega \ni \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \text{ and } \pi_{\gamma}(\bar{x}) = x \}$, where $\pi_{\gamma} : \mathbb{R}^m \to \mathbb{R}^{m-1}$ is defined by $\pi_{i_0}(x) = (\pi_1(x), \ldots, \pi_{i_0-1}(x), \pi_{i_0+1}(x), \ldots, \pi_m(x))$. $\forall x \in \mathbb{R}^m$, $F_{i_0} : D_{i_0} \to \gamma$ is defined by $F_{i_0}(x) = F(\bar{x})$, $\forall x \in D_{i_0}$, $\forall \bar{x} \in \Omega$ with $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{\gamma}(\bar{x}) = x$. The function $F_{i_0}$ is absolutely continuous at $\pi_{\gamma}(x_0)$.
Then, we will prove the result using mathematical induction on $F$. Suppose $F$ is absolutely continuous at $x_0 \in \Omega$. $F$ is said to be absolutely continuous at any $x_0 \in \Omega$.

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, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $\mathcal{Y}$ be a normed linear space over $\mathcal{K}$, and $F : \Omega \rightarrow \mathcal{Y}$ 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, suppose $F_i : \bar{D}_{i_0} \rightarrow \mathcal{Y}$ is absolutely continuous at $x_{i_0}(x_0)$, where $F_{i_0}$ and $D_{i_0}$ are defined in Definition 12.58. Clearly, $\bar{D}_{i_0} \in \mathcal{B}_B(\mathbb{R}^{n-1})$ is a rectangle since $\bar{\Omega}$ is a rectangle. By the inductive assumption, the definition for general $\mathbb{R}$ case subsumes the definition for $\mathbb{R}$ case. We explicitly specify the $\mathbb{R}$ case for its direct implications.
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, $\Omega \in B_{\mathbb{B}}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $y$ be a normed linear space over $K$, $F : \Omega \to y$, 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 $\pi_J(x) = (\pi_i(x))_{i \in J}$, $\forall \bar{x} \in \mathbb{R}^m$, $\pi_J \circ \pi_J : \mathbb{R}^m \to \mathbb{R}^{m-i}$ be such that $\pi_I \circ \pi_J(x) = (\pi_I(x))_{i \in J \cup J}$, $\forall \bar{x} \in \mathbb{R}^m$, $M : \mathbb{R}^i \times \mathbb{R}^{m-i} \to \mathbb{R}^m$ be such that $M(x_a, x_b) = \bar{x}$ with $x_a = \pi_I(x)$ and $x_b = \pi_J(x)$, $\forall x_a \in \mathbb{R}^i$ and $\forall x_b \in \mathbb{R}^{m-i}$, $x_f := \pi_I(x_0)$, $\bar{x}_f := \pi_I(x_0)$, $\bar{\Omega} := \pi_I(\Omega)$, and $F_f : \bar{\Omega} \to y$ be defined by $F_f(\bar{x}) = F(M(x_a, x_b))$, $\forall \bar{x} \in \bar{\Omega}$, $\exists a_i \in (0, \infty) \subset \mathbb{R}$, $\forall \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_f(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^n$, $\forall (\bar{x}_j)_{j=1}^n \subseteq \mathbb{R}_x - a_{j_1} \cdots a_{j_m} - \bar{\Omega}^{ij}$ with $\bar{x}_j \equiv \bar{x}_j$, $\forall j \in \{1, \ldots, n\}$, $(r_{x_j, x_j})_{j=1}^n$ be pairwise disjoint, and $\sum_{j=1}^n \mu_B(m-i)(r_{x_j, x_j}) < \delta_f(\epsilon)$, we have $\sum_{j=1}^n \|\Delta F_f(r_{x_j, x_j})\| < \epsilon$.

**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 $i \in \{0, \ldots, m-1\}$ 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, $\Omega \in B_{\mathbb{B}}(\mathbb{R}^m)$ be a rectangle with subset topology $\mathcal{O}$, $y$ and $Z$ be normed linear spaces over $K$, $f : \Omega \to y$, $g : \Omega \to Z$, $x_0 \in \Omega$, and $h : \Omega \to y \times Z$ be defined by $h(x) = (f(x), g(x))$, $\forall x \in \Omega$. Then, $h$ is
absolutely continuous at \( x_0 \) if, and only if, \( f \) and \( g \) are absolutely continuous at \( x_0 \).

Proof. We will prove the result using mathematical induction on \( m \).

1° \( m = 1 \). "Sufficiency" By \( f \) being absolutely continuous at \( x_0 \), \( \exists a \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta f(\epsilon) \in (0, a) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall x_0 - a f \leq x_1 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq x_n \leq x_0 + a f \) with \( x_i, \bar{x}_i \in I, i = 1, \ldots, n \), and \( \sum_{i=1}^{n}(\bar{x}_i - x_i) < \delta_f(\epsilon) \), we have \( \sum_{i=1}^{n}||f(\bar{x}_i) - f(x_i)|| < \epsilon \). By \( g \) being absolutely continuous at \( x_0 \), \( \exists a_g \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta g(\epsilon) \in (0, a_g) \subset \mathbb{R} \), \( \forall n \in \mathbb{N}, \forall x_0 \leq a_g \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq x_n \leq x_0 + a_g \) with \( x_i, \bar{x}_i \in I, i = 1, \ldots, n \), and \( \sum_{i=1}^{n}(\bar{x}_i - x_i) < \delta_g(\epsilon) \), we have \( \sum_{i=1}^{n}||g(\bar{x}_i) - g(x_i)|| < \epsilon \). Let \( a := \min\{a_f, a_g\} \in (0, \infty) \subset \mathbb{R} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \forall \epsilon \leq \epsilon \) and \( \epsilon \leq \epsilon \). Hence, \( h \) is absolutely continuous at \( x_0 \).

"Necessity" By \( h \) being absolutely continuous at \( x_0 \), \( \exists a \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta h(\epsilon) \in (0, a) \subset \mathbb{R} \), \( \forall n \in \mathbb{N}, \forall x_0 - a \leq x_1 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq x_n \leq x_0 + a \) with \( x_i, \bar{x}_i \in I, i = 1, \ldots, n \), and \( \sum_{i=1}^{n}(\bar{x}_i - x_i) < \delta_h(\epsilon) \), we have \( \epsilon > \sum_{i=1}^{n}||h(\bar{x}_i) - h(x_i)|| = \sum_{i=1}^{n}||f(\bar{x}_i) - f(x_i)|| + ||g(\bar{x}_i) - g(x_i)|| < \epsilon \). Hence, \( f \) is absolutely continuous at \( x_0 \). By symmetry, \( g \) is absolutely 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 \( m = k \in \{2, 3, \ldots \} \).

"Sufficiency" Assume that \( f \) and \( g \) are absolutely continuous at \( x_0 \), \( \forall i \in J := \{1, \ldots, m\} \), let \( D_i \subseteq \mathbb{R}^{m-1}, \pi_i : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}, f_i : D_i \rightarrow \mathcal{Y}, g_i : D_i \rightarrow \mathcal{Z} \), and \( h_i : D_i \rightarrow Y \times Z \) be defined as in Definition 12.58. By (i) of Definition 12.58, \( f_i \) and \( g_i \) are absolutely continuous at \( \pi_i(x_0) \). By the inductive assumption, \( h_i \) is absolutely continuous at \( \pi_i(x_0) \). Then, \( h \) satisfies (i) of Definition 12.58.

By (ii) of Definition 12.58, \( \exists a_{f_i} \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta f_i(\epsilon) \in (0, a_{f_i}) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta g_i(\epsilon) \in (0, a_{g_i}) \subset \mathbb{R}, \forall \epsilon \leq \epsilon \) and \( \epsilon \leq \epsilon \). Let \( a := a_f \wedge a_g \in (0, \infty) \subset \mathbb{R} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta(\epsilon) := \min\{\delta f_i(\epsilon/2), \delta g_i(\epsilon/2)\} \in (0, \infty) \subset \mathbb{R} \). \( \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \leq \epsilon \) and \( \epsilon \leq \epsilon \). Hence, \( \delta(\epsilon) \) is absolutely continuous at \( x_0 \).
12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

\[ \sum_{i=1}^{n} \| (\Delta f(\bar{x}_i, \bar{x}_i), \Delta g(\bar{x}_i, \bar{x}_i)) \| = \sum_{i=1}^{n} (\| \Delta f(\bar{x}_i, \bar{x}_i) \|^2 + \| \Delta g(\bar{x}_i, \bar{x}_i) \|^2)^{1/2} \leq \sum_{i=1}^{n} (\| \Delta f(\bar{x}_i, \bar{x}_i) \| + \sum_{i=1}^{n} \| \Delta g(\bar{x}_i, \bar{x}_i) \| < \epsilon. \]

Hence, \( h \) satisfies (ii) of Definition 12.58. Then, \( h \) is absolutely continuous at \( x_0 \).

"Necessity" Let \( h \) be absolutely continuous at \( x_0 \). \( \forall \epsilon > 0 \), let \( D_{\epsilon} \), \( \pi_{\epsilon} \), \( f_{\epsilon} \), \( g_{\epsilon} \), and \( h_{\epsilon} \) be defined as in Definition 12.58. By (i) of Definition 12.58, \( h_{\epsilon} \) is absolutely continuous at \( \pi_{\epsilon}(x_0) \). By the inductive assumption, \( f_{\epsilon} \) and \( g_{\epsilon} \) are absolutely continuous at \( x_0 \). Hence, \( f \) and \( g \) satisfy (i) of Definition 12.58.

By (ii) of Definition 12.58, \( \exists \alpha \in (0, \infty) \subset \mathbb{R} \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall \bar{x}_i \in \mathbb{R} \), \( \bar{x}_i = \pi_{\epsilon}(x_0) \cap \Omega \) with \( \bar{x}_i \leq x_i \), \( \forall i \in \{1, \ldots, n\} \), \( (r_{\bar{x}_i, \bar{x}_i})_{i=1}^{n} \) being pairwise disjoint, and \( \sum_{i=1}^{n} \mu_{BM}(r_{\bar{x}_i, \bar{x}_i}) < \delta(\epsilon) \), we have \( \epsilon > \sum_{i=1}^{n} 2 \| \Delta h(r_{\bar{x}_i, \bar{x}_i}) \| \geq \sum_{i=1}^{n} 2 \| \Delta f(r_{\bar{x}_i, \bar{x}_i}) \| \). 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.

\[ \square \]

Proposition 12.62 Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_\emptyset(\mathbb{R}^m) \) be a rectangle with subset topology \( \mathcal{O} \), \( \mathcal{Y} \) be a normed linear space over \( \mathbb{K} \), \( F : \Omega \to \mathcal{Y} \), and \( \Omega_1, \Omega_2 \subseteq \Omega \) 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 \( \Omega \subseteq \Omega \), then \( F|_{\Omega_1} \) is absolutely continuous on \( \Omega \cap \Omega_1 \).

(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) \( \forall x_0 \in \Omega \cap \Omega_1 \), 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 \( \Omega \subseteq \Omega_1 \).

(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 \( \exists \alpha \in (0, \infty) \subset \mathbb{R} \) such that \( (x_0 + a, x_0 + a) \cap \Omega_1 = (x_0, x_0 + a) \cap \Omega \). By \( F|_{\Omega_1} \) being absolutely continuous at \( x_0 \), we can easily show that \( F \) is absolutely continuous at \( x_0 \). This completes the first step in the induction process.

2° Assume that the result holds for \( m < k - 1 \in \mathbb{N} \).

3° Consider the case when \( m = k \in \{2, 3, \ldots\} \). Without loss of generality, assume \( x_0 \in \Omega_1 \). \( \forall i_0 \in \tilde{J} := \{1, \ldots, m\} \), let \( D_{i_0} \) and \( F_{i_0} : D_{i_0} \to \mathcal{Y} \) be as
defined in Definition 12.58. Let \( D_{i_0,j} := \{ x \in \mathbb{R}^{m-1} \mid \exists \bar{x} \in \Omega_j \exists \cdot \pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0) \} \) and \( \pi_{i_0}'(\bar{x}) = \pi_{i_0}'(x_0) \) and \( F_{i_0,j} : D_{i_0,j} \to \mathbb{R} \) 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 \( \pi_{i_0}'(\bar{x}) = x, j = 1, 2 \). 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 \( \pi_{i_0}'(x_0) \) when \( \pi_{i_0}'(x_0) \in D_{i_0,j} \). 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 \( \pi_{i_0}'(x_0) \in D_{i_0} \). Hence, \( F \) satisfies (i) of Definition 12.58.

Since \( \Omega_1 \) is relative open in \( \Omega \), then \( \exists a \in (0, \infty) \subseteq \mathbb{R} \) such that \( \cap_{x_0 - a_1 x_0 + a_1} x_0 = \cap_{x_0 - a_1 x_0 + a_1} = \Omega_1 \cap \Omega \). By \( F|_{\Omega_1} \) being absolutely continuous at \( x_0 \) and (ii) of Definition 12.58, \( \exists a \in (0, a) \subseteq \mathbb{R} \), \( \forall \in (0, \infty) \subseteq \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall (\bar{\pi}_1)^m_{i_0=1} \), \( \bar{x} \in \Omega_1 \cap \Omega \) with \( \bar{x}_i \leq \bar{x} \), \( \forall i \in \{1, \ldots, n\} \), \( (r_{x_i}, \bar{x}_i)_{i=1}^n \) being pairwise disjoint, and \( \Sigma_{i=1}^n \mu_{\Omega_1}(r_{x_i}, \bar{x}_i) < \delta(\epsilon) \), we have \( \Sigma_{i=1}^n \mu_{\Omega_1}(r_{x_i}, \bar{x}_i) < \delta(\epsilon) \). Hence, \( F \) satisfies (ii) of Definition 12.58.

Hence, \( F \) 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 1.

Case 2: \( \Omega_1 \) and \( \Omega_2 \) are relatively closed in \( \Omega \). Fix any \( x_0 \in \Omega \). We will prove the statement that "if \( F|_{\Omega} \) is absolutely continuous at \( x_0 \) when \( x_0 \in \Omega_1 \), and \( \Omega_1 \) and \( \Omega_2 \) are relatively closed in \( \Omega \), then \( F \) is absolutely continuous at \( x_0 \) by mathematical induction on \( m \).

1° \( m = 1 \). Without loss of generality, assume \( x_0 \in \Omega_1 \). We will further distinguish two exhaustive and mutually exclusive cases: Case 2a: \( x_0 \notin \Omega_2 \); Case 2b: \( x_0 \in \Omega_2 \). Case 2a: \( x_0 \notin \Omega_2 \). Since \( \Omega_2 \) is relatively closed in \( \Omega \) and \( x_0 \in \Omega \setminus \Omega_2 \), then, by Proposition 4.10, \( \exists a \in (0, \infty) \subseteq \mathbb{R} \) such that \( (x_0 - a, x_0 + a) \cap \Omega \subseteq \Omega \setminus \Omega_2 \). \( F|_{\Omega_2} \) 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 \( x_0 \), \( \exists a \in (0, \infty) \subseteq \mathbb{R} \), \( \forall \in (0, \infty) \subseteq \mathbb{R} \), \( \exists \delta(\epsilon) \in (0, a) \subseteq \mathbb{R} \), \( \forall n \in \mathbb{N} \), \( \forall x_0 - a_1 \leq x_1 \leq \cdots \leq x_n \leq x_0 + a \) with \( x_j, \bar{x}_j \in \Omega_1 \), \( j = 1, \ldots, n \), and \( \Sigma_{j=1}^n (x_j - \bar{x}_j) \leq \delta(\epsilon) \), we have \( \Sigma_{j=1}^n \mu_{\Omega_1}(x_j, \bar{x}_j) < \delta(\epsilon) \). Without loss of generality, assume \( \exists a \in \{1, \ldots, n - 1\} \) such that \( x_{a} \leq x_0 \leq x_{a+1} \). (In case \( j_0 \) does not exist, add a pair of \( x_0 \)'s into the collection of \( x_0 \)'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 \( \Sigma_{j=1}^{j_0} (x_j - \bar{x}_j) < \delta(\epsilon) \leq \delta(\epsilon/2) \). Then, \( \Sigma_{j=1}^{j_0} \mu_{\Omega_1}(x_j, \bar{x}_j) < \epsilon/2 \). By a similar argument, we have
12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

Let $\sum_{j=0}^{n-1} \| F(\bar{x}_j) - F(x_j) \| < \epsilon/2$. 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 $x_0 \in \Omega_1$. $\forall \delta_0 \in \mathcal{J}$, let $D_{i_0} \cap D_{i_0} = y$ be as defined in Definition 12.58. Let $D_{i_{0,j}} := \{ x \in \mathbb{R}^{m-1} | \exists \bar{x} \in \Omega_j \ni \pi_{i_0}(\bar{x}) = \pi_{i_0,0}(x_0) \}$ and $\pi_{i_0}^\tau(\bar{x}) = x$ and $F_{i_{0,j}} : D_{i_{0,j}} \rightarrow y$ be defined by $F_{i_{0,j}}(x) = F_{\Omega_j}(\bar{x})$, $\forall \bar{x} \in D_{i_{0,j}}$, where $\bar{x} \in \Omega_j$ is such that $\pi_{i_0}(\bar{x}) = \pi_{i_0}(x_0)$ and $\pi_{i_0}^\tau(\bar{x}) = x$, $j = 1, 2$. 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 $F|_{\Omega_j}$ being absolutely continuous at $x_0$ when $x_0 \in \Omega_j$, we have $F_{i_{0,j}}$ is absolutely continuous at $\pi_{i_0}^\tau(x_0)$ when $\pi_{i_0}^\tau(x_0) \in D_{i_{0,j}}$. 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 $\pi_{i_0}^\tau(x_0) \in D_{i_0}$. 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, $\exists \alpha \in (0, \infty) \subset \mathbb{R}$ such that $r_{x_0-a,0^{+},0} \cap \Omega \subset \Omega \setminus \Omega_2 \subseteq \Omega_1$. By $F|_{\Omega_1}$ being absolutely continuous at $x_0$, $\exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_i)_{i=1}^{n} \subseteq \mathbb{R}^{m-1}$, $\forall j \in \{1, \ldots, n\}$, $(\bar{x}_j)_{j=1}^{n} = \pi_{i_0}^\tau(x_0) \cap \Omega$ with $\bar{x}_j \equiv x_j$, $\forall j \in \{1, \ldots, n\}$, and $\| \Delta F(x_i, \bar{x}_i) \| < \epsilon$. 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 $F|_{\Omega_1}$ being absolutely continuous at $x_0$ and (ii) of Definition 12.58, $\exists \alpha_i \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_i(\epsilon) \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^{n} \subseteq \mathbb{R}^{m-1}$, $\forall j \in \{1, \ldots, n\}$, $(\bar{x}_j)_{j=1}^{n} = \pi_{i_0}^\tau(x_0) \cap \Omega$ with $\bar{x}_j \equiv x_j$, $\forall j \in \{1, \ldots, n\}$, $(\bar{x}_j)_{j=1}^{n}$ being pairwise disjoint, and $\sum_{j=1}^{n} \mu_{\Omega_1}(\bar{x}_j, x_j) < \delta_i(\epsilon)$, we have $\sum_{j=1}^{n} \| \Delta F(x_i, \bar{x}_i) \| < \epsilon$.

Take $\alpha := \min\{a_1, a_2\} \in (0, \infty) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) := \min\{\delta_1(\epsilon/2), \delta_2(\epsilon/2)\} \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall (\bar{x}_j)_{j=1}^{n} \subseteq \mathbb{R}^{m-1}$, $\forall j \in \{1, \ldots, n\}$, $(\bar{x}_j)_{j=1}^{n}$ being pairwise disjoint, and $\sum_{j=1}^{n} \mu_{\Omega_1}(\bar{x}_j, x_j) < \delta(\epsilon)$. $\forall j \in \{1, \ldots, n\}$, since $\Omega = \Omega_1 \cup \Omega_2$ and $\Omega_1$ and $\Omega_2$ are relatively closed rectangles in the rectangle $\Omega$, then $r_{\bar{x}_j, x_j} = r_{\bar{x}_j, k_{\bar{x}_j, x_j}} \cap \Omega$ is a rectangle, $k = 1, 2$. If $r_{\bar{x}_j, x_j} \subseteq \Omega_2$, for some $l \in \{1, 2\}$, then let $\bar{x}_{j,l} = \bar{x}_j$, $\bar{x}_{j,l} = \bar{x}_j$, $\bar{x}_{j,3-l} = x_0$ and $\bar{x}_{j,3-l} = x_0$. The desired result holds. Consider the remaining pos-
sibility, we must have $\bar{x}_j \in \Omega_1 \setminus \Omega_3 \setminus \Omega_l$ and $\bar{x}_j \in \Omega_3 \setminus \Omega_l$, 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 $\pi_l(\Omega_1) = \pi_l(\Omega_2) = \pi_l(\Omega), \ \forall i \in \mathcal{J} \setminus \{i_0\}$, for some $i_0 \in \mathcal{J}$, and $\pi_{i_0}(\Omega_1)$ and $\pi_{i_0}(\Omega_2)$ are overlapping relatively closed intervals in $\pi_{i_0}(\Omega)$. Let $\bar{x}_{j,1} = \bar{x}_j \in \Omega_1 \setminus \Omega_2$. Since $\bar{r}_{x,j,\bar{x}_j} \subseteq \Omega$, then $\bar{r}_{x,j,\bar{x}_j} \cap \Omega_1 \subseteq \Omega$ is a relatively closed rectangle with respect to $\bar{r}_{x,j,\bar{x}_j}$. Then, $\emptyset \neq \bar{r}_{x,j,\bar{x}_j} \cap \Omega_1 =: \bar{r}_{x,j,1,\bar{x}_j,1} \subseteq \bar{r}_{x,j,\bar{x}_j} \subseteq \Omega$. This implies that $\bar{x}_{j,1,\bar{x}_j,1} \in \bar{r}_{x_0-a_1m,x_0+a_1m} \cap \Omega_1$ and $\bar{x}_{j,1} \equiv \bar{x}_{j,1,\bar{x}_j,1}$. Since $\bar{x}_j \in \Omega_2 \setminus \Omega_1$, then $\emptyset \neq \bar{r}_{x,j,\bar{x}_j} \cap \Omega_2 =: \bar{r}_{x,j,2,\bar{x}_j,2} \subseteq \bar{r}_{x,j,\bar{x}_j} \subseteq \Omega$, where $\bar{x}_j \in \bar{r}_{x_0-a_1m,x_0+a_1m} \cap \Omega_2$ and $\bar{x}_j \equiv \bar{x}_{j,1}$. Clearly, $\bar{x}_{j,1} \in \Omega_1 \cap \Omega_2$ since $\Omega_1$ is relatively closed in $\Omega$ and $\bar{x}_j \in \Omega_1 \cap \Omega_2$ since $\Omega_1$ is relatively closed in $\Omega$. Then, $\bar{x}_j \in \bar{r}_{x,j,\bar{x}_j} \cap \Omega_1 = \bar{r}_{x,j,1,\bar{x}_j,1}$ and $\bar{x}_j \equiv \bar{x}_{j,1}$. This implies that $\bar{r}_{x,j,\bar{x}_j} \cap \Omega_2 = \bar{r}_{x,j,2,\bar{x}_j,2} \subseteq \bar{r}_{x,j,\bar{x}_j} \subseteq \Omega$. Let $\bar{x}_{j,2} = \bar{x}_j$ and $\bar{x}_{j,2} \subseteq \bar{r}_{x,j,\bar{x}_j}$, be such that $\pi_{i}(\bar{x}_{j,2}) = \begin{cases} \pi_{i}(\bar{x}_j) & \text{if } \pi_{i}(\bar{x}_{j,1}) = \pi_{i}(\bar{x}_j) \\ \pi_{i}(\bar{x}_{j,1}) & \text{if } \pi_{i}(\bar{x}_{j,1}) < \pi_{i}(\bar{x}_j) \end{cases} \forall i \in \mathcal{J}$. Then, $\bar{x}_{j,2} \equiv \bar{x}_{j,2}$. Clearly, $\bar{r}_{x,j,1,\bar{x}_j,1} \cap \bar{r}_{x,j,2,\bar{x}_j,2} = \emptyset$ and $\bar{r}_{x,j,1,\bar{x}_j,1} \cup \bar{r}_{x,j,2,\bar{x}_j,2} = \bar{r}_{x,j,\bar{x}_j}$, since $\bar{r}_{x,j,\bar{x}_j} \subseteq \bar{r}_{x,j,\bar{x}_j}$. These $\bar{x}_{j,1}, \bar{x}_{j,2}, \bar{x}_{j,1}$, and $\bar{x}_{j,2}$ satisfy the conditions specified above.

Then, $\sum_{j=1}^{n} \mu_{B_{l}}(\bar{r}_{x,j,\bar{x}_j}) = \sum_{k=1}^{2} \sum_{j=1}^{m} \mu_{B_{l}}(\bar{r}_{x,j,k,\bar{x}_j,k}) < \delta(\epsilon)$ and $\sum_{k=1}^{2} \sum_{j=1}^{m} \mu_{B_{l}}(\bar{r}_{x,j,k,\bar{x}_j,k}) < \delta_{k}(\epsilon/2)\ \forall k = 1,2$. This implies that $\sum_{k=1}^{2} \sum_{j=1}^{m} \left| \Delta_{F}(\bar{r}_{x,j,\bar{x}_j}) \right| = \sum_{k=1}^{2} \sum_{j=1}^{m} \left| \Delta_{F}(\bar{r}_{x,j,1,\bar{x}_j,1}) + \Delta_{F}(\bar{r}_{x,j,2,\bar{x}_j,2}) \right| \leq \sum_{k=1}^{2} \sum_{j=1}^{m} \left| \Delta_{F}(\bar{r}_{x,j,k,\bar{x}_j,k}) \right| < \epsilon$, 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. \hfill \qed

**Definition 12.63** Let $X$ be a normed linear space over $K_1$, $Y$ be a normed linear space over $K_2$, $U \subseteq X$, $f : U \rightarrow Y$, and $x_0 \in U$. $f$ is said to be locally Lipschitz at $x_0$ (with Lipschitz constant $L_0 \in [0,\infty) \subseteq \mathbb{R}$) if $\exists \delta \in (0,\infty) \subseteq \mathbb{R}$, $\forall x_1, x_2 \in U \cap B_X(x_0, \delta)$, we have $\left\| f(x_1) - f(x_2) \right\| \leq L_0 \left\| x_1 - x_2 \right\|$. $f$ is said to be locally Lipschitz on $U$ if it is locally Lipschitz at $x$, $\forall x \in U$. $f$ is said to be Lipschitz on $U$ (with Lipschitz constant $L \in [0,\infty) \subseteq \mathbb{R}$) if $\forall x_1, x_2 \in U$, we have $\left\| f(x_1) - f(x_2) \right\| \leq L \left\| x_1 - x_2 \right\|$.

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 $U$. 

12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

**Definition 12.64** Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $Y$ be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \subseteq \mathcal{X}$, $f : U \times Y \to \mathcal{Z}$, and $x_0 \in U$. $f$ is said to be locally Lipschitz on $x_0$ (with Lipschitz constant $L_0 \in [0, \infty) \subseteq \mathbb{R}$) uniformly over $Y$ if $\exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in U \cap B_X(x_0, \delta)$, $\forall y \in Y$, 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 $L \in [0, \infty) \subset \mathbb{R}$ 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 $\mathbb{K}$, $Y$ be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \subseteq \mathcal{X}$ be compact, $f : U \times Y \to \mathcal{Z}$ 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** $\forall x \in U$, by the local Lipschitz continuity of $f$, $\exists L_x \in [0, \infty) \subset \mathbb{R}$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$, $\forall x_1, x_2 \in O_x := U \cap B_X(x, \delta_x)$, $\forall y \in Y$, we have $\| f(x_1, y) - f(x_2, y) \| \leq L_x \| x_1 - x_2 \|$. Let $\hat{O}_x := B_X(x, \delta_x/2)$. 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 $L_M := \max \{ \max_{x \in X_N} L_x, \frac{2}{\min_{x \in X_N} \delta_x} \max_{\hat{x}_1, \hat{x}_2 \in X_N} (L_{\hat{x}_1} \delta_{\hat{x}_1}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 + \sup_{y \in Y} \| f(\hat{x}_1, y) \|) \} \in [0, \infty) \subset \mathbb{R}$. $\forall x_1, x_2 \in U$, $\forall y \in Y$, $\exists \hat{x}_1 \in X_N$ 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, $x_1, x_2 \in O_{\hat{x}_1}$ and $\| f(x_1, y) - f(x_2, y) \| \leq L_{\hat{x}_1} \| x_1 - x_2 \| \leq L_M \| x_1 - x_2 \|$. Case 2: $x_2 \notin O_{\hat{x}_1}$. Then, $\| x_2 - \hat{x}_1 \| \geq \delta_{\hat{x}_1}$ and $\| x_2 - x_1 \| \geq \| x_2 - \hat{x}_1 \| - \| x_1 - \hat{x}_1 \| \geq \delta_{\hat{x}_1}/2 \geq \frac{\min_{x \in X_N} \delta_x}{2} > 0$. $\exists \hat{x}_2 \in X_N$ such that $x_2 \in O_{\hat{x}_2}$. Then, $\| f(x_1, y) - f(x_2, y) \| \leq \| f(x_1, y) - f(\hat{x}_1, y) \| + \| f(\hat{x}_1, y) - f(\hat{x}_2, y) \| + \| f(\hat{x}_2, y) - f(x_2, y) \| \leq L_{\hat{x}_1} \| x_1 - \hat{x}_1 \| + \| f(\hat{x}_1, y) \| + L_{\hat{x}_2} \| x_2 - \hat{x}_2 \| \leq \sup_{y \in Y} \| f(\hat{x}_1, y) \| + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 + L_{\hat{x}_2} \delta_{\hat{x}_2}/2 \leq L_M \| x_1 - x_2 \|$. 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. $\square$

**Proposition 12.66** Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $Y$ be a set, $\mathcal{Z}$ be a normed linear space over $\mathbb{K}$, $U \subseteq \mathcal{X}$ be convex, $f : U \times Y \to \mathcal{Z}$ be partial differentiable with respect to $x$. Assume that $\exists L \in [0, \infty) \subset \mathbb{R}$ 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 $Y$.  

For all $x_1, x_2 \in U$, $\forall y \in Y$, by Mean Value Theorem 9.23, $\exists t \in (0, 1) \subset \mathbb{R}$ such that $\|f(x_1, y) - f(x_2, y)\| \leq \|\frac{\partial f}{\partial x}(x_1 + (1-t)x_2, y)\| \cdot \|x_1 - x_2\|$. Hence, $f$ is Lipschitz on $U$ with Lipschitz constant $L$ uniformly over $Y$. This completes the proof of the proposition.

**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 $\exists L_0 \in (0, \infty) \subset \mathbb{R}$, $\exists \delta_g \in (0, \infty) \subset \mathbb{R}$, $\forall y_1, y_2 \in U \cap B(y_0, \delta_g)$, 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, $\exists \alpha_0 \in (0, \infty) \subset \mathbb{R}$ such that $\forall x \in B_E(x_0, \alpha_0) \cap I$, we have $f(x) \in U \cap B_y(y_0, \delta_g)$. By the absolute continuity of $f$ at $x_0$, $\exists a_f \in (0, \alpha_0) \subset \mathbb{R}$, $\forall \epsilon \in (0, (\alpha_0, \infty) \subset \mathbb{R}$, $\exists \delta_f(\epsilon) \in (0, a_f) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $\forall x_0 - a_f \leq x_1 \leq \tilde{x}_1 \leq x_2 \leq \cdots \leq x_n \leq \tilde{x}_n \leq x_0 \leq a_f$ with $x_i, \tilde{x}_i \in I$, $i = 1, \ldots, n$, and $\sum_{i=1}^n (\tilde{x}_i - x_i) < \delta_f(\epsilon)$, we have $\sum_{i=1}^n f(\tilde{x}_i) - f(x_i) < \epsilon$. For the function $g \circ f$, take $a = a_f \in (0, \alpha_0) \subset \mathbb{R}$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, let $\delta(\epsilon) = \delta_f(\epsilon/(1 + L_0)) \in (0, a_f) \subset \mathbb{R}$. $\forall n \in \mathbb{N}$, $\forall x_0 - a \leq x_1 \leq \tilde{x}_1 \leq x_2 \leq \cdots \leq x_n \leq \tilde{x}_n \leq x_0 + a$ with $x_i, \tilde{x}_i \in I$, $i = 1, \ldots, n$, and $\sum_{i=1}^n (\tilde{x}_i - x_i) < \delta(\epsilon)$, we have $f(\tilde{x}_i), f(x_i) \in U \cap B_y(y_0, \delta_g)$ and $\sum_{i=1}^n \|g \circ f(\tilde{x}_i) - g \circ f(x_i)\| \leq \sum_{i=1}^n L_0 \|f(\tilde{x}_i) - f(x_i)\| \leq L_0 \frac{\delta(\epsilon)}{1 + L_0} < \epsilon$. Hence, $g \circ f$ is absolutely continuous at $x_0$. This completes the proof of the proposition.

**Proposition 12.68** Let $I \subseteq \mathbb{R}$ be an interval, $Y$ be a normed linear space over $\mathbb{K}$, $Z$ be a normed linear space over $\mathbb{K}$, $f : I \rightarrow Y$ 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.

**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 $\mathbb{K}$, $Z$ be a normed linear space over $\mathbb{K}$, $F : \Omega \rightarrow Y$ 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 $m$.

$1^o$ $m = 1$. 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 \( i_0 \in \bar{J} := \{1, \ldots, n\} \), let \( D_{i_0}, \pi_{i_0}, \) and \( F_{i_0} \) be defined as in (i) of Definition 12.58. Let \( h_{i_0} : D_{i_0} \rightarrow \mathbb{Z} \) be defined by \( h_{i_0}(x) = h(\bar{x}) = AF(\bar{x}) = A\bar{F}(x), \ \forall x \in D_{i_0} \subseteq \mathbb{R}^k \).

Consider two exhaustive and mutually exclusive cases: Case 1: \( \exists h(\bar{x}) \in U \subseteq \mathbb{R}^k \) such that \( \exists \pi_{i_0}(x) = \pi_{i_0}(x_0) \) and \( \pi_{i_0}(\bar{x}) = x_0 \). By the assumption, \( F \) is absolutely continuous at \( x_0 \), then \( F_{i_0} \) is absolutely continuous at \( \pi_{i_0}(x_0) \). By inductive assumption, \( h_{i_0} \) 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, \( \exists \alpha \in (0, \infty) \subset \mathbb{R}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta(\epsilon) \in (0, \infty) \subset \mathbb{R}, \forall n \in \mathbb{N}, \forall (\bar{x}_i)_{i=1}^n, (\bar{x}_i)_{i=1}^n \subseteq \mathbb{R} \times \mathbb{R} \cap \Omega \) such that \( \bar{x}_i \subseteq \bar{x}_i \), \( \forall i \in \{1, \ldots, n\} \), \( (\bar{r}_{x_i, \bar{x}_i})_{i=1}^n \) being pairwise disjoint, and \( \sum_{i=1}^n |\Delta_x(t_{x_i, \bar{x}_i})| < \delta(\epsilon) \), we have \( \sum_{i=1}^n \|\Delta_x(t_{x_i, \bar{x}_i})\| < \epsilon \). It is easy to see that \( \Delta_h(t_{x_i, \bar{x}_i}) = A\Delta_x(t_{x_i, \bar{x}_i}) \), \( \forall i \in \{1, \ldots, n\} \). Then, \( \sum_{i=1}^n \|\Delta_h(t_{x_i, \bar{x}_i})\| \leq \|A\| \sum_{i=1}^n \|\Delta_x(t_{x_i, \bar{x}_i})\| < \epsilon \). 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 \( \mathcal{X} \) be a normed linear space over \( \mathbb{K}_1 \), \( Y \) be a set, \( \mathcal{Z} \) be a normed linear space over \( \mathbb{K}_2 \), \( U \subseteq \mathcal{X} \) be endowed with subset topology \( \mathcal{O} \), \( f : U \times Y \rightarrow \mathcal{Z}, g : U \rightarrow \mathbb{R} \) be continuous, \((c_i)_{i \in \mathbb{Z}} \subset \mathbb{R} \) be a sequence of strictly increasing real numbers with \( \lim_{i \rightarrow -\infty} c_i = -\infty \) and \( \lim_{i \rightarrow \infty} c_i = \infty \), \( U_i := \{x \in U \mid c_i \leq g(x) \leq c_{i+1}\}, \forall i \in \mathbb{Z} \). Assume that

(i) \( U \) is locally convex according to Definition 7.12;

(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 \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that \( B_X(x_0, \delta_1) \cap U \subseteq U_i \). By \( f|_{U_{i \times Y}} \) being locally Lipschitz on \( U_i \) uniformly over \( Y \), we have \( f|_{U_{i \times Y}} \) is locally Lipschitz at \( x_0 \) uniformly over \( Y \). By Definition 12.64, \( \exists L_1 \in [0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \delta_1] \subset \mathbb{R}, \forall x_1, x_2 \in B_x(x_0, \delta) \cap U \subseteq B_X(x_0, \delta_1) \cap U \cap U \), \( \forall y \in Y \), 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 \( g \), \( \exists x_0 \in (0, \infty) \subset \mathbb{R} \) such that \( U \cap B_X(x_0, \delta_{x_0}) \subseteq U_{i-1} \cap U_i \) is convex. By \( f|_{U_{i \times Y}} \) being locally Lipschitz on \( U_i \) uniformly over \( Y \), we have \( f|_{U_{j \times Y}} \) is locally Lipschitz at \( x_0 \) uniformly over \( Y \), \( j = i - 1, i \).
By Definition 12.64, \( \exists L_j \in [0, \infty) \subset \mathbb{R}, \exists \delta_j \in (0, \delta_{x_0}] \subset \mathbb{R}, \forall x_1, x_2 \in B_{X} (x_0, \delta_j) \cap U_j, \forall y \in Y, \) we have \( \| f(x_1, y) - f(x_2, y) \| \leq L_j \| x_1 - x_2 \|, \) \( j = i - 1, i. \) Take \( L := L_{i-1} \cap L_i \in [0, \infty) \subset \mathbb{R} \) and \( \delta := \delta_1 \vee \delta_2 \in (0, \delta_{x_0}] \subset \mathbb{R}. \) \( \forall x_1, x_2 \in B_{X} (x_0, \delta) \cap U := D, \forall y \in Y, \) 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
\[
t_0 := \sup \{ t \in [0, 1] \subset \mathbb{R} \mid x_t \in U_i \} =: \sup A
\]
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, \( g(x_{t_n}) = \lim_{n \to \infty} g(x_{t_n}) \geq c_i, x_{t_n} \in U_i, \) 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, \( \| f(x_1, y) - f(x_2, y) \| \leq \| f(x_1, y) - f(x_{t_0}, y) \| + \| f(x_{t_0}, y) - f(x_2, y) \| \leq L_{i-1} \| x_1 - x_{t_0} \| + L_i \| x_2 - x_{t_0} \| = L_{i-1} (1 - t_0) \| x_1 - x_2 \| + L_i \| x_1 - x_2 \| \leq L \| x_1 - x_2 \|. \)

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, \( \Gamma := ((I, |·|), \mathcal{B}, \mu) \) 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
\[
\int_a^b f(x) \, dx := \int_a^b f \, d\mu := \begin{cases} \int_a^b f \, d\mu & \text{if } b \geq a \\ -\int_b^a f \, d\mu & \text{if } b < a \end{cases}
\]
whenever the right-hand-side makes sense.

By Lemma 11.73, the interval \( I \) may be \( [a, b], \) \( (a, b], \) \( [a, b), \) or \( (a, b), \) which will not change the value of the integral.
2.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

Fact 12.72 Let \( a, b, c \in \mathbb{R} \), \( I := r_{a,b} \cap (a \lor b, c \lor b) \subset \mathbb{R} \), \( I := ((I, | \cdot |), \mathcal{B}, \mu) \) 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, \( \int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx \in \mathcal{Y} \).

Proof We will distinguish six exhaustive and mutually exclusive cases:

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: \( b \leq a \leq c \). Let \( J_1 := r_{a,c} \subset \mathbb{R} \), \( J_2 := r_{e,b} \subset \mathbb{R} \). By Proposition 11.92, \( \int_a^b f(x) \, dx = \int_{J_1} f \, d\mu_B = \int_{J_2} f \, d\mu_B + \int_{J_2} f \, d\mu_B = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \in \mathcal{Y} \). Then, \( \int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx \in \mathcal{Y} \).

Case 2: \( a \leq b \leq c \). By Case 1, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \). Then, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \).

Case 3: \( b \leq a \leq c \). By Case 1, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx \). Then, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_b^c f(x) \, dx \).

Case 4: \( b \leq c \leq a \). By Case 1, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx \). Then, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_b^c f(x) \, dx \).

Case 5: \( c \leq a \leq b \). By Case 1, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \). Then, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_c^b f(x) \, dx \).

Case 6: \( b \leq a \leq c \). By Case 1, \( \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx \). Then, \( \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_b^c f(x) \, dx \).

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, \( \Omega \in \mathcal{B}(\mathbb{R}^m) \) be a rectangle with the subset topology \( \mathcal{O} \), \( (\mathbb{P}(\Omega), | \cdot |, \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure space 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 \((\mathbb{P}(\Omega), \mathcal{Y}, \nu)\) 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 \( x_1 \triangleq x_2 \), and \( \mathcal{P} \circ \nu(\mathbb{P}(\Omega)) = T_F \). \( \mathcal{P} \circ \nu \ll \mu \).

Proof (i) By Proposition 12.52, \( \Omega \) is a region and \( \mathbb{P}(\Omega) \in \mathcal{B}(\mathbb{R}^m) \) is a rectangle. By Proposition 12.59, \( F \) is continuous. Then, (i) of Definition 12.41 holds.
∀x₁, x₂ ∈ Ω with x₁ ≤ x₂, consider the closed rectangle rₓ₁,x₂ ⊂ Ω. ∀x₀ ∈ rₓ₁,x₂, F is absolutely continuous at x₀. By (ii) of Definition 12.58, ∃x_{x₀} ∈ (0, ∞) ⊂ R, ∀ε ∈ (0, ∞) ⊂ R, ∃δ_{x₀}(ε) ∈ (0, ∞) ⊂ R, ∀n ∈ N, ∀(x_{i})_{i=1}^{n} ⊂ rₓ₀-a_{1}m_{x}+a_{1}m_{n} ∩ Ω with x_{i} ≤ x_{j}, ∀i ∈ {1, ..., n}, (rₓ_{i}, x_{j})_{i=1}^{n} being pairwise disjoint, and \( \sum_{i=1}^{n} \varnothing \mu(Bm(rₓ_{i}, x_{j})) < δ_{x₀}(ε) \), we have \( \sum_{i=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| < ε \). Then, \( \varnothing \sum_{i=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| < ε \). Hence, \( \varnothing \) is compact. ∀x₁, x₂, ∃ a finite set N ⊂ rₓ₁,x₂ such that \( \varnothing \sum_{i=1}^{n} \varnothing \sum_{j=1}^{n} \| rₓ_{i}, x_{j} \| ≤ \sum_{x∈N} \sum_{i=1}^{n} \sum_{j=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| \leq \sum_{x∈N} \sum_{i=1}^{n} \sum_{j=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| \). Hence, F satisfies (ii) of Definition 12.41.

∀x₁, x₂ ∈ Ω with x₁ < x₂, define \( G : rₓ₁,x₂ → (0, ∞) ⊂ R \) by \( G(x) = T_F(rₓ₁,x₂) \), ∀x ∈ rₓ₁,x₂. We will show that \( G \) is continuous on the right. Fix any x₀ ∈ rₓ₁,x₂. Consider the sets S_{J} : \{ x ∈ rₓ₁,x₂ : \| rₓ_{i}(x_{j}) \| = \| rₓ(x_{j}) \| + \| rₓ_{i}(x_{j}) \| \}, ∀J ⊂ J : \{ 1, ..., m \} with J ≠ \emptyset. Fix any J ⊂ J with J ≠ \emptyset. Clearly, S_{J} is compact. ∀x ∈ S_{J}, by (ii) of Definition 12.58, ∃x₀ ∈ (0, ∞) ⊂ R, ∀ε ∈ (0, ∞) ⊂ R, ∃δ_{x_{0}}(ε) ∈ (0, ∞) ⊂ R, ∀n ∈ N, ∀(x_{i})_{i=1}^{n} ⊂ rₓ₀-a_{1}m_{x}+a_{1}m_{n} ∩ rₓ₁,x₂ with x_{i} ≤ x_{j}, ∀i ∈ {1, ..., n}, (rₓ_{i}, x_{j})_{i=1}^{n} being pairwise disjoint, and \( \sum_{i=1}^{n} \varnothing \mu(Bm(rₓ_{i}, x_{j})) < δ_{x_{0}}(ε) \), we have \( \sum_{i=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| < ε \). Then, \( S_{J} ⊂ \bigcup_{x∈S_{J}} \sum_{i=1}^{n} \sum_{j=1}^{n} \| \Delta F(rₓ_{i}, x_{j}) \| \). By the compactness of S_{J}, ∃ a finite set N_{J} ⊂ S_{J} such that S_{J} ⊂ \( \bigcup_{x∈N_{J}} rₓ_{i}, x_{j} \bigcup x_{a_{1}m_{x}+a_{1}m_{n}} \). Let N := \( \bigcup_{J⊂J, J≠\emptyset} N_{J} \). Let \( a := \varnothing \min_{x∈N} a_{x} ∈ (0, ∞) ⊂ R \), ∀ε ∈ (0, ∞) ⊂ R, let \( δ := \varnothing a_{x} ε/(1 + \varnothing \text{card}(N)) \) ∈ (0, ∞) ⊂ R. Let \( δ \) be the minimum positive real root of \( \prod_{i=1}^{n} (\pi_{i}(x_{0}) + \pi_{i}(x_{1}) + \delta) - \prod_{i=1}^{n} (\pi_{i}(x_{0}) - \pi_{i}(x_{1})) = \delta / 2 \) and \( δ := \varnothing a_{x} ε/(1 + \varnothing \text{card}(N)) \) ∈ (0, ∞) ⊂ R. ∀x ∈ B_{R}(m, x_{0}, δ) ∩ rₓ₁,x₂ with x_{0} ≤ x, we have \( G(x) - G(x_{0}) = | T_F(rₓ₁,x₂) - T_F(rₓ₁,x₀) | = \sum_{J⊂J, J≠\emptyset} T_F(rₓ_{i}, x_{j}) | x_{j} ∈ J, x_{j} ≤ x, x_{j} ∈ rₓ₁,x₂. \)
\( \forall i \in \mathcal{J} \setminus J; \) and \( \hat{x}_J \in \mathbb{T}_{r_0, \bar{x}} \) is such that \( \pi_i(\hat{x}_J) = \pi_i(\bar{x}), \) \( \forall i \in J, \)
\( \pi_i(\hat{x}_J) = \pi_i(x_0) \), \( \forall i \in J \setminus J; \) and the second equality follows from Proposition 12.47. \( \forall J \subseteq J \) with \( J \not= \emptyset \), \( r_{x_J, \hat{x}_J} \subseteq \bigcup_{x \in I_0, r^2_{x-a_1m, x+a_1m} \cap r_{x_J, \hat{x}_J}} \). Then, we have \( |\frac{G(x)}{G(x_0)}| = 1 \) for any \( x \in I_0 \).

Then, \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \), \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \).

All we need to show is \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \). Note that, \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \).

Then, \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \).

By Definition 12.41, \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \).

Thus, \( \exists \gamma < \delta \), \( \forall x \in \mathbb{T}_{r_0, \bar{x}} \).
2^{n-1} \epsilon$. Fix any finite subset $N_f \subseteq N$, we have $\sum_{i \in N_f} P \circ \nu(r_{x_{i,1},x_{i,2}}) \leq \sum_{i \in N_f} \sum_{j=1}^{n_i} \left\| \Delta P(r_{x_{i,j},x_{i,j+1}}) \right\| + \epsilon/2 =: \sum_{i=1}^{n} \left\| \Delta P(r_{x_{i,1},x_{i,2}}) \right\| + \epsilon/2$, where $n \in \mathbb{Z}_+$ and $(x_{i,1})_{i=1}^{n}, (x_{i,2})_{i=1}^{n} \subseteq r_{x_{i,1},x_{i,2}}$ with $\sum_{i=1}^{n} \mu(r_{x_{i,1},x_{i,2}}) = \sum_{i=1}^{n} \mu B_m(r_{x_{i,1},x_{i,2}}) \leq \mu B_m(\Omega \cap r_{x_{1,1},x_{2,2}}) < \delta(\epsilon)$. We may rearrange and partition the collection into $p$ segments:

$r_{x_{1,1},x_{1,2}}, ... , r_{x_{i,1},x_{i,2}} \subseteq r_{\tilde{x}_{1},\lambda_{p_{1}},1_{1},\tilde{x}_{1},\lambda_{p_{1}},1_{1}}; r_{x_{i+1,1},x_{i+1,2}}; \cdots; r_{x_{p_{1}+1,1},x_{p_{1}+1,2}}, r_{x_{p_{1}+2,1},x_{p_{1}+2,2}} \subseteq r_{\tilde{x}_{1},\lambda_{p_{1}},1_{1},\tilde{x}_{1},\lambda_{p_{1}},1_{1}}; \cdots; r_{x_{p_{1}+1,1},x_{p_{1}+1,2}}, r_{x_{p_{1}+2,1},x_{p_{1}+2,2}} \subseteq r_{\tilde{x}_{1},\lambda_{p_{1}},1_{1},\tilde{x}_{1},\lambda_{p_{1}},1_{1}}$. Note that $\forall j \in \{1, ..., p\}$, $\sum_{i=j+1}^{n} \mu B_m(r_{x_{i,1},x_{i,2}}) < \delta(\epsilon) \leq \delta_{\epsilon}(\frac{\epsilon}{2p})$, which implies that $\sum_{i=j+1}^{n} \left\| \Delta P(r_{x_{i,1},x_{i,2}}) \right\| < \frac{\epsilon}{2p}$. Then, we have $\sum_{i \in N_f} P \circ \nu(r_{x_{i,1},x_{i,2}}) < \epsilon$. By the arbitrariness of $N_f$, we have $P \circ \nu(E) \leq P \circ \nu(O \cap r_{x_{1,1},x_{1,2}}) = \sum_{i \in N} P \circ \nu(r_{x_{i,1},x_{i,2}}) = \sum_{i \in N} P \circ \nu(r_{x_{i,1},x_{i,2}}) \leq \epsilon$. By the arbitrariness of $\epsilon$, we have $P \circ \nu(E) = 0$. By the arbitrariness of $E$, we have $P \circ \nu \ll \mu$. This proves that $P \circ \nu \ll \mu$ in the special case.

Consider the general case when $\Omega$ is a rectangle. By Proposition 12.39, $\Omega$ is a region and $P(\Omega)$ is a rectangle. By Definition 12.38, $\exists N \subseteq \mathbb{N}$ and $(\tilde{x}_{i})_{i \in N} \subseteq \Omega$ with $\tilde{x}_{i} \leq \tilde{x}_{i}, \forall i \in N$, such that $\Omega = \bigcup_{i \in N} r_{\tilde{x}_{i},\tilde{x}_{i}}$. For the special case, we have $P \circ \nu(E \cap r_{x_{1,1},x_{1,2}}) = 0$, $\forall i \in N$. Then, we have $P \circ \nu(E) = \sum_{i \in N} P \circ \nu(E \cap r_{x_{i,1},x_{i,2}}) = 0$. By the arbitrariness of $E$, we have $P \circ \nu \ll \mu$.

This completes the proof of the proposition. □

Lemma 12.74 Let $X := (X, B, \mu)$ be a measure space and $\nu$ be a measure on $(X, B)$. The following statements hold.

(i) $\nu \ll \mu$ and $\nu$ is finite implies that $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}$, $\forall E \in B$ with $\mu(E) < \delta$, we have $\nu(E) < \epsilon$.

(ii) $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}, \forall E \in B$ with $\mu(E) < \delta$, we have $\nu(E) < \epsilon$. Then, $\nu \ll \mu$.

Proof (i) We will prove the result by an argument of contradiction. Suppose the result does not hold. Then, $\exists \epsilon_0 \in (0, \infty) \subset \mathbb{R}, \forall \in N$, $\exists E_{n} \in B$ with $\mu(E_{n}) < 2^{-n-1}$ such that $\nu(E_{n}) > \epsilon_0$. Let $E_{n} := \bigcup_{j=n}^{\infty} E_{j}$, $\forall n \in N$. Then, $E_{n} \supseteq \bigcap_{n=1}^{\infty} E_{n}$ and $\mu(E_{n}) \leq \sum_{j=n}^{\infty} \mu(E_{j}) < 2^{-n}$. By Proposition 11.5, we have $0 = \lim_{n \to \infty} \mu(E_{n}) = \mu(\bigcap_{n=1}^{\infty} E_{n})$. Then, by $\nu \ll \mu$ and Proposition 11.5, we have $0 = \nu(\bigcap_{n=1}^{\infty} E_{n}) \leq \lim_{n \to \infty} \nu(E_{n})$. This contradicts with the fact that $\nu(E_{n}) \geq \nu(E_{n}) > \epsilon_0, \forall n \in N$. Hence, the result holds.

(ii) $\forall E \in B$ with $\mu(E) = 0$. By the assumption, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall E \in B$ with $\mu(E) < \delta$, we have $\nu(E) < \epsilon$. In particular, $\mu(E) = 0 < \delta$, and this implies that $\nu(E) < \epsilon$. By the arbitrariness of $\epsilon$, we have $\nu(E) = 0$. By the arbitrariness of $E$, we have $\nu \ll \mu$. 

610
12.5. ABSOLUTE AND LIPSCHITZ CONTINUITY

This completes the proof of the lemma. □

Proposition 12.75 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 \( \mathcal{O} \), \( y \) be a normed linear space over \( K \), \( \nu \) be a \( \sigma \)-finite \( y \)-valued measure on the measurable space \( (\mathbb{P}(\Omega), \mathcal{B}(\mathbb{P}(\Omega)), \mathbb{X}) := (\mathbb{P}(\Omega), \mathcal{B}(\mathbb{P}(\Omega)), \mu) \) be the metric measure subspace of \( \mathbb{R}^m \). Assume that \( \mathcal{P} \circ \nu \ll \mu \) and \( \mathcal{P} \circ \nu (r_{x_{1},x_{2}}) \) be an arbitrary fixed vector in \( \Omega \) if \( \Omega \neq \emptyset \). Then, \( \mathcal{P} \) is of locally bounded variation and is absolutely continuous.

In particular, if \( y \) be a Banach space over \( K \), \( M \subseteq y \) be a separable subspace, and \( f : \mathbb{P}(\Omega) \rightarrow M \) be \( \mathcal{B} \)-measurable. We may define \( \nu \) to be the \( y \)-valued measure with kernel \( f \) over \( \mathbb{X} \) 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 \( x_{1} \cong x_{2} \), \( f \) is absolutely integrable on \( r_{x_{1},x_{2}} \). Then, \( \nu \) satisfies the assumption of the previous paragraph. Let \( f : \Omega \rightarrow y \) be the cumulative distribution function of \( \nu \) with origin \( x_{s} \), where \( x_{s} \) is an arbitrary fixed vector in \( \Omega \).

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 \( J \subseteq \mathcal{J} := \{1,\ldots,m\} \) with \( \text{card}(J) = i \). Let \( \pi_{J} : \mathbb{R}^m \rightarrow \mathbb{R}^{i} \), \( \pi_{J \setminus \{i\}} : \mathbb{R}^m \rightarrow \mathbb{R}^{m-i} \), \( M : \mathbb{R}^{i} \times \mathbb{R}^{m-i} \rightarrow \mathbb{R} \), \( J \) with \( \tilde{x}_{f} \), \( \tilde{\Omega}_{f} \), and \( \bar{F}_{J} : \tilde{\Omega}_{f} \rightarrow y \) be as defined in (a) Proposition 12.60. Define \( \bar{x} := M(\pi_{J}(x_{i}),\pi_{J \setminus \{i\}}(x_{i})) \in \Omega \), \( E := r_{x_{f} \cap \pi_{J}(x_{i}) \cup \pi_{J \setminus \{i\}}(x_{i})} \), and \( E_{n} := r_{x_{f} - 1/n, \tilde{x}_{f} + 1/n} \cap \tilde{\Omega}_{f} \), \( \forall n \in \mathbb{N} \). Clearly, \( V_{n} := M(E \times E_{n}) \subseteq \Omega \) and \( \mu_{\mathcal{B}m}(P(V_{n})) = \mu_{\mathcal{B}m}(V_{n}) = \mu_{\mathcal{B}m}(E) \mu_{\mathcal{B}m}(E_{n}) \leq (2/n)^{m-i} \mu_{\mathcal{B}m}(E) < \infty \), \( \forall n \in \mathbb{N} \). Then, \( \lim_{n \rightarrow \infty} \mu_{\mathcal{B}m}(P(V_{n})) = \mu_{\mathcal{B}m}(\bigcap_{n=1}^{\infty} P(V_{n})) = 0 \), 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 \( \exists x_{0} \in \mathcal{N} \) such that \( P(V_{x_{0}}) = r_{x_{1},x_{2}} \) for some \( x_{1},x_{2} \in \Omega \) with \( x_{1} \cong x_{2} \). Then, \( \mathcal{P} \circ \nu (P(V_{x_{0}})) \).
CHAPTER 12. DIFFERENTIATION AND INTEGRATION

\[ P(V_{n_0}), \] where the second equality, the first set containment, the third equality, and the second set containment follow from Proposition 12.39. Hence, \( \mathcal{P} \circ \nu(M(E \times (\bigcup_{j=1}^n r_{\hat{x}_j}, \hat{x}_j)))) < \epsilon \). Therefore, \( \sum_{j=1}^n \| \Delta F_{r_{\hat{x}_j}, \tilde{x}_j} \| < \epsilon \).

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, \( (P(\Omega), \mathcal{B}_{\mathcal{P}(\Omega)}, \nu) \) is a \( \sigma \)-finite \( Y \)-valued measure space and \( \mathcal{P} \circ \nu = \bar{\nu} \), where \( \bar{\nu} \) is a \( \sigma \)-finite measure on \( (P(\Omega), \mathcal{B}_{\mathcal{P}(\Omega)}) \) defined by \( \bar{\nu}(E) = \int_E \mathcal{P} \circ f \, d\mu \in [0, \infty] \subset \text{IR}^e \), \( \forall E \in \mathcal{B}_{\mathcal{P}(\Omega)} \). 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 \( x_1 \leq x_2 \), 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. \( \Box \)

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 \( \text{dia}(S) = \max\{\sup_{x_1, x_2 \in S} \rho(x_1, x_2), 0\} \in \text{IR}^e \).
12.6. FUNDAMENTAL THEOREM OF CALCULUS

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., $\mu_{\text{Box}}(U) > 0$, $\forall U \in \mathcal{I}$. We will say that $\mathcal{I}$ covers $E$ in the sense of Vitali with index $c \in (0, \frac{1}{\sqrt{m}}] \subseteq \mathbb{R}$ if $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, $\forall x \in E$, $\exists U \in \mathcal{I}$ such that $x \in U$, $\text{dia}(U) < \epsilon$, 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 $\mu_{\text{Box}}(E) < +\infty$ 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, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$, there exists a finite pairwise disjoint subcollection $\{U_1, \ldots, U_n\} \subseteq \mathcal{I}$, where $n \in \mathbb{Z}_+$, such that $\mu_{\text{Box}}(E \setminus \bigcup_{i=1}^{n} U_i) < \epsilon$.

Proof Let $\hat{\mathcal{I}} := \{\bigcup_{U \in \mathcal{I}} \subseteq \mathbb{R}^m \mid U \in \mathcal{I}\}$. 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 $\epsilon \in (0, \infty) \subseteq \mathbb{R}$.

By Example 12.56, $\mu_{\text{Box}}(E) = \inf_{(O)_{\subseteq \mathbb{R}^m} \subseteq \mathcal{I}, \sum_{i=1}^{\infty} \mu_{\text{Box}}(O_i) < +\infty}$. Then, $\exists O \in \mathcal{O}_{\mathbb{R}^m}$ such that $E \subseteq O$ and $\mu_{\text{Box}}(O) < +\infty$. 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 $c \text{dia}(U)$, $\forall U \in \mathcal{I}$.

$\forall k \in \mathbb{Z}_+$, assume that pairwise disjoint $U_1, \ldots, U_k \in \hat{\mathcal{I}}$ has already be 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: Let $\hat{\mathcal{I}} = \bigcup_{i=1}^{k} U_i$. Clearly, $k \leq \sqrt[2m]{\mu_{\text{Box}}(E)} < +\infty$. Then, $\exists x_0 \in E \setminus \bigcup_{i=1}^{k} U_i$. Since $\bigcup_{i=1}^{k} U_i$ is closed, then, by Proposition 4.10, $d := \inf_{x \in (\bigcup_{i=1}^{k} U_i)} |x - x_0| > 0$. Then, by the assumption, $\exists \hat{U} \in \hat{\mathcal{I}}$ such that $x_0 \in \hat{U}$ and $\text{dia}(\hat{U}) < d$. Then, $\hat{U} \cap (\bigcup_{i=1}^{k} U_i) = \emptyset$. This shows that $\hat{k} \geq \text{dia}(\hat{U}) > 0$. 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 $\text{dia}(U_{k+1}) \geq \frac{l_k}{2}$. Inductively, we either have the result holds for $\hat{\mathcal{I}}$ or $\exists (U_k)_{k=1} \subseteq \hat{\mathcal{I}}$ such that $\text{dia}(U_{k+1}) \geq \frac{l_k}{2}$, $\forall k \in \mathbb{Z}_+$ and the sequence is pairwise disjoint.

In the latter case, we have $O \supseteq \bigcup_{k=1}^{\infty} U_k$ and $+\infty > \mu_{\text{Box}}(O) \geq \mu_{\text{Box}}(\bigcup_{k=1}^{\infty} U_k) = \sum_{k=1}^{\infty} \mu_{\text{Box}}(U_k) \geq \sum_{k=0}^{\infty} 2^{-m} c^m p_k$. 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 $\sum_{i=n+1}^{\infty} \mu_{\text{Box}}(U_i) < \pi^{-m/2} \Gamma\left(\frac{m}{2} + 1\right)\epsilon$, 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 \( x_k \), \( \forall k \in \mathbb{N} \). \( \forall \hat{x} \in R \), since \( \bigcup_{k=1}^n U_k = \mathbb{R} \) is closed, then, by Proposition
4.10, \( d := \inf_{x \in \bigcup_{k=1}^n U_k} |x - \hat{x}| > 0 \). Thus, the assumption, \( \exists U \in \mathcal{I} \)
such that \( \hat{x} \in U \) and \( 0 < \text{dia}(U) < d \). Then, \( U \cap (\bigcup_{k=1}^n U_k) = \emptyset \).
By the fact that \( \lim_{k \in \mathbb{N}} I_k = 0 \), 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, \( \text{dia}(U) \leq I_{i_0-1} \leq 2 \text{dia}(U_{i_0}) = 2p_{i_0} \). Let \( x \in U \cap U_{i_0} \).
Then, \( |\hat{x} - x| \leq |\hat{x} - \hat{x}| + |x - x_{i_0}| \leq \text{dia}(U) + \frac{1}{2}p_{i_0} \leq \frac{1}{2}p_{i_0} \). Then, \( \hat{x} \in V_{i_0} := \mathcal{B}_{\mathbb{R}^m}(x_{i_0}, \frac{1}{2}p_{i_0}) \subset \mathbb{R}^m \).
Then, \( \hat{x} \in \bigcup_{k=n+1}^{\infty} V_k := \bigcup_{k=n+1}^{\infty} \mathcal{B}_{\mathbb{R}^m}(x_k, \frac{1}{2}p_k) \subset \mathbb{R}^m \). Hence, \( R \subset \bigcup_{k=n+1}^{\infty} V_k \).
Then, \( \mu_{\text{LMo}}(R) \leq \mu_{\text{LMo}}(\bigcup_{k=n+1}^{\infty} V_k) \leq \sum_{k=n+1}^{\infty} \mu_{\text{BM}}(V_k) = \sum_{k=n+1}^{\infty} \frac{\pi^{m/2}}{\Gamma(\frac{m}{2}+1)} \frac{\pi^{m/2}p_k^{m}}{\Gamma(\frac{m}{2}+1)} \sum_{k=n+1}^{\infty} c^m p_k^m \leq \frac{\pi^{m/2}p_1^{m}}{\Gamma(\frac{m}{2}+1)} \sum_{k=n+1}^{\infty} c^m p_k^m \). \( \sum_{k=n+1}^{\infty} \mu_{\text{BM}}(U_k) < \epsilon \), where the second equality follows from (Mathematics Handbook Editors Group, 1979, pp. 320). Hence, the result holds for \( \mathcal{I} \).
Then, \( \exists \in \mathbb{Z}_+, \exists U_1, \ldots, U_n \subseteq \mathcal{I} \) such that \( \mu_{\text{LMo}}(E \setminus (\bigcup_{i=1}^n U_i)) < \epsilon \). Note that \( \mu_{\text{LMo}}(J) := \mu_{\text{LMo}}((\bigcup_{i=1}^n U_i) \setminus (\bigcup_{i=1}^n U_i)) = 0 \). Then, \( J \in \mathcal{B}_1 \) and \( \mu_{\text{LMo}}(E \setminus (\bigcup_{i=1}^n U_i)) = \mu_{\text{LMo}}((E \setminus (\bigcup_{i=1}^n U_i)) \setminus J) + \mu_{\text{LMo}}((E \setminus (\bigcup_{i=1}^n U_i)) \setminus J) < \epsilon \), where the equality follows from Definition 11.15; and the first inequality follows from fact that \( \mu_{\text{LMo}}((E \setminus (\bigcup_{i=1}^n U_i)) \setminus J) = 0 \) and the fact that \( E \setminus (\bigcup_{i=1}^n U_i) \setminus J = E \setminus (\bigcup_{i=1}^n U_i) \).
This completes the proof of the lemma. \( \square \)

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 \( \mu_{\text{BM}}(E) > 0 \), \( \text{E} := ((E, |\cdot|), \mathcal{B}, \mu) \) 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 \( \mathbb{K} \), \( c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R} \), and \( f \in L_1(E, \text{Y}) \). A point \( x \in E \) is said to be a rectangular Lebesgue point with regularity \( c \) of \( f \) if \( \limsup_{r \to 0} \sup_{x_1, x_2 \in \mathbb{R}^m} \left[ \frac{1}{\mu((x_1, x_2) \in \text{E})} \int_{x_1, x_2 \in \text{E}} \|f(y) - f(x)\| \ d\mu(y) \right] = 0 \).

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 \( \mu_{\text{BM}}(E) > 0 \), \( \text{E} := ((E, |\cdot|), \mathcal{B}, \mu) \) 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 \( \mathbb{K} \), \( c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R} \), \( f \in L_1(E, \text{Y}) \), and \( A := \{ x \in E \mid x \) is a rectangular Lebesgue point of \( f \) with regularity \( c \} \). Then, \( A \in \mathcal{B}_{\text{LM}} \) and \( \mu_{\text{LM}}(E \setminus A) = 0 \).

Proof Since \( \text{E} \) is \( \sigma \)-finite, then \( \exists \bigcup_{n=1}^{\infty} E_n \subseteq \mathcal{B} \) 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, \( \forall x \in E \),

\[
(T_{r,c}(f))(x) = \sup_{\frac{x_1 + x_2}{x_2 - x_1} = x_{1,2}, x_1, x_2 \in \mathbb{R}^m} \frac{1}{\mu((x_1, x_2) \in \text{E})} \int_{x_1, x_2 \in \text{E}} \|f(y) - f(x)\| \ d\mu(y)
\]

\[
(M_{r,c} : E \to [0, \infty) \subset \mathbb{R} \) by, \( \forall x \in E \),
Clearly, \( \mu \) measures, we have \( \hat{\mu}(A) \geq \mu(A) \) for any \( A \subseteq \mathbb{R} \). Hence, \( \hat{\mu} \) is a measure on \( \mathcal{B} \). Let \( \Delta = \mathbb{N} \) and any \( \epsilon > 0 \). Suppose \( \mu_{\Delta}(\mathbb{N}) > 0 \). By monotonicity of outer measures, we have \( \mu_{\Delta}(A_{n,k}) - \mu_{\Delta}(E_n) = \mu_{\Delta}(E_n) < +\infty \). Let \( \epsilon_0 = \frac{\mu_{\Delta}(E_n)}{2k} \in (0, +\infty) \subset \mathbb{R} \). By Propositions 11.182 and 4.11, there exists a continuous function \( g : E \to \mathbb{Y} \) such that \( g \in L_1(\mathbb{E}, \mathcal{Y}) \) and \( g - f < \epsilon_0 \). Let \( h := f - g \). Then, \( \|h\|_1 < \epsilon_0 \). Note that \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x) + (T_{r,c}h)(x), \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E \). Then, by Proposition 3.85, \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x), \forall x \in E \). Since \( g \) is continuous, we have \( (T_{r,c}g)(x) = 0, \forall x \in E \). Then, \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x), \forall x \in E \). Note that

\[
(T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \|h(x)\|, \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E
\]

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, \( (T_{r,c}h)(x) \leq (M_{r,c}h)(x), \forall x \in E \). Then, \( \hat{\mu}_{\Delta}(A_{n,k}) \leq \mu_{\Delta}(\mathbb{N}) \). By monotonicity of outer measures, we have \( \mu_{\Delta}(A_{n,k}) - \mu_{\Delta}(E_n) = \mu_{\Delta}(E_n) < +\infty \). Let \( \Delta = \mathbb{N} \) and any \( \epsilon > 0 \). Suppose \( \mu_{\Delta}(\mathbb{N}) > 0 \). By monotonicity of outer measures, we have \( \mu_{\Delta}(A_{n,k}) - \mu_{\Delta}(E_n) = \mu_{\Delta}(E_n) < +\infty \). Let \( \epsilon_0 = \frac{\mu_{\Delta}(E_n)}{2k} \in (0, +\infty) \subset \mathbb{R} \). By Propositions 11.182 and 4.11, there exists a continuous function \( g : E \to \mathbb{Y} \) such that \( g \in L_1(\mathbb{E}, \mathcal{Y}) \) and \( g - f < \epsilon_0 \). Let \( h := f - g \). Then, \( \|h\|_1 < \epsilon_0 \). Note that \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x) + (T_{r,c}h)(x), \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E \). Then, by Proposition 3.85, \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x), \forall x \in E \). Since \( g \) is continuous, we have \( (T_{r,c}g)(x) = 0, \forall x \in E \). Then, \( (T_{r,c}h)(x) \leq (T_{r,c}g)(x), \forall x \in E \). Note that

\[
(T_{r,c}h)(x) \leq (M_{r,c}h)(x) + \|h(x)\|, \forall r \in (0, \infty) \subset \mathbb{R}, \forall x \in E
\]
Define \( I := \{ r_{x_1, x_2} \subseteq \mathbb{R}^m \mid x = \frac{1}{2}(x_1 + x_2) \in E, |x_2 - x_1| = r > 0, \min(x_2 - x_1) \leq c r, \int_{r_{x_1, x_2} \cap E} P \circ h \, d\mu > \frac{1}{k} \mu(r_{x_1, x_2} \cap E) \} \). Then, \( I \) covers \( \tilde{A}_{n, k} \) in the sense of Vitali with index \( c \). By Vitali’s Lemma 12.78, there exists pairwise disjoint rectangles \( \{ r_{x_{i,1,1,2}} \}_{i=1}^m \subseteq I \) with \( m \in \mathbb{Z}_+ \) such that \( \mu_{\tilde{A}_{n, k}}(\tilde{A}_{n, k} \setminus \left( \bigcup_{i=1}^m r_{x_{i,1,1,2}} \right)) < \epsilon_0 \). Let \( B = (\bigcup_{i=1}^m r_{x_{i,1,1,2}}) \cap E \in \mathcal{B} \). Then, \( \mu(B) = \mu_{\tilde{A}_{n, k}}(B \cap \tilde{A}_{n, k}) = \mu_{\tilde{A}_{n, k}}(\tilde{A}_{n, k} \setminus B) > (3k - 1)\epsilon_0 \), where the first inequality follows from the monotonicity of outer measures; and the second equality follows from the measurability of \( B \).

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 \( D_i F : U_i \rightarrow B(\mathbb{K}, \mathcal{Y}) \) 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 \( F^{(1)} : U \rightarrow B(\mathbb{K}^m, \mathcal{Y}) \) 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 \( \mathcal{X} := \mathbb{K}, D := \pi_1(\Omega) \in \mathcal{B}_B(\mathbb{R}) \), which is an open set. \( \forall s \in D \), clearly \( \text{span}(A_{D}(s)) = \mathbb{K} = \mathcal{X} \). Then, \( D_i F : U_i \rightarrow B(\mathbb{K}, \mathcal{Y}) \), where \( U_i := \text{dom}(D_i F) \subseteq \Omega \). Clearly, \( B(\mathbb{K}, \mathcal{Y}) \) 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** \( U_i = \bigcap_{n=1}^{\infty} \bigcup_{\delta > 0} \bigcup_{0 < |h_1| < \delta} \bigcup_{0 < |h_2| < \delta} \bigcup_{h_2 \in \mathcal{Q}} U_{i,n,h_1,h_2} := U_i \subseteq \mathcal{B}_B(\mathbb{R}^m) \), where \( U_{i,n,h_1,h_2} := \{ x \in \Omega \mid \| (F(x + he_{m,i}) - F(x))/h_1 - (F(x + he_{m,i}) - F(x))/h_2 \| < 1/n \}, \forall n \in \mathbb{N}, \forall h_1, h_2 \in \mathbb{Q} \) 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, \( \forall n \in \mathbb{N}, \exists \delta > 0, \exists \mathcal{Q} \) with \( \delta > 0 \), such that \( x + he_{m,i} \in \Omega, \forall h \in \mathcal{B}_R(0, \delta) \), and, \( \forall h_1, h_2 \in \mathcal{B}_R(0, \delta) \) \( \{0\} \), we have \( \| (F(x + he_{m,i}) - F(x))/h \| < 1/n \).
By Propositions 7.23, 11.38, and 11.39, and \( \delta > 0 \), such that, \( \forall h \in \mathcal{B}_\mathbb{R}(0, \delta) \), \( x + h e_{m,i} \in \Omega \), and \( h_1, h_2 \in \mathbb{Q} \) with \( 0 < |h_1| < \delta \) and \( 0 < |h_2| < \delta \), we have \( \| F(x + h_1 e_{m,i}) - F(x) - F(x + h_2 e_{m,i}) - F(x)/h_2 \| < 1/n \). 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, \( \forall h \in \mathcal{B}_\mathbb{R}(0, \delta) \), \( x + h e_{m,i} \in \Omega \), and \( h_1, h_2 \in \mathbb{Q} \) with \( 0 < |h_1| < \delta \) and \( 0 < |h_2| < \delta \), we have \( \| (F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2 \| < 1/n \). \( \forall h_1, h_2 \in \mathcal{B}_\mathbb{R}(0, \delta) \setminus \{0\} \), \( \exists (h_j)_{j=1}^\infty \subseteq (\mathbb{Q} \cap \mathcal{B}_\mathbb{R}(0, \delta)) \setminus \{0\} \) such that \( h_j < h_{j+1} \), \( \forall j \in \mathbb{N} \), and \( \lim_{j \to \infty} h_j = h, j = 1, 2, \forall \ell \in \mathbb{N} \), \( \| (F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2 \| < 1/n \). Then,

\[
\| (F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2 \| = \lim_{i \to \infty} \left\| \frac{(F(x + h_1 e_{m,i}) - F(x))}{h_1} - \frac{(F(x + h_2 e_{m,i}) - F(x))}{h_2} \right\| = \lim_{i \to \infty} \left\| \frac{(F(x + h_1 e_{m,i}) - F(x))/h_1 - (F(x + h_2 e_{m,i}) - F(x))/h_2}{h_1} \right\| \leq 1/n
\]

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, \( U_{i,n,h_1,h_2} \in \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \), \( \forall n \in \mathbb{N} \), \( \forall h_1, h_2 \in \mathbb{Q} \) with \( h_1 \neq 0 \) and \( h_2 \neq 0 \). Then, \( U_i = \bar{U}_{i} \in \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \).

\( \forall n \in \mathbb{N} \), define \( \mathcal{H}_n : \Omega \to \mathcal{Y} \) by, \( \forall x \in \Omega \),

\[
\mathcal{H}_n(x) = \left\{ \begin{array}{ll}
\frac{n}{\vartheta}(F(x + \frac{1}{n} e_{m,i}) - F(x)) & x \in O := \Omega \cap (\Omega - \frac{1}{n} e_{m,i}) \\
0 & x \in \Omega \setminus O
\end{array} \right.
\]

By Propositions 7.23, 11.38, 11.39, and 11.41, \( \mathcal{H}_n \) is \( \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \)-measurable. Then, \( \mathcal{H}_n|_{U_i} \) is \( \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \)-measurable. Note that, \( \forall x \in U_i, D_1 F(x) = \lim_{n \to \mathbb{R}^m} \mathcal{H}_n(x) = \lim_{n \to \mathbb{R}^m} \mathcal{H}_n|_{U_i}(x) \). By Proposition 11.48, \( D_1 F \) is \( \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \)-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 \( \text{span}(\mathcal{A}_\Omega(x)) = \mathbb{K}^m \).

Claim 12.81.2 \( U = \bigcap_{n=1}^{\infty} \bigcup_{\delta > 0} \bigcap_{0 < |h_1| < \delta} \bigcap_{m \in \mathbb{Q}} \bigcap_{|h_m| < \delta} V_{n,h_1,\ldots,h_m} =: \bar{V} \in \mathcal{B}_\mathbb{R}(\mathbb{R}^m) \),

where \( V_{n,h_1,\ldots,h_m} := \{ x \in \bigcap_{i=1}^{m} U_i \bigcap_{|h_j| < \delta} \left\| F(x + \sum_{i=1}^{m} h_i e_{m,i}) - F(x) - \sum_{i=1}^{m} h_i D_i F(x) \right\| \leq \frac{1}{n} \sqrt{\sum_{i=1}^{m} h_i^2} \} \), \( \forall n \in \mathbb{N} \), \( \forall h_1, \ldots, h_m \in \mathbb{Q} \).
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 \( \left[ D_1F(x) \cdots D_mF(x) \right] := L \in B(\mathbb{K}^m, \mathbb{Y}). \) By \( x \in \text{dom}(F^{(1)}), \) we have \( \forall n \in \mathbb{N}, \exists \delta \in \mathbb{Q} \) with \( \delta > 0, \) such that \( B_{\mathbb{R}^m}(x, \sqrt{m\delta}) \subseteq \Omega, \) and \( \forall h \in B_{\mathbb{R}^m}(0, \sqrt{m\delta}), \) 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 \( \|F(x + h) - F(x) - Lh\| = \|F(x + \sum_{i=1}^{m} h_ie_{m,i}) - F(x) - \sum_{i=1}^{m} h_i D_i F(x)\| \leq \frac{1}{n} \|h\| = \frac{1}{n} \sqrt{\sum_{i=1}^{m} h_i^2}. \) 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, \( |h_i| < \delta, \) \( i = 1, \ldots, m. \) Then, \( \exists \{h_{i,j}\}_{j=1}^{\infty} \subseteq \mathbb{Q} \) with \( |h_{i,j}| < \delta \) such that \( \bar{h}_i < h_{i,j}, \forall \bar{v} \in \mathbb{N}, \) and \( \lim_{i \in \mathbb{N}} h_{i,j} = h_i, i = 1, \ldots, m. \) By the assumption,

\[
\|F(\bar{x}) - F(x) - L(\bar{x} - x)\| = \left\| \lim_{i \in \mathbb{N}} \left( F(x + \sum_{i=1}^{m} h_{i,j}e_{m,i}) - F(x) - \sum_{i=1}^{m} h_{i,j} D_i F(x) \right) \right\| = \lim_{i \in \mathbb{N}} \left\| F(x + \sum_{i=1}^{m} h_{i,j}e_{m,i}) - F(x) - \sum_{i=1}^{m} h_{i,j} D_i F(x) \right\| \leq \lim_{i \in \mathbb{N}} \frac{1}{n} \left( \sum_{i=1}^{m} h_{i,j}^2 \right)^{1/2} = \frac{1}{n} \left( \sum_{i=1}^{m} h_i^2 \right)^{1/2} = \frac{1}{n} |\bar{x} - x|,
\]

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, \( x \in \text{dom}(F^{(1)}) = U. \) By the arbitrariness of \( x, \) we have \( V \subseteq U. \) Therefore, \( U = V. \)

By Propositions 7.21, 7.23, 11.38, and 11.39, \( V_{n,h_1,\ldots,h_m} \subseteq B_B(\mathbb{R}^m), \) \( \forall n \in \mathbb{N}, \forall h_1, \ldots, h_m \in \mathbb{Q}. \) Then, \( U = V \subseteq B_B(\mathbb{R}^m). \) \( \square \)

Note that, \( \forall x \in U, F^{(1)}(x) = [D_1F(x) \cdots D_mF(x)] = \left[ (D_1F)|_U(x) \cdots (D_mF)|_U(x) \right]. \) By Propositions 11.38, 11.39, 11.41, and 11.139, \( F^{(1)} \) is \( B_B(\mathbb{R}^m) \)-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 \( \mathbb{K}, \) \( f : I \rightarrow \mathbb{Y} \) be absolutely integrable over \( I \) with respect to \( \mu_s, \) and \( F : I \rightarrow \mathbb{X} \) be defined by \( F(x) = \int_x^a f(t) \, dt, \forall x \in I. \) 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 \( F^{(1)}. \))
Proof. Clearly, $F$ is well-defined by Proposition 11.92. Note that span$(A_1(x_0)) = K$. For all $\epsilon \in (0, \infty) \subset \mathbb{R}$, 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,

$$\|F(x) - F(x_0) - f(x_0)(x - x_0)\|
= \left\| \int_a^x f(t) \, dt - \int_{x_0}^x f(t) \, dt - \int_{x_0}^x f(x_0) \, dt \right\|
= \left\| \int_{x_0}^x f(t) \, dt - \int_{x_0}^x f(x_0) \, dt \right\|
= \left\| \int_{x_0}^x (f(t) - f(x_0)) \, dt \right\|
\leq \epsilon \mu_B(r_{x_0} \land x, x_0 \lor x) = \epsilon |x - x_0|$$

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. \qed

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 $\mathbb{R}$, $\mathbb{Y}$ be a Banach space over $K$, and $F : I \to \mathbb{Y}$ be $C_1$. (Note that, when $K = C$, $I$ is viewed as a subset of $C$ in calculations of $F(1)$. Then, $F(1) : I \to \mathbb{Y}$ is absolutely integrable over $I$ and $F(b) - F(a) = \int_a^b F(1)(t) \, dt = \int_a^b F(1) \, d\mu_B$.)

Proof. Since $F(1) : I \to \mathbb{Y}$ 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, $F(1) : I \to N := \text{span}(F(1)(I)) \subseteq \mathbb{Y}$ and $N$ is a separable Banach subspace of $\mathbb{Y}$. Define $F_a : I \to N \subseteq \mathbb{Y}$ by $F_a(x) = \int_a^x F(1)(t) \, dt = \int_a^x F(1) \, d\mu_B, \forall x \in I$. 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 $g : I \to \mathbb{Y}$ by $g = F - F_a$. Then, by Proposition 9.15, $g$ is Fréchet differentiable and $g^{(1)}(x) = \partial_y, \forall x \in I$. 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, $g(b) = F(a) = F(b) - F_a(b) = F(b) - \int_a^b F(1)(t) \, dt$. Therefore, $F(b) - F(a) = \int_a^b F(1)(t) \, dt$. This completes the proof of the theorem. \qed

Proposition 12.84 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, $X := (\mathcal{P}(\Omega), |\cdot|, B, \mu)$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$, $\mathbb{Y}$ be a finite-dimensional Banach space over $K$, $f : \mathcal{P}(\Omega) \to \mathbb{Y}$ be $\mathcal{B}_B(\mathbb{R}^m)$-measurable, $J := \{1, \ldots, m\}$, and $x_0 := (x_{0,1}, \ldots, x_{0,m}) \in \Omega$. Then, we may define the $\sigma$-finite $\mathbb{Y}$-valued measure $\nu$ with kernel $f$ over $X$ according to Proposition 11.116. By Proposition 11.167, $f$ is the unique Radon-Nikodym derivative of $\nu$ with respect
to μ. Assume that, ∀x₁, x₂ ∈ Ω with x₁ ≤ x₂ and rₓ₁,x₂ ⊆ Ω, we have f is bounded over rₓ₁,x₂ ⊆ P(Ω). Let F : Ω → Y be a cumulative distribution function of ν. Then, ∀x := (x₁, ..., xₘ) ∈ Ω with rₓ,x ⊆ Ω, where ½ := x₀ ∧ x and ½ := x₀ ∨ x, we have

\[ F(x) = \int_{x₀,1}^{x₁} \cdots \int_{x₀,m}^{xₘ} f(s₁, ..., sₘ) \, dsₘ \cdots ds₁ - \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J)} F(x_J) \tag{12.1} \]

and the order of integration can be arbitrary, where \( x_J \in rₓ,x \subseteq Ω \) is defined by \( π_i(x_J) = \begin{cases} π_i(x₀) = x₀,i & i \in J \\ π_i(x) = x_i & i \in J \setminus J, \forall i \in J \end{cases} \).

**Proof** \ Y is separable and is σ-compact, since it is finite-dimensional. Fix any \( x \in Ω \) with rₓ,x ⊆ Ω. Then, rₓ,x ⊆ P(Ω). By Definition 12.42 and Definition 11.166, we have \( \Delta_F(rₓ,x) = ν(rₓ,x) = \int_{rₓ,x} f \, dμ = \int_{rₓ,x} f \, dμ_{Bm} \).

By the assumption and Fubini’s Theorem 12.31, we have

\[ \int_{rₓ,x} f \, dμ_{Bm} = \int_{x₀,1 \wedge x₁}^{x₁} \cdots \int_{x₀,m \wedge xₘ}^{xₘ} f(s₁, ..., sₘ) \, dsₘ \cdots ds₁ = (−1)^{n(x)} \int_{x₀,1}^{x₁} \cdots \int_{x₀,m}^{xₘ} f(s₁, ..., sₘ) \, dsₘ \cdots ds₁ = Δ_F(rₓ,x) \]

and the order of integration can be arbitrary, where \( n(x) = \text{card} \{ i ∈ J | π_i(x₀) > π_i(x) \} =: \text{card}(J) \). By Definition 12.41,

\[ Δ_F(rₓ,x) = \sum_{J \subseteq J} (-1)^{\text{card}(J)} F(x_J) \]

\[ = (-1)^{n(x)} \int_{x₀,1}^{x₁} \cdots \int_{x₀,m}^{xₘ} f(s₁, ..., sₘ) \, dsₘ \cdots ds₁ \]

where \( x_J \in rₓ,x \subseteq ℝ^m \) is defined by, ∀i ∈ J,

\[ π_i(x_J) = \begin{cases} π_i(½) & ∀i ∈ J \\ π_i(x) & ∀i ∈ J \setminus J \end{cases} \]

\[ = \begin{cases} π_i(x) & ∀i ∈ (J \cap J) ∪ ((J \setminus J) \cap (J \setminus J)) \\ π_i(x₀) & \text{otherwise} \end{cases} \]

\[ = \begin{cases} π_i(x) & ∀i ∈ J \setminus (J ∧ J) \\ π_i(x₀) & ∀i ∈ J ∧ J \end{cases} = π_i(x_{J \triangle J}) \]

This leads to

\[ Δ_F(rₓ,x) = \sum_{J \subseteq J} (-1)^{\text{card}(J)} F(x_{J \triangle J}) \]
= \sum_{J \subseteq J} (-1)^{\text{card}(J)} \sum_{J \subseteq J} (-1)^{\text{card}(J) + \text{card}(J)} F(x_{J \Delta J})

= (-1)^{n(x)} \sum_{J \subseteq J} (-1)^{\text{card}(J \Delta J)} F(x_{J \Delta 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

Then, we can conclude

\[ \sum_{J \subseteq J} (-1)^{\text{card}(J)} F(x_J) = \int_{x_{0,1}}^{x_{1}} \cdots \int_{x_{0,m}}^{x_{m}} f(s_1, \ldots, s_m) \, ds_m \cdots ds_1 \]

The above equation is equivalent to (12.1). This completes the proof of the proposition. \( \square \)

**Proposition 12.85** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_\mathbb{R}(\Omega) \) be an open rectangle, \( \mathcal{X} := ((\Omega, | \cdot |), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathcal{Y} \) be a separable Banach space, \( f : \Omega \to \mathcal{Y} \) be \( \mathcal{B} \)-measurable, \( x_0 \in \Omega \), and \( J := \{1, \ldots, m\} \). Assume that \( f \) is absolutely integrable over \( r_{x_0, x_b} \) with respect to \( \mu \), \( x_a \leq x_b \). Define \( \pi_J : \mathbb{R}^m \to \mathbb{R}^{\text{card}(J)} \) by \( \pi_J(x) = (\pi_i(x))_{i \in J} \), \( \forall x \in \mathbb{R}^m \), \( \forall J \subseteq J ; (\pi_J(\Omega), \mathcal{B}_J, \mu_J) \) be the measure subspace of \( \mathbb{R}^{\text{card}(J)} \), \( \forall J \subseteq J \) with \( J \neq \emptyset ; M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \to \mathbb{R}^m \) by \( M_J(s, t) \in \mathbb{R}^m \) is such that \( \pi_J(M_J(s, t)) = s \) and \( \pi_{J \setminus J}(M_J(s, t)) = t \), \( \forall s \in \mathbb{R}^{\text{card}(J)} \), \( \forall t \in \mathbb{R}^{m-\text{card}(J)} \), \( \forall J \subseteq J \). Then, the following statements hold.

(i) \( \forall \emptyset \neq J \subseteq \mathcal{J}, \exists U_J \in \mathcal{B} \) with \( \mu(\Omega \setminus U_J) = 0 \) such that \( f(M_J(\cdot, \pi_{J \setminus J}\,(\cdot))) : \pi_J(\Omega) \to \mathcal{Y} \) is absolutely integrable over \( \pi_J(r_{x, x}) \) with respect to \( \mu_J \), \( \forall x \in U_J \); and \( F : \Omega \to \mathcal{Y} \), defined by, \( \forall x \in \Omega \), \( F(x) = \left\{ \begin{array}{ll}
(-1)^{\text{card}(J \cap J)} \int_{\pi_{J \setminus J}(x)} f(M_J(s, \pi_{J \setminus J}\,(x))) \, d\mu_J(s) & x \in U_J \\
\emptyset & x \in \Omega \setminus U_J
\end{array} \right. , \)
where \( \hat{x} := x_0 \land x, \hat{x} := x_0 \lor x, \) and \( \mathcal{J} := \{ i \in J \mid \pi_i(x_0) > \pi_i(x) \} \), is \( \mathcal{B} \)-measurable.

(ii) Furthermore, if, in addition, \( \mathcal{Y} \) is finite dimensional then \( U_J = \{ x \in \Omega \mid f(M_J(\cdot, \pi_{J \setminus J}\,(\cdot))) : \pi_J(\Omega) \to \mathcal{Y} \) is absolutely integrable over \( \pi_J(r_{x, x}) \) with respect to \( \mu_J \} \).

**Proof** First consider the special case when \( \mu(\Omega) < \infty \) and \( f \) is absolutely integrable over \( \mathcal{X} \). Since \( \Omega \) is a nonempty open rectangle, then \( \Omega \) is bounded. Define \( \hat{f} : \pi_J(\Omega) \times \Omega \to \mathcal{Y} \) by \( \hat{f}(s, x) = (-1)^{\text{card}(J \cap J)} f(M_J(s, \pi_{J \setminus J}\,(x)))) \chi_{\pi_{J \setminus J}(x)}(s), \forall (s, x) \in \pi_J(\Omega) \times \Omega \). By Propositions 7.23, 11.39, and 11.38, \( \hat{f} \) is \( \mathcal{B}_\mathcal{Y}(\mathbb{R}^{\text{card}(J) + m}) \)-measurable. Note
the following derivation,

\[
\int_{\pi_j(\Omega) \times \Omega} \mathcal{P} \circ \tilde{f}(s, x) \, d(\mu_j \times \mu)(s, x)
\leq \int_{\pi_j(\Omega)} \int_{\Omega} \| f(M_j(s, \pi_{J \setminus j}(x))) \| \, d\mu(x) \, d\mu_j(s)
= \int_{\pi_j(\Omega)} \int_{\pi_{J \setminus j}(\Omega)} \int_{\pi_j(\Omega)} \| f(M_j(s, t)) \| \, d\mu_j(t) \, d\mu_j(s)
\leq \mu_j(\pi_j(\Omega)) \int_{\pi_j(\Omega)} \int_{\pi_{J \setminus j}(\Omega)} \| f(x) \| \, d\mu_j(\pi_{J \setminus j}(x)) \, d\mu_j(\pi_j(x))
= \mu_j(\pi_j(\Omega)) \int_{\Omega} \mathcal{P} \circ f \, d\mu < \infty
\]

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 \( \tilde{f} \) is absolutely integrable over \( \pi_j(\Omega) \times \Omega \) with respect to \( \mu_j \times \mu \). By Fubini’s Theorem 12.30, \( \exists p_j : \Omega \to \mathbb{Y} \) and \( U_j \in \mathcal{B} \) such that \( \mu(\Omega \setminus U_j) = 0 \); \( \tilde{f}(\cdot, x) : \pi_j(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( (\pi_j(\Omega), \mathcal{B}_{j, \mu_j}) \), \( \forall x \in U_j \); \( p_j : \Omega \to \mathbb{Y} \), defined by, \( \forall x \in \Omega \), \( p_j(x) = \{ \int_{\pi_j(\Omega)} \tilde{f}(s, x) \, d\mu_j(s) \mid x \in U_j \} \) is \( \mathcal{B} \)-measurable; \( p_j \) is absolutely integrable over \( X \), and \( \int_{\pi_j(\Omega) \times \Omega} \tilde{f}(s, x) \, d(\mu_j \times \mu)(s, x) = \int_{\Omega} p_j(x) \, d\mu(x) \). Clearly, \( F(x) = p_j(x) \), \( \forall x \in \Omega \). Hence, \( F \) is \( \mathcal{B} \)-measurable. Then, \( f(M_j(\cdot, \pi_{J \setminus j}(x))) : \pi_j(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_j(\Omega) \subseteq \pi_j(\Omega) \) with respect to \( \mu_j \), \( \forall x \in U_j \).

If, in addition, \( \mathbb{Y} \) is finite dimensional, let \( U_j \) be defined as in the statement (ii). \( \forall x \in \Omega \), \( f(M_j(\cdot, \pi_{J \setminus j}(x))) : \pi_j(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_j(\mathbb{X}) \) with respect to \( \mu_j \) if, and only if, \( \tilde{f}(\cdot, x) \) is absolutely integrable over \( \pi_j(\Omega) \subseteq \pi_j(\Omega) \). Then, \( U_j := \{ x \in \Omega \mid f(\cdot, x) : \pi_j(\Omega) \to \mathbb{Y} \) 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 \( (\hat{x}_j)_{j=1}^{\infty} \subseteq \Omega \) such that \( \cdots < \hat{x}_j < \cdots < \hat{x}_1 < x_0 < \hat{x}_1 < \cdots < \hat{x}_j < \cdots \), and \( \Omega = \bigcup_{j=1}^{\infty} E_j := \bigcup_{j=1}^{\infty} \hat{x}_j \). 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 \( f(M_j(\cdot, \pi_{J \setminus j}(x))) : \pi_j(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_j(\mathbb{X}) \) with respect to \( \mu_j \), \( \forall x \in U_{j,j} \); and \( F_j : E_j \to \mathbb{Y} \), defined by, \( \forall x \in E_j \),

\[
F_j(x) = \left\{ \begin{array}{ll}
(-1)^{\text{card}(J \setminus j)} \int_{\pi_j(\mathbb{X})} f(M_j(s, \pi_{J \setminus j}(x))) \, d\mu_j(s) & x \in U_{j,j} \\
\phi_y & x \in E_j \setminus U_{j,j}
\end{array} \right.
\]
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 \( x \in U_J; F(x) = \nu_y, \forall x \in \Omega \setminus U_J \). By Proposition 11.41, \( F \) is \( \mathcal{B} \)-measurable. Clearly, \( f(M_J(\cdot, \pi_J(x))) : \pi_J(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_J(r_{x,y}) \subseteq \pi_J(\Omega) \) with respect to \( \mu_J, \forall x \in U_J \).

If, in addition, \( \mathbb{Y} \) is finite dimensional, then, by the special case, \( U_{j,j} = \{ x \in E_j \mid f(M_J(\cdot, \pi_J(x))) : \pi_J(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_J(r_{x,y}) \) with respect to \( \mu_J \}, \forall j \in \mathbb{N} \). Then, \( U_J = \bigcup_{j=1}^{\infty} U_{j,j} = \{ x \in \Omega \mid f(M_J(\cdot, \pi_J(x))) : \pi_J(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_J(r_{x,y}) \) with respect to \( \mu_J \}. \) This completes the proof of the proposition.

\[ \square \]

**Theorem 12.86 (Fundamental Theorem of Calculus I)** Let \( m \in \mathbb{N} \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega \in \mathcal{B}_X(\mathbb{R}^m) \) be an open rectangle, \( \mathbb{X} := ((\Omega, \| \|), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( \mathbb{Y} \) be a separable Banach space over \( \mathbb{K} \), \( f : \Omega \to \mathbb{Y} \) be \( \mathcal{B} \)-measurable, \( \hat{J} := \{ 1, \ldots , m \} \), and \( x_0 \in \Omega \). Assume that \( f \) is absolutely integrable over \( r_{x_a,x_b} \) with respect to \( \mu \), \( \forall x_a, x_b \in \Omega \) with \( x_a \leq x_b \). Let \( \nu \) be the \( \sigma \)-finite \( \mathbb{Y} \)-valued measure with kernel \( f \) over \( \mathbb{X} \) according to Proposition 11.116. Define \( F : \Omega \to \mathbb{Y} \) 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 \( f = \frac{d\mu}{d\nu} \) a.e. in \( \mathbb{X} \). Let \( D_iF : U_i \to \mathbb{Y}, \forall i \in \hat{J}, \) 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 \( \pi_i : \mathbb{R}^m \to \mathbb{R}^{m-1} \) by \( \pi_i(x) = (\pi_j(x))_{j \in \hat{J} \setminus \{i\}} \), \( \forall x \in \mathbb{R}^m \); also define \( M_i : \mathbb{R} \times \mathbb{R}^{m-1} \to \mathbb{R}^m \) by \( \pi_i(M_i(x,y)) = x \) and \( \pi_j(M_i(x,y)) = y, \forall x \in \mathbb{R}, \forall y \in \mathbb{R}^{m-1} \). Then, the following statements hold.

(i) \( F(x) = (-1)^{\text{card}(\hat{J} \setminus \{i\})} \int_{r_{x_a,x_b}} f \, d\mu, \forall x \in \Omega \), where \( \hat{x} := x_0 \wedge x, \hat{x} := x_0 \lor x, \) and \( J := \{ i \in \hat{J} \mid \pi_i(x_0) > \pi_i(x) \} \).

(ii) \( U_i \in \mathcal{B}, \mu(\Omega \setminus U_i) = 0, D_iF \) is \( \mathcal{B} \)-measurable; \( \exists \tilde{U}_i \in \mathcal{B} \) with \( \mu(\Omega \setminus \tilde{U}_i) = 0, \forall x \in \tilde{U}_i, f(M_i(\pi_i(x), \cdot)) : \pi_i(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_i(r_{x,y}) \) with respect to \( \mu_i \), and \( \exists p_i : \Omega \to \mathbb{Y} \) defined by, \( \forall x \in \Omega, \)

\[
p_i(x) = \begin{cases} (-1)^{\text{card}(\hat{J} \setminus \{i\})} \int_{\pi_i(r_{x,y})} f(M_i(\pi_i(x), s)) \, d\mu_i(s) & x \in \tilde{U}_i \\ \nu_y & x \in \Omega \setminus \tilde{U}_i \end{cases}
\]

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 \( D_iF(x) = p_i(x) \) a.e. \( x \in \mathbb{X} \), \( \forall i \in \hat{J} \).

(iii) If, in addition, \( \mathbb{Y} \) is finite dimensional, then \( \tilde{U}_i = \{ x \in \Omega \mid f(M_i(\pi_i(x), \cdot)) : \pi_i(\Omega) \to \mathbb{Y} \) is absolutely integrable over \( \pi_i(r_{x,y}) \) with respect to \( \mu_i \).
Proof. By Proposition 12.75, \( F \) is absolutely continuous. By Definition 12.41 and Definition 11.166, we have \( \Delta F(x) = \nu(x) = \int_{x^-}^x f \mu \). Since \( F \) is the cumulative distribution function of \( \nu \) with origin at \( x_0 \), then 
\[
\Delta F(x) = (-1)^\text{card}(J) F(x).
\]
Thus, we have \( F(x) = (-1)^\text{card}(J) \int_{x^-}^x f \mu, \forall x \in \Omega \). This establishes (i).

(ii) By Proposition 12.81, \( D_i F \) is \( \mathcal{B} \)-measurable, \( \forall i \in \mathcal{I} \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( m = 1 \); Case 2: \( m > 1 \). Case 1: \( m = 1 \). Then, \( i = 1 \). \( F(x) = (-1)^\text{card}(J) \int_{x^-}^x f \mu, \forall x \in \Omega \). Since \( \Omega \) is a nonempty open interval, then \( \exists (E_j)_{j=1}^\infty \subseteq \mathcal{O}_\mathbb{R} \) 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 \( E_j := (E_j, \mathcal{B}_j, \mu_j) \) 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 \( A_j := \{ x \in E_j \mid x \text{ is a rectangular Lebesgue point of } f |_{E_j} \text{ with regularity } c \} \).

By Proposition 12.80, \( A_j \in \mathcal{B}_L \) and \( \mu_L(E_j \setminus A_j) = 0 \). Fix any \( x \in A_j \).

We have \( \limsup_{r \to 0^+} \sup_{r_{x_1,x_2} \subseteq \mathbb{R} \text{ with } \frac{1}{2}(x_1+x_2) = x, \frac{|x_2-x_1|}{|x_2-x_1|} \geq r} \left| \int_{r_{x_1,x_2} \subseteq \mathbb{R}} f(y) - f(x) \right| \mu(y) = 0. \forall \epsilon \in (0,\infty) \subset \mathbb{R}, \exists r_0 \in (0,\infty) \subset \mathbb{R}, \forall r \in (0,r_0) \subset \mathbb{R}, \text{ we have } 0 \leq \sup_{r_{x_1,x_2} \subseteq \mathbb{R} \text{ with } \frac{1}{2}(x_1+x_2) = x, \frac{|x_2-x_1|}{|x_2-x_1|} \geq r} \left| \int_{r_{x_1,x_2} \subseteq \mathbb{R}} f(y) - f(x) \right| \mu(y) < \epsilon. \text{ Let } r_{x_1,x_2} \subseteq \mathbb{R} \text{ be such that } \frac{1}{2}(x_1+x_2) = x, x_2-x_1 < r_0/2, \text{ and } r_{x_1,x_2} \subseteq E_j, \forall x \in r_{x_1,x_2} \text{ with } x \neq x, \text{ we have, where } h := x - x,
\]
\[
0 \leq \left| \int_{x}^{x+h} f(y) - f(x) \right| \mu(y) = \left| \int_{x}^{x+h} f(y) \mu(y) - \int_{x}^{x+h} f(x) \mu(y) \right| = \left| \int_{x}^{x+h} (f(y) - f(x)) \mu(y) \right| = \left| \int_{x \wedge (x+h), x \vee (x+h)} (f(y) - f(x)) \mu(y) \right| \leq \int_{x \wedge (x+h), x \vee (x+h)} \left| f(y) - f(x) \right| \mu(y) = 2 \frac{|h|}{r_B(x_- - |h|, x_+ + |h|)} \int_{x \wedge (x-h), x \vee (x+h)} \left| f(y) - f(x) \right| \mu_B(y) \leq 2 \epsilon |x - x|
\]
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, \( x \in \text{dom}(D_i F) = U_1 \) and \( D_i F(x) = f(x) \). 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 $0 \leq \mu(\Omega \setminus U_1) = \mu_B(\Omega \setminus U_1) = \mu_L(\Omega \setminus U_1) \leq \mu_L(\bigcup_{j=1}^{\infty} (E_j \setminus (\bigcup_{j=1}^{\infty} A_j))) \leq \mu_L(\bigcup_{j=1}^{\infty} E_j) \leq 0$. Hence, $\mu(\Omega \setminus U_1) = 0$. By $D_1 F$ being $\mathcal{B}$-measurable and $f$ being $\mathcal{B}$-measurable, we have $D_1 F(x) = f(x)$ a.e. $x \in \mathbb{X}$. This case is proved.

Case 2: $m > 1$. Without loss of generality, consider $i = 1$. Let $\mathbb{X}_1 := (\pi_1(\Omega), \mathcal{B}, \hat{\mu})$ be the measure subspace of $\mathbb{R}$. By Proposition 12.85, $\exists \bar{U}_1 \in \mathcal{B}$ with $\mu(\Omega \setminus \bar{U}_1) = 0$, and $p_1$ as defined in the statement of the theorem, such that $p_1$ is $\mathcal{B}$-measurable. By Fubini’s Theorem 12.30, we have $F(x) = \int_{\pi_1(\Omega)} p_1(M_1(s, \pi_1(x))) \, d\mu(s), \forall x \in \Omega$. Then, by Case 1, we have $D_1 F(x) = p_1(x)$ a.e. $\pi_1(x) \in \mathbb{X}_1, \forall \pi_1(x) \in \pi_1(\Omega)$. Since $D_1 F$ and $p_1$ are $\mathcal{B}$-measurable and $\mathbb{Y}$ is separable, then, by Propositions 7.23, 11.39, and 11.38, $D_1 F - p_1$ is $\mathcal{B}$-measurable. Then, $\hat{U}_1 := \{x \in \Omega \mid D_1 F(x) = p_1(x)\} \in \mathcal{B}$. By Lemma 12.27, $\mu(\Omega \setminus \hat{U}_1) =: \mu(E) = \int_{\pi_1(\Omega)} \hat{\mu}(E_{\pi_1(x)}) \, d\mu_1(\pi_1(x)) = \int_{\pi_1(\Omega)} 0 \, d\hat{\mu}_1 = 0$, where $E_{\pi_1(x)} := \{s \in \pi_1(\Omega) \mid M_1(s, \pi_1(x)) \in E\}$, and $\hat{\mu}(E_{\pi_1(x)}) = 0$ follows from $D_1 F(x) = p_1(x)$ a.e. $\pi_1(x) \in \mathbb{X}_1, \forall \pi_1(x) \in \pi_1(\Omega)$. Hence, $D_1 F = p_1$ a.e. in $\mathbb{X}$, and $U_1 \in \mathcal{B}$ 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. □

Proposition 12.87 Let $m \in \mathbb{N}$, $\mathbb{R}^m$ be endowed with the usual positive cone, $\Omega \in \mathcal{B}_B(\mathbb{R}^m)$ be region, $\mathbb{Z}$ be a normed linear space, $G : \Omega \to \mathbb{Z}$, $\bar{J} := \{1, \ldots, m\}$, and $x_0 \in \Omega$. $\forall J \subseteq \bar{J}$, define $\pi_J : \mathbb{R}^m \to \mathbb{R}^{\text{card}(J)}$ by $\pi_J(x) = (\pi_j(x))_{j \in \bar{J}}, \forall x \in \mathbb{R}^m$; define $M_J : \mathbb{R}^{\text{card}(J)} \times \mathbb{R}^{m-\text{card}(J)} \to \mathbb{R}^m$ by $M_J(s, t) \in \mathbb{R}^m$ such that $\pi_J(M_J(s, t)) = s$ and $\pi_{\bar{J} \setminus J}(M_J(s, t)) = t$, $\forall s \in \mathbb{R}^{\text{card}(J)}, \forall t \in \mathbb{R}^{m-\text{card}(J)}$; also define $G_J : \pi_J(\Omega) \to \mathbb{Y}$ by $G_J(s) = G(M_J(s, \pi_{\bar{J} \setminus J}(x_0))), \forall s \in \pi_J(\Omega)$. Then, we have, $\forall x \in \Omega$,

$$(-1)^{\text{card}(J)} \Delta_{\bar{G}}(x_\bar{x}, \bar{x}) = \Delta_{\bar{G}}(x_\bar{x}, \bar{x}) - \sum_{J \subset \bar{J}, J \neq \emptyset} (-1)^{\text{card}(J \setminus J)} \Delta_{\bar{G}_J}(\pi_J(x_\bar{x}, \bar{x})) \quad (12.2)$$

where $\bar{x} := x \wedge x_0$, $\bar{x} := x \vee x_0$; and $\bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\}$.

Proof We will prove the result using mathematical induction on $m$.

1° $m = 1$. This is obvious.

2° Assume the claim holds for $m \leq k \in \mathbb{N}$.

3° Consider the case when $m = k + 1 \in \{2, 3, \ldots\}$. Fix any $x \in \Omega$. Define $\bar{G} : \pi_J(\Omega) \to \mathbb{Z}$ by $\bar{G}(s) = G(M_J(s, \pi_{k+1}(x))), \forall s \in \pi_J(\Omega)$ and $\bar{G}_J : \pi_J(\Omega) \to \mathbb{Y}$ in terms of $\bar{G}$ in a similar manner as
$G_J$ in terms of $G$, $\forall J \subseteq \bar{J} \setminus \{k + 1\}$. Then, we have the following sequence of arguments.

$(-1)^{\text{card}(J)} \Delta_G(r_{\bar{x}, \hat{x}})$

$= (-1)^{\text{card}(J \setminus \{k + 1\})} \left( \Delta_G(\pi_{J \setminus \{k + 1\}}(r_{\bar{x}, \hat{x}})) - \Delta_G(\pi_{J \setminus \{k + 1\}}(r_{\bar{x}, \hat{x}})) \right)$

$= G(\pi_{J \setminus \{k + 1\}}(x)) - G(\pi_{J \setminus \{k + 1\}}(x_0))$

$- \sum_{J \subseteq J \setminus \{k + 1\}, J \neq \emptyset} (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G(\pi_J(r_{\bar{x}, \hat{x}}))$

$- (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G(\pi_{J \setminus \{k + 1\}}(r_{\bar{x}, \hat{x}}))$

$= G(x) - G_{\{k + 1\}}(\pi_{k + 1}(x)) - (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G(\pi_{J \setminus \{k + 1\}}(r_{\bar{x}, \hat{x}}))$

$- \sum_{J \subseteq J \setminus \{k + 1\}, J \neq \emptyset} (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G(\pi_J(r_{\bar{x}, \hat{x}}))$

$= G(x) - G_{\{k + 1\}}(\pi_{k + 1}(x))$

$- \sum_{J \subseteq J \setminus \{k + 1\}, J \neq \emptyset} (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G(\pi_J(r_{\bar{x}, \hat{x}}))$

$- \sum_{J \subseteq J, J \cap \{k + 1\} \neq \emptyset} (-1)^{\text{card}(J \cap \{k + 1\})} \Delta_G_J(\pi_J(r_{\bar{x}, \hat{x}}))$

$= G(x) - G_{\{k + 1\}}(\pi_{k + 1}(x))$

$- \sum_{J \subseteq J \setminus \{k + 1\}, J \neq \emptyset} (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_G_J(\pi_J(r_{\bar{x}, \hat{x}}))$

$+ (-1)^{\text{card}(J \setminus \{k + 1\})} \Delta_{G(\{k + 1\})}(\pi_{\{k + 1\}}(r_{\bar{x}, \hat{x}}))$

$= G(x) - G_{\{k + 1\}}(\pi_{k + 1}(x))$

$- \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cup \{k + 1\})} \Delta_G_J(\pi_J(r_{\bar{x}, \hat{x}}))$

$+ (-1)^{\text{card}(J \cup \{k + 1\})} \Delta_{G(\{k + 1\})}(\pi_{\{k + 1\}}(r_{\bar{x}, \hat{x}}))$

$= G(x) - G(x_0) - \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cup \{k + 1\})} \Delta_G_J(\pi_J(r_{\bar{x}, \hat{x}}))$

where the first equality follows from Definition 12.41; the second equality follows from the inductive assumption $2^2$; 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, \( \Omega \in \mathcal{B}_{\mathbb{R}}(\mathbb{R}^m) \) be an open rectangle, \( X := ((\Omega, | \cdot |), \mathcal{B}, \mu) \) be the \( \sigma \)-finite metric measure subspace of \( \mathbb{R}^m \), \( Y \) be a separable reflexive Banach space over \( K \) with \( Y^* \) being separable, \( F : \Omega \to Y \) be absolutely continuous, \( \bar{J} := \{1, \ldots, m\} \), \( F(1) : U \to B(\mathbb{R}^m, Y) \), \( D_i F : U_i \to Y, \forall i \in \bar{J} \), 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 \( \pi_j : \mathbb{R}^m \to \mathbb{R}^{|\bar{J}\setminus J|} \) by \( \pi_j(x) = (\pi_i(x))_{i \in \bar{J}, \forall x \in \mathbb{R}^m} \); define \( M_j : \mathbb{R}^{|\bar{J}\setminus J|} \times \mathbb{R}^{m-|\bar{J}\setminus J|} \to \mathbb{R}^m \) by \( M_j(s, t) \in \mathbb{R}^m \) is such that \( \pi_j(M_j(s, t)) = s \) and \( \pi_j \downarrow (M_j(s, t)) = t \), \( \forall s \in \mathbb{R}^{|\bar{J}\setminus J|}, \forall t \in \mathbb{R}^{m-|\bar{J}\setminus J|} \); also define \( F_j : \pi_j(\Omega) \to Y \) by \( F_j(s) = F(M_j(s, \pi_j^{-1}(x_0))) \), \( \forall s \in \pi_j(\Omega) \); and let \( \mathbb{X}_j := (\pi_j(\Omega), \mathcal{B}, \mu_j) \) be the measure subspace of \( \mathbb{R}^{|\bar{J}\setminus J|} \) when \( J \neq \emptyset \). Then, the following statements hold.

(i) There exists a set of functions \( f_{J,x_0} : \pi_j(\Omega) \to Y \), which is \( \mathcal{B}_{\pi_j} \)-measurable, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), such that \( f_{J,x_0} \)'s are absolutely integrable on \( \pi_j(r_{x_0,x_b}) \) with respect to \( \mu_j \), \( \forall x_a, x_b \in \Omega \) with \( x_a \triangleq x_b \), and we have, \( \forall \bar{x} \in \Omega \),

\[
F(x) = \sum_{J \subseteq \bar{J}, J \neq \emptyset} (-1)^{|\bar{J}\setminus J|} \int_{\pi_j(r_{\bar{x},\bar{x}})} f_{J,x_0} \, d\mu_j + F(x_0) \tag{12.3}
\]

where \( \bar{x} := x \setminus x_0 \), \( \hat{x} := x \setminus x_0 \); and \( \bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(x)\} \).

The set of functions \( f_{J,x_0} \) are \( \mathcal{B}_{\pi_j} \)-measurable, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), is unique in the sense that if \( \{g_{J,x_0} \} \subseteq \bar{J}, J \neq \emptyset \) is another set of such functions satisfying (12.3), then \( f_{J,x_0} = g_{J,x_0} \) a.e. in \( X_j \), \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \).

(ii) If we expand absolutely continuous function \( F \) at \( \bar{x} \in \Omega \), instead of \( x_0 \), then we will have a set of functions \( f_{J,\bar{x}} : \pi_j(\Omega) \to Y \), which is \( \mathcal{B}_{\pi_j} \)-measurable, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), satisfying the conclusion of (i) with \( x_0 \) replaced by \( \bar{x} \). The \( f_{J,\bar{x}} \)'s relate to \( f_{J,x_0} \)'s according to the following formula, \( \forall J \subseteq \bar{J} \) with \( J \neq \emptyset \), \( \exists U_j \in \mathcal{B}_{\bar{J}} \) with \( \mu_j(\pi_j(\Omega) \setminus U_j) = 0 \), such that

\[
f_{J,\bar{x}}(s) = \sum_{J \subseteq \bar{J}} (-1)^{|\bar{J}\setminus J|} \int_{\pi_j(r_{\bar{x},\bar{x}})} f_{J,x_0}(\pi_j(M_j(s, \pi_j \downarrow (t)))) \, d\mu_j(\pi_j \downarrow (t)) + f_{J,x_0}(s), \quad \forall s \in U_j \tag{12.4}
\]

and \( f_{J,\bar{x}} = \emptyset_Y \), \( \forall s \in \pi_j(\Omega) \setminus U_j \), where \( \bar{x} := \bar{x} \setminus x_0 \), \( \tilde{x} := \bar{x} \setminus x_0 \), and \( \bar{J} := \{i \in \bar{J} \mid \pi_i(x_0) > \pi_i(\bar{x})\} \); 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, $\forall i \in J$, $D_i F(x) = f_{(i)}(x,\pi_{(i)}(x))$ a.e. $x \in X$.

(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 $F^{(1)}(x) = [D_{1}F(x) \ldots D_{m}F(x)]$. If, in addition, $\forall J \subseteq J$ with $J \neq \emptyset$, $\forall \delta(J,x) \in (0,\infty) \subseteq \mathbb{R}$, $\exists (J,x) \in [0,\infty) \subseteq \mathbb{R}$ such that $\operatorname{ess sup}_{y \in \mathcal{B}_{RM}^\prime}(y,\delta(J,x)) \| f_{J,x}(\pi_{J}(y)) \| \leq c(J,x)$ and $\mathcal{B}_{RM}^\prime (x,\delta(J,x)) \subseteq \Omega$, then $\mu(\Omega \setminus U) = 0$.

(v) If, in addition, $Y$ is finite dimensional, then $\bar{U}_J = \left\{ s \in \pi_{J}(\Omega) \mid f_{J,x}(\pi_{J}(M_{J}(s,\pi_{J \setminus J}(t)))) \text{ is absolutely integrable over } \pi_{J \setminus J}(r_{x,z}) \right\}$ with respect to $\mu_{J \setminus J}(\pi_{J \setminus J}(t))$, $\forall J \subseteq J$ with $J \neq \emptyset$.

**Proof**

(i) By $F = F_{J}$ being absolutely continuous and Definition 12.41, $\forall J \subseteq J$ 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 $Y$ 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 $X_{J}$. $\forall x_{a},x_{b} \in \Omega$ with $x_{a} \leq x_{b}$, we have $\mathcal{P} \circ \nu_{J}(\pi_{J}(r_{x_{a},x_{b}})) = \int_{\pi_{J}(r_{x_{a},x_{b}})} \mathcal{P} \circ f_{J,x_{0}} d\mu_{J} = T_{F_{J}}(\pi_{J}(r_{x_{a},x_{b}})) < \infty$, 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 $\pi_{J}(r_{x_{a},x_{b}})$ with respect to $\mu_{J}$. Then, $\forall x \in \Omega$, we have $\Delta_{F_{J}}(\pi_{J}(r_{x,z})) = \nu_{J}(\pi_{J}(r_{x,z})) = \int_{\pi_{J}(r_{x,z})} f_{J,x_{0}}(s) d\mu_{J}(s)$.

By Proposition 12.87, we have

$(-1)^{\operatorname{card}(J)} \Delta_{F_{J}}(r_{x,z}) = F(x) - F(x_{0}) - \sum_{J \subseteq J} (-1)^{\operatorname{card}(J \cap J)} \Delta_{F_{J}}(\pi_{J}(r_{x,z}))$

which implies that

$(-1)^{\operatorname{card}(J)} \int_{\pi_{J}(r_{x,z})} f_{J,x_{0}}(s) d\mu_{J}(s)

= F(x) - F(x_{0}) - \sum_{J \subseteq J} (-1)^{\operatorname{card}(J \cap J)} \int_{\pi_{J}(r_{x,z})} f_{J,x_{0}}(s) d\mu_{J}(s)$

Reorganizing the above formula, we have

$F(x) = F(x_{0}) + \sum_{J \subseteq J} (-1)^{\operatorname{card}(J \cap J)} \int_{\pi_{J}(r_{x,z})} f_{J,x_{0}}(s) d\mu_{J}(s)$
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 $F_i$’s and the uniqueness of Radon-Nikodym derivatives. Hence, (i) holds.

(ii) By (i), there exists a set of functions $f_{J,x} : \pi_J(\Omega) \to \mathbb{Y}$, which is $\mathcal{B}$-measurable, $\forall J \subseteq \tilde{J}$ with $\tilde{J} \neq \emptyset$, such that $f_{J,x}$’s are absolutely integrable on $\pi_J(r_{x_a,x_b})$ with respect to $\mu_J$, $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$, and we have

$$ F(x) = \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J = \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J $$

This establishes (12.3). The uniqueness of the set of functions $(f_{J,x_0})_{J \subseteq \tilde{J}, \tilde{J} \neq \emptyset}$ is clear from the uniqueness of $F_i$’s and the uniqueness of Radon-Nikodym derivatives. Hence, (i) holds.

(ii) By (i), there exists a set of functions $f_{J,x} : \pi_J(\Omega) \to \mathbb{Y}$, which is $\mathcal{B}$-measurable, $\forall J \subseteq \tilde{J}$ with $\tilde{J} \neq \emptyset$, such that $f_{J,x}$’s are absolutely integrable on $\pi_J(r_{x_a,x_b})$ with respect to $\mu_J$, $\forall x_a, x_b \in \Omega$ with $x_a \leq x_b$, and we have

$$ F(x) = \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J = \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J $$

Rearrange terms in the above, we obtain

$$ \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J = \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J $$

$$ f_{J,x_0} \, d\mu_J - \sum_{J \subseteq \tilde{J}, J \neq \emptyset} (-1)^{\text{card}(J \cap \tilde{J})} \int_{\pi_J(r_{\tilde{J},x})} f_{J,x_0} \, d\mu_J $$

$$ \sum_{J \subseteq \tilde{J}, J \neq \emptyset} \prod_{J \subseteq \tilde{J}, j \in J, j \neq \emptyset} \left( \int_{\pi_J(x_0)} \pi_J(x) \right) f_{J,x_0}(s_1, \ldots, s_{\text{card}(J)}) \, d\mu_B(s_{\text{card}(J)}) \cdots \, d\mu_B(s_1) $$

$$ - \sum_{J \subseteq \tilde{J}, J \neq \emptyset} \prod_{J \subseteq \tilde{J}, j \in J, j \neq \emptyset} \left( \int_{\pi_J(x_0)} \pi_J(x) \right) f_{J,x_0}(s_1, \ldots, s_{\text{card}(J)}) \, d\mu_B(s_{\text{card}(J)}) \cdots \, d\mu_B(s_1) $$

$$ \sum_{J \subseteq \tilde{J}, J \neq \emptyset} \prod_{J \subseteq \tilde{J}, j \in J, j \neq \emptyset} \left( \int_{\pi_J(x_0)} \pi_J(x) \right) f_{J,x_0}(s_1, \ldots, s_{\text{card}(J)}) $$

$$ \cdots \, d\mu_B(s_{\text{card}(J)}) \cdots \, d\mu_B(s_1) $$

$$ f_{J,x_0}(s_1, \ldots, s_{\text{card}(J)}) \, d\mu_B(s_{\text{card}(J)}) \cdots \, d\mu_B(s_1) $$
\[
\sum_{J \subseteq J, J \neq \emptyset} \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(\tilde{J} \cap J)} \pi_j \int_{\tilde{J} \cap J} \frac{\pi_j(\tilde{x})}{\pi_j(x)} \ prod_j \left( \int_{\pi_j(x)} \right) dJ_{j, x_0} (s_1, \ldots, s_{\text{card}(J)}) d\mu_B(s_{\text{card}(J)}) \ldots d\mu_B(s_1)
\]

\[
\frac{1}{\prod_{J \neq \emptyset} J \cap J} \sum_{J \subseteq J, J \neq \emptyset} \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(\tilde{J} \cap J)} \pi_j \int_{\tilde{J} \cap J} \frac{\pi_j(\tilde{x})}{\pi_j(x)} \ prod_j \left( \int_{\pi_j(x)} \right) dJ_{j, x_0} d\mu_j
\]

\[
\frac{1}{\prod_{J \neq \emptyset} J \cap J} \sum_{J \subseteq J, J \neq \emptyset} \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(\tilde{J} \cap J)} \pi_j \int_{\tilde{J} \cap J} \frac{\pi_j(\tilde{x})}{\pi_j(x)} \ prod_j \left( \int_{\pi_j(x)} \right) dJ_{j, x_0} d\mu_j
\]

\[
\frac{1}{\prod_{J \neq \emptyset} J \cap J} \sum_{J \subseteq J, J \neq \emptyset} \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(\tilde{J} \cap J)} \pi_j \int_{\tilde{J} \cap J} \frac{\pi_j(\tilde{x})}{\pi_j(x)} \ prod_j \left( \int_{\pi_j(x)} \right) dJ_{j, x_0} d\mu_j
\]

where \( \tilde{x}_{j, x, x, x_0} := x_{1, J, x, x, x_0} \land x_{2, J, x, x, x_0} \); \( \tilde{x}_{j, x, x, x_0} := x_{1, J, x, x, x_0} \lor x_{2, J, x, x, x_0} \); \( x_{1, J, x, x, x_0} := M_j(\pi_j(x), \pi_j(x), \tilde{x}_{j, x, x, x_0}, \tilde{x}_{j, x, x, x_0}) \); \( x_{2, J, x, x, x_0} := M_j(\pi_j(x), \pi_j(x), \tilde{x}_{j, x, x, x_0}, \tilde{x}_{j, x, x, x_0}) \); 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 \((-1)^{\text{card}(\tilde{J} \cap J)} \int_{\tilde{J} \cap J} d\mu_j \) into pieces along the intermediate point \( \tilde{x} \); the sixth equality follows from the idea that the corresponding integral will be zero if its domain of integration is an empty set, which will happen since \( \pi_j(x) = \pi_j(x) \); \( \forall i \in \tilde{J} \setminus J \); the seventh equality follows from the interchange of summing variables \( J \) and \( \tilde{J} \); the ninth equality follows from Fubini’s Theorem 12.30, and \( U_{j, j} \in B_j \) with \( \mu_j(\pi_j(\Omega) \setminus \tilde{U}_{j, j}) = 0 \), and \( p_{j, j} : \pi_j(\Omega) \to y, \forall \tilde{J} \subseteq J \), defined by \( p_{j, j}(s) := \left\{ \begin{array}{ll} \int_{\tilde{U}_{j, j}} f_j_{j, x_0} (s, \pi_j(M_j(s, \pi_j(\tilde{x}_{j, x, x, x_0})))) \ d\mu_j_{j, j}(\pi_j(\tilde{x}_{j, x, x, x_0})) & s \in \tilde{U}_{j, j} \\
\partial_y & s \in \pi_j(\Omega) \setminus \tilde{U}_{j, j} \end{array} \right. \)
is $B_J$-measurable; and the tenth equality follows from Proposition 11.92.
Let $\bar{U}_J := \bigcap_{J \subseteq J} \bar{U}_{J,i}$, $\forall J \subseteq J$ with $J \neq \emptyset$. Clearly, $\bar{U}_j \in B_J$ and $\mu_J(\pi(\Omega) \setminus \bar{U}_j) = 0$. The definition of $f_{J,x}$'s (12.4) satisfies the above equalities, $\forall J \subseteq J$ with $J \neq \emptyset$. Then, by (i), this set of functions $(f_{J,x})_{J \subseteq J, J \neq \emptyset}$ is the unique one we seek.

(ii) 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 $\mathbb{X}$. $\forall J \subseteq J$ with $1 \in J$, $\exists U_{J,1} \in B$ with $\mu(\Omega \setminus U_{J,1}) = 0$, $D_1 \left( (-1)^{\text{card}(J \cap \Omega)} \int_{\pi_{J,(x)}} f_{J,x_0} \, d\mu_J \right) : U_{J,1} \to \mathbb{Y}$ is $B$-measurable, $\exists \hat{U}_{J,1} \in B$ with $\mu(\Omega \setminus \hat{U}_{J,1}) = 0$, and $\exists p_{J,1} : \Omega \to \mathbb{Y}$ defined by $p_{J,1} = \left\{ (-1)^{\text{card}(J \cap \Omega \setminus \{1\})} \int_{\pi_{J,(x)}} f_{J,x_0}(\pi_J(\pi_J(\{1\})) d\mu_J_{\{1\}}(\pi_J(\{1\}))) x \in \hat{U}_{J,1}, \forall x \right\} x \in \Omega \setminus \hat{U}_{J,1}$ in $\Omega$, when $J \neq \{1\}$, $p_{J,1} = f_{J,x_0}(\pi_J(x))$, $\forall x \in \Omega$, $\hat{U}_{J,1} = \Omega$, when $J = \{1\}$, such that $p_{J,1}$ is $B$-measurable and $D_1 \left( (-1)^{\text{card}(J \cap \Omega \setminus \{1\})} \int_{\pi_{J,(x)}} f_{J,x_0} \, d\mu_J \right) = p_{J,1}(x)$ a.e. $x \in \mathbb{X}$. $\forall J \subseteq J$ with $1 \notin J$, we have $D_1 \left( (-1)^{\text{card}(J \cap \Omega \setminus \{1\})} \int_{\pi_{J,(x)}} f_{J,x_0} \, d\mu_J \right) = \vartheta_y, \forall x \in \Omega$. In this case, we have $\hat{U}_{J,1} = \hat{U}_J = \Omega$. Let $\bar{U}_J := \bigcap_{J \subseteq J, \emptyset \neq J} \bar{U}_{J,i}$. Then, $\bar{U}_J \in B$ and $\mu(\Omega \setminus \bar{U}_J) = 0$. Hence, $D_1 F(x) = \sum_{J \subseteq J, \emptyset \neq J} p_{J,1}(x) = f_{J,x}(\pi_J(x))$ a.e. $x \in \mathbb{X}$. This implies that $\mu(\Omega \setminus \bar{U}_J) = 0$ and (iii) holds.

(iv) By Proposition 12.81, $U \subseteq \bigcap_{i=1}^m U_i$, $F^{(1)}$ is $B$-measurable, and $\forall x \in U$, we have $F^{(1)}(x) = \left[ D_1 F(x) \cdots D_m F(x) \right]$. All we need to show is that $\mu(\Omega \setminus U) = 0$ under the additional assumption that $f_{J,x_0}$'s are locally bounded.

$\forall i \in J$, $\forall J \subseteq J$ with $i \in J$, by the proof of (iii), $\exists \hat{U}_{J,i} \in B$ with $\mu(\Omega \setminus \hat{U}_{J,i}) = 0$, such that $D_i H_{J,x}(x) := D_i \left( (-1)^{\text{card}(J \cap \Omega \setminus \{i\})} \int_{\pi_{J,(x)}} f_{J,x_0} \, d\mu_J \right)(x) = p_{J,i}(x)$, $\forall x \in \hat{U}_{J,i}$. Then, let $\hat{U}_{J,i} := \hat{U}_{J,i} \cap \hat{U}_J \cap \hat{U}_{J,i} \in B$, we have $\mu(\Omega \setminus \hat{U}_{J,i}) = 0$ and $D_i H_{J,x} = (-1)^{\text{card}(J \cap \Omega \setminus \{i\})} \int_{\pi_{J,(x)}} f_{J,x_0}(\pi_J(\pi_{J,(x)}) d\mu_{J,(x)}(\pi_J(x))) = H_{J,i}(x)$, $\forall x \in \hat{U}_{J,i}$.

Let $\bar{U} := \left( \bigcap_{J \subseteq J} \bigcap_{i \in J \subseteq J} \bar{U}_{J,i} \right) \cap \left( \bigcap_{J \subseteq J, \emptyset \neq J} \bar{U}_{J,i} \right) \in B$. Clearly, $\mu(\Omega \setminus \bar{U}) = 0$. Fix any $\bar{x} \in \bar{U}$. Then, by (ii), there exists a unique set of functions $(f_{J,x})_{J \subseteq J, \emptyset \neq J}$ such that $F(x) = \sum_{J \subseteq J, \emptyset \neq J} (-1)^{\text{card}(J \cap \Omega \setminus \{i\})} \int_{\pi_{J,(x)}} f_{J,x} \, d\mu_J + \tilde{F}(\bar{x})$, $\forall x \in \Omega$. Furthermore, $f_{J,x}$'s are related to $f_{J,x}$'s by (12.4). Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$. By $f_{J,x}$'s being locally bounded on their domain of definition and (12.4), we can conclude that $f_{J,x}$'s are also locally bounded on their domain of definition. $\forall J \subseteq J$ with $J \neq \emptyset$, $\forall x \in \Omega$, $\exists \delta(J,x) \in (0, \infty) \subset \mathbb{R}$, $\exists \tilde{c}(J,x) \in [0, \infty) \subset \mathbb{R}$ such that
Chapter 12. Differentiation and Integration

\[ \text{ess sup}_{y \in B_{\mathbb{R}^m}(x, \delta(J,x))} \| f_i(x, y) \| \leq \bar{c}(J, x) \quad \text{and} \quad B_{\mathbb{R}^m}(x, \delta(J,x)) \subseteq \Omega. \]

Let \( \delta_0 := \min_{J \subseteq J} \bigcup_{J \neq \emptyset} \delta(J, x) \in (0, \infty) \subseteq \mathbb{R} \) and \( \bar{c}_0 := \max_{J \subseteq J} \delta(J, x) \in [0, \infty) \subseteq \mathbb{R} \). Then, \( \text{ess sup}_{y \in B_{\mathbb{R}^m}(x, \delta_0)} \| f_i(x, y) \| \leq \bar{c}_0 \) and \( B_{\mathbb{R}^m}(\bar{x}, \delta_0) \subseteq \Omega \). \( \forall i \in J, \) we have

\[
D_i F(\bar{x}) = f_{(1), x}(\pi_i(\bar{x})) + \sum_{J \subseteq J, \{i\} \subset J} (-1)^{\text{card}(J \cap \{i\})} \int_{\pi_J(x_i)} f_{J,x}(\pi_J(x_i)) \, d\mu_J
\]

where the first equality follows from the preceding discussion and the second equality follows from (12.4). \( \forall i \in J, \forall J \subseteq J \) with \( i \in J, \exists \alpha_{J,i}(\bar{x}) \in (0, \infty) \subseteq \mathbb{R} \) such that \( \| H_J(\bar{x} + h \pi_i e_{m,i}) - H_J(\bar{x}) - H_J(\bar{x})h_i \| \leq m^{-1}2^{-m} | h_i | \). Let \( \delta_1 := \left( \min_{i \in J} \min_{J \subseteq J} \delta_{J,i}(\bar{x}) \right) \wedge \delta_0 \wedge \frac{\bar{c}_1}{2^{m-1}} \wedge 1 \in (0, \infty) \subseteq \mathbb{R} \).

This leads to, \( \forall x \in B_{\mathbb{R}^m}(\bar{x}, \delta_1), \)

\[
\| F(x) - F(\bar{x}) - \sum_{i=1}^{m} D_i F(\bar{x}) \pi_i(x - \bar{x}) \|
\]

\[
= \left\| \sum_{J \subseteq J, J \neq \emptyset} (-1)^{\text{card}(J \cap \{i\})} \int_{\pi_J(x_i)} f_{J,x} \, d\mu_J - \sum_{i=1}^{m} f_{(1), x}(\pi_i(\bar{x})) \pi_i(x - \bar{x}) \right\|
\]

\[
+ \sum_{i=1}^{m} (-1)^{\text{card}(\hat{J} \cap \{i\})} \int_{\pi_i(x_i)} f_{(1), x} \, d\mu_i(t)
\]

\[
- \sum_{i=1}^{m} (-1)^{\text{card}(\hat{J} \cap \{i\})} \int_{\pi_i(x_i)} f_{(1), x} \, d\mu_i(t)
\]

\[
\leq \sum_{J \subseteq J, \text{card}(J) \geq 2} \int_{\pi_J(x_i)} P \circ f_{J,x} \, d\mu_J + \left\| \sum_{i=1}^{m} (-1)^{\text{card}(\hat{J} \cap \{i\})} \int_{\pi_i(x_i)} (f_{(1), x}(t) - f_{(1), x}(\pi_i(\bar{x}))) \, d\mu_i(t) \right\|
\]

\[
\leq \sum_{J \subseteq J, \text{card}(J) \geq 2} \int_{\pi_J(x_i)} P \circ f_{J,x} \, d\mu_J
\]

\[
+ \sum_{i=1}^{m} \sum_{i \in J} \| H_J(\bar{x} + \pi_i(x - \bar{x}) e_{m,i}) - H_J(\bar{x}) - H_J(\bar{x}) \pi_i(x - \bar{x}) \|
\]

\[
\leq \sum_{J \subseteq J, \text{card}(J) \geq 2} \bar{c}_0 \mu_J(\pi_J(x_i)) + \sum_{i=1}^{m} \sum_{i \in J} m^{-1}2^{-m} | \pi_i(x - \bar{x}) | \]
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, $\bar{x} \in \text{dom } (F^{(1)})$. By the arbitrariness of $\bar{x}$, we have $\bar{U} \subseteq \text{dom } (F^{(1)}) = U$. 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. 

\begin{theorem}[Integration by Parts]
Let $I \subseteq \mathbb{R}$ be a nonempty open or closed interval, $\mathbb{I} := (\mathbb{P}(I), \langle \cdot, \cdot \rangle), \mathbb{B}, \mu$ be the $\sigma$-finite metric measure subspace of $\mathbb{R}$, $\mathbb{Y}$ be a separable Banach space over $\mathbb{K}$, $\mathbb{Z}$ be a separable reflexive Banach space over $\mathbb{K}$ with $\mathbb{Z}^*$ being separable, and $\mathbb{W} := \mathbb{B}(\mathbb{Y}, \mathbb{Z})$ be separable, $a : \mathbb{P}(I) \rightarrow \mathbb{W}$ be $\mathbb{B}$-measurable and absolutely integrable over $\mathbb{I}$, $\nu_A$ be the $\mathbb{W}$-valued measure with kernel $a$ over $\mathbb{I}$, and $A : \mathbb{I} \rightarrow \mathbb{W}$ be a cumulative distribution function for $\nu_A$, $f : \mathbb{P}(I) \rightarrow \mathbb{Y}$ be $\mathbb{B}$-measurable and absolutely integrable over $\mathbb{I}$, $\nu_F$ be the $\mathbb{Y}$-valued measure with kernel $f$ over $\mathbb{I}$, and $F : \mathbb{I} \rightarrow \mathbb{Y}$ be a cumulative distribution function for $\nu_F$. Then, $H : \mathbb{I} \rightarrow \mathbb{Z}$ defined by $H(x) = A(x)F(x)$, $\forall x \in \mathbb{I}$, is absolutely continuous and is a cumulative distribution function of $\nu_H$ which is the $\mathbb{Z}$-valued measure with kernel $h : \mathbb{P}(I) \rightarrow \mathbb{Z}$ over $\mathbb{I}$, where $h$ is defined by $h(x) = A(x)f(x) + a(x)F(x)$, $\forall x \in \mathbb{P}(I)$, and $Af, aF,$ and $h$ are absolutely integrable over $\mathbb{I}$. As a consequence, $\forall b, c \in \mathbb{I}$ with $b \leq c$, we have

$$\int_b^c A(x)f(x) \, dx + \int_b^c a(x)F(x) \, dx = A(c)F(c) - A(b)F(b)$$

\begin{proof}
By Proposition 7.66, $\mathbb{W}$ is a separable Banach space. Fix any $x_0 \in \mathbb{I}$. Let $F_0 : \mathbb{I} \rightarrow \mathbb{Y}$ be the cumulative distribution function of $\nu_F$ with origin $x_0$, and $A_0 : \mathbb{I} \rightarrow \mathbb{W}$ 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 $\mathbb{I}$, then $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|F_0(x)\|_Y \leq M$ and $\|A_0(x)\|_W \leq M$, $\forall x \in \mathbb{I}$. By the assumption, we have $F(x) = F(x_0) + F_0(x)$ and $A(x) = A(x_0) + A_0(x)$, $\forall x \in \mathbb{I}$. Then, $F$ and $A$ are absolutely continuous, therefore continuous, and $\exists M \in [0, \infty) \subset \mathbb{R}$ such that $\|F(x)\|_Y \leq M$ and $\|A(x)\|_W \leq M$, $\forall x \in \mathbb{I}$. By Proposition 11.37, $F$ and $A$ are $\mathbb{B}$-measurable. By Propositions 7.23, 7.65, 11.38, and 11.39, $Af, aF,$ and $h$ are $\mathbb{B}$-measurable. Then, $Af, aF$ are absolutely integrable over $\mathbb{I}$. Hence, $h$ is absolutely integrable over $\mathbb{I}$ by Proposition 11.83. By Proposition 12.68, $H$ is absolutely continuous.

$$\sum_{J \subseteq J, \text{card}(J) \geq 2} \bar{c}_0 \prod_{j \in J} |\pi_j(x - \bar{x})| + \frac{\epsilon}{2} |x - \bar{x}| \leq (2^n - 1 - m)\bar{c}_0 \delta_1^{\text{card}(J)-1} |x - \bar{x}| + \frac{\epsilon}{2} |x - \bar{x}| \leq \epsilon |x - \bar{x}|$$
We will first consider the case that $I$ is an open interval. By Fundamental Theorem of Calculus I 12.86, $F^{(1)} = D_1 F : U_F \rightarrow \mathbb{Y}$ and $A^{(1)} = D_1 A : U_A \rightarrow W$ 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 $\mathcal{B}$-measurable, $F^{(1)} = f$ a.e. in $I$, and $A^{(1)} = a$ a.e. in $I$. Then, $\exists \hat{U}_F, \hat{U}_A \in \mathcal{B}$ such that $\mu(I \setminus \hat{U}_F) = 0 = \mu(I \setminus \hat{U}_A)$ and $F^{(1)}(x) = f(x)$, $\forall x \in \hat{U}_F$, and $A^{(1)}(x) = a(x)$, $\forall x \in \hat{U}_A$.

$\forall x \in \hat{U}_F \cap \hat{U}_A$, 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, $H^{(1)} = D_1 H : U_H \rightarrow \mathbb{Z}$ is such that $U_H \in \mathcal{B}$, $\mu(I \setminus U_H) = 0$, and $H^{(1)}$ is $\mathcal{B}$-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, $H(x) = \int_x^\hat{h} d\mu + H(x_0)$, $\forall x \in I$, $\hat{h}$ is $\mathcal{B}$-measurable and absolutely integrable over $r_{x_a, x_b}$, $\forall x_a, x_b \in I$ with $x_a \leq x_b$, and $H^{(1)}(x) = \hat{h}(x)$ a.e. $x \in I$. Thus, by Lemma 11.44, $h = \hat{h}$ a.e. in $I$. Then, by Proposition 11.92, we have and $H(x) - H(x_0) = \int_{x_0}^x h(s) \, ds$, $\forall x \in I$. By Proposition 11.92 and Fact 12.72, $\forall b, c \in I$ with $b \leq c$, we have $A(c)F(c) - A(b)F(b) = H(c) - H(x_0) - (H(b) - H(x_0)) = \int_{x_0}^x h(x) \, dx - \int_{x_0}^b h(x) \, dx = \int_b^c (A(x)f(x) + a(x)F(x)) \, dx + \int_b^c a(x)F(x) \, dx$.

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 $\hat{I}$ to be an open interval such that $I \subseteq \hat{I}$. Define $\hat{f} : \hat{I} \rightarrow \mathbb{Y}$ by $\hat{f}(x) = \begin{cases} f(x) & x \in P(I) \\ \varnothing & x \in I \setminus P(I) \end{cases}$. Similarly define $\hat{a}$ in terms of $a$. Define $\hat{F} : \hat{I} \rightarrow \mathbb{Y}$ by $\hat{F}(x) = \begin{cases} F(x) & x \in I \\ F(\min_{s \in I} s) & x < \min_{s \in I} s \\ F(\max_{s \in I} s) & x > \max_{s \in I} s \end{cases}$. $\hat{A}$. Then, the result can be applied for $\hat{I} := ((\hat{I}, |\cdot|), \hat{\mathcal{B}}, \hat{\mu})$ for the functions $\hat{f}$, $\hat{F}$, $\hat{a}$, $\hat{A}$, 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. □

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 $\mu_{\nu, m}(E) < +\infty$, $\mathcal{I} \subseteq \mathcal{B}(\mathbb{R}^m)$ be a collection of nondegenerate closed rectangles in $\mathbb{R}^m$ that covers $E$ in the sense of Vitali with index $c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R}$, and $\mathcal{I} := \{ V \subseteq \mathbb{R}^m \mid V$ is a nondegenerate rectangle in $\mathbb{R}^m$ with center $x \in \mathbb{R}^m$, such that the rectangle $V(x) = x \in \mathcal{I} \}$, $\widehat{\mathcal{I}}$ be a collection of nondegenerate closed rectangles in $\mathbb{R}^m$ that covers $E$ in the sense of Vitali with index $c$. Then, $\forall c \in (0, \infty) \subset \mathbb{R}$, there exists a countable subcollection $(\mathcal{V}_i)_{i \in N} \subseteq \mathcal{I}$ with $N \subseteq \mathbb{N}$, (each $\mathcal{V}_i$ corresponds to $U_i \in \mathcal{I}$ as outlined in the definition of $\mathcal{I}$) and $\exists n \in \mathbb{Z}$ such that $\{ 1, \ldots, n \} \subseteq N$. 
Note that, for any closed rectangle $\emptyset$ length at least $\mu$, or $\exists$ length at least $\mu$ or $\exists$ length at least $\mu$. By neglecting rectangles in $\tilde{T}$, we may without loss of generality assume that $V \subseteq O$ and the shortest side of $V$ has length at least $c \varepsilon$. 

**Proof** Fix any $\epsilon \in (0, \infty) \subseteq \mathbb{R}$. By Example 12.56, $\mu_{\text{Bm}}(E) = \inf_{(O_i)_{i=1}^\infty \subseteq \mathbb{R}^m, E \subseteq \bigcup_{i=1}^\infty O_i} \sum_{i=1}^\infty 2^{-i} \mu_{\text{Bm}}(O_i) < +\infty$. Then, $\exists O \in \mathcal{O}_{\mathbb{R}^m}$ such that $E \subseteq O$ and $\mu_{\text{Bm}}(O) < +\infty$. By neglecting rectangles in $\tilde{T}$, we may without loss of generality assume that $V \subseteq O$ and the shortest side of $V$ has length at least $c \varepsilon$.

Case 1: $E \subseteq \bigcup_{i=1}^k V_i$. Then, by Proposition 4.10, $d := \inf_{x \in \bigcup_{i=1}^k V_i} |x - x_0| > 0$. Then, by the assumption, $\exists \tilde{V} \subseteq \tilde{T}$ such that $x_0 \in \tilde{V}$ and $\text{dia}(\tilde{V}) < d$ and the shortest side of the rectangle $\tilde{V}$ has length at least $c \varepsilon$. Then, $\tilde{V} \cap (\bigcup_{i=1}^k V_i) = \emptyset$. This shows that $l_k \geq \text{dia}(\tilde{V}) > 0$. Hence, $\exists V_{k+1} \subseteq \tilde{T}$ such that $V_1, \ldots, V_{k+1} \subseteq \tilde{T}$ are pairwise disjoint and $\text{dia}(V_{k+1}) \geq l_k/2$. Inductively, we either have the result holds or $\exists (V_k)_{k=1}^\infty \subseteq \tilde{T}$ such that $V_1, \ldots, V_{k+1} \subseteq \tilde{T}$ are pairwise disjoint.

In the latter case, we have $O \subseteq \bigcup_{i=1}^k V_i$ and $\infty > \mu_{\text{Bm}}(O) \geq \mu_{\text{Bm}}(\bigcup_{i=1}^k V_i) = \sum_{k=1}^\infty \mu_{\text{Bm}}(V_k) \geq \sum_{k=0}^\infty 2^{-k} \mu_{\text{Bm}}(V_k)$. 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 $\sum_{i=1}^n \mu_{\text{Bm}}(V_i) < 5^{-m} \epsilon^m$. Let $R := E \setminus \bigcup_{i=1}^n V_i$. Without loss of generality, assume $\text{dia}(V_k) =: p_k \in (0, \infty) \subseteq \mathbb{R}$ and the center of $V_k$ is $x_k$, $\forall k \in \mathbb{N}$. Then, $\exists \tilde{V} \subseteq \tilde{T}$ such that $\tilde{V} \subseteq \tilde{T}$ and $\text{dia}(\tilde{V}) < d$ and the shortest side of the rectangle $\tilde{V}$ has length at least $c \varepsilon$. Then, $V \cap (\bigcup_{i=1}^n V_i) = \emptyset$. By the fact that $\lim_{k \to \infty} l_k = 0$, we have $\exists \tilde{V} \subseteq \bigcup_{i=1}^n \tilde{V}_i$ such that $V \cap \tilde{V}_i \neq \emptyset$ and $V \cap \tilde{V}_i = \emptyset$, $\forall i \in \{1, \ldots, i_0 - 1\}$. Then, $\text{dia}(V) \leq l_{i_0 - 1} \leq 2 \text{dia}(V_{i_0}) = 2p_{i_0}$.

Let $\tilde{x} \in \bigcup_{i=1}^n \tilde{V}_i$. Then, $|\tilde{x} - x_{i_0}| \leq |\tilde{x} - \tilde{x}| + |\tilde{x} - x_{i_0}| \leq \text{dia}(V) + \frac{2}{5} p_{i_0} \leq \frac{6}{5} p_{i_0}$. Therefore, $x_{i_0} \in \tilde{T} x_{i_0} - \frac{2}{5} p_{i_0}.x_{i_0} + \frac{2}{5} p_{i_0} x_{i_0} \leq \frac{6}{5} \text{dia}(V) + x_{i_0} \subseteq \tilde{V}(V_{i_0} - x_{i_0}) + x_{i_0} = U_{i_0} \subseteq \tilde{T}$. Then, $x_{i_0} \in \bigcup_{k=1}^\infty U_k$. Hence, $R \subseteq \bigcup_{k=1}^\infty U_k$. Then, $\mu_{\text{Bm}}(R) \leq \mu_{\text{Bm}}(\bigcup_{k=1}^\infty U_k) \leq \sum_{k=1}^\infty \mu_{\text{Bm}}(U_k) = \sum_{k=1}^\infty \mu_{\text{Bm}}(V_k) = 5^m e^{-m} \mu_{\text{Bm}}(V_k) = 5^m e^{-m} \sum_{k=1}^\infty \mu_{\text{Bm}}(V_k) < \epsilon$. Note that
\[
E \subseteq \left( \bigcup_{i=1}^{n} V_i \right) \cup \left( \bigcup_{i=n+1}^{\infty} U_i \right) \text{ and } \sum_{i \in N, i > n} \mu_{BM}(U_i) < \epsilon. \text{ Hence, the result holds.}
\]

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, \(\Omega \in \mathcal{O}_{\mathbb{R}^m}\) be an open subset, \(\tilde{\Omega} = (\Omega, \cdot, | \cdot |, B, \tilde{\mu})\) be the \(\sigma\)-finite metric measure subspace of \(\mathbb{R}^m\), and \(\tilde{\mu}\) be the induced measure on \(\tilde{\Omega}\) and \(\mu\) be the induced measure on \((\Omega, B)\) and \(\mu\) be the induced measure on \((\tilde{\Omega}, B)\) as defined in Proposition 12.9. Then, \(\tilde{\mu}\) and \((\tilde{\Omega}, \cdot, | \cdot |, B, \tilde{\mu})\) is homeomorphic isomeric under \(F\), and \(\tilde{\mu}\) and \((\tilde{\Omega}, \cdot, | \cdot |, B, \tilde{\mu})\) is homeomorphic isomeric under \(F\). Assume that \(F^{(1)} : \tilde{\Omega} \to \mathbb{R}^m, \mathbb{R}^m\) with \(\tilde{\mu}(\tilde{\Omega} \setminus \tilde{U}) = 0\) satisfies, \(\forall x, x \in \Omega\) with \(x \equiv x + x\) and \(\tilde{\tau}_{x, x} \subseteq \tilde{\Omega}\), we have \(\sup_{x \in \tilde{\tau}_{x, x} \cap U} |det(F^{(1)}(x))| \leq M_{x, x} \in (0, \infty) \subset \mathbb{R}\), \(F_{x, x}^{(1)} : \tilde{\Omega} \to \mathbb{R}^m, \mathbb{R}^m\) with \(\tilde{\mu}(\tilde{\Omega} \setminus \tilde{U}) = 0\) satisfies, \(\forall x, x \in \Omega\) with \(x \equiv x + x\) and \(\tilde{\tau}_{x, x} \subseteq \tilde{\Omega}\), we have \(\sup_{x \in \tilde{\tau}_{x, x} \cap U} |det(F^{(1)}(x))| \leq M_{x, x} \in (0, \infty) \subset \mathbb{R}\), and \(\tilde{\mu}(\tilde{\Omega} \setminus \tilde{U}) = 0 = \mu(\tilde{\Omega} \setminus \tilde{U}) := \mu(\tilde{\Omega} \setminus F_{x, x}^{(1)}(\tilde{U})).\) Then, \(F_{x, x} \ll \mu\) and \(\frac{dF_{x, x}}{d\mu} =: g_{1} a.e. in \tilde{\mu}, g_{1} \ll \mu\) and \(\frac{dF_{x, x}}{d\mu} = g_{1} a.e. in \tilde{\mu}\), where \(g : \Omega \to (0, \infty) \subset \mathbb{R}\) is given by \(g(\bar{x}) = \left\{ \begin{array}{ll}
|det(F^{(1)}(\bar{x}))| & \bar{x} \in \tilde{U} \\
1 & \bar{x} \in \tilde{\Omega} \setminus \tilde{U} \end{array} \right.\) and \(g_{1}(x) = \frac{1}{g(F(x))}\), \(\forall x \in \Omega\). Furthermore, let \(Y\) be a separable Banach space, \(\forall f : \Omega \to Y\) that is absolutely integrable over \(\tilde{\mu}\), we have

\[
\int_{\tilde{\Omega}} f \, d\tilde{\mu} = \int_{\tilde{\Omega}} f(F(x))g(x) \, d\tilde{\mu}(x)
\]

and the right-hand-side integrand is absolutely integrable over \(\tilde{\mu}\); \(\forall \tilde{f} : \tilde{\Omega} \to Y\) that is absolutely integrable over \(\tilde{\mu}\), we have

\[
\int_{\tilde{\Omega}} \tilde{f} \, d\tilde{\mu} = \int_{\tilde{\Omega}} \tilde{f}(F(x))g_{1}(x) \, d\mu(x)
\]

and the right-hand-side integrand is absolutely integrable over \(\tilde{\mu}\);

**Proof** Fix \(c \in (0, \frac{1}{\sqrt{m}}) \subset \mathbb{R}\). The result is trivial if \(\Omega = \emptyset\). Consider the case \(\tilde{\Omega} \neq \emptyset\). Then, \(\tilde{\Omega} = F_{x, x}(\tilde{\Omega}) \neq \emptyset, \mu(\tilde{\Omega}) > 0, \text{ and } \tilde{\mu}(\tilde{\Omega}) > 0\). First, we need the following intermediate result.

**Claim 12.91.1** \(\forall \bar{x} \in \tilde{U}, \forall \epsilon \in (0, \infty) \subset \mathbb{R}\), \(\exists \delta \in (0, \infty) \subset \mathbb{R}\) such that, \(\forall h \in \mathcal{B}_{\mathbb{R}^m}(0, m, \delta)\) with \(h > 0\) and \(\min h \geq c |h|\), we have \(\mathcal{F}_{x, h}(\tilde{x} - \bar{h}, \tilde{x} + \bar{h}) \subseteq \tilde{\Omega}\) and \(F_{x, h}(\mathcal{F}_{x, h}(\tilde{x} - \bar{h}, \tilde{x} + \bar{h})) \leq (1 + \epsilon)g(\bar{x})\tilde{\mu}(\mathcal{F}_{x, h}(\tilde{x} - \bar{h}, \tilde{x} + \bar{h})).\)
12.6. FUNDAMENTAL THEOREM OF CALCULUS

Symmetrically, \( \forall x \in \tilde{U} = F_i(\tilde{U}) \), \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in B_{\mathbb{R}^{m}}(0_m, \delta) \) with \( h \succ 0_m \) and \( \min h \geq c |h| \), we have \( \tilde{\tau}_x - h, x+h \subseteq \Omega \), and \( F_i \mu(\tilde{\tau}_x - h, x+h) \leq (1 + \epsilon) (g(F(x)))^{-1} \mu(\tilde{\tau}_x - h, x+h) = (1 + \epsilon) g_1(x) \mu(\tilde{\tau}_x - h, x+h) \).

**Proof of claim:** \( \forall \tilde{x} \in \tilde{U} \), we have \( F_i^{(1)}(\tilde{x}) \) exists and \( \det(F_i^{(1)}(\tilde{x})) \neq 0 \). This implies that \( F_i^{(1)}(F_i(\tilde{x})) = (F_i^{(1)}(\tilde{x}))^{-1} \) and \( \det(F_i^{(1)}(F_i(\tilde{x}))) = \left( \det(F_i^{(1)}(\tilde{x})) \right)^{-1} \). By Definition 9.3, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta_1 \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall \tilde{h} \in B_{\mathbb{R}^{m}}(0_m, \delta_1) \), we have \( \tilde{x} + \tilde{h} \in \Omega \) and

\[
\left| F_i(x + \tilde{h}) - F_i(x) - F_i^{(1)}(\tilde{x})\tilde{h} \right| \leq \epsilon |\tilde{h}|
\]

Then, \( \forall \tilde{h} \in B_{\mathbb{R}^{m}}(0_m, \delta) \) with \( \tilde{h} \succ 0_m \) and \( \min \tilde{h} \geq c |\tilde{h}| \), we have \( F_i(\tilde{\tau}_x - h, \tilde{x} + \tilde{h}) \subseteq F_i(\tilde{x}) + F_i^{(1)}(\tilde{x}) \tilde{\tau}_x - h, \tilde{h} \) with \( \tilde{x} \in B_{\mathbb{R}^{m}}(0_m, \epsilon |\tilde{h}|) \). This implies that \( F_i \mu(\tilde{\tau}_x - h, \tilde{x} + \tilde{h}) = \mu(F_i(\tilde{\tau}_x - h, \tilde{x} + \tilde{h})) \leq \mu(B_{\mathbb{R}^{m}}(0_m, \epsilon |\tilde{h}| + B_{\mathbb{R}^{m}}(0_m, \epsilon |\tilde{h}|)) \).

Clearly, \( \mu(B_{\mathbb{R}^{m}}(0_m, \epsilon |\tilde{h}|)) = g(\tilde{x}) \mu(B_{\mathbb{R}^{m}}(\tilde{x}, \tilde{h})) \), which can be proved by decomposing the matrix \( F_i^{(1)}(\tilde{x}) \) into basic row operations, and the equality holds for each row operation. Then, for sufficiently small \( \epsilon > 0 \), we have \( F_i \mu(\tilde{\tau}_x - h, \tilde{x} + \tilde{h}) \leq (1 + \epsilon) g(\tilde{x}) \mu(B_{\mathbb{R}^{m}}(\tilde{x}, \tilde{h})) = (1 + \epsilon) g(\tilde{x}) \mu(\tilde{\tau}_x - h, \tilde{x} + \tilde{h}) \), since \( F_i^{(1)}(\tilde{x}) \) is a closed parallelogram with all of its sides has length of at least some constant proportion of \( |\tilde{h}| \). This proves the first paragraph of the claim.

\( \forall x \in \tilde{U} \), we have \( F(x) \in F(\tilde{U}) = \tilde{U} \). Then, \( F_i^{(1)}(x) = \left( F_i^{(1)}(F_i(x)) \right)^{-1} \) and \( \left| \det(F_i^{(1)}(x)) \right| = \left| \det(F_i^{(1)}(F_i(x))) \right|^{-1} = (g(F(x)))^{-1} \). By symmetry, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in B_{\mathbb{R}^{m}}(0_m, \delta) \) with \( h \succ 0_m \) and \( \min h \geq c |h| \), we have \( \tilde{\tau}_x - h, x+h \subseteq \Omega \), and \( \mu(F(\tilde{\tau}_x - h, x+h)) = F_i \mu(\tilde{\tau}_x - h, x+h) \leq (1 + \epsilon) \mu(\tilde{\tau}_x - h, x+h) = (1 + \epsilon) g(F(x))^{-1} \).

This completes the proof of the claim.

We first show that \( F_i \mu \ll \tilde{\mu} \) under the condition that \( \tilde{\Omega} \) is a nonempty bounded open set and \( \left| \det(F_i^{(1)}(\tilde{x})) \right| \leq M \in (0, \infty) \subset \mathbb{R} \), \( \forall \tilde{x} \in \tilde{U} \). Then, \( 0 < \tilde{\mu}(\tilde{\Omega}) < \infty \). We need the following intermediate result.

**Claim 12.91.2** \( \forall \tilde{\Omega} \in \mathcal{B} \text{ with } \tilde{\Omega} \text{ being open and } \tilde{\mu}(\tilde{\Omega}) > 0 \), we have \( \tilde{\mu}(\tilde{\Omega}) \geq \mu(\tilde{\Omega})/M := \mu(F_i(\tilde{\Omega}))/M \).

**Proof of claim:** Fix any \( \tilde{\Omega} \in \mathcal{B} \) with \( \tilde{\Omega} \) being open and \( \tilde{\mu}(\tilde{\Omega}) > 0 \). Then, \( \tilde{\Omega} \in \mathcal{B} \) is open and \( \forall \tilde{x} \in \tilde{\Omega} := \tilde{\Omega} \cap \tilde{U} \), by Claim 12.91.1, \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists \delta(\tilde{x}, \epsilon) \in (0, \infty) \subset \mathbb{R} \) such that, \( \forall h \in B_{\mathbb{R}^{m}}(0_m, \delta(\tilde{x}, \epsilon)) \) with \( h \succ 0_m \) and \( \min h \geq c |h| \), we have \( \tilde{\tau}_x - h, \tilde{x} + h \subseteq \Omega \) and \( \mu(F(\tilde{\tau}_x - h, \tilde{x} + h)) = F_i \mu(\tilde{\tau}_x - h, \tilde{x} + h) \leq (1 + \epsilon) g(\tilde{x}) \mu(\tilde{\tau}_x - h, \tilde{x} + h) \). The collection of such \( \tilde{\tau}_x - h, \tilde{x} + h \)
is denoted $\mathcal{I}$. We will restrict $\mathcal{I}$ to consist of all of the above mentioned $\mathbb{R}^{m}$'s which are subsets of $\tilde{O}$. Clearly, $\mathcal{I}$ covers $\tilde{O}$ in the sense of Vitali with index $c$. Define $\tilde{I} := \{ V \subseteq \mathbb{R}^{m} \mid V$ is a nondegenerate rectangle in $\mathbb{R}^{m}$ with center $x \in \mathbb{R}^{m}$, such that the rectangle $(\tilde{\omega}(V - x) + x) \in \mathcal{I} \}$. Clearly, $\tilde{I} \subseteq \mathcal{I}$. It is easy to see that $\tilde{I}$ covers $\tilde{O}$ in the sense of Vitali with index $c$. By Vitali’s Lemma 12.90, $\exists N \subseteq \mathbb{N}$, $\exists \{ V_{i} \}_{i \in N} \subseteq \tilde{I}$, $\exists n \in \mathbb{Z}_{+}$ such that $\{ 1, \ldots, n \} \subseteq N$, $\{ V_{1}, \ldots, V_{n} \}$ is pairwise disjoint, $\tilde{\mu}(\tilde{O} \setminus (\bigcup_{i=1}^{n} V_{i})) < \frac{\epsilon}{4} \tilde{\mu}(\tilde{O})$, $\tilde{O} \subseteq (\bigcup_{i=1}^{n} V_{i}) \cup (\bigcup_{i \in N, i > n} U_{i})$, and $\sum_{i \in N, i > n} \tilde{\mu}(U_{i}) \leq \frac{\epsilon}{4} \tilde{\mu}(\tilde{O})$, where $V_{i} = \mathbb{T}_{x_{i} - h_{i}, \bar{x}_{i}, \bar{h}_{i}}, \bar{x}_{i} \in \tilde{I} \subseteq \mathcal{I}$ and $U_{i} = \mathbb{T}(V_{i} - \bar{x}_{i}) + \bar{x}_{i} = \mathbb{T}_{x_{i} - \bar{h}_{i}, \bar{x}_{i}, \bar{h}_{i}}, \bar{h}_{i} \in \tilde{I}, \forall i \in N$. Then, we have the following line of arguments.

$$
\mu(O) = \mu(\tilde{O} \cup (O \setminus \tilde{O})) = \mu(\tilde{O}) + \mu(O \setminus \tilde{O})
$$

$$
= \mu(F_{i}(\tilde{\omega})) \leq \mu \left( F_{i} \left( \left( \bigcup_{i=1}^{n} V_{i} \right) \cup \left( \bigcup_{i \in N, i > n} U_{i} \right) \right) \right)
$$

$$
\leq \sum_{i=1}^{n} \mu(F_{i}(V_{i})) + \sum_{i \in N, i > n} \mu(F_{i}(U_{i})) = \sum_{i=1}^{n} F_{i} \mu(V_{i}) + \sum_{i \in N, i > n} F_{i} \mu(U_{i})
$$

$$
\leq \sum_{i=1}^{n} (1 + \epsilon)g(\bar{x}_{i})\tilde{\mu}(\mathbb{T}_{x_{i} - h_{i}, \bar{x}_{i} + \bar{h}_{i}}) + \sum_{i \in N, i > n} (1 + \epsilon)g(\bar{x}_{i})\tilde{\mu}(\mathbb{T}_{x_{i} - \bar{h}_{i}, \bar{x}_{i} + \bar{h}_{i}})
$$

$$
\leq (1 + \epsilon)\tilde{M} \sum_{i=1}^{n} \tilde{\mu}(V_{i}) + (1 + \epsilon)\tilde{M} \sum_{i \in N, i > n} \tilde{\mu}(U_{i})
$$

$$
\leq (1 + \epsilon)\tilde{M} \tilde{\mu}(\bigcup_{i=1}^{n} V_{i}) + (1 + \epsilon)\tilde{M} \frac{\epsilon}{4} \tilde{\mu}(\tilde{O})
$$

$$
\leq (1 + \epsilon)\tilde{M} \tilde{\mu}(\tilde{O}) + (1 + \epsilon)\tilde{M} \frac{\epsilon}{4} \tilde{\mu}(\tilde{O}) = (1 + \epsilon) \frac{\epsilon}{4} \tilde{M} \tilde{\mu}(\tilde{O})
$$

where $\tilde{O} := O \cap \tilde{U} = F_{i}(\tilde{O} \cap \tilde{U}) = F_{i}(\tilde{O})$; the third equality follows from the fact that $O \setminus \tilde{O} \subseteq \Omega \setminus \tilde{U}$; 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 $\tilde{I}$; the fifth inequality follows from the countable additivity of measure and from the preceding discussion; and the last inequality follows from the monotonicity of measure. Since $\epsilon > 0$ is arbitrary, then we have $\mu(O) \leq \tilde{M} \tilde{\mu}(\tilde{O})$.

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

$\forall E \in \mathcal{B}$ with $\tilde{\mu}(E) = 0$, $\forall \epsilon > 0$, By the fact that $\tilde{X}$ is a metric measure space, there exists $\tilde{O} \in \mathcal{B}$ with $E \subseteq \tilde{O}$ and $\tilde{O}$ being open, such that $0 < \tilde{\mu}(\tilde{O}) < \epsilon/\tilde{M}$. Then, $F_{i} \tilde{\mu}(E) = \mu(F_{i}(\tilde{E})) \leq \mu(F_{i}(\tilde{O})) \leq \tilde{M} \tilde{\mu}(\tilde{O}) < \epsilon$. By
the arbitrariness of $\epsilon$, we have $F_\mu(\tilde{E}) = 0$. By the arbitrariness of $\tilde{E}$, we have $F_\mu \ll \tilde{\mu}$. This proves that $F_\mu \ll \tilde{\mu}$ in the special case.

Now, consider the general case as stipulated in the theorem statement. Since $\bar{\Omega}$ is an nonempty open set and $\mathbb{R}^m$ is second countable, then \( \exists (\hat{x}_i)_{i=1}^\infty, (\hat{\hat{x}}_i)_{i=1}^\infty \subseteq \bar{\Omega} \) such that $\hat{x}_i < \hat{\hat{x}}_i$ and $\mathbf{r}_{\hat{x}_i, \hat{\hat{x}}_i} \subseteq \bar{\Omega}, \forall i \in \mathbb{N}$, and $\bar{\Omega} = \bigcup_{i=1}^\infty \mathbf{r}_{\hat{x}_i, \hat{\hat{x}}_i}$.

By symmetry, $F_{\mu, \bar{\Omega}} \ll \mu$.

By Radon-Nikodym Theorem 11.169, \( \frac{dF_\mu}{d\mu} =: \tilde{g} \) exists and is unique almost everywhere on $\bar{\mathbb{X}}$, and $\frac{dF_\mu}{d\mu} =: \tilde{\tilde{g}}$ exists and is unique almost everywhere on $\mathbb{X}$, where $\tilde{g} : \bar{\Omega} \to [0, \infty) \subset \mathbb{R}$ is $\mathcal{B}$-measurable (or simply $\mathcal{B}_\mathbb{R}(\mathbb{R}^m)$-measurable); and $\tilde{\tilde{g}} : \Omega \to [0, \infty) \subset \mathbb{R}$ is $\mathcal{B}$-measurable (or simply $\mathcal{B}_\mathbb{R}(\mathbb{R}^m)$-measurable). By Definitions 12.7 and 12.8 and Proposition 12.168, it is clear that $\tilde{g}(x) = \frac{1}{\tilde{g}(F(x))}$ a.e. $x \in \mathbb{X}$ and $\tilde{\tilde{g}}(x) = \frac{1}{\tilde{\tilde{g}}(F(x))}$ a.e. $\tilde{x} \in \bar{\mathbb{X}}$. Then, by changing $\tilde{g}$ on a set of measure zero in $\tilde{\mu}$ and changing $\tilde{\tilde{g}}$ on a set of measure zero in $\mu$, we have $\tilde{g}(x) = \frac{1}{\tilde{g}(F(x))}$, $\forall x \in \bar{\Omega}$.

By Propositions 11.92, 12.11, and 11.168, we have (12.5) holds with $g$ substituted by $\tilde{g}$ and $g_1$ substituted by $\tilde{g}$. We will only need to show that $\tilde{g} = g$ a.e. in $\bar{\mathbb{X}}$. Then, the theorem is proved.

First, consider the special case when $\Omega$ is bounded. Then, $\chi_{\Omega, \Omega}$ is absolutely integrable over $\mathbb{X}$. By the preceding paragraph, we have $\int_{\bar{\Omega}} \chi_{\Omega, \Omega} \, d\mu = \int_{\bar{\Omega}} \chi_{\Omega, \Omega} (F(x)) \tilde{g}(\tilde{x}) \, d\tilde{\mu}(\tilde{x}) = \int_{\Omega} \chi_{\Omega, \Omega}(x) \tilde{g}(x) \, d\tilde{\mu}(x)$.

Claim 12.91.3 $\tilde{B} \supseteq \tilde{\tilde{A}} \cap \tilde{\tilde{U}}$.

Proof of claim: Fix any $\tilde{x} \in \tilde{\tilde{A}} \cap \tilde{\tilde{U}}$. Fix any $\epsilon \in (0, \infty) \subset \mathbb{R}$.

By Claim 12.91.1, $\exists \delta \in (0, \infty) \subset \mathbb{R}$ such that $\forall \tilde{h} \in \mathcal{B}_{\mathbb{R}^m}(0_m, \delta)$ with $\tilde{h} > 0_m$ and $\min \tilde{h} \geq c \vert \tilde{h} \vert$, we have $\mathbf{r}_{\tilde{x} - \tilde{h}, \tilde{x} + \tilde{h}} \subseteq \tilde{\tilde{A}}$ and $F_\mu(\mathbf{r}_{\tilde{x} - \tilde{h}, \tilde{x} + \tilde{h}})$.

By Definition 12.79, $\exists \tau_0 \in (0, \infty) \subset \mathbb{R}$ such that $\forall \tau \in (0, \tau_0)$, we have $\mathcal{B}_{\mathbb{R}^m}(\tilde{x}, \tau) \subseteq \tilde{\tilde{U}}$ and $\sup_{\tilde{h} \in \mathbb{R}^m, \tilde{h} > 0_m} \frac{1}{\tilde{\mu}(\mathbf{r}_{\tilde{x} - \tilde{h}, \tilde{x} + \tilde{h}})} \int_{\mathbf{r}_{\tilde{x} - \tilde{h}, \tilde{x} + \tilde{h}}} \vert \tilde{g}(y) - \tilde{\tilde{g}}(\tilde{x}) \vert \, d\tilde{\mu}(y) < \epsilon$.
Take \( \tilde{h} \in \mathbb{R}^m \) with \( |\tilde{h}| < r_0 \wedge \delta, \tilde{h} > 0, \) and \( \min \tilde{h} \geq c|\tilde{h}|, \) we have

\[
\hat{g}(\tilde{x}) - \epsilon \leq \hat{g}(\tilde{x}) - \frac{1}{\mu(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} \int_{\tilde{r}_{\tilde{x}-h,\tilde{x}+h}} |\hat{g}(\tilde{y}) - \hat{g}(\tilde{x})| \, d\tilde{\mu}(\tilde{y})
\]

\[
\leq \hat{g}(\tilde{x}) - \frac{1}{\tilde{\mu}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} \int_{\tilde{r}_{\tilde{x}-h,\tilde{x}+h}} \hat{g}(\tilde{y}) - \hat{g}(\tilde{x}) \, d\tilde{\mu}(\tilde{y})
\]

\[
\leq \frac{1}{\tilde{\mu}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} \int_{\tilde{r}_{\tilde{x}-h,\tilde{x}+h}} \hat{g} \, d\tilde{\mu} = \frac{1}{\tilde{\mu}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} \int_{F}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h}) 1 \, d\mu
\]

\[
= \frac{\mu(F_i(\tilde{r}_{\tilde{x}-h,\tilde{x}+h}))}{\tilde{\mu}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} = \frac{F_i \mu(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})}{\tilde{\mu}(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})} \leq \frac{F_i \mu(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})}{\mu(\tilde{r}_{\tilde{x}-h,\tilde{x}+h})}
\]

\[
\leq (1 + \epsilon)\hat{g}(\tilde{x})
\]

where the first inequality follows from the fact \( |\tilde{h}| < r_0; \) the second inequality follows from Proposition 11.92; the third inequality follows from basic algebra; the first equality follows from (12.5a) 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 \( |\tilde{h}| < \delta. \) By the arbitrariness of \( \epsilon, \) we have \( \hat{g}(\tilde{x}) \leq g(\tilde{x}) \) and \( \tilde{x} \in \tilde{B}. \) By the arbitrariness of \( \tilde{x}, \) we have \( \tilde{A} \cap \tilde{U} \subseteq \tilde{B}. \)

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

By Claim 12.91.3, we have \( 0 \leq \hat{\mu}(\tilde{A} \setminus \tilde{B}) \leq \mu_{\text{Leb}}(\tilde{A} \setminus \tilde{U}) \leq \mu_{\text{Leb}}(\tilde{A} \setminus \tilde{\Omega}) = 0. \) Thus, \( \hat{\mu}(\tilde{A} \setminus \tilde{B}) = 0. \)

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 \( \tilde{G} := \tilde{B} \cap F(C), \) we have \( \hat{g}(\tilde{x}) \leq g(\tilde{x}) \) and \( \hat{g}(F_i(\tilde{x})) \leq g_1(F_i(\tilde{x})). \) This implies that \( \hat{g}(\tilde{x}) \leq g(\tilde{x}) = \frac{1}{g_1(F_i(\tilde{x}))} \leq \frac{1}{\hat{g}(F_i(\tilde{x}))} = \hat{g}(\tilde{x}), \forall \tilde{x} \in \tilde{G}. \) Note that \( 0 \leq \hat{\mu}(\tilde{G}) \leq \hat{\mu}(\tilde{A} \setminus \tilde{B}) + \tilde{\mu}(\tilde{A} \setminus \tilde{C}) = \hat{\mu}(F(\tilde{G}) \setminus F(C)) = \mu(F(\tilde{G}) \setminus F(C)) = F_i \mu(\tilde{A} \setminus \tilde{C}). \) By \( \mu(\tilde{A} \setminus \tilde{C}) = 0 \) and \( F_i \mu \ll \mu, \) we have \( F_i \mu(\tilde{A} \setminus \tilde{C}) = 0. \) Then, we have \( \hat{\mu}(\tilde{G}) = 0. \) Hence, \( g = \hat{g} \) a.e. in \( \tilde{X}. \) This proves the special case.

Finally, consider the general case. Since \( \tilde{\Omega} \) is an nonempty open set and \( \mathbb{R}^m \) is second countable, then \( \exists \{ \tilde{x}_i \}_{i=1}^{\infty}, \{ \tilde{x}_i \}_{i=1}^{\infty} \subseteq \tilde{\Omega} \) such that \( \tilde{x}_i < \tilde{x}_i \) and \( \tilde{\Omega}_{\tilde{x}_i} \subseteq \tilde{\Omega}, \forall i \in \mathbb{N}, \) and \( \Omega = \bigcup_{i=1}^{\infty} \tilde{r}_{\tilde{x}_i}^2 \tilde{x}_i. \) For \( \tilde{\Omega} = (r_{\tilde{x}_i}^2 \tilde{x}_i, \tilde{B}, \tilde{\mu}) \) be the finite metric measure subspace of \( \tilde{X}. \) By the special case, we have \( \hat{g}(\tilde{x}) = g(\tilde{x}) \) a.e. in \( \tilde{X}. \) Let \( \tilde{D} := \{ \tilde{x} \in \tilde{\Omega} \mid \hat{g}(\tilde{x}) \neq g(\tilde{x}) \}. \) By Propositions 7.23, 11.38, 11.39, and 11.35, \( \tilde{D} \in \tilde{B}. \) Then, we have \( 0 \leq \tilde{\mu}(\tilde{D}) = \tilde{\mu}((\bigcup_{i=1}^{\infty} (\tilde{D} \cap r_{\tilde{x}_i}^2 \tilde{x}_i))) \leq \sum_{i=1}^{\infty} \tilde{\mu}(\tilde{D} \cap r_{\tilde{x}_i}^2 \tilde{x}_i) = 0. \) Hence, \( \hat{g} = g \) a.e. in \( \tilde{X}. \)

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, \(\forall x \in \Omega\), \(\exists \delta_x \in (0, \infty) \subseteq \mathbb{R}\) such that \(\Omega \cap B_X(x, \delta_x) - x\) is a conic segment and \(\text{span}(A_X(x)) = X\). Let \(Z := C_0(\Omega, Y)\) be the Banach space over \(K\) as prescribed in Example 9.70. 

\[
\forall f \in Z^*, \exists \mu := (\mu_0, \ldots, \mu_k) \in \prod_{i=0}^k (C(\Omega, B_i(X, Y)))^* \text{ such that } f(z) = \sum_{i=0}^k \langle \langle \mu_i, z^{(i)} \rangle \rangle, \forall z \in Z, \text{ with } \|\mu\| = \left(\sum_{i=0}^k \|\mu_i\|^2\right)^{1/2} = \|f\|_Z. \quad \text{Furthermoe, } \forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^k (C(\Omega, B_i(X, Y)))^*, \text{ define } h : Z \to K \text{ by } h(z) = \sum_{i=0}^k \langle \langle \nu_i, z^{(i)} \rangle \rangle, \forall z \in Z. \quad \text{Then, } h \in Z^* \text{ and } \|h\|_Z \leq \|\nu\|.
\]

**Proof** Let \(W := \prod_{i=0}^k C(\Omega, B_i(X, Y)) =: \prod_{i=0}^k W_i\). By Proposition 7.66 and Example 7.32, \(W_i\) is a Banach space over \(K\), \(\forall i \in \{0, \ldots, k\}\). Then, by Propositions 7.22 and 4.31, \(W\) is a Banach space over \(K\). \(\forall z \in Z, \forall i \in \{0, \ldots, k\}\), we have \(z^{(i)} \in W_i\). Then, we may define a mapping \(T : Z \to W\) by \(T(z) = (z^{(0)}, \ldots, z^{(k)})\), \(\forall z \in Z\). Clearly, \(T\) is linear and norm-preserving. Hence, \(T \in B(Z, W)\) and is injective.

**Claim 12.92.1** \(R(T) \subseteq W\) is closed.

**Proof of claim:** Clearly, we have \(R(T) \subseteq \overline{R(T)} \subseteq W\). \(\forall w \in \overline{R(T)}\), \(\exists (z_n)_{n=1}^\infty \subseteq Z\) such that \(w = \lim_{n \to \infty} T(z_n)\) in \(W\). Then, \((T(z_n))_{n=1}^\infty \subseteq W\) is a Cauchy sequence. Since \(T\) is norm-preserving, then \((z_n)_{n=1}^\infty \subseteq Z\) is a Cauchy sequence. By the completeness of \(Z\), \(\exists z_0 \in Z\) such that \(\lim_{n \to \infty} z_n = z_0\) in \(Z\). By the continuity of \(T\), we have \(w = \lim_{n \to \infty} T(z_n) = T(z_0) \in W\). 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.

By Claim 12.92.1 and Proposition 4.39, \(R(T) \subseteq W\) is a Banach space over \(K\). By Open Mapping Theorem 7.103, \(T_{\text{inv}} : R(T) \to Z\) and \(T_{\text{inv}} \in B(R(T), Z)\).

\[
\forall f \in Z^*, \text{ define } g : R(T) \to K \text{ by } g(w) = f(T_{\text{inv}}(w)), \forall w \in R(T).\quad \text{Clearly, } g \text{ is a bounded linear functional on } R(T). \text{ By Hahn-Banach Extension Theorem 7.83, } \exists \mu \in W^* \text{ such that } \mu(w) = g(w), \forall w \in R(T), \text{ and } \|\mu\|_{W^*} = \|g\|_{R(T)} \leq \|f\|_Z \cdot \|T_{\text{inv}}\|_{B(R(T), Z)} \leq \|f\|_Z, \text{ where the last inequality follows from the fact that } T \text{ is norm-preserving and so is } T_{\text{inv}}.\quad \text{By Example 7.76, } W^* = \prod_{i=0}^k W_i^* = \prod_{i=0}^k (C(\Omega, B_i(X, Y)))^* \text{. Then, } \mu = (\mu_0, \ldots, \mu_k) \in W^* \text{ and } f(z) = g(T(z)) = \langle \langle \mu, T(z) \rangle \rangle = \sum_{i=0}^k \langle \langle \mu_i, z^{(i)} \rangle \rangle, \forall z \in Z.\quad \text{By Example 7.76, } W^* = \prod_{i=0}^k W_i^* = \prod_{i=0}^k (C(\Omega, B_i(X, Y)))^* \text{. Then, } \mu = (\mu_0, \ldots, \mu_k) \in W^* \text{ and } f(z) = g(T(z)) = \langle \langle \mu, T(z) \rangle \rangle = \sum_{i=0}^k \langle \langle \mu_i, z^{(i)} \rangle \rangle, \forall z \in Z. \text{ Then, } \langle \langle \nu, T(z) \rangle \rangle = \langle \langle \nu, z^{(i)} \rangle \rangle, \forall z \in Z. \quad \text{Then, } h \in Z^* \text{ and } h_{|Z^*} \leq \|\nu\|_{W^*} \cdot \|T\|_{B(Z, W)} \leq \|\nu\|_{W^*}, \text{ where the last inequality follows from the fact that } T \text{ is norm-preserving.}
\]

Then, \(\|f\|_{Z^*} \leq \|\mu\|_{W^*} \leq \|f\|_{Z^*}\). Hence, \(\|f\|_{Z^*} = \|\mu\|_{W^*}\). This completes the proof of the theorem. 

\(\square\)
Theorem 12.93 (Riesz Representation Theorem) Let $m \in \mathbb{Z}_+$, $X := K^m$ be the Euclidean space, $I_1, \ldots, I_m \subset \mathbb{R}$ be nondegenerate compact intervals, $\Omega := \prod_{j=1}^m I_j \subset X$ with the subset topology $\mathcal{O}$, $\gamma$ be a Banach space over $K$, $k \in \mathbb{N}$, and $\mathcal{Z} := \mathcal{C}_k(\Omega, \gamma)$ be the Banach space over $K$ as prescribed in Example 9.70. For each $f \in \mathbb{Z}$, $\exists \mu := (\mu_0, \ldots, \mu_k) \in \prod_{i=0}^k \mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle d\mu_i, z(i) \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left( \sum_{i=0}^k (\mathcal{P} \circ \mu_i(\Omega))^2 \right)^{1/2} = \|f\|_{\mathcal{Z}}$. Furthermore, $\forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^k \mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*)$, define $h : \mathcal{Z} \rightarrow K$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle d\nu_i, z(i) \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathbb{Z}$ and $\|h\|_{\mathcal{Z}} \leq \|\nu\|$. 

Proof Let $\mathcal{W} := \prod_{i=0}^k \mathcal{C}(\Omega, B_i(\Omega, \gamma)) := \prod_{i=0}^k W_i$. By Proposition 7.66 and Example 9.70, $W_i$ is a Banach space over $K$, $\forall i \in \{0, \ldots, k\}$. It is easy to show that $\Omega \subseteq X$ satisfies the assumption of Theorem 12.92. By Riesz Representation Theorem 11.204, we have $\mathcal{W}^* = \prod_{i=0}^k \mathcal{C}(\Omega, B_i(\Omega, \gamma))^* = \prod_{i=0}^k \mathcal{M}_f(\Omega, B_i(\Omega, \gamma))^*$ \(= \prod_{i=0}^k \mathcal{M}_f(\Omega, B_i(\Omega, \gamma))^* \). Then, the result follows immediately from Theorem 12.92. This completes the proof of the theorem. \(\square\)

Theorem 12.94 (Riesz Representation Theorem) Let $X$ and $\gamma$ be finite-dimensional Banach spaces over $K$, $\Omega \subseteq X$ be compact with the subset topology $\mathcal{O}$, and $k \in \mathbb{N}$. Assume that, $\forall x \in X$, $\exists \delta_x \in (0, \infty) \subset \mathbb{R}$ such that $\Omega \cap B_X(x, \delta_x) - x$ is a conic segment and $\text{span} \{A_{\Omega}(x)\} = X$. Let $\mathcal{Z} := \mathcal{C}_k(\Omega, \gamma) := \prod_{i=0}^k \mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*)$ such that $f(z) = \sum_{i=0}^k \int_{\Omega} \langle d\mu_i, z(i) \rangle$, $\forall z \in \mathcal{Z}$, with $\|\mu\| = \left( \sum_{i=0}^k (\mathcal{P} \circ \mu_i(\Omega))^2 \right)^{1/2} = \|f\|_{\mathcal{Z}}$. Furthermore, $\forall \nu := (\nu_0, \ldots, \nu_k) \in \prod_{i=0}^k \mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*)$, define $h : \mathcal{Z} \rightarrow K$ by $h(z) = \sum_{i=0}^k \int_{\Omega} \langle d\nu_i, z(i) \rangle$, $\forall z \in \mathcal{Z}$. Then, $h \in \mathbb{Z}$ and $\|h\|_{\mathcal{Z}} \leq \|\nu\|$. 

Proof Let $\mathcal{W} := \prod_{i=0}^k \mathcal{C}(\Omega, B_i(\Omega, \gamma)) := \prod_{i=0}^k W_i$. The result follows immediately from Theorem 12.92 if we can prove that $\mathcal{W}^* = \prod_{i=0}^k \mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*) = \prod_{i=0}^k \mathcal{M}_f(\Omega, B_i(\Omega, \gamma))^*$. The first equality follows from Example 7.66 and Riesz Representation Theorem 11.205. Since $\Omega \subseteq X$ is compact then it is locally compact. Since $X$ 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 $\mathcal{M}_f(\Omega, (B_i(\Omega, \gamma))^*) = \mathcal{M}_f(\Omega, B_i(\Omega, \gamma))^*$, $\forall i \in \{0, \ldots, k\}$. This completes the proof of the theorem. \(\square\)

Example 12.95 Let $\mathbb{R}^2$ be endowed with the usual positive cone, $\Omega := \{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 > 0, \angle(x_1 + ix_2) \neq 0 \}$, $\overline{\Omega} := \{ (\rho, \theta) \in \mathbb{R}^2 \mid \rho > 0, \theta \in (0, 2\pi) \}$, and $F : \Omega \rightarrow \Omega$ be defined by $F(x_1, x_2) :=$
Let \( \Omega \rightarrow \Omega \) be defined by \( F_i(\rho, \theta) = F_i^{\text{inv}}(\rho, \theta) = (\rho \cos(\theta), \rho \sin(\theta)) \). It is easy to check that \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^2, \mathbb{R}^2) \) and \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^2, \mathbb{R}^2) \). \( F_i^{(1)}(\rho, \theta) = \begin{bmatrix} \cos(\theta) & -\rho \sin(\theta) \\ \sin(\theta) & \rho \cos(\theta) \end{bmatrix} \)

and \( \left| \det(F_i^{(1)}(\rho, \theta)) \right| = \rho > 0, \forall (\rho, \theta) \in \tilde{\Omega} \). Hence, \( \tilde{U} = \tilde{\Omega} \) as defined in Change of Variable Theorem 12.16. Then, all assumption of that theorem are satisfied. Let \( Y \) be a separable Banach space, \( \forall f : \Omega \rightarrow Y \) that is absolutely integrable, we have \( \int_\Omega f(x_1, x_2) d\mu_{B^2}(x_1, x_2) = \int_{\tilde{\Omega}} f(F_i(\rho, \theta)) \rho d\mu_{B^2}(\rho, \theta) \). Furthermore, \( \forall f : \mathbb{R}^2 \rightarrow Y \) that is absolutely integrable over \( \mathbb{R}^2 \), we have \( \int_{\mathbb{R}^2} f(x_1, x_2) d\mu_{B^2}(x_1, x_2) = \int_{\mathbb{R}^2} f(x_1, x_2) d\mu_{B^2}(x_1, x_2) = \int_{\tilde{\Omega}} f(F_i(\rho, \theta)) \rho d\mu_{B^2}(\rho, \theta) \), where the first equality follows from the fact that \( \mu_{B^2}(\mathbb{R}^2 \setminus \Omega) = 0 \) and Proposition 11.92.

**Example 12.96** Let \( m \in \mathbb{N} \) with \( m \geq 3 \), \( \mathbb{R}^m \) be endowed with the usual positive cone, \( \Omega := \{ (x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_3 \neq 0, \ldots, x_m \neq 0, x_1^2 + x_2^2 > 0, \angle(\mathbb{R}^2 \setminus \bar{\Omega}) \neq 0 \} \), \( \bar{\Omega} := \{ (\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) \in \mathbb{R}^m \mid \rho > 0, \theta_i \in (0, \pi), i = 1, \ldots, m-2, \varphi \in (0, 2\pi) \} \), and \( F : \Omega \rightarrow \Omega \) be defined by \( F(x_1, \ldots, x_m) := \left( \sqrt{\sum_{i=1}^{m} x_i^2}, \arccos\left( \frac{x_m}{\sqrt{\sum_{i=1}^{m} x_i^2}} \right), \ldots, \arccos\left( \frac{x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \right) \right) \).\( \angle(\mathbb{R}^2 \setminus \bar{\Omega}) \neq 0 \), and \( F_i : \Omega \rightarrow \Omega \) be defined by \( F_i(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) = F_i^{\text{inv}}(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi) = (\rho \sin(\theta_1) \cdots \sin(\theta_{m-2}) \cos(\varphi), \rho \sin(\theta_1) \cdots \cos(\theta_{m-2}) \sin(\varphi)) \). It is easy to check that \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^m, \mathbb{R}^m) \) and \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^m, \mathbb{R}^m) \) and \( \left| \det(F_i^{(1)}(\rho, \theta_1, \ldots, \theta_{m-2}, \varphi)) \right| = \rho^{m-1}(\sin(\theta_1))^2 \).

**Example 12.97** Let \( \mathbb{R}^3 \) be endowed with the usual positive cone, \( \Omega := \{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 > 0, \angle(\mathbb{R}^2 \setminus \bar{\Omega}) \neq 0 \} \), \( \bar{\Omega} := \{ (\rho, \varphi, z) \in \mathbb{R}^3 \mid \rho > 0, \varphi \in (0, 2\pi) \} \), and \( F : \Omega \rightarrow \Omega \) be defined by \( F(x_1, x_2, x_3) := (\sqrt{x_1^2 + x_2^2}, \angle(\mathbb{R}^2 \setminus \bar{\Omega}), \angle(\mathbb{R}^2 \setminus \bar{\Omega})) \), and \( F_i : \Omega \rightarrow \Omega \) be defined by \( F_i(\rho, \varphi, z) = F_i^{\text{inv}}(\rho, \varphi, z) = (\rho \cos(\varphi), \rho \sin(\varphi), z) \). It is easy to check that \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^3, \mathbb{R}^3) \) and \( F_i^{(1)} : \Omega \rightarrow B(\mathbb{R}^3, \mathbb{R}^3) \), \( \begin{bmatrix} \cos(\varphi) & -\rho \sin(\varphi) & 0 \\ \sin(\varphi) & \rho \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(\varphi) & -\rho \sin(\varphi) & 0 \\ \sin(\varphi) & \rho \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \).
Proof
Since \( µ \) is absolutely integrable over \( \Omega \) as defined in Change of Variable Theorem 12.91. Then, all assumption of that theorem are satisfied. Let \( y \) be a separable Banach space, \( ∀f : Ω \rightarrow y \) that is absolutely integrable, we have
\[
\int_{Ω}(f(x_1, x_2, x_3))dµ_{B3}(x_1, x_2, x_3) = \int_{Ω}(F_i(ρ, ϕ, z))ρdµ_{B3}(ρ, ϕ, z).
\]
Furthermore, \( ∀f : \mathbb{R}^3 \rightarrow y \) that is absolutely integrable over \( \mathbb{R}^3 \), we have
\[
\int_{\mathbb{R}^3}(f(x_1, x_2, x_3))dµ_{B3}(x_1, x_2, x_3) = \int_{\mathbb{R}^3}(F_i(ρ, ϕ, z))ρdµ_{B3}(ρ, ϕ, z),
\]
where the first equality follows from the fact that \( µ_{B3}(\mathbb{R}^3 \setminus Ω) = 0 \) and Proposition 11.92.

**Theorem 12.98 (Riemann-Lebesgue)** Let \( y \) be a separable Banach space over \( \mathbb{K} \), \( f \in L_1(\mathbb{R}, y) \). Then, \( \lim_{|t| \rightarrow \infty} \int_{\mathbb{R}} f(t) \sin(pt) dµ_B(t) = \vartheta_y \).

**Proof**
Since \( f \in L_1(\mathbb{R}, y) \), then \( f \) is absolutely integrable over \( \mathbb{R} \). Then, \( \bar{f}_p : \mathbb{R} \rightarrow y \) defined by \( \bar{f}_p(t) := f(t) \sin(pt) \), \( ∀t \in \mathbb{R} \), is absolutely integrable over \( \mathbb{R} \), \( ∀p \in \mathbb{R} \), by Lebesgue Dominated Convergence Theorem 11.91.

First, we consider the special case when \( f = yχI_{\mathbb{R}} \) with \( I \) being an interval and \( y \in \mathbb{Y} \) 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 \( T := [a, b] \) with \( a, b \in \mathbb{R} \) and \( a \leq b \). Then, \( \lim_{|p| \rightarrow \infty} \int_{\mathbb{R}} f(t) \sin(pt) dµ_B(t) = \lim_{|p| \rightarrow \infty} \int_{a}^{b} y \sin(pt) dµ_B(t) = \lim_{|p| \rightarrow \infty} \int_{a}^{b} \frac{yu}{p}(\cos(pa) - \cos(pb)) dµ_B(t) = \vartheta_y \), where the first equality follows from \( I \) being an interval and \( T = [a, b] \); the second equality follows from Proposition 11.92 and Theorem 12.83. This special case is proved.

Next, we consider the general case when \( f \in L_1(\mathbb{R}, y) \). By Proposition 11.66, there exists a sequence of simple functions \( (ϕ_n)_{n=1}^{∞} \), \( ϕ_n : \mathbb{R} \rightarrow y \), \( ∀n \in \mathbb{N} \), such that \( \lim_{n \in \mathbb{N}} ϕ_n(x) = f(x) \), \( a.e. \in \mathbb{R} \), and \( \|ϕ_n(x)\| < \|f(x)\| \), \( ∀x \in \mathbb{R} \), \( ∀n \in \mathbb{N} \). Then, \( \lim_{n \in \mathbb{N}} \int_{\mathbb{R}} \|f - ϕ_n\| dµ_B = 0 \), by Lebesgue Dominated Convergence Theorem 11.91. \( ∀ε ∈ (0, ∞) ⊆ \mathbb{R} \), \( ∃ \delta \in \mathbb{N} \) such that \( \int_{\mathbb{R}} \|f - ϕ_n\| dµ_B < \frac{ε}{3} \). Let \( ϕ_{no} \) admit the canonical representation \( ϕ_{no} = \sum_{i=1}^{n} y_iχ_{E_i} \), where \( n ∈ \mathbb{Z}_+ \), \( y_1, \ldots, y_n \in y \) are distinct and none equals to \( \vartheta_y \), \( E_1, \ldots, E_n \in \mathcal{B}_B(\mathbb{R}) \) are nonempty, pairwise disjoint and of finite measure, \( \int_{\mathbb{R}} ϕ_{no} dµ_B = \sum_{i=1}^{n} \|y_i\| µ_B(E_i) \). By Proposition 11.7, \( ∃a, b \in \mathbb{R} \) with \( a < b \) such that \( \int_{\mathbb{R}}[a, b] \|ϕ_{no}\| dµ_B < \frac{ε}{3} \). Let \( f_1 := ϕ_{no}χ_{[a, b]} \). Then, \( \int_{\mathbb{R}} \|f - f_1\| dµ_B < \frac{ε}{3} \). By Proposition 11.182, \( ∃ \) a continuous function \( g : [a, b] \rightarrow y \) such that \( g \in L_1([a, b], y) \) and \( \|g - f_1\|_{[a, b]} < \frac{ε}{4(b-a)} \). By Proposition 5.39, \( g \) is uniformly continuous. \( ∃ m \in \mathbb{N} \), \( ∀x_1, x_2 \in [a, b] \) with \( |x_1 - x_2| ≤ \frac{b-a}{m} \), we have \( \|g(x_1) - g(x_2)\| < \frac{ε}{4(b-a)} \). Construct a step function \( h : \mathbb{R} \rightarrow y \) by
\[
h(x) = \begin{cases} 0 & x \in \mathbb{R} \setminus [a, b] \\
g(a + \left\lfloor \frac{x - a}{b-a} \right\rfloor \frac{b-a}{m}) & x \in [a, b] \end{cases}
\]
Then, we have
\[
\int_{[a, b]} \|h_{[a, b]} - g\| dµ_B < \frac{ε}{3} \text{.}
\]
Hence, \( \int_{\mathbb{R}} \|f - h\| dµ_B ≤ \int_{\mathbb{R}} \|f - f_1\| dµ_B + \int_{\mathbb{R}} \|h_{[a, b]} - g\| dµ_B + \int_{\mathbb{R}} \|f - g\| dµ_B < \frac{ε}{3} + \frac{ε}{3} + \frac{ε}{3} = ε \), as desired.
12.7. REPRESENTATION OF \((C_K(\Omega, Y))^*\)

\[
\int_{\mathbb{R}} \| f_1 - h \| \, d\mu_B < \frac{\epsilon}{2} + \int_{[a,b]} \| f_1|_{[a,b]} - h|_{[a,b]} \| \, d\mu_B \leq \frac{\epsilon}{2} + \int_{[a,b]} \| f_1|_{[a,b]} - g \| \, d\mu_B + \int_{[a,b]} \| g - h|_{[a,b]} \| \, d\mu_B < \epsilon. \]  

Then, \(0 \leq \liminf_{|p| \to \infty} \int_{\mathbb{R}} f_p \, d\mu_B \leq \limsup_{|p| \to \infty} \int_{\mathbb{R}} f_p \, d\mu_B \leq \limsup_{|p| \to \infty} \int_{\mathbb{R}} f(t) \sin(pt) \, d\mu_B(t) \leq \limsup_{|p| \to \infty} \int_{\mathbb{R}} f(t) - h(t) \| \, d\mu_B(t) < 0 + \epsilon, \) where the first three inequalities follow from Proposition 3.83; the fourth inequality follows from Propositions 3.83 and 11.92; and the last inequality follows from the preceding discussion and the special case. Hence, by the arbitrariness of \(\epsilon,\) we have \(\lim_{|p| \to \infty} \int_{\mathbb{R}} f(t) \sin(pt) \, d\mu_B(t) = 0,\) by Proposition 3.83. Then, \(\lim_{|p| \to \infty} \int_{\mathbb{R}} f(t) \sin(pt) \, d\mu_B(t) = \vartheta_Y.\)

This completes the proof of the theorem. \(\square\)
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 settings in Hilbert spaces.

13.1 Fundamental Notions

Definition 13.1 A pre-Hilbert space $X$ over $\mathbb{K}$ is a vector space $X$ over $\mathbb{K}$, together with an inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ satisfying, $\forall x, y, z \in X$, $\forall \lambda \in \mathbb{K}$,

(i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$;

(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, $x = \vartheta_X$.

Proposition 13.2 Let $X$ be a pre-Hilbert space over $\mathbb{K}$. Define $\| \cdot \| : X \to [0, \infty) \subset \mathbb{R}$ by $\| x \| = \sqrt{\langle x, x \rangle}$, $\forall x \in X$. Then, $\forall x, y, z \in X$, $\forall \alpha, \beta \in \mathbb{K}$, 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) $\langle \vartheta_X, x \rangle = 0 = \langle x, \vartheta_X \rangle$;

(iv) $| \langle x, y \rangle | \leq \| x \| \| y \|$, where equality holds if, and only if, $y = \vartheta_X$ 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, $y = \vartheta_X$ or $x = \lambda y$, for some $\lambda \in [0, \infty) \subset \mathbb{R}$. 

647
CHAPTER 13. HILBERT SPACES

(vi) \( \| \cdot \| \) is a norm of \( X \). (Induced norm)

Proof

(i) \( \langle \alpha x + \beta y, z \rangle = \langle \alpha x, z \rangle + \langle \beta y, z \rangle \), where the first equality follows from (ii) of Definition 13.1; and the second equality follows from (iii) of Definition 13.1.

(ii) \( \langle z, \alpha x + \beta y \rangle = \langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle = \overline{\alpha \langle x, z \rangle + \beta \langle z, y \rangle} \), where the first equality follows from (i) of Definition 13.1; the second equality follows from (i); and the fourth equality follows from (i) of Definition 13.1.

(iii) \( \langle \vartheta X, x \rangle = \langle y - y, x \rangle = \langle y, x \rangle - \langle y, x \rangle = 0 \), 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, \( |\langle x, y \rangle| = 0 = \| x \| \cdot \| y \| \). This case is proved.

Case 2: \( y \neq \vartheta X \). By (iv) of Definition 13.1, we have \( \langle y, y \rangle > 0 \). Let \( \lambda \in \mathbb{K} \), we have \( 0 \leq \langle x - \lambda y, x - \lambda y \rangle \leq \langle x, x \rangle - \lambda \langle x, y \rangle + \lambda \langle y, y \rangle \), 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, \( 0 \leq \langle x, x \rangle - \lambda \langle x, y \rangle + \lambda \langle y, y \rangle \), where the equality follows from (i) of Definition 13.1. Hence, \( |\langle x, y \rangle| \leq \| x \| \| y \| \). Equality holds if, and only if, \( x - \lambda y = \vartheta X \).

(v) Note that \( \| x + y \|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \| x \|^2 + \| y \|^2 + 2 \text{Re} \langle x, y \rangle \leq \| x \|^2 + 2 \| x \| \| y \| \leq \| x \|^2 + 2 \| x \| \| y \| + \| y \|^2 = (\| x \| + \| y \|)^2 \), 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, \( \text{Re} \langle x, y \rangle = |\langle x, y \rangle| = \| x \| \| y \| \). 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, \( \| x \| \in [0, \infty) \subset \mathbb{R} \). \( \| x \| = 0 \Leftrightarrow \langle x, x \rangle = 0 \Leftrightarrow x = \vartheta X \).

Hence, (i) of Definition 7.1 holds. By (v), (ii) of Definition 7.1 holds. Note also that \( \| \alpha x \| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \overline{\alpha} \langle x, x \rangle} = |\alpha| \| x \| \), where the second equality follows from (i) and (ii); and the third equality follows from the fact that \( \overline{\alpha \overline{\alpha}} = |\alpha|^2 \). Hence, (iii) of Definition 7.1 holds. Therefore, \( \| \cdot \| \) defines a norm of \( X \). \( \square \)

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. \( \square \)

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 $\mathcal{X}$ be a pre-Hilbert space. Then, the inner product is a continuous function of the product space $\mathcal{X} \times \mathcal{X}$.

Proof We will show that the inner product is continuous at any $(x_0, y_0) \in \mathcal{X} \times \mathcal{X}$. For $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, 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 $(x, y) \in B_{\mathcal{X} \times \mathcal{X}}((x_0, y_0), \delta)$. Then, we have $|\langle x, y \rangle - \langle x_0, y_0 \rangle| = |\langle x - x_0, y \rangle + \langle x_0, y - y_0 \rangle| \leq \|x - x_0\| \cdot \|y\| + \|y - y_0\| \cdot \|x_0\| \leq \|x - x_0\| \cdot M + \|y - y_0\| \cdot \|x_0\| \leq 2M \delta \leq \epsilon$, 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.

Definition 13.5 A complete pre-Hilbert space is called a Hilbert space.

Definition 13.6 Let $\mathcal{X}$ be a pre-Hilbert space, $x, y \in \mathcal{X}$, and $S \subseteq \mathcal{X}$. 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 $x \perp s$, $\forall s \in S$. We will then write $x \perp S$.

Theorem 13.7 (Pythagorean) Let $\mathcal{X}$ be a pre-Hilbert space over $\mathbb{K}$ and $x, y \in \mathcal{X}$. 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.

Example 13.8 Let $n \in \mathbb{Z}^+$ and $\mathcal{X} = \mathbb{K}^n$. Define the inner product by $\langle x, y \rangle = \sum_{i=1}^{n} \xi_i \overline{\eta_i} \in \mathbb{K}$, $\forall x := (\xi_1, \ldots, \xi_n) \in \mathbb{K}^n$ and $\forall y := (\eta_1, \ldots, \eta_n) \in \mathbb{K}^n$. It is easy to check that the pair $\mathcal{X} := (\mathcal{X}, \langle \cdot, \cdot \rangle)$ is a pre-Hilbert space over $\mathbb{K}$. The induced norm is clearly the Euclidean norm. By Example 7.29, $\mathcal{X}$ is complete. Therefore, $\mathcal{X}$ is a Hilbert space, which is to be denoted simply by $\mathbb{K}^n$.

Example 13.9 Let $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle$, $i = 1, 2$. Let $\mathcal{Z} := \mathcal{X}_1 \times \mathcal{X}_2$. Define inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}} : \mathcal{Z} \times \mathcal{Z} \to \mathbb{K}$ 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 $\mathbb{K}$.
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 $\mathbb{K}$, $\forall i \in \{1, \ldots, n\}$.

Example 13.10 Let $X$ be a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_X$ and $\mathcal{Z} := L_2(X)$. We will show that $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ defined by $\langle x, y \rangle_{\mathcal{Z}} = \sum_{i=1}^{\infty} \langle \xi_i, \eta_i \rangle$, $\forall x := (\xi_1, \xi_2, \ldots) \in \mathcal{Z}$ and $\forall y := (\eta_1, \eta_2, \ldots) \in \mathcal{Z}$. To see that $\langle x, y \rangle_{\mathcal{Z}} \in \mathbb{K}$, we note that $x \in \mathcal{Z}$ 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 $y \in \mathcal{Z}$, $\exists n_2 \in \mathbb{N}$ 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 $|\sum_{i=n}^{m} \langle \xi_i, \eta_i \rangle| \leq \left( \sum_{i=n}^{m} \| \xi_i \|^2 \right)^{1/2} \left( \sum_{i=n}^{m} \| \eta_i \|^2 \right)^{1/2} \leq \left( \sum_{i=0}^{\infty} \| \xi_i \|^2 \right)^{1/2} \left( \sum_{i=0}^{\infty} \| \eta_i \|^2 \right)^{1/2} < \epsilon$, where the first inequality follows from Hölder’s Inequality 7.8. Hence, $\left( \sum_{i=1}^{\infty} \langle \xi_i, \eta_i \rangle \right)_{n=1}^{\infty} \subseteq \mathbb{K}$ is a Cauchy sequence, which converges since $\mathbb{K}$ is complete. Hence, $\langle x, y \rangle_{\mathcal{Z}} \in \mathbb{K}$. It is easy to show that (i) – (iv) of Definition 13.1 are satisfied. Hence, $\mathcal{Z}$ is a pre-Hilbert space. It is easy to see that the induced norm is equal to the norm of $\mathcal{Z}$ as defined in Example 7.10. By Example 7.33, we have $\mathcal{Z}$ is complete and therefore a Hilbert space.

Example 13.11 Let $\mathcal{X} := (X, \mathcal{B}, \mu)$ be a $\sigma$-finite measure space, $Y$ be a separable Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_Y$ and $\mathcal{Z} := L_2(\mathcal{X}, Y)$. We will show that $\mathcal{Z}$ is a Hilbert space over $\mathbb{K}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ defined by $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} = \int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_Y \, d\mu(x)$, $\forall z_1, z_2 \in L_2(\mathcal{X}, Y)$.

First, we show that $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ is well-defined. $\forall z_1, z_2 \in L_2(\mathcal{X}, Y)$, we have $z_1$ and $z_2$ are $\mathcal{B}$-measurable. By Propositions 11.39, 11.38, and 13.4, the function $h : \mathcal{X} \to \mathbb{K}$ defined by $h(x) = \langle z_1(x), z_2(x) \rangle_Y$, $\forall x \in \mathcal{X}$, 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, $\int_{\mathcal{X}} \mathcal{P} \circ h \, d\mu \leq \int_{\mathcal{X}} \mathcal{P} \circ z_1 \mathcal{P} \circ z_2 \, d\mu \leq \left( \int_{\mathcal{X}} \mathcal{P} \circ z_1 \, d\mu \right)^{1/2} \left( \int_{\mathcal{X}} \mathcal{P} \circ z_2 \, d\mu \right)^{1/2} = \| z_1 \| \| z_2 \| < \infty$, where the first inequality follows from Definition 11.79; and the second inequality follows from Hölder’s Inequality 11.178. Hence, $h$ is absolutely integrable over $\mathcal{X}$. By Proposition 11.92, $h$ is integrable over $\mathcal{X}$ and $\int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_Y \, d\mu(x) \in \mathbb{K}$. $\forall \tilde{z}_1, \tilde{z}_2 \in L_2(\mathcal{X}, Y)$ 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 $\tilde{h} : \mathcal{X} \to \mathbb{K}$ by $\tilde{h}(x) = \langle \tilde{z}_1(x), \tilde{z}_2(x) \rangle_Y$, $\forall x \in \mathcal{X}$. 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, $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} \in \mathbb{K}$ is well-defined.

Next, we show that $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ satisfies (i) – (iv) of Definition 13.1. Fix any $z_1, z_2, z_3 \in L_2(\mathcal{X}, Y)$ and any $\lambda \in \mathbb{K}$. (i) $\langle [z_1], [z_2] \rangle_{\mathcal{Z}} = \int_{\mathcal{X}} \langle z_1(x), z_2(x) \rangle_Y \, d\mu(x)$, (ii) $\langle [z_1] + [z_2], [z_3] \rangle_{\mathcal{Z}} = \langle [z_1 + z_2], [z_3] \rangle_{\mathcal{Z}} = \int_{\mathcal{X}} \langle z_1(x) + z_2(x), z_3(x) \rangle_Y \, d\mu(x) = \int_{\mathcal{X}} \langle z_1(x), z_3(x) \rangle_Y \, d\mu(x)$,
\[ \langle z_2(x), z_3(x) \rangle_Y \, d\mu(x) = \int_X \langle z_1(x), z_3(x) \rangle_Y \, d\mu(x) + \int_X \langle z_2(x), z_3(x) \rangle_Y \, d\mu(x) = \langle [z_1], [z_3] \rangle_Z + \langle [z_2], [z_3] \rangle_Z, \] where the first equality follows from Proposition 7.43; the third equality follows from the fact that \( Y \) is a Hilbert space; and the fourth equality follows from Proposition 11.92.

(iii) \( \langle \lambda[z_1], [z_2] \rangle_Z = \int_X \lambda(z_1(x), z_2(x))_Y \, d\mu(x) = \lambda \int_X \langle z_1(x), z_2(x) \rangle_Y \, d\mu(x) = \lambda \langle [z_1], [z_2] \rangle_Z, \) where the first equality follows from Proposition 7.43; the third equality follows from the fact that \( Y \) is a Hilbert space; and the fourth equality follows from Proposition 11.92.

(iv) \( \langle [z_1], [z_1] \rangle_Z = \int_X \langle z_1(x), z_1(x) \rangle_Y \, d\mu(x) = \int_X \| z_1(x) \|_Y^2 \, d\mu(x) = \| z_1 \|_Z^2 \in [0, \infty) \subset \mathbb{R}. \) Clearly, \( \langle [z_1], [z_1] \rangle_Z = 0 \iff \| z_1 \|_Z = 0 \iff [z_1] = \vartheta_Z. \) Hence, \( Z \) is a pre-Hilbert space over \( \mathbb{K}. \)

Note that the induced norm of \( Z \) is equal to the norm of \( Z \) as defined in Example 11.173. By Example 11.179, \( Z \) is complete. Hence, \( Z \) is a Hilbert space over \( \mathbb{K}. \)

\[ \Diamond \]

### 13.2 Projection Theorems

The basic concept of the Projection Theorem is illustrated in the Figure 13.1.

**Figure 13.1**: Projection onto a subspace.

**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 13.2, $\|x - m_1\|^2 = (x - m_0 - \delta m, x - m_0 - \delta m) = \|x - m_0\|^2 - (x - m_0, \delta m) - (\delta m, x - m_0) + \|\delta m\|^2 = \|x - m_0\|^2 - |\delta|^2 + |\delta|^2 \|m\|^2 = \|x - m_0\|^2 - |\delta|^2 < \|x - m_0\|^2$. 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. $\forall m \in M$, 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. $\forall x \in X$, 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 $\delta := \inf_{m \in M} \|x - m\| \in [0, \infty) \subset \mathbb{R}$. Then, $\exists (m_n)_{n=1}^{\infty} \subseteq M$ such that $\lim_{n \to \infty} \|x - m_n\| = \delta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}, \exists n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$ with $n_0 \leq n$, we have $\delta \leq \|x - m_n\| < \sqrt{\delta^2 + \epsilon^2} / 2$. 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 $\|x - m_0\| = \lim_{n \to \infty} \|x - m_n\| = \delta = \inf_{m \in M} \|x - m\|$. 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. $\forall x \in X$, 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 $X$ be a Hilbert space over $K$. Then, the following statements hold.

(i) $\forall f \in X^*$, there exists a unique $y_0 \in X$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in X$, and $\|f\|_{X^*} = \|y_0\|_X$. Therefore, we may define a mapping $\Phi: X^* \to X$ by $\Phi(f) = y_0$.

(ii) $\forall y \in X$, define $g: X \to K$ by $g(x) = \langle x, y \rangle$, $\forall x \in X$, then $g \in X^*$.

(iii) The mapping $\Phi$ is bijective, uniformly continuous, norm preserving, and conjugate linear (that is $\Phi(\alpha f_1 + \beta f_2) = \alpha \Phi(f_1) + \beta \Phi(f_2)$, $\forall f_1, f_2 \in X^*$, $\forall \alpha, \beta \in K$).

(iv) If $K = \mathbb{R}$, then $\Phi$ is an isometrical isomorphism between $X^*$ and $X$.

(v) If $K = \mathbb{C}$, let $\phi: X \to X^{**}$ be the natural mapping as defined in Remark 7.88, then $\phi$ is surjective and $X$ is reflexive.

(vi) If $K = \mathbb{C}$, then $X^*$ with the inner product $\langle \cdot, \cdot \rangle_{X^*}$, defined by $\langle f, g \rangle_{X^*} := \langle \Phi(g), \Phi(f) \rangle$, $\forall f, g \in X$, is a Hilbert space, and it is reflexive.

Henceforth, we will denote $\Phi_{\text{inv}}(x) =: x^*$, $\forall x \in X$. Furthermore, the following statement holds.

(vii) When $K = \mathbb{C}$, let $\Phi_*: X^{**} = X \to X^*$ be the mapping of $\Phi$ if $X$ is replaced by $X^*$. Then, $\Phi_* = \Phi_{\text{inv}}$. This leads to the identity $(x^*)^* = x$, $\forall x \in X$.

**Proof** We will first show that, $\forall y \in X$, the functional $g$ defined in the statement (ii) of the theorem is a bounded linear functional with $\|g\|_{X^*} = \|y\|_X$. By Definition 13.1, $g$ is a linear functional. By Proposition 13.2, we have $|g(x)| = \|\langle x, y \rangle\| \leq \|x\|_X \|y\|_X$. Hence, $g \in X^*$ and $\|g\|_{X^*} \leq \|y\|_X$. Note that $|g(y)| = \|\langle y, y \rangle\| = \|y\|_X^2$. Hence, $\|g\|_{X^*} \geq \|y\|_X$. Hence, $\|g\|_{X^*} = \|y\|_X$.

Next, we will show that, for any $f \in X^*$, there exists a $y_0 \in X$ such that $f(x) = \langle x, y_0 \rangle$, $\forall x \in X$. Let $M := N(f)$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $M = X$; Case 2: $M \subset X$.

Case 1: $M = X$. Take $y = \vartheta_X$. Then, $f(x) = 0 = \langle x, y \rangle$, $\forall x \in X$. This case is proved.

Case 2: $M \subset X$. Then, $\exists x_0 \in X \setminus M$. Since $f \in X^*$, then $M$ is a closed subspace of $X$ by Proposition 7.68. By Projection Theorem 13.13, $\exists m_0 \in M$ such that $(x_0 - m_0) \perp M$. Let $z_0 := x_0 - m_0 \in X \setminus M$. Then, $f(z_0) \neq 0$. Without loss of generality, we may assume that $f(z_0) = 1$. $\forall x \in X$, 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 $0 = \langle x - f(x)z_0, z_0 \rangle = \langle x, z_0 \rangle - f(x)\|z_0\|^2$. 


Then, \( f(x) = \langle x, z_0 \rangle / \| z_0 \|^2 = \langle x, z_0 \rangle / \| z_0 \|^2 \). 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, 

\[
0 = f(x) - f(x) = \langle x, y_0 \rangle - \langle x, y \rangle = \langle x, y_0 - y \rangle, \forall x \in \mathcal{X} \text{. Then, } 0 = \langle y_0 - y, y_0 - y \rangle = \|y_0 - y\|. 
\]

Thus, \( y = y_0 \). Hence, \( y_0 \) is unique.

Based on the above, we have (i) and (ii) holds.

(iii) Note that, \( \forall f_1, f_2 \in \mathcal{X}^* \), \( \forall \alpha, \beta \in \mathbb{K} \), \( (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha (x, \Phi(f_1)) + \beta (x, \Phi(f_2)) = \langle x, \alpha \Phi(f_1) + \beta \Phi(f_2) \rangle \), \( \forall x \in \mathcal{X} \), 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 (ii), \( \| \Phi(f) \|_{\mathcal{X}} = \| f \|_{\mathcal{X}^*}, \forall f \in \mathcal{X}^* \). 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 \( \langle \langle x, x \rangle \rangle = \langle \langle \Phi(x), x_* \rangle \rangle \). Then, \( \forall y \in \mathcal{X} \), we have \( \langle x, y \rangle = \langle \langle \Phi(y), x \rangle \rangle = \langle \langle \Phi(x_*), \Phi(x) \rangle \rangle \). 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. \( \forall x_* \in \mathcal{X}^{**} \), \( \langle \langle x_*, x \rangle \rangle = \langle \langle x_*, \Phi(x) \rangle \rangle \). Define \( G : \mathcal{X} \rightarrow \mathcal{K} \) by \( G(y) = \langle \langle x_*, \Phi(y) \rangle \rangle, \forall y \in \mathcal{X} \). By (iii), it is easy to show that \( G \in \mathcal{X}^* \). By (i), \( G(y) = \langle \langle y, \Phi(G) \rangle \rangle, \forall y \in \mathcal{X} \). Then, \( \langle \langle x_*, x \rangle \rangle = \langle \langle x_*, \Phi(x) \rangle \rangle = G(\Phi(x)) = \langle \langle \Phi(x), \Phi(x) \rangle \rangle = \langle \langle \phi(\Phi(G)), x \rangle \rangle, \forall x \in \mathcal{X}^* \). Thus, we have \( x_* = \phi(\Phi(G)) \in \mathcal{K}(\phi) \). 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 \( \mathcal{X}^* \), \( \forall f, g, h \in \mathcal{X}^* \), \( \forall \lambda \in \mathcal{K} \). (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) \( \langle \lambda f, g \rangle_{\mathcal{X}^*} = \langle \Phi(g), \Phi(\lambda f) \rangle = \langle \Phi(g), \lambda \Phi(f) \rangle = \lambda \langle \Phi(g), \Phi(f) \rangle = \lambda \langle f, g \rangle_{\mathcal{X}^*} \), where the second equality follows from (ii); and the third equality follows from Definition 13.1. (d) \( \langle f, f \rangle_{\mathcal{X}^*} = \langle \Phi(f), \Phi(f) \rangle = \| \Phi(f) \|_{\mathcal{X}}^2 = \| f \|_{\mathcal{X}}^2, \in [0, \infty) \subset \mathbb{R} \), 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.
13.3. DUAL OF HILBERT SPACES

Let \( \mathbb{K} = \mathbb{C} \). \( \forall x \in \mathcal{X} \). Let \( f = \Phi_*(x) = \Phi_*(\phi(x)) \). Then, \( f(x) = \langle x, \Phi(f) \rangle = \langle \phi(x)(f) \rangle = \langle f, \Phi_*(\phi(x)) \rangle_{\mathcal{X}^*} = \langle f, \Phi_*(x) \rangle_{\mathcal{X}^*} = \langle f, f \rangle_{\mathcal{X}^*} = \| f \|^2_{\mathcal{X}^*} \), where the first equality follows from (i); the second equality follows from Remark 7.88; the third equality follows from (i) as applied to \( \mathcal{X}^* \); and the fourth equality follows from (v). By (iii), \( \| f \|_{\mathcal{X}^*} = \| x \|_{\mathcal{X}} \) and \( \| f \|_{\mathcal{X}^*} = \| \Phi(f) \|_{\mathcal{X}} \). This implies that \( \langle x, \Phi(f) \rangle = \| x \|_{\mathcal{X}} \| \Phi(f) \|_{\mathcal{X}} \). By (iv) of Proposition 13.2, \( \Phi(f) = \vartheta_{\mathcal{X}} \) or \( x = \lambda \Phi(f) \), for some \( \lambda \in \mathbb{C} \).

If \( \Phi(f) = \vartheta_{\mathcal{X}} \), by (iii), \( f = \vartheta_{\mathcal{X}^*} \). Thus, by (iii), \( x = \vartheta_{\mathcal{X}} = \Phi(f) \) since \( f = \Phi_*(x) \).

If \( \Phi(f) \neq \vartheta_{\mathcal{X}} \), we must have \( x = \lambda \Phi(f) \), for some \( \lambda \in \mathbb{C} \). Then, \( \| f \|^2_{\mathcal{X}^*} = \langle x, \Phi(f) \rangle = \lambda \langle \Phi(f), \Phi(f) \rangle = \lambda \| \Phi(f) \|^2_{\mathcal{X}} = \lambda \| f \|^2_{\mathcal{X}^*} \). 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 \( \Phi \circ \Phi_* = \text{id}_{\mathcal{X}} \). By (iii) and Proposition 2.4, we have \( \Phi_* = \Phi_\text{inv} \).

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

Definition 13.16 Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{K} \), \( S \subseteq \mathcal{X} \). By Definition 7.95, \( S^\perp \subseteq \mathcal{X}^* \) is the set \( S^\perp = \{ x_\ast \in \mathcal{X}^* \mid \langle x_\ast, x \rangle = 0, \ \forall x \in S \} \).

We will abuse the notation and denote \( S^\perp = \{ y \in \mathcal{X} \mid \langle x, y \rangle = \langle y^\ast, x \rangle = 0, \ \forall x \in S \} \), which is said to be the orthogonal complement of \( S \).

Proposition 13.17 Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{K} \) and \( S, T \subseteq \mathcal{X} \). Then,

(i) \( S^\perp \subseteq \mathcal{X} \) 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-Frêchet Theorem 13.15, \( \mathcal{X} \) is reflexive. Then, \( \mathcal{X} = \mathcal{X}^{**} \). By Definition 7.95, \( \mathcal{X}^{**} \subseteq \mathcal{X}^* \) is well-defined and \( \mathcal{X}^{**} = S^{\perp} \). Then, by Proposition 7.98, we have \( S^{\perp \perp} = \mathcal{X}^{**} = \text{span}(S) \).

(iii) By (iv), we have \( S \subseteq \text{span}(S) = S^{\perp \perp} \). By (ii), we have \( S^{\perp \perp} \subseteq S^{\perp \perp} \). Again by (iv), we have \( S^{\perp \perp} \subseteq \text{span}(S^{\perp \perp}) = S^{\perp \perp} \). Therefore, \( S^{\perp \perp} = S^{\perp \perp} \). This completes the proof of the proposition. \( \square \)

Definition 13.18 Let \( \mathcal{X} \) be a vector space over the field \( \mathbb{F} \) and \( M, N \subseteq \mathcal{X} \) are subspaces. We say that \( \mathcal{X} \) is the direct sum of \( M \) and \( N \) if \( \forall x \in \mathcal{X} \), \( \exists m \in M \) and \( \exists n \in N \), such that \( x = m + n \). In this case, we will write \( \mathcal{X} = M \oplus N \).

Theorem 13.19 Let \( M \) be a closed subspace of a Hilbert space \( \mathcal{X} \). Then, \( \mathcal{X} = M \oplus M^{\perp} \).
Proposition 13.20 Let $X$ be a real Hilbert space, $f : X \times X \to \mathbb{R}$ be given by $f(x, y) = \langle x, y \rangle$, $\forall (x, y) \in X \times X$. Then, $f$ is analytic with arbitrarily large analytic radius, $f^{(1)}(x, y)(h_x, h_y) = \langle h_x, y \rangle + \langle x, h_y \rangle$, $f^{(2)}(x, y)(h_{x_1}, h_{y_1}) = \langle h_{x_1}, h_{y_1} \rangle + \langle h_{x_2}, h_{y_2} \rangle$, and $f^{(i+2)}(x, y) = \vartheta_{\mathcal{B}_i}^{(2)}(X \times \mathcal{X}, \mathcal{Y})$, $\forall (x, y) \in X \times X$, $\forall i \in \mathbb{N}$, $\forall (h_x, h_y) \in X \times X$, $\forall (h_{x_1}, h_{y_1}) \in X \times X$.

Proof This follows immediately from Propositions 9.41 and 9.27 and Riesz-Frèchet Theorem 13.15. □

13.4 Hermitian Adjoints

Definition 13.21 Let $X$ and $Y$ be Hilbert spaces over $\mathbb{K}$, $A \in \mathcal{B}(X, Y)$, and $\Phi_X : X^* \to X$ and $\Phi_Y : Y^* \to Y$ as defined in Riesz-Frèchet Theorem 13.15. Define the Hermitian adjoint of $A$ by $A^* := \Phi_X \circ A^T \circ \Phi_Y$, $\forall \lambda \in \mathbb{K}$. $A$ is said to be Hermitian if $A^* = A$.

Proposition 13.22 Let $X$, $Y$, and $Z$ be Hilbert spaces over $\mathbb{K}$, $A, B \in \mathcal{B}(X, Y)$, $C \in \mathcal{B}(Y, Z)$, $x \in X$, $y \in Y$, and $\lambda \in \mathbb{K}$. Then,

(i) $\text{id}_X^* = \text{id}_X$;
(ii) $(A + B)^* = A^* + B^*$;
(iii) $(\lambda A)^* = \overline{\lambda} A^*$;
(iv) $(CA)^* = A^* C^*$;
(v) if $A$ is bijective, then $(A^*)^{-1} = (A^{-1})^* =: A^{-*}$;
(vi) $\langle Ax, y \rangle_Y = \langle x, A^* y \rangle_X$;
(vii) $\langle y, Ax \rangle_Y = \langle A^* y, x \rangle_X$;
(viii) $(A^*)^* = A$;
(ix) $\|A^*\|_{\mathcal{B}(Y, X)} = \|A\|_{\mathcal{B}(X, Y)}$;
(x) $(\mathcal{R}(A))^\perp = \mathcal{N}(A^*)$ and $(\mathcal{R}(A^*))^\perp = \mathcal{N}(A)$;
(xi) if $\mathcal{R}(A) \subseteq X$ is closed, then $\mathcal{R}(A) = (\mathcal{N}(A^*))^\perp$ and $\mathcal{R}(A^*) = (\mathcal{N}(A))^\perp$. 
(xii) Let \((X, \mathcal{B})\) be a measurable space and \(\mu\) be a \(\mathcal{B}(\mathbb{Y}, \mathbb{Z})\)-valued measure on \((X, \mathcal{B})\). Define \(\mu^*\) to be a mapping from \(\mathcal{B}\) to \(\mathcal{B}(\mathbb{Z}, \mathbb{Y})\) 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 \(\mathcal{B}(\mathbb{Z}, \mathbb{Y})\)-valued measure on \((X, \mathcal{B})\) with \(\mathcal{P} \circ \mu^* = \mathcal{P} \circ \mu\).

Proof Let \(\Phi_X : X^* \rightarrow \mathbb{X}\), \(\Phi_Y : Y^* \rightarrow \mathbb{Y}\), and \(\Phi_Z : Z^* \rightarrow \mathbb{Z}\) be as defined in Riesz-Fréchet Theorem 13.15.

(i) \(\text{id}_X^* = \Phi_X \circ \text{id}_{X^*} \circ \Phi_{X^{-1}} = \Phi_X \circ \text{id}_{X^*} \circ \Phi_{X^{-1}} = \text{id}_X\), where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(ii) \((A + B)^* = \Phi_X \circ (A + B)^* \circ \Phi_{Y^{-1}} = \Phi_X \circ (A^* + B^*) \circ \Phi_{Y^{-1}} = \Phi_X \circ A^* \circ \Phi_{Y^{-1}} + \Phi_X \circ B^* \circ \Phi_{Y^{-1}} = A^* + B^*\), 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) \((\lambda A)^* = \Phi_X \circ (\lambda A)^* \circ \Phi_{Y^{-1}} = \Phi_X \circ (\lambda A^*) \circ \Phi_{Y^{-1}} = \lambda \Phi_X \circ A^* \circ \Phi_{Y^{-1}} = \lambda A^*\), 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) \((CA)^* = \Phi_X \circ (CA)^* \circ \Phi_{Y^{-1}} = \Phi_X \circ (C^* A) \circ \Phi_{Y^{-1}} = \Phi_X \circ A^* \circ C^* \circ \Phi_{Y^{-1}} = \Phi_X \circ A^* \circ C^* \circ \Phi_{Y^{-1}} = A^* C^*\), 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, \(A^{-1} \in \mathcal{B}(\mathbb{Y}, \mathbb{X})\). Then, we have \((A^{-1})^* = \Phi_Y \circ (A^{-1})^* \circ \Phi_{X^{-1}} = \Phi_Y \circ (A')^{-1} \circ \Phi_{X^{-1}} = (\Phi_X \circ A^* \circ \Phi_{Y^{-1}})^{-1} = (A^*)^{-1}\), where the first equality follows from Definition 13.21; and the second equality follows from Proposition 7.110.

(vi) \(\langle Ax, y \rangle_y = \langle (y^*, Ax) \rangle_y = \langle (A y^*, x) \rangle_X = \langle x, \Phi_X(A y^*) \rangle_X = \langle x, \Phi_X(A^* y^*) \rangle_X = \langle x, A^* y, x \rangle_X\), 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) \(\langle y, Ax \rangle_y = \langle (Ax, y^*)_y \rangle_X = \langle (x, A^* y, x) \rangle_X\), 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) \(\langle y, Ax \rangle_y = \langle (A^* y, x) \rangle_X = \langle y, (A^*)^* x \rangle_y\), where the first equality follows from (vii); and the second equality follows from (vi). By the arbitrariness of \(x\) and \(y\), we have \(A = (A^*)^*\).

(ix) \(\|A^*\|_{\mathcal{B}(\mathbb{Y}, \mathbb{X})} = \sup_{\|y\|_{\mathbb{Y}} \leq 1} \|A^* y\|_{\mathbb{X}} = \sup_{\|y\|_{\mathbb{Y}} \leq 1} \|\Phi_X \circ A^* \circ \Phi_{Y^{-1}}(y)\|_{\mathbb{X}} = \|A\|_{\mathcal{B}(\mathbb{Y}, \mathbb{X})} = \|A\|_{\mathcal{B}(\mathbb{X}, \mathbb{Y})}\), 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, \((\mathcal{R}(A^*))^\perp = \mathcal{N}((A^*)^*) = \mathcal{N}(A), \) where the first equality follows from the previous sentence; and the second equality
follows from (viii).

(xii) By Proposition 7.114 and Definition 13.16, we have \( \mathcal{R}(A) = \langle \mathcal{N}(A') \rangle = \langle \Phi_y(\mathcal{N}(A')) \rangle = \langle \Phi_y(\mathcal{N}(A')) \rangle^\perp = (\mathcal{N}(A^*))^\perp\). Furthermore, \( \mathcal{R}(A^*) = (\mathcal{N}(A^*))^\perp = (\mathcal{N}(A))^\perp\), where the first equality follows from the previous sentence; the second equality follows from (viii).

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 \( K \) and \( S \subseteq X \) be an orthogonal set. Assume that \( \vartheta_X \notin S \). Then, \( S \) is a linearly independent set.
Hence, we have \( \text{span} \left( \{ x_i \} \right) \). Then, \( \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K}, x_0 = \sum_{i=1}^{n} \alpha_i x_i. \) 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 \( x_0 \neq \varnothing \). 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 \( n \in \mathbb{N} \cup \{ \infty \}, X \) be a pre-Hilbert space over \( \mathbb{K} \), and \( \{ x_i \}_{i=1}^{n} \subseteq X \) be a linearly independent set and vectors in the sequence are distinct. Then, there is an orthonormal sequence \( \{ e_i \}_{i=1}^{n} \subseteq X \) 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 \( k = 1 \), clearly \( x_1 \neq \varnothing \) since \( \{ x_i \}_{i=1}^{n} \) is a linearly independent set. Let \( e_1 := x_1 / \| x_1 \| \in X \). 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 \( z_{k+1} := x_{k+1} - \sum_{i=1}^{k} \langle x_{k+1}, e_i \rangle e_i \in X \). It is easy to show that \( \langle z_{k+1}, e_i \rangle = 0, \forall i \in \{ 1, \ldots, k \} \). Then, \( z_{k+1} \perp e_i. \) By the assumption and Lemma 6.49, the vectors \( x_1, \ldots, x_{k+1} \) are linearly independent. Then, \( z_{k+1} \neq \varnothing \). Define \( e_{k+1} := z_{k+1} / \| z_{k+1} \| \in X \). 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 \( \{ e_i \}_{i=1}^{n} \subseteq X \) 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, \( x \in \text{span} \{ ( x_1, \ldots, x_m ) \} = \text{span} \{ ( e_1, \ldots, e_m ) \} \subseteq \text{span} \{ ( e_i )_{i=1}^{\infty} \}. \) 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 $K$, $n \in \mathbb{N}$, $y_1, \ldots, y_n \in X$. 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 $K$, $y_1, \ldots, y_n \in X$. 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, $\exists \alpha_1, \ldots, \alpha_n \in K$, 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 $\exists \alpha_1, \ldots, \alpha_n \in K$, which are not all zeros, such that

$$0_{1 \times n} = \begin{bmatrix}
\alpha_1 & \cdots & \alpha_n
\end{bmatrix} \text{Gram}(y_1, \ldots, y_n)$$

Then,

$$0 = \begin{bmatrix}
\langle \sum_{i=1}^n \alpha_i y_i, y_1 \rangle & \cdots & \langle \sum_{i=1}^n \alpha_i y_i, y_n \rangle
\end{bmatrix} \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}$$

$$= \sum_{j=1}^n \alpha_j \left( \sum_{i=1}^n \alpha_i y_i, y_j \right) = \sum_{i=1}^n \alpha_i y_i, \sum_{j=1}^n \alpha_j y_j$$

$$= \| \sum_{i=1}^n \alpha_i y_i \|^2$$

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. $\Box$
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

$$\begin{align*}
(G(y_1, \ldots, y_n))' \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix} &= \begin{bmatrix}
\langle x, y_1 \rangle \\
\vdots \\
\langle x, y_n \rangle
\end{bmatrix} \\
(\text{Gram}(y_1, \ldots, y_n))' \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix} &= \begin{bmatrix}
\langle x, y_1 \rangle \\
\vdots \\
\langle x, y_n \rangle
\end{bmatrix} \tag{13.1}
\end{align*}$$

(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, $(x - \sum_{i=1}^n \alpha_i y_i, y_j) = 0$, $j = 1, \ldots, n$. 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 $\delta^2 = \|x - m_0\|^2 = \langle x - m_0, x - m_0 \rangle = \langle x - m_0, x \rangle - \sum_{i=1}^n \alpha_i \langle y_i, x \rangle$, 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.

$$\begin{align*}
\begin{bmatrix}
\langle y_1, x \rangle & \cdots & \langle y_n, x \rangle
\end{bmatrix} \begin{bmatrix}
0_{n \times 1} \\
\vdots \\
1
\end{bmatrix} &= \begin{bmatrix}
\langle x, y_1 \rangle \\
\vdots \\
\langle x, y_n \rangle
\end{bmatrix} \\
\langle y_1, x \rangle & \cdots & \langle y_n, x \rangle
\end{align*}$$

By Cramer’s Rule, we have $\delta^2 = \frac{\det(G_1)}{\det(G)}$, where

$$G = \begin{bmatrix}
\langle y_1, y_1 \rangle & \cdots & \langle y_n, y_1 \rangle & 0 \\
\vdots & \ddots & \vdots & \vdots \\
\langle y_1, y_n \rangle & \cdots & \langle y_n, y_n \rangle & 0 \\
\langle y_1, x \rangle & \cdots & \langle y_n, x \rangle & 1
\end{bmatrix} ; \quad G_1 = (\text{Gram}(y_1, \ldots, y_n, x))'$$

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. \qed
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 \( \mathcal{X} \) be Hilbert space over \( \mathbb{K} \), \( (e_i)_{i=1}^{\infty} \subseteq \mathcal{X} \) be orthonormal, and \( (\xi_i)_{i=1}^{\infty} \subseteq \mathbb{K} \). Then, \( x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X} \) 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 \( s_i := \sum_{j=1}^{i} \xi_j e_j \in \mathcal{X} \), \( \forall i \in \mathbb{N} \).

"Sufficiency" Let \( \sum_{i=1}^{\infty} |\xi_i|^2 < \infty \). \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists n_0 \in \mathbb{N} \) such that \( \sum_{i=n_0+1}^{\infty} |\xi_i|^2 < \varepsilon^2 \). \( \forall n, m \in \mathbb{N} \) with \( n_0 \leq n \leq m \), \( \|s_n - s_m\|^2 = \langle \sum_{i=n_0+1}^{m} \xi_i e_i, \sum_{i=n_0+1}^{m} \xi_i e_i \rangle = \sum_{i=n_0+1}^{m} |\xi_i|^2 < \varepsilon^2 \), where the second equality follows from the assumption that \( (e_i)_{i=1}^{\infty} \) is orthonormal. Hence, \( \|s_n - s_m\| < \varepsilon \) and \( (s_i)_{i=1}^{\infty} \subseteq \mathcal{X} \) is a Cauchy sequence. By the completeness of \( \mathcal{X} \), \( \sum_{i=1}^{\infty} \xi_i e_i = \lim_{n \in \mathbb{N}} s_n =: x \in \mathcal{X} \).

"Necessity" Let \( x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X} \). Then, \( x = \lim_{n \in \mathbb{N}} s_n \) and \( (s_n)_{n=1}^{\infty} \subseteq \mathcal{X} \) is a Cauchy sequence. \( \forall \varepsilon \in (0, \infty) \subset \mathbb{R} \), \( \exists n_0 \in \mathbb{N} \), \( \forall n, m \in \mathbb{N} \) with \( n_0 \leq n \leq m \), we have \( \|s_n - s_m\| < \sqrt{\varepsilon} \). Then, \( \sum_{i=n_0+1}^{\infty} |\xi_i|^2 = \langle \sum_{i=n_0+1}^{\infty} \xi_i e_i, \sum_{i=n_0+1}^{\infty} \xi_i e_i \rangle = \|s_n - s_m\|^2 < \varepsilon \). Hence, \( \sum_{i=1}^{\infty} |\xi_i|^2 < \infty \).

When \( x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X} \), by Propositions 13.4 and 3.6, 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. \( \square \)

**Proposition 13.30** (Bessel’s Inequality) Let \( \mathcal{X} \) be a pre-Hilbert space over \( \mathbb{K} \), \( x \in \mathcal{X} \), and \( (e_i)_{i=1}^{\infty} \subseteq \mathcal{X} \) be orthonormal. Then, \( \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 \).

**Proof** Let \( \alpha_i := \langle x, e_i \rangle \in \mathbb{K}, \forall i \in \mathbb{N} \). \( \forall n \in \mathbb{N} \), \( 0 \leq \|x - \sum_{i=1}^{n} \alpha_i e_i\|^2 = \|x\|^2 - \sum_{i=1}^{n} \alpha_i \langle e_i, x \rangle - \sum_{i=1}^{n} \overline{\alpha_i} \langle x, e_i \rangle + \sum_{i=1}^{n} |\alpha_i|^2 = \|x\|^2 - \sum_{i=1}^{n} |\alpha_i|^2 \). Hence, \( \sum_{i=1}^{\infty} |\langle x, e_i \rangle|^2 \leq \|x\|^2 \). \( \square \)

**Proposition 13.31** Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{K} \), \( x \in \mathcal{X} \), and \( (e_i)_{i=1}^{\infty} \subseteq \mathcal{X} \) be orthonormal. Then, \( m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in \text{span}(\{e_i\}_{i=1}^{\infty}) =: M, \langle x - m_0 \rangle \perp M \), 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, \( m_0 := \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in \mathcal{X} \). Since \( \sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \in M \), \( \forall n \in \mathbb{N} \), then \( m_0 \in M \) by Proposition 4.13. By Propositions 13.4 and 3.6, \( \langle x - m_0, e_j \rangle = \lim_{n \in \mathbb{N}} \langle x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle \|e_j\|^2 = 0 \), \( \forall j \in \mathbb{N} \). Hence, \( x - m_0 \in \text{span}(\{e_i\}_{i=1}^{\infty}) = (\{e_i\}_{i=1}^{\infty})_{\perp \perp} \subseteq (\text{span}(\{e_i\}_{i=1}^{\infty}))_{\perp} = M_{\perp} \), 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. \( \square \)
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 \( \text{span} \left( \{ e_i \}_{i=1}^n \right) = X \).

**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, \((e_i)_{i=1}^n \) is an orthonormal set.

**Proof** By Proposition 13.17, \((e_i)_{i=1}^n \) is orthonormal if, and only if, \((e_i)_{i=1}^n \) is complete.

**Example 13.34** Let \( I := [-1, 1] \subset \mathbb{R} \), \( I := ((J, |\cdot|), \mathcal{B}, \mu) \) be the complete metric measure subspace of \( \mathbb{R} \), \( Z := L_2(I, K) \), \( z_i \in L_2(I, K) \) be defined by \( z_i(t) = t^i \), \( \forall t \in I \), \( \forall i \in \mathbb{Z}_+ \). It is easy to show that \( \{ [z_i] \}_{i=0}^\infty \subseteq Z \) is a linearly independent set. Then, the Gram-Schmidt Procedure 13.25 can be applied to this sequence to produce an orthonormal sequence \((e_i)_{i=1}^\infty \subseteq Z \), where \( e_i \in Z \), \( \forall i \in \mathbb{Z}_+ \), are \( i \)-th order polynomials. Let \( p_i \in Z \) be defined by \( p_i(t) = \frac{(-1)^i}{2i!} (1 - t^2)^i \), \( \forall t \in I \), \( \forall i \in \mathbb{Z}_+ \), and \( P_i \in Z \) 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 \( \{ [P_i] \}_{i=0}^\infty \subseteq Z \) is a complete orthonormal sequence. As defined, \( \{ P_i \}_{i=0}^\infty \subseteq Z \) 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

\[
\langle [P_n], [P_m] \rangle = \int_{-1}^{1} P_n(t) P_m(t) \, dt = \int_{-1}^{1} p_n^{(n)}(t) p_m^{(m)}(t) \, dt
\]

\[
= \left. p_n^{(n-1)}(t) p_m^{(m)}(t) \right|_{-1}^{1} - \int_{-1}^{1} p_n^{(n-1)}(t) p_m^{(m+1)}(t) \, dt
\]

\[
= - \int_{-1}^{1} p_n^{(n-1)}(t) p_m^{(m+1)}(t) \, dt
\]

\[
\vdots
\]

\[
= (-1)^n \int_{-1}^{1} p_n^{(0)}(t) p_m^{(m+n)}(t) \, dt
\]

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 \( p_n^{(0)}(x) = 0 \), \( \forall i \in \mathbb{Z}_+ \) 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 \( p_m^{(m+n)} = \vartheta_z \) since \( p_m \) is a polynomial of order \( 2m < m + n \). Then, \( \langle [P_n], [P_m] \rangle = 0 \). Hence, \( \langle [P_i] \rangle_{i=0}^\infty \) is an orthogonal set.

Next, we show that \( \| [P_i] \|_2 \leq \frac{\sqrt{2}}{2^i+1} \), \( \forall i \in \mathbb{Z}_+ \). Clearly, \( P_0(t) = 1 \), \( \forall t \in I \) and \( \| [P_0] \|_2^2 = \langle [P_0], [P_0] \rangle = 2 = \frac{2}{2^0+1} \). Recursively, assume that \( \| [P_{n-1}] \|_2 \leq \frac{\sqrt{2}}{2^{n-1}} \), for some \( n - 1 \in \mathbb{Z}_+ \). Then,

\[
\| [P_n] \|_2^2 = \langle [P_n], [P_n] \rangle = (-1)^n \int_{-1}^1 p_n^{(0)}(t)p_n^{(2n)}(t) \, dt
\]

\[
= \frac{(2n)!}{2^{2n}n!} \int_{-1}^1 (1 - t^2)^n \, dt
\]

\[
= \frac{(2n)!}{2^{2n}n!} \left( \frac{1}{2} \right) + \frac{2n(2n)!}{2^{2n}n!} \int_{-1}^1 t^2(1 - t^2)^{n-1} \, dt
\]

\[
= \frac{2n(2n)!}{2^{2n}n!} \left( \int_{-1}^1 (1 - t^2)^{n-1} \, dt - \int_{-1}^1 (1 - t^2)^n \, dt \right)
\]

\[
= (2n - 1) \| [P_{n-1}] \|_2^2 - 2n \| [P_n] \|_2^2
\]

where the second equality follows from the previous paragraph; the third equality follows from the definition of \( p_n \in \mathcal{Z} \); 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 \( \| [P_n] \|_2^2 \leq \frac{\sqrt{2}}{2^{n+1}} \). 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 \( \forall z \in \mathcal{Z} \) with \( \{ [z], [z_i] \} = 0, \forall i \in \mathbb{Z}_+ \), we have \( [z] = \vartheta_z \). Since \( \mu(I) = 2 < \infty \), by Cauchy-Schwarz Inequality, we have \( \int_{-1}^1 \mathcal{P} \circ z \, dt \leq (\int_{-1}^1 \mathcal{P}_2 \circ z \circ dt)^{1/2}(\int_{-1}^1 1 \, dt)^{1/2} = \sqrt{2} \| [z] \|_2 < \infty \). Then, \( z \in \mathcal{L}_1(\mathbb{I}, \mathcal{K}) \). Define \( F : I \rightarrow \mathcal{K} \) by \( F(t) = \int_{-1}^t z(r) \, dr \). \( F \) is absolutely continuous and therefore continuous. Since \( [z] \perp [z_0] \), then \( F(-1) = 0 = F(1) = \langle [z], [z_0] \rangle \). By Proposition 11.37, \( F \) is \( \mathcal{B} \)-measurable. Then, \( F \in \mathcal{Z} \). \( \forall i \in \mathbb{Z}_+ \), we have

\[
\langle [F], [z_i] \rangle = \int_{-1}^1 F(t) z_i(t) \, dt = \frac{t^{i+1}}{i+1} F(t) \bigg|_{-1}^1 - \int_{-1}^1 \frac{t^{i+1}}{i+1} z(t) \, dt = 0
\]

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 \( \| F - Q \|_{c(I, K)} < \epsilon \). Then,

\[
\| [F] \|^2 = \int_{-1}^{1} |F(t)|^2 \, dt = \int_{-1}^{1} F(t)(F(t) - Q(t)) \, dt \\
\leq \| [F] \| \| [F(t) - Q(t)] \| \leq \| [F] \| \sqrt{2} \epsilon
\]

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 \( [F] = \theta_z \). Since \( F \) is continuous, then \( F(t) = 0 \), \( \forall t \in I \). Then, by Fundamental Theorem of Calculus I 12.86, \( z = F^{(1)} \) a.e. in \( I \). Then, \( [z] = \theta_z \).

Then, by Proposition 13.33, \( \{ [e_i] \}_{i=0}^{\infty} \) is complete.

\( \diamond \)

Proposition 13.35 Let \( S_1 = [0, 2\pi] \subset \mathbb{R} \) and \( z \in C(S_1, \mathbb{C}) = : Z \). Assume that \( z(0) = z(2\pi) \). Let \( M = \{ \tilde{z} \in Z \mid \tilde{z}(x) = \cos(nx), \forall x \in S_1 \text{ or } \tilde{z}(x) = \sin(nx), \forall x \in S_1 \text{, where } n \in \mathbb{Z}_+ \} \subset Z \). Let \( \mathcal{M} = \text{span}(M) \subset Z \). Then, \( z \in \overline{M} \).

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 \( \Phi : S_2 \to \mathbb{C} \) such that \( \Phi \circ \Psi = z \). 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, \( \Phi \in C(S_2, \mathbb{C}) \). Let \( \Phi_r := \text{Re} \circ \Phi \in C(S_2, \mathbb{R}) \) and \( \Phi_i := \text{Im} \circ \Phi \in C(S_2, \mathbb{R}) \). 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 \( \Phi_r |_{S_2} = \Phi_r \) and \( \Phi_i |_{S_2} = \Phi_i \). Let \( \Phi := \Phi_r + i\Phi_i \in C(I^2, \mathbb{C}) \). Then, \( \Phi |_{S_2} = \Phi \). 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 \( \| \Phi(y_1, y_2) - P(y_1, y_2) \| < \epsilon \), \( \forall y_1, y_2 \in I^2 \). 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, \( \forall \gamma, \theta \in \mathbb{R} \),

\[
(s^2) = \frac{1}{2} (1 - \cos(2\theta)); \quad (\cos(\theta))^2 = \frac{1}{2} (1 + \cos(2\theta)); \\
\sin(\gamma) \cos(\theta) = \frac{1}{2} (\sin(\gamma + \theta) + \sin(\gamma - \theta)); \\
\sin(\gamma) \sin(\theta) = \frac{1}{2} (\cos(\gamma - \theta) - \cos(\gamma + \theta)); \\
\cos(\gamma) \cos(\theta) = \frac{1}{2} (\cos(\gamma - \theta) + \cos(\gamma + \theta))
\]

Hence, \( z \in \overline{M} \), by the arbitrariness of \( \epsilon \). This completes the proof of the proposition.

\( \square \)
Example 13.36 Let $\mathcal{I} := (\mathbb{R} \times (-\pi, \pi], \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$; $\mathcal{Z} := L_2(\mathcal{I}, \mathbb{K})$ be the Hilbert space with inner product $\langle \cdot, \cdot \rangle_\mathcal{Z}$ defined by $\langle [g], [h] \rangle_\mathcal{Z} = \frac{1}{\pi} \int g(t)h(t) dt$, $\forall g, h \in \mathcal{Z}$; $\mathcal{Z} := L_2(\mathcal{I}, \mathbb{K})$; $(e_n)_{n=0}^\infty \in \mathcal{Z}$ be defined by $e_0(t) = \frac{1}{\sqrt{2\pi}}$, $\forall t \in \mathcal{I}$, $e_{2n-1}(t) = \sin(nt), \forall t \in \mathcal{I}$, $e_{2n}(t) = \cos(nt), \forall t \in \mathcal{I}$, $\forall n \in \mathbb{N}$; and $f \in \mathcal{Z}$. 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, $\sum_{n=1}^\infty a_n e_n$ is called the Fourier series for $f$.

By the finiteness of measure $\mu$, we have $L_2(\mathcal{I}, \mathbb{Y}) \subseteq L_1(\mathcal{I}, \mathbb{Y})$. Then, $f$ is absolutely integrable over $\mathcal{I}$. By Propositions 7.23, 11.38, and 11.39, $f e_n$ is $\mathcal{B}$-measurable and therefore absolutely integrable over $\mathcal{I}$, $\forall n \in \mathbb{Z}_+$. 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. We will show that it is complete by Proposition 13.33.

Fix any $[z] \in (\{e_n\}_{n=0}^\infty)_{\perp}^\perp$, where $z \in \mathcal{Z}$. Then, $z$ is $\mathcal{B}$-measurable and absolutely integrable over $\mathcal{I}$. Define $F : \mathcal{T}_{-\pi, \pi} \rightarrow \mathbb{K}$ by $F(t) = \int_{-\pi}^{\pi} z(t) dt$, $\forall t \in \mathcal{T}_{-\pi, \pi}$. By Proposition 12.75, $F$ is absolutely continuous. Clearly $F(-\pi) = 0$. Since $[z] \perp [e_0]$, then $F(\pi) = 0. \forall n \in \mathbb{N}$,

$$\langle [F], [e_{2n-1}] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \sin(nt) dt$$

$$= -\frac{1}{n\pi} \cos(nt) F(t)|_{t=-\pi}^{t=\pi} + \frac{1}{n\pi} \int_{-\pi}^{\pi} \cos(nt) z(t) dt = 0$$

where the second equality follows from Integration by Parts Theorem 12.89; and the last equality follows from the preceding discussion and the fact $[z] \perp [e_{2n}], \forall n \in \mathbb{N}$,

$$\langle [F], [e_{2n}] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \cos(nt) dt$$

$$= \frac{1}{n\pi} \sin(nt) F(t)|_{t=-\pi}^{t=\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} \sin(nt) z(t) dt = 0$$

where the second equality follows from Integration by Parts Theorem 12.89; and the last equality follows from the preceding discussion and the fact $[z] \perp [e_{2n-1}]$. Note also that

$$\langle [F], [e_0] \rangle_\mathcal{Z} = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \frac{1}{\sqrt{2}} dt$$

$$= \frac{1}{\sqrt{2\pi}} \left[ F(t) \right]_{t=-\pi}^{t=\pi} - \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} t z(t) dt = -\frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} t z(t) dt$$

Consider the function $g : \mathbb{R} \rightarrow \mathbb{K}$ defined by $g(t) = t$, $\forall t \in \mathbb{R} \times (-\pi, \pi] + 2n\pi, \forall n \in \mathbb{Z}$. Then $g \in PC(2\pi)$ as discussed in Bartle (1976). It is an odd function, then its Fourier series has only $a_{2n-1} \in \mathbb{K}$ associate with $e_{2n-1}$, $\forall n \in \mathbb{N}$. The rest of the constants are 0. Let its Fourier series be given by
\[ \sum_{n=1}^{\infty} \bar{a}_{2n-1} e_{2n-1}. \]

By Norm Convergence Theorem 38.10 of Bartle (1976), we have \( \lim_{k \in \mathbb{N}} \| \left[ g|_I - \sum_{n=1}^{k} \bar{a}_{2n-1} e_{2n-1} \right] \|_Z = 0. \) Then, we have the following line of arguments. Clearly \( g|_I \in \hat{Z}, \) then \( \sum_{n=1}^{k} \bar{a}_{2n-1} e_{2n-1} \in \hat{Z}, \) \( \forall k \in \mathbb{N}. \) \( \forall \in \mathbb{R}^+, \exists k_0 \in \mathbb{N} \) such that \( \forall k \in \mathbb{N} \) with \( k_0 \leq k, \) we have \( \| \left[ g|_I - \sum_{n=1}^{k} \bar{a}_{2n-1} e_{2n-1} \right] \|_Z < \epsilon. \) Then,

\[
\begin{align*}
\frac{1}{\pi} \left| \int_{-\pi}^{\pi} tz(t) \, dt \right| &= \left| \langle \left[ z \right], \left[ g|_I \right] \rangle_Z - \left[ \left[ z \right], \sum_{n=1}^{k_0} \bar{a}_{2n-1} [e_{2n-1}] \right] \right|_Z \\
&= \left| \left[ \left[ z \right], \left[ g|_I - \sum_{n=1}^{k_0} \bar{a}_{2n-1} e_{2n-1} \right] \right] \right|_Z \\
&\leq \| \left[ z \right] \|_Z \left[ g - \sum_{n=1}^{k_0} \bar{a}_{2n-1} e_{2n-1} \right] \|_Z \leq \| \left[ z \right] \|_Z \epsilon
\end{align*}
\]

where the first equality follows from the fact that \( \left[ z \right] \perp [e_n], \) \( \forall n \in \mathbb{Z}^+. \) the second equality follows from Definition 13.1; and the first inequality follows from Cauchy-Schwarz Inequality. By the arbitrariness of \( \epsilon, \) we have \( \langle \left[ F|_I \right], [e_n] \rangle_Z = 0 \) and \( \left[ F|_I \right] \perp [e_0]. \)

In the above, we have shown that \( \left[ F|_I \right] \in \left( ([e_n])_{n=0}^{\infty} \right)^{\perp}. \) Clearly, \( F|_I \) may be periodically extended to a continuous function of \( \mathbb{R} \) to \( K \) that is periodic with period \( 2\pi. \) By Norm Convergence Theorem 38.10 of Bartle (1976) (applied to the real and imaginary part of \( F), \) the partial sum of the Fourier series for \( F, \) denoted by \( \sum_{n=1}^{k} \hat{a}_n e_n \) converges to \( F \) in \( \| \cdot \|_Z \) as \( k \to \infty. \) \( \forall \in \mathbb{R}^+, \exists k_1 \in \mathbb{N}, \forall k \in \mathbb{N} \) with \( k_1 \leq k, \) we have \( \left[ \left[ F|_I - \sum_{n=1}^{k} \hat{a}_n e_n \right] \right] \|_Z < \epsilon. \) Then,

\[
\begin{align*}
\| \left[ F|_I \right] \|_Z^2 &= \langle \left[ F|_I \right], \left[ F|_I \right] \rangle_Z = \left[ \left[ F|_I \right], \left[ F|_I \right] - \sum_{n=1}^{k} \hat{a}_n [e_n] \right] \rangle_Z \\
&= \left[ \left[ F|_I \right], \left[ F|_I - \sum_{n=1}^{k} \hat{a}_n e_n \right] \right] \leq \| \left[ F|_I \right] \|_Z \left[ \left[ F|_I - \sum_{n=1}^{k} \hat{a}_n e_n \right] \right] \|_Z \\
&\leq \| \left[ F|_I \right] \|_Z \epsilon
\end{align*}
\]

where the second equality follows from the fact that \( \left[ F|_I \right] \in \left( ([e_n])_{n=0}^{\infty} \right)^{\perp}; \) and the first inequality follows from Cauchy-Schwarz Inequality. By the arbitrariness of \( \epsilon, \) we have \( \left[ F|_I \right] = \theta_Z. \) Then, \( F|_I = 0 \) a.e. in \( \mathcal{I}. \) Since \( F \) is absolutely continuous, then it is continuous by Proposition 12.59, and then \( F(t) = 0, \forall t \in \mathcal{I} \) (otherwise, we can always find a nonempty interval that is a subset of \( \mathcal{I} \) such that \( F \) is nonzero on the interval. Then, by Fundamental Theorem of Calculus I, Theorem 12.86, \( z(t) = 0 \) a.e. \( t \in \mathcal{I}. \) Hence, \( \left[ z \right] = \theta_Z. \)
We have shown that \( (\{e_n\})_{n=0}^{\infty} \perp = \{\vartheta\} \). Then, by Proposition 13.33, \( \text{span} (\{e_n\})_{n=0}^{\infty} = \mathcal{Z} \), and \( (\{e_n\})_{n=0}^{\infty} \) is a complete orthonormal sequence. Then \( f = \sum_{n=0}^{\infty} a_n [e_n] \) in \( \mathcal{Z} \).

\[ e_1 = \frac{y_1}{\|y_1\|}; \quad e_k = \frac{y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i}{\|y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i\|}; \quad \forall k > 1 \]

The vector \( y_k - \sum_{i=1}^{k-1} \langle y_k, e_i \rangle e_i \) 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 \( \mathcal{M} \) 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
13.6. OTHER MINIMUM NORM PROBLEMS

\[ \| x - v_0 \| = \min_{v \in V} \| x - v \|. \] 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. \( \square \)

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 \( V = \{ v \in \mathcal{X} \mid \langle v, y_i \rangle = c_i, \ i = 1, \ldots, n \} \), where \( y_1, \ldots, y_n \in \mathcal{X} \) 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 \( M^{\perp \perp} = M \) is finite-dimensional by Proposition 13.17.

**Theorem 13.38** Let \( \mathcal{X} \) be a Hilbert space over \( \mathbb{K} \), \( n \in \mathbb{N} \), \( y_1, \ldots, y_n \in \mathcal{X} \), \( c_1, \ldots, c_n \in \mathbb{K} \), \( V = \{ v \in \mathcal{X} \mid \langle v, y_i \rangle = c_i, \ i = 1, \ldots, n \} \neq \emptyset \), \( x \in \mathcal{X} \).
Then, the following equation
\[(\text{Gram}(y_1, \ldots, y_n))' \begin{bmatrix}
\beta_1 \\
\vdots \\
\beta_n
\end{bmatrix} = \begin{bmatrix}
\langle x, y_1 \rangle - c_1 \\
\vdots \\
\langle x, y_n \rangle - c_n
\end{bmatrix} \quad (13.2)\]
adopts 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 \(\min_{v \in V} \| x - v \| \).

**Proof** Let \(M := \text{span} (\{y_1, \ldots, y_n\})\). 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, \((x - v_0) \in M^\perp = M\), 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 \((x - v_0) \in M = M^\perp\). 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. \(\square\)
13.6. OTHER MINIMUM NORM PROBLEMS

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 $\langle x - k_0, k - k_0 \rangle \leq 0$, $\forall k \in K$.

![Figure 13.4: Projection to a convex set.](image)

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°.

**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 $\lim_{i \in \mathbb{N}} \|x - k_i\| = \delta$. $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall i \in \mathbb{N}$ with $n_0 \leq i$, we have $\delta \leq \|x - k_i\| < \sqrt{\delta^2 + \epsilon^2}/4$. $\forall i, j \in \mathbb{N}$ with $n_0 \leq i \leq j$, 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 $\bar{k} \in K$ be such that $\delta = \|x - \bar{k}\|$. The sequence $(k_i)_{i=1}^{\infty} := (k_0, \bar{k}, k_0, \bar{k}, \ldots) \subseteq K$ satisfies $\lim_{i \in \mathbb{N}} \|x - k_i\| = \delta$. 

![Figure 13.4: Projection to a convex set.](image)
By the proof of the existence, \((k_i)_{i=1}^{\infty}\) is convergent. Then, we must have \(k = k_0\). 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 \(\|x - k_0\|^2 = \|(x - k_0) - \alpha (k_1 - k_0)\|^2 = \|x - k_0\|^2 - 2\alpha \langle x - k_0, k_1 - k_0 \rangle + \alpha^2 \|k_1 - k_0\|^2 = \delta^2 - 2\alpha\lambda + \alpha^2 \|k_1 - k_0\|^2\). 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 \(k \in K\) satisfies \(\langle x - k, k - k \rangle \leq 0\), \(\forall k \in K\), then \(\|x - k\| = \min_{k \in K} \|x - k\|\), \(\forall k \in K\), we have \(\|x - k\|^2 = \|(x - \tilde{k}) - (k - \tilde{k})\|^2 = \|x - \tilde{k}\|^2 - 2 \langle x - \tilde{k}, k - \tilde{k} \rangle + \|k - \tilde{k}\|^2 \geq \|x - \tilde{k}\|^2\). Hence, \(\|x - \tilde{k}\| \leq \|x - k\|^2\), \(\forall k \in K\).

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

### 13.7 Positive Definite Operators on Hilbert Spaces

Let \(X\) be a Hilbert space over \(\mathbb{R}\). In Chapter 10, we defined the symmetric operators \(S_X\) to be \(B_{S_2}(X, \mathbb{R}) \subseteq B(X, X^*)\). By Riesz-Fréchet Theorem 13.15, \(X^* = X\). Then, \(S_X \subseteq B(X, X)\). Fix any \(A \in S_X\). \(\forall x, y \in X\), we have \(\langle y, Ax \rangle = A(x)(y) = A(y)(x) = \langle x, Ay \rangle = \langle A^* x, y \rangle = \langle y, A^* x \rangle\), where the first equality follows since \(A(x)\) is a linear functional on \(X\) and by Riesz-Fréchet Theorem 13.15; the second equality follows since \(A \in B_{S_2}(X, \mathbb{R})\); 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 \(X\) 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 \(X\) is a real Hilbert space. Here, we will generalize the definition of \(S_X, S_{+X}, S_{\text{psd}X}, S_{-X}\), and \(S_{\text{ncd}X}\) to the case where \(X\) is a complex Hilbert space.

**Definition 13.40** Let \(X\) be a Hilbert space over \(\mathbb{K}\) and \(A \in B(X, X)\). We will write \(A \in S_X\), if \(A\) is Hermitian. We will write \(A \in S_{+X}\) if \(A \in S_X\) and \(\exists m \in (0, \infty) \subset \mathbb{R}\) such that \(\langle x, Ax \rangle \geq m\|x\|^2\), \(\forall x \in X\). We will write \(A \in S_{\text{psd}X}\) if \(A \in S_X\) and \(\exists m \in [0, \infty) \subset \mathbb{R}\) such that \(\langle x, Ax \rangle \geq m\|x\|^2\), \(\forall x \in X\). We will write \(A \in S_{-X}\) if \(A \in S_X\) and \(\exists m \in (0, \infty) \subset \mathbb{R}\) such that \(\langle x, Ax \rangle \leq -m\|x\|^2\), \(\forall x \in X\). We will write \(A \in S_{\text{ncd}X}\) if \(A \in S_X\) and \(\exists m \in [0, \infty) \subset \mathbb{R}\) such that \(\langle x, Ax \rangle \leq -m\|x\|^2\), \(\forall x \in X\).

**Proposition 13.41** Let \(X\) be a Hilbert space, \(A \in B(X, X)\) be Hermitian (that is, \(A \in S_X)\). Assume that \(\exists \delta \in (0, \infty) \subset \mathbb{R}\) such that \(\langle x, Ax \rangle \geq \delta \|x\|^2\), \(\forall x \in X\). Then, \(A \in S_{+X}\).
Then, $A$ is bijective and $A^{-1} \in B(X, X)$. Furthermore, $A^{-1}$ is Hermitian and $\forall x \in \mathcal{X}$, we have $\langle x, A^{-1}x \rangle \geq \frac{\delta}{\|B(x, x)\|} \|x\|^2$, $\forall \epsilon \in (0, \infty) \subset \mathbb{R}$.

Proof By the assumption, $\forall x \in \mathcal{X}$ with $x \neq \vartheta_X$, $A(x) \neq \vartheta_X$ since $\langle x, A(x) \rangle \geq \delta \|x\|^2 > 0$. Hence, $N(A) = \{\vartheta_X\}$ 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, $\mathcal{R}(A) = \mathcal{X}$. 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 $x = \lim_{i \rightarrow \infty} x_i$. 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 $x_i \in \mathcal{X}$. Clearly, $x_i \neq \vartheta_X$ since $x_i \neq \vartheta_X$, $\forall i \in \mathbb{N}$. Then, we have

$$\|\bar{x}_i\| \geq \|\bar{x}_i, x_i\| = \|\bar{x}_i, A\bar{x}_i\| \geq \delta \|\bar{x}_i\|^2 \quad \forall \bar{x}_i \in \mathcal{N}$$

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 $c \in (0, \infty) \subseteq \mathbb{R}$ such that $\|x_i\| \leq c$, $\forall i \in \mathbb{N}$. Then, $\|\bar{x}_i\| \leq c/\delta =: c_1 \in (0, \infty) \subseteq \mathbb{R}$, $\forall i \in \mathbb{N}$. This shows that $(\bar{x}_i)_{i=1}^\infty \subseteq \overline{\mathcal{B}_X}(\bar{x}_i, c_1) =: S_1 \subseteq \mathcal{X} = \mathcal{X}^{**}$, where the last equality follows from Riesz-Frêchet Theorem 13.15. By Alaoglu Theorem 7.122, $S_1 \subseteq \mathcal{X}^{**}$ is weak* compact. By Propositions 5.22 and 5.26, $(\bar{x}_i)_{i=1}^\infty$ has a cluster point $\bar{x} \in \mathcal{X}^{**} = \mathcal{X}$ in weak* topology. Since $\mathcal{X}$ is reflexive, the weak* topology on $\mathcal{X}^{**} = \mathcal{X}$ is identical to the weak topology on $\mathcal{X}$. By Proposition 7.123, $A : \mathcal{X} = \mathcal{X}^{\text{weak}} \rightarrow \mathcal{X}^{\text{weak}}$. 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 $x = \lim_{i \rightarrow \infty} x_i$ and $x = \lim_{i \rightarrow \infty} x_i$ 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

$$\langle x, A^{-1}x \rangle = \langle A\bar{x}, \bar{x} \rangle = \langle \bar{x}, A\bar{x} \rangle \geq \delta \|\bar{x}\|^2$$

Then, $x = A\bar{x}$ and $\|x\| \leq \|A\|_{\mathcal{B}(X, X)} \|\bar{x}\|$. We will distinguish two exhaustive and mutually exclusive cases: Case 1: $\|A\|_{\mathcal{B}(X, X)} = 0$; Case 2: $\|A\|_{\mathcal{B}(X, X)} > 0$.

Case 1: $\|A\|_{\mathcal{B}(X, X)} = 0$, we have $\|x\| = 0$, $\forall x \in \mathcal{X}$. Hence, $\mathcal{X} = \{\vartheta_X\}$. Then, $\langle x, A^{-1}x \rangle = 0 \geq \frac{\delta}{\|B(x, x)\|} \|x\|^2 = 0$, $\forall \epsilon \in (0, \infty) \subseteq \mathbb{R}$.
Case 2: \( \|A\|_{B(X,Y)} > 0 \). Then, we have

\[
\langle x, A^{-1}x \rangle \geq \delta \|x\|^2 \geq \frac{\delta}{\|A\|^2_{B(X,Y)}} \|x\|^2 \\
\geq \frac{\delta}{\epsilon + \|A\|^2_{B(X,Y)}} \|x\|^2, \quad \forall \epsilon \in (0, \infty) \subset \mathbb{R}
\]

Hence, we have shown, in both cases, that \( \langle x, A^{-1}x \rangle \geq \frac{\delta}{\epsilon + \|A\|^2_{B(X,Y)}} \|x\|^2 \), \( \forall x \in X \) and \( \forall \epsilon \in (0, \infty) \subset \mathbb{R} \).

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

**Proposition 13.42** Let \( X \) be a Hilbert space over \( K \) and \( A \in S_X \). Then, the following statements hold.

(i) \( \forall x \in X, \langle x, Ax \rangle \in \mathbb{R} \).

(ii) \( S_X \) is a closed subset of \( B(X,X) \). \( \forall A_1, A_2 \in S_X \), \( \forall \lambda \in \mathbb{R} \), we have \( A_1 + A_2 \in S_X \) and \( \lambda A_1 \in S_X \). If \( K = \mathbb{R} \), \( S_X \) is a closed subspace of \( B(X,X) \), and hence a Banach subspace. If \( K = \mathbb{C} \), 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) \( S_{psdX} \) and \( S_{nsdX} \) are closed convex cones in \( S_X \) (in the subset topology of \( S_X \subseteq B(X,X) \)); and \( S_{psdX} = -S_{nsdX} \).

(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 \( S_{psdX} \) relative to \( S_X \) is \( S_{+X} \), and the interior of \( S_{nsdX} \) relative to \( S_X \) is \( S_{-X} \).

(viii) \( A \in S_{+X} \), \( B, C \in S_{psdX} \), \( \alpha \in (0, \infty) \subset \mathbb{R} \), and \( \beta \in [0, \infty) \subset \mathbb{R} \), implies that \( A + B \in S_{+X} \), \( B + C \in S_{psdX} \), \( \alpha A \in S_{+X} \), and \( \beta B \in S_{psdX} \).

(ix) Let \( Y \) be a Hilbert space over \( K \), and \( B \in B(X,Y) \). Then, \( BB^* \in S_{psdY} \) and \( B^*B \in S_{psdX} \).

**Proof**  
(i) \( \forall x \in X, \langle x, Ax \rangle = \langle A^*x, x \rangle = \langle Ax, x \rangle = \overline{\langle x, Ax \rangle} \), where the first equality follows from Proposition 13.22; the second equality follows from \( A \in S_X \); and the last equality follows from Definition 13.1. Hence, \( \langle x, Ax \rangle \in \mathbb{R} \).

(ii) \( \forall A \in S_X \), by Proposition 4.13, there exists \( (A_i)_{i=1}^\infty \subseteq S_X \) such that \( \lim_{i \in \mathbb{N}} A_i = A \) in \( B(X,X) \). Then, \( \forall x, y \in X \), we have \( \langle Ax, y \rangle = \langle x, Ay \rangle = \overline{\langle x, Ay \rangle} = \overline{\langle Ax, y \rangle} \). The proof is completed.
\[ \lim_{n \to \infty} \langle A_n x, y \rangle = \lim_{n \to \infty} \langle x, A_n^* y \rangle = \lim_{n \to \infty} \langle x, A_n y \rangle = \langle x, A y \rangle = \langle A^* x, y \rangle, \]
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 \( A_n \in \mathcal{S}_X \) and therefore is Hermitian; the fourth equality 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 \( A \in \mathcal{S}_X \). Thus, we have shown that \( \overline{\mathcal{S}_X} \subseteq \mathcal{S}_X \subseteq \overline{\mathcal{S}_X} \).

By Proposition 3.3, \( \mathcal{S}_X \) is a closed subset of \( B(\mathcal{X}, \mathcal{X}) \).

\[ \forall A_1, A_2 \in \mathcal{S}_X \text{ and } \forall \alpha \in \mathbb{R}, \text{ we have } \langle A_1 + A_2 \rangle = \langle A_1 + A_2 \rangle^* = A_1^* + A_2^* = A_1 + A_2, \]
where the first equality follows from Proposition 13.22. Then, \( A_1 + A_2 \in \mathcal{S}_X \). In addition, \( \langle A_1 \rangle = \overline{\mathcal{S}_X} = \lambda A_1 \), where the first equality follows from Proposition 13.22. Hence, \( \lambda A_1 \in \mathcal{S}_X \).

Note that, the preceding paragraph is enough to conclude that \( \mathcal{S}_X \) is a subspace if \( \mathcal{X} \) is a real Hilbert space. In this case, \( \mathcal{S}_X \) is a closed subspace and therefore a Banach subspace of \( B(\mathcal{X}, \mathcal{X}) \) by Proposition 4.39. When \( \mathbb{K} = \mathbb{C} \), we only have \( \mathcal{S}_X \) is a closed convex cone and hence a complete metric subspace of \( B(\mathcal{X}, \mathcal{X}) \) 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 \( \mathcal{S}_X \). Fix any \( B \in \mathcal{S}_+ \mathcal{X} \). Then, \( \exists m \in (0, \infty) \subseteq \mathbb{R} \) such that \( \langle x, B x \rangle \geq m \| x \|^2 \), \( \forall x \in \mathcal{X} \). \( \forall C \in \mathcal{S}_X \cap B(\mathcal{X}, \mathcal{X})(B, \frac{1}{2}m) \), we have \( \langle x, C x \rangle \in \mathbb{R} \), \( \forall x \in \mathcal{X} \), by (i). Then, \( \langle x, C x \rangle \geq \langle x, B x \rangle - \| x, C x \| \geq m \| x \|^2 - \| x, (B - C) x \| \geq \frac{1}{2}m \| x \|^2 \), \( \forall x \in \mathcal{X} \), 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 \( \| B - C \|_{B(\mathcal{X}, \mathcal{X})} \leq \frac{1}{2}m \). By the arbitrariness of \( x, C \in \mathcal{S}_+ \mathcal{X} \). By the arbitrariness of \( C \), we have \( B(\mathcal{X}, \mathcal{X})(B, \frac{1}{2}m) \cap \mathcal{S}_X \subseteq \mathcal{S}_+ \mathcal{X} \). Hence, \( \mathcal{S}_+ \mathcal{X} \) is an open subset of \( \mathcal{S}_X \). By the relation \( \mathcal{S}_+ \mathcal{X} = -\mathcal{S}_- \mathcal{X} \), we have \( \mathcal{S}_- \mathcal{X} \) is also an open subset of \( \mathcal{S}_X \).

(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 \( \mathcal{S}_X \). \( \forall B \in \mathcal{S}_\text{psd} \mathcal{X} \), by Proposition 4.13, there exists \( (B_i)_{i=1}^\infty \subseteq \mathcal{S}_\text{psd} \mathcal{X} \) such that \( \lim_{i \to \infty} B_i = B \in B(\mathcal{X}, \mathcal{X}) \). By (ii), \( B \in \mathcal{S}_X \). \( \forall x \in \mathcal{X}, \langle x, B x \rangle = \lim_{i \to \infty} \langle x, B_i \rangle \geq 0 \), 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 \( \mathcal{S}_\text{psd} \mathcal{X} \subseteq \mathcal{S}_\text{psd} \mathcal{X} \subseteq \mathcal{S}_\text{psd} \mathcal{X} \). By Proposition 3.3, \( \mathcal{S}_\text{psd} \mathcal{X} \) is a closed subset of \( \mathcal{S}_X \).

\[ \forall A_1, A_2 \in \mathcal{S}_\text{psd} \mathcal{X}, \forall \alpha \in [0,1] \subseteq \mathbb{R}. \text{ Let } B := \alpha A_1 + (1 - \alpha) A_2. \text{ By (ii), } B \in \mathcal{S}_\text{psd} \mathcal{X}. \forall x \in \mathcal{X}, \langle x, B x \rangle = \alpha \langle x, A_1 x \rangle + (1 - \alpha) \langle x, A_2 x \rangle = \langle x, \alpha A_1 x \rangle + (1 - \alpha) \langle x, A_2 x \rangle = \alpha \langle x, A_1 x \rangle + (1 - \alpha) \langle x, A_2 x \rangle \geq 0 = 0 \| x \|^2, \]
where the second and third equalities follow from Definition 13.1; and the inequality follows from \( A_1, A_2 \in \mathcal{S}_\text{psd} \mathcal{X} \) by the arbitrariness of \( x \), we have \( B \in \mathcal{S}_\text{psd} \mathcal{X} \), hence, \( \mathcal{S}_\text{psd} \mathcal{X} \) is convex.
Obviously, \( \vartheta_{B(X,X)} \in S_{\text{psd}} X \). \( \forall A_1 \in S_{\text{psd}} X \), \( \forall \lambda \in [0, \infty) \subset \mathbb{R} \), we have \( \lambda A_1 \in S_X \) by (ii). \( \forall x \in X \), \( \langle x, \lambda A_1 x \rangle = \lambda \langle x, A_1 x \rangle \geq 0 \), 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 \( S_{\text{psd}} X = -S_{\text{npsd}} X \), we have \( S_{\text{npsd}} X \) 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 \( P \not\subseteq S_+ X \). Then, there exists \( B \in P \setminus S_+ X \). \( B \in P \) implies that \( \exists \delta \in (0, \infty) \subset \mathbb{R} \) such that \( B_{(X,X)} (B, \delta) \cap S_X \subseteq P \subseteq S_{\text{psd}} X \). Since \( B \not\in S_+ X \), then \( \exists x_0 \in X \) such that \( \langle x_0, Bx_0 \rangle < \frac{\delta}{2} \| x_0 \| ^2 \). Clearly, \( x_0 \neq \vartheta_X \).

Consider the operator \( \hat{B} := B - \frac{\delta}{2} \text{id}_X \). Clearly, \( \hat{B} \in B_{(X,X)} (B, \delta) \). 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, \( \langle x_0, Bx_0 \rangle = \langle x_0, B_x 0 \rangle = \langle x_0, \text{id}_X x_0 \rangle < \frac{\delta}{2} \| x_0 \| ^2 - \frac{\delta}{2} \| x_0 \| ^2 = -\frac{\delta}{4} \| x_0 \| ^2 < 0 \), 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 \( x_0 \neq \vartheta_X \). 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 \( S_{\text{psd}} X \) 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, \( B^* \in B(Y,Y) \). Then, \( BB^* \in B(Y,Y) \). By Proposition 13.22, we have \( (BB^*)^* = (B^*)^* B^* = BB^* \). Hence, \( BB^* \in S_Y \). \( \forall y \in Y \), \( \langle y, BB^* y \rangle_Y = \langle B^* y, B^* y \rangle_X = \| B^* y \| ^2_X \geq 0 \), 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 \( A - B \in S_+ X \).

### 13.8 Pseudoinverse Operator

**Proposition 13.43** Let \( X \) be a Hilbert space, \( M \subseteq X \) be a closed subspace. Then, we may define a mapping \( P : X \rightarrow M \) by \( P(x) = x_0 \), \( \forall x \in X \), where
13.8. PSEUDOINVERSE OPERATOR

\( x_0 \) 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 \( X \) to \( M \). Then, we have the following properties for the projection operator.

(i) \( P \in B(X,X) \).

(ii) \( \| P \|_{B(X,X)} \leq 1 \).

(iii) \( P^2 = P \) (idempotent)

(iv) \( P^* = P \) (Hermitian)

(v) \( \forall A \in B(X,X) \) that satisfies (iii) and (iv) is a projection operator.

**Proof**

(i) By Theorem 13.19, \( X = M \oplus M^\perp \). Then, \( \forall x \in X \), 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. \( \forall x, y \in X, \forall \lambda \in \mathbb{K} \), 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, \( P(\lambda x) = \lambda P(x) \). Hence, \( P \) is linear.

(iii) This follows from the preceding paragraph.

(iii) This is obvious.

(iv) \( \forall x, y \in X, \langle x, Py \rangle = \langle P^* x, y \rangle = \langle x_1 + x_2, y_1 \rangle = \langle x_1, y_1 \rangle = \langle x_1, y_1 + y_2 \rangle = \langle P(x), y \rangle \), 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) \( \forall A \in B(X,X) \) that satisfies (iii) and (iv). Let \( M := \mathcal{R}(A) \). Then, \( M^\perp = \mathcal{N}(A^*) = \overline{\mathcal{N}(A)} \), by Proposition 13.22 and (iv). By Theorem 13.19, \( X = M \oplus (M^\perp)^\perp \). By Proposition 13.17, \( (M^\perp)^\perp = \overline{\mathcal{R}(A)} \). Let \( P : X \to \overline{\mathcal{R}(A)} \) be the projection operator. We will show that \( P = A \). \( \forall x, y \in X \), we have

\[
\langle Ax, y - Ay \rangle = \langle Ax, y \rangle - \langle Ax, Ay \rangle = \langle Ax, y \rangle - \langle A^* Ax, y \rangle
\]

\[
= \langle Ax, y \rangle - \langle A^2 x, y \rangle = \langle Ax, y \rangle - \langle Ax, y \rangle = 0
\]

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 \overline{\mathcal{R}(A)}. Then, we have \(Ay = Py\). By the arbitrariness of \(y\), we have \(A = P\).

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

**Proposition 13.44** Let \(\mathcal{X}\) and \(\mathcal{Y}\) be Hilbert spaces over \(\mathbb{K}\), \(A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})\), and \(y \in \mathcal{Y}\). Assume that \(\mathcal{R}(A) \subseteq \mathcal{Y}\) is closed. Then, there exists a unique vector \(x_0 \in \mathcal{X}\) such that \(x_0 = \arg\min_{x_1} \|x_1\|_\mathcal{X}\), where \(x_1\) satisfies \(\|Ax_1 - y\| = \min_{x \in \mathcal{X}} \|Ax - y\|\). We will denote \(x_0 =: A^!y\), where \(A^! : \mathcal{Y} \to \mathcal{X}\) is called the pseudoinverse of \(A\). Furthermore, we have the following properties for \(A^!\).

(i) \(A^! \in \mathcal{B}(\mathcal{Y}, \mathcal{X})\).

(ii) \((A^!)^* = A^!\).

(iii) \((A^*)^* = (A^!)^*\).

(iv) \(A^!AA^! = A^!\).

(v) \(AA^!A = A\).

(vi) \((A^!A)^* = A^!A\).

(vii) \((AA^!)^* = AA^!\).

(viii) \(A^! = (A^*A)^!\).

(ix) \(A^! = A^*(AA^*)^!\).

(x) \(A^! = \lim_{\epsilon \to 0^+} A^*(\epsilon \text{id}_\mathcal{Y} + AA^*)^{-1} = \lim_{\epsilon \to 0^+} (\epsilon \text{id}_\mathcal{X} + A^*A)^{-1}A^*\).

(xi) \(A\) can be uniquely expressed as \(A = P_1A_\perp P_2\), where \(P_1 : \mathcal{Y} \to \mathcal{R}(A)\) and \(P_2 : \mathcal{X} \to (\mathcal{N}(A))_\perp\) are projection operators, and \(A_\perp \in \mathcal{B}(\mathcal{N}(A), (\mathcal{N}(A))_\perp)\) is bijective with \(A_\perp^{-1} \in \mathcal{B}(\mathcal{R}(A), (\mathcal{N}(A))_\perp)\).

(xii) \(AA^! = P_1\) and \(A^!A = P_2\).

**Proof** Since \(\mathcal{R}(A)\) is closed. Then, by the Classical Projection Theorem 13.13, \(\exists y_0 \in \mathcal{R}(A)\) such that \(\|y - y_0\|_\mathcal{Y} = \min_{y \in \mathcal{R}(A)} \|y - \bar{y}\|_\mathcal{Y}\), furthermore, \(y_0 \in \mathcal{R}(A)\) is defined to be the unique vector such that \(y - y_0 \in (\mathcal{R}(A))_\perp\). Then, the set of \(x_1\)'s defined in the proposition statement is given by \(\mathcal{N}(A) + x_1\), where \(x_1\) is any vector such that \(Ax_1 = y_0\). Then, by the Classical Projection Theorem 13.13 and Proposition 13.22, \(\exists x_0 \in (\mathcal{N}(A))_\perp = \mathcal{R}(A^*)\) such that \(\|x_0\|_\mathcal{X} = \min_{x \in \mathcal{N}(A)} \|x_1 - x\|_\mathcal{X}\). This proves the existence and uniqueness of \(x_0 \in \mathcal{X}\) for any given \(y \in \mathcal{Y}\). Thus, \(A^!\) is well-defined.

(i) By Theorem 13.19, \(\mathcal{X} = \mathcal{N}(A) \oplus (\mathcal{N}(A))_\perp =: M \oplus M_\perp\). Define the projection mapping \(P_2 : \mathcal{X} \to M_\perp\). \(\forall x \in \mathcal{X}\), there is unique pair \(x_1 \in M\) and \(x_2 \in M_\perp\) such that \(x = x_1 + x_2 = x_1 + P_2x\). Then, \(Ax = AP_2x\). This
allows us to define a mapping $A_r : M^\perp \to N := \mathcal{R}(A)$ by $A_r(x_2) = A(x_2)$, $\forall x_2 \in M^\perp$. Clearly, $A_r \in B(M^\perp, N)$ and $A_r$ is surjective. It is easy to show that $\mathcal{N}(A_r) = \{\vartheta_{M^\perp} = \vartheta_X\}$ since the domain of $A_r$ is $M^\perp$. Then, $A_r$ is bijective. Note that $\mathcal{N}$ is a closed subspace of $\mathcal{Y}$ by assumption. Then, by Proposition 4.39, $\mathcal{N}$ is a Banach subspace of $\mathcal{Y}$. By Proposition 13.17, $M^\perp$ is a closed subspace of $X$. Then, by Proposition 4.39, $M^\perp$ is a Banach subspace of $X$. Then, $A_r^{-1} \in B(N, M^\perp)$ by Open Mapping Theorem 7.103.

Then, $A = A_rP_2$. By Theorem 13.19, $\mathcal{Y} = N \oplus N^\perp$. Define $P_1 : \mathcal{Y} \to N$ to be the projection mapping. Clearly, $P_1 \in B(\mathcal{Y}, N)$, $\forall y \in \mathcal{Y}, y_0 = P_1y$.

Let $x_1 := A_r^{-1}y_0 \in M^\perp \subset X$. 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, $A^\dagger = A_r^{-1}P_1 = P_2A_r^{-1}P_1 \in B(\mathcal{Y}, X)$.

(iii) (i) This follows immediately from the preceding discussion in (i).

(ii) For $A \in B(X, \mathcal{Y})$, $A_r \in B(M^\perp, N)$, $A = A_rP_2 = P_1AP_r$ and $A^\dagger = A_r^{-1}P_1 = P_2A_r^{-1}P_1$. Then, $(A^\dagger)^\dagger = P_1AP_r = A$.

(iii) $(A^\dagger)^* = (A_r^{-1}P_1)^* = (P_2A_r^{-1}P_1)^* = P_1^*(A_r^{-1})^*P_2^* = P_1A_r^{-1}P_2$, where the first two equalities follow from (i); the third equality follows from Proposition 13.22; and the last equality follows from Proposition 13.43. Now, $\mathcal{N}(A^\dagger)^\dagger = \mathcal{R}(A) = N$ and $\mathcal{R}(A^\dagger)^\dagger = \mathcal{N}(A)^\perp = M^\perp$, by Proposition 13.22. Therefore, $(A^\dagger)^\dagger = (P_2^*A_r^*P_1^*)^\dagger = (P_2A_r^*P_1)^\dagger = P_1(A_r^*)^{-1}P_2 = P_1A_r^{-1}P_2$, where the first equality follows from (ii); the second equality follows from Proposition 13.43; the third equality follow from (iii); and the last equality follows from Proposition 13.22. Hence, $(A^\dagger)^* = (A^\dagger)^\dagger$.

(iv) By (xii), we have $A^\dagger A = P_2A_r^{-1}P_1P_1A_rP_2 = P_2A_r^{-1}A_rP_2 = P_2$, where the second and third equalities follow from the fact that $P_1$ and $P_2$ are projection operators. Similarly, $AA^\dagger = P_1A_rP_2A_r^{-1}P_1 = P_1A_rA_r^{-1}P_1 = P_1$, where the second and third equalities follow from the fact that $P_1$ and $P_2$ are projection operators.

(v) By (xii) and (xiii), we have $AA^\dagger A = P_1P_1A_rP_2 = P_1A_rP_2 = A$, where the second equality follows from Proposition 13.43.

(vi) By (xii) and (xiii), we have $(A^\dagger A)^* = P_2^* = P_2 = A^\dagger A$, where the second equality follows from Proposition 13.43.

(vii) By (xii) and (xiii), we have $(AA^\dagger)^* = P_1^* = P_1 = AA^\dagger$, where the second equality follows from Proposition 13.43.

(viii) By (xii), we have $A^\dagger A = (P_1A_rP_2)^*P_1A_rP_2 = P_2A_r^*P_1A_rP_2 = P_2A_r^*A_rP_2$, 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 $(A^\dagger A)^\dagger A^\dagger = P_2(A^\dagger A)^{\dagger}P_1(P_1A_rP_2)^* = P_2A_r^{-1}A_rA_r^{-1}P_2P_1 = P_2A_r^{-1}A_rA_r^{-1}P_1 = P_2$, by Theorem 13.19, where the first equality follows from (xii) applied to the operator $A^\dagger A$: the second equality follows from Proposition 13.43; the third equality follows from Proposition 13.22; the fourth equality follows from the fact $P_2$ is a projection operator; and the last equality follows from (xii).
(ix) By (xi), we have $AA^* = P_1 A P_2 (P_1 A P_2)^* = P_1 A P_2 P_2^* A^* P_1^* = P_1 A R A P_1$, 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 $A^* (AA^*)^! = (P_1 A P_2)^* P_1 (A R A^* P_1) = P_2 A R A P_2 P_1 A^* A^* P_1 = P_2 A R A P_2 P_1 A^* A^* P_1 = P_2 A R A P_2 P_1 = P_2$, where the first equality follows from (xi) applied to the operator $AA^*$; the second equality follows from Proposition 13.22; the third equality follows from Proposition 13.43; the fourth equality follows from the fact $P_2$ is a projection operator; and the last equality follows from (xi).

(x) Note that $e i d_X + A^* A = e (P_2 + P_2) + P_2 A^* A P_2 = e P_2 e i d_M P_2 + P_2 (e i d_M + A^* A) = (P_2 e i d_M + A^* A)^! P_2 = e i d_M + A^* A P_2 = \forall e \in (0, \infty) \subset \mathbb{R}$, where $P_2 : X \rightarrow M$ is the projection operator; the first equality follows from Theorem 13.19 and the proof of (viii); the second equality follows from Proposition 13.43. Then, $(e i d_X + A^* A)^! P_2 = (P_2 e i d_M P_2 + P_2 (e i d_M + A^* A)^! P_2, \forall e \in (0, \infty) \subset \mathbb{R}$, where the equality follows from the fact $P_2 P_2 = e i d_B X$ and Proposition 13.43; and the invertibility of $e i d_M + A^* A$ follows from Proposition 13.41. This yields $(e i d_X + A^* A)^! A^* = (P_2 e i d_M P_2 + P_2 (e i d_M + A^* A)^! P_2, \forall e \in (0, \infty) \subset \mathbb{R}$. Since $A^* A$ is invertible, by Proposition 9.55, we have $\lim_{\epsilon \rightarrow 0^+} (e i d_X + A^* A)^! A^* = \lim_{\epsilon \rightarrow 0^+} P_2 (e i d_M + A^* A)^! A^* P_2 = P_2 A^* A^* A^* P_2 = P_2 A^* A^* A^* P_2 = P_2 A^* A^* P_2 = A^!$. The other equality in (x) can be proved by symmetry.

This completes the proof of the proposition. □

Proposition 13.45 Let $X$ and $Y$ be Hilbert spaces over $\mathbb{K}$, and $B \in B(X, Y)$. If $B$ is surjective, then $BB^* \in S_+ Y$. If $B$ is injective and $R(B^*) \subseteq X$ is closed, then $B^* B \in S_+ X$.

Proof. If $B$ is surjective, then $R(B) = Y$ and it is a closed set in $Y$. By Proposition 13.44, we have $B = P_1 B P_2$, where $P_1 : Y \rightarrow R(B)$ and $P_2 : X \rightarrow (N(B))^!$ are projection operators, and $B e \in B((N(B))^!, R(B))$ is bijective. Clearly, $P_1 = e i d_Y$. Then, $BB^* = B e B^*$ is bijective, and then $(BB^*)^{-1} \in B(Y, Y)$ by Open Mapping Theorem 7.103. Fix any $y \in Y$. Let $y := (BB^*)^{-1} y$. This yields $\langle y, (BB^*)^{-1} y \rangle = \langle BB^* y, y \rangle = \langle B^* y, B^* y \rangle^! = \|B^* y\|_Y^2$. Define $\bar{x} := B^* y = B(B^*)^{-1} y = B^! y \in X$, where the second equality follows from Proposition 13.44. This leads to $B \bar{x} = BB^! y = P_1 y = y$, 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 $\langle y, (BB^*)^{-1} y \rangle = \|\bar{x}\|_X^2 \geq \frac{\|B\|_{B(X, Y)}^2}{\|B\|_{B(X, Y)}^2} \|\bar{x}\|_X^2 \geq \frac{1}{\epsilon + \|B\|_{B(X, Y)}^2} \|y\|_Y^2$, where the first equality follows from the preceding discussion; the first inequality holds for any $\epsilon \in (0, \infty) \subset \mathbb{R}$; and the last inequality follows the preceding discussion. This yields that $(BB^*)^{-1} \in S_+ Y$ (clearly, $(BB^*)^{-1} \in S_y$). By (v) of Proposition 13.42, $BB^* \in S_+ Y$. 


13.9. SPECTRAL THEORY OF LINEAR OPERATORS

If $B$ is injective and $\mathcal{R}(B^*) \subseteq \mathcal{X}$ is closed, then $\mathcal{N}(B) = \{0\}$. By Proposition 13.22, we have $\mathcal{R}(B^*) = (\mathcal{N}(B))^\perp = \mathcal{X}$. Hence, $B^*$ is surjective. By the preceding paragraph, we have $B^*(B^*)^* = B^*B \in \mathcal{S}_{\mathcal{X}}$, where the equality follows from Proposition 13.22.

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

13.9 Spectral Theory of Linear Operators

**Definition 13.46** Let $\mathcal{X}$ be a normed linear space over $\mathbb{K}$, $A \in B(\mathcal{X}, \mathcal{X})$, $\lambda \in \mathbb{K}$, and $x_0 \in \mathcal{X}$ with $x_0 \neq 0$. If $\lambda x_0 = Ax_0$, or equivalently $(\lambda \text{id}_\mathcal{X} - A)x_0 = 0$, we will say that $\lambda$ is an eigenvalue of $A$, $x_0$ is an eigenvector of $A$ associated with the eigenvalue $\lambda$. If $(\lambda \text{id}_\mathcal{X} - A)^k x_0 \neq 0$ and $(\lambda \text{id}_\mathcal{X} - A)^k x_0 = 0$, 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 $\mathcal{X}$ be a Hilbert space over $\mathbb{K}$, $Q \in B(\mathcal{X}, \mathcal{X})$ be Hermitian, and $A \in B(\mathcal{X}, \mathcal{X})$. Then, the following statements hold.

(i) If $\lambda \in \mathbb{K}$ is an eigenvalue of $Q$, then $\lambda \in \mathbb{R}$;

(ii) If $S_\gamma \subseteq \mathcal{X}$ be an linearly independent set of eigenvectors of $Q$ associated with the 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 $\lambda_i \in \mathbb{K}$, $i = 1, 2$, 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 $\mathcal{R}(Q) \subseteq \mathcal{X}$ is closed, then $0 \in \mathbb{K}$ is not an eigenvalue of $Q$ if, and only if, $Q$ is bijective and $Q^{-1} \in B(\mathcal{X}, \mathcal{X})$;

(vi) If $Q \in \mathcal{S}_{\mathcal{X}}$ then $\exists \theta \in (0, \infty) \subseteq \mathbb{R}$ such that $\forall \lambda \in \mathbb{R}$ with $\lambda$ being an eigenvalue of $Q$, we have $\lambda \geq \theta$;

(vii) If $\lambda \in \mathbb{K}$ is an eigenvalue of $A$, then $|\lambda| \leq \|A\|_{B(\mathcal{X}, \mathcal{X})} \|\text{id}_\mathcal{X}\|_{B(\mathcal{X}, \mathcal{X})} \leq \|A\|_{B(\mathcal{X}, \mathcal{X})}$.

(viii) If $\mathcal{R}(Q) \subseteq \mathcal{X}$ is closed, then $Q \in \mathcal{S}_{psd, \mathcal{X}}$ implies that all of the eigenvalues of $Q$ are nonnegative.

**Proof**

(i) Let $x_0 \in \mathcal{X}$ be the eigenvector of $Q$ associated with the eigenvalue $\lambda$. Then, $\lambda x_0 = Qx_0$. This implies that $\lambda \|x_0\|^2 = \langle x_0, \lambda x_0 \rangle = \langle x_0, Qx_0 \rangle = \langle Q^*x_0, x_0 \rangle = \langle Qx_0, x_0 \rangle = \langle \lambda x_0, x_0 \rangle = \lambda \|x_0\|^2$, 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 eigenvalue \( \lambda \); and the seventh equality follows from Definition 13.1 and Proposition 13.2. Note that \( x_0 \neq \vartheta \chi \) since it is an eigenvector, then \( \| x_0 \| ^2 > 0 \). Thus, the above implies that \( \chi = \lambda \).

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 \( \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K} \setminus \{ 0 \} \) 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 \( x_0 = \vartheta \chi \). 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, \( x_0 \neq \vartheta \chi \). 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 \( \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K} \setminus \{ 0 \} \) such that \( x_0 = \sum_{i=1}^{n} \alpha_i x_i \). We will distinguish two exhaustive and mutually exclusive cases: Case 1: \( x_0, \ldots, x_n \in S_{\gamma_0} \) for some \( \gamma_0 \in \Lambda \); Case 2: \( x_0 \in S_{\gamma_0} \) with \( \gamma_0 \in \Gamma \), and \( \exists \alpha_1 \in \{ 1, \ldots, n \} \) such that \( x_{\alpha_1} \in S_{\gamma_1} \), with \( \gamma_1 \in \Gamma \) and \( \gamma_1 \neq \gamma_0 \).

Case 1: \( x_0, \ldots, x_n \in S_{\gamma_0} \) for some \( \gamma_0 \in \Gamma \). Then, \( x_0, \ldots, x_n \) 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 \( \exists \alpha_1 \in \{ 1, \ldots, n \} \) such that \( x_{\alpha_1} \in S_{\gamma_1} \) with \( \gamma_1 \in \Gamma \) and \( \gamma_1 \neq \gamma_0 \). Without loss of generality, assume that \( x_i \in S_{\gamma_i} \), \( i = 1, \ldots, n \), where \( \gamma_i \in \Gamma \). Then, \( \vartheta \chi = (\lambda_{\gamma_0} \text{id}_{\chi} - Q) x_0 = \sum_{i=1}^{n} \lambda_{\gamma_0} (\lambda_{\gamma_0} \text{id}_{\chi} - Q) (\alpha_i x_i) = \sum_{i=1}^{n} \alpha_i (\lambda_{\gamma_0} - \lambda_{\gamma_1} x_i). By the preceding discussion and (i), \( 0 \neq \lambda_{\gamma_0} - \lambda_{\gamma_1} \in \mathbb{R} \). Then, we have \( \vartheta \chi = \sum_{i=1}^{n} \frac{(\lambda_{\gamma_0} - \lambda_{\gamma_1}) \alpha_i}{\lambda_{\gamma_0} - \lambda_{\gamma_1}} x_i \). Rearranging the equation, we have \( x_i = \sum_{i=1, i 
eq i_1}^{n} \frac{(\lambda_{\gamma_0} - \lambda_{\gamma_1}) \alpha_i}{\lambda_{\gamma_0} - \lambda_{\gamma_1}} x_i \). 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 \( \lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = (Q x_1, x_2) = \langle x_1, Q^* x_2 \rangle = \langle x_1, \lambda_2 x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle \), 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 sixth 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, \( B^{k-1}x_0 \neq 0 \) and \( B^kx_0 = \vartheta \chi \), where \( B := \lambda \mathbb{I} - Q \). Let \( x_{k-2} := B^{k-2}x_0 \) and \( x_{k-1} := B^{k-1}x_0 \). Then, \( B^2x_{k-2} = \vartheta \chi \). This implies that \( 0 = \langle x_{k-2}, B^2x_{k-2} \rangle = \langle B^*x_{k-2}, Bx_{k-2} \rangle = \langle Bx_{k-2}, Bx_{k-2} \rangle = \langle x_{k-1}, x_{k-1} \rangle = \| x_{k-1} \|^2 \), 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, \( x_{k-1} = \vartheta \chi \). This contradicts with \( B^{k-1}x_0 \neq 0 \). Hence, the result holds.

(v) “Necessity” Since \( 0 \) is not an eigenvalue of \( Q \), then \( \mathcal{N} (Q) = \{ \vartheta \chi \} \).

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 \( \mathcal{N} (Q) = \{ \vartheta \chi \} \). Then, \( 0 \in \mathbb{R} \) is not an eigenvalue of \( Q \), since there is no eigenvector associated with \( 0 \).

(vi) By Definition 13.40, \( \exists x \in (0, \infty) \subset \mathbb{R} \) 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 \( x_0 \neq \vartheta \chi \) and \( \| x_0 \| > 0 \). Then, we have \( \lambda \geq m \).

(vii) Fix any \( \lambda \in \mathbb{K} \) with \( |\lambda| > \| A \|_{B(\mathcal{X}, \mathcal{X})} \| \mathbb{I} \|_{B(\mathcal{X}, \mathcal{X})} \geq 0 \). Then, \( \| \mathbb{I} - (\lambda \mathbb{I} - A) \|^2_{B(\mathcal{X}, \mathcal{X})} \| \mathbb{I} - (\lambda \mathbb{I} - A) \|^2_{B(\mathcal{X}, \mathcal{X})} = \frac{1}{\lambda^2} \| A \|^2_{B(\mathcal{X}, \mathcal{X})} \| \mathbb{I} \|^2_{B(\mathcal{X}, \mathcal{X})} < 1 \).

By Proposition 9.55, \( \mathbb{I} - (\lambda \mathbb{I} - A) \) is bijective and admits continuous inverse. Then, \( \lambda \mathbb{I} - A \) is bijective and admits continuous inverse since \( \lambda \neq 0 \). This implies that \( \mathcal{N} (\lambda \mathbb{I} - A) = \{ \vartheta \chi \} \). Hence, \( \lambda \) is not an eigenvalue of \( A \). Hence, the result holds noting \( \| \mathbb{I} \|_{B(\mathcal{X}, \mathcal{X})} \leq 1 \).

(viii) Since \( Q \in \mathcal{S}_{\text{psd}, \mathcal{X}} \), 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 \( x_0 \neq \vartheta \chi \) 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 \( \mathbb{K} \), \( A \in \mathcal{B} (\mathcal{X}, \mathcal{X}) \) be a Hermitian operator. Then, \( \lambda := \sup_{x \in \mathcal{X}, \| x \| \leq 1} | \langle Ax, x \rangle | = \| A \|_{B(\mathcal{X}, \mathcal{X})} \).

**Proof** Note that \( | \langle Ax, x \rangle | \leq \| Ax \| \| x \| \leq \| A \|_{B(\mathcal{X}, \mathcal{X})} \| x \|^2 \), \( \forall x \in \mathcal{X} \), where the first inequality follows from Cauchy-Schwarz Inequality; and the second inequality follows from Proposition 7.64. Hence, \( \lambda \leq \| A \|_{B(\mathcal{X}, \mathcal{X})} \).

Fix any \( u \in \mathcal{B}_\mathcal{X} (\vartheta \chi, 1) \), let \( v_1 := \alpha u + \alpha^{-1} A u \) and \( v_2 := \alpha u - \alpha^{-1} A u \), where \( \alpha \in (0, \infty) \subset \mathbb{R} \). Note also that

\[
\| Au \|^2 = \langle Au, Au \rangle = \langle A^* Au, u \rangle = \langle A^2 u, u \rangle = \langle u, A^2 u \rangle
\]
Proposition 13.48 and Definition 13.40, \( \exists \lambda \) follows from Proposition 13.22; the third equality follows from the fact that \( \lim Q \) 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 \( u \in \mathbb{F}_X (\vartheta_X, 1) \).

If \( \|Au\| = 0 \), then \( \|Au\| \leq \lambda \). On the other hand, if \( \|Au\| > 0 \), let \( \alpha^2 := \|Au\| \), we have \( \|Au\|^2 \leq \frac{1}{2} \alpha^2 \|Au\| \), which leads to \( \|Au\| \leq \lambda \). In both cases, we have \( \|Au\| \leq \lambda \), \( \forall u \in \mathbb{F}_X (\vartheta_X, 1) \). Hence, \( \|A\| \leq \lambda \).

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

**Definition 13.49** Let \( X \) be a Banach space over \( \mathbb{K} \) and \( Y \) be a Banach space over \( \mathbb{K} \), \( A \in B(X, Y) \). \( A \) is said to be compact if \( \forall x_0 \in X \) and \( \forall \varepsilon > 0 \in \mathbb{R} \), \( A(\mathbb{F}_X(x_0, \varepsilon)) \subseteq Y \) is a compact set.

**Theorem 13.50 (Spectral Theory)** Let \( X \) be a Hilbert space over \( \mathbb{K} \), and \( Q \in \mathbb{S}_{\text{psd}}X \) be Hermitian. Assume that \( Q \) is a compact operator. Then, the following statements hold.

(i) \( Q \neq \vartheta_{B(X, X)} \) if, and only if, \( Q \) admits an eigenvector \( x_0 \in X \) with \( \|x_0\| = 1 \) associated with the eigenvalue \( \lambda = \|Q\|_{B(X, X)} \in (0, \infty) \subset \mathbb{R} \).

(ii) \( Q \) either has finite number of eigenvectors \( x_i \in X \) with \( \|x_i\| = 1 \), each associated with an eigenvalue \( \lambda_i \in (0, \infty) \subset \mathbb{R} \), \( i = 1, \ldots, n \), \( n \in \mathbb{N}_+ \), where \( (x_i^n)_{i=1}^n \subseteq X \) is an orthonormal sequence, and \( Q = \sum_{i=1}^n \lambda_i x_i x_i^* \); or has countably many eigenvectors \( x_i \in X \) with \( \|x_i\| = 1 \), each associated with an eigenvalue \( \lambda_i \in (0, \infty) \subset \mathbb{R} \), \( i \in \mathbb{N} \), where \( (x_i)_{i=1}^\infty \subseteq X \) is an orthonormal sequence, and \( Q \geq \sum_{i=1}^\infty \lambda_i x_i x_i^* \).

(iii) In the latter case of (ii), \( \lim_{i\to\infty} \lambda_i = 0 \) and, if \( (x_i)_{i=1}^\infty \) is complete, (i.e., span \( (x_i)_{i=1}^\infty \) = \( X \)) then we have \( Q = \sum_{i=1}^\infty \lambda_i x_i x_i^* \).

**Proof** (i) “Necessity” \( \lambda = \|Q\|_{B(X, X)} > 0 \) since \( Q \neq \vartheta_{B(X, X)} \). By Proposition 13.48 and Definition 13.40, \( \exists (u_i)_{i=1}^\infty \subseteq X \) with \( \|u_i\| = 1 \) such that \( \lim_{i\to\infty} \langle Q(u_i), u_i \rangle = \lambda > 0 \). Since \( Q \) is a compact operator, then \( (Q(u_i))_{i=1}^\infty \subseteq Q(\mathbb{F}_X(\vartheta_X, 1)) \), which is a compact set. By Borel-Lebesgue Theorem 5.37, the set \( Q(\mathbb{F}_X(\vartheta_X, 1)) \subseteq X \) is sequentially compact. Then, there exists a subsequence \( (Q(u_{i_k}))_{k=1}^\infty \subseteq Q(\mathbb{F}_X(\vartheta_X, 1)) \) such that \( \lim_{k\to\infty} Q(u_{i_k}) = u_0 \). Note that \( 0 \leq \|Q(u_i) - \lambda u_i\| = \langle Q(u_i) - \lambda u_i, Q(u_i) - \lambda u_i \rangle = \langle Q(u_i), Q(u_i) \rangle - \lambda \langle u_i, Q(u_i) \rangle - \lambda \langle Q(u_i), u_i \rangle - \lambda^2 \langle u_i, u_i \rangle \leq \|Q(u_i)\|^2 - \lambda^2 - \lambda (\langle Q(u_i), u_i \rangle + \langle u_i, Q(u_i) \rangle - 2\lambda) \leq \|Q(u_i)\|^2 - \lambda^2 - \lambda (\langle Q(u_i), u_i \rangle + \langle u_i, Q(u_i) \rangle) \leq \|Q(u_i)\|^2 - \lambda^2 \).
\[ \lambda^2 - \lambda(\langle Q(u_i), u_i \rangle + \langle Q^* u_i, u_i \rangle - 2\lambda) = -2\lambda(\langle Q(u_i), u_i \rangle - \lambda) \to 0, \] 
where the first equality follows from Proposition 13.2; the second equality follows from Definition 13.1; the second inequality follows from simple algebra; and the last equality follows from \( Q \) is Hermitian. Then, we have \( 0 = \lim_{k \to \infty} \| Q(u_i_k) - \lambda u_i_k \| = \lim_{k \to \infty} \| u_0 - \lambda u_i_k \|. \) Hence, \( \lim_{k \to \infty} u_i_k = \frac{1}{\lambda} u_0. \)

By continuity of \( Q \), we have \( Q\left( \frac{1}{\lambda} u_0 \right) = \lim_{k \to \infty} Q(u_i_k) = u_0. \) Furthermore, \( \| \frac{1}{\lambda} u_0 \| = \lim_{k \to \infty} \| u_i_k \| = 1. \) Hence, \( \frac{1}{\lambda} u_0 =: x_0 \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda = \| Q \|_{B(H,X)} > 0. \)

Let 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 \( P_1 : X \to M^\perp \) be projection operators. We have \( P_1 + P_1 = \text{id}_X \). This implies that \( Q = (P_1 + \bar{P}_1)Q(P_1 + \bar{P}_1) = P_1QP_1 + \bar{P}_1QP_1 + \bar{P}_1QP_1 + \bar{P}_1QP_1 \).

Note that \( P_1 = x_0x_0^* \). \( \forall x \in X, P_1(x) = x_0(x_0^*, x) = x_0(x, x_0). \) This implies that \( P_1QP_1(x) = (x, x_0) P_1Q(x_0) = \lambda \langle x, x_0 \rangle x_0 = \lambda x_0(\langle x_0^*, x \rangle) \) and \( \bar{P}_1QP_1(x) = (x, x_0) \bar{P}_1Q(x_0) = \bar{\lambda} \langle x, x_0 \rangle x_0 = \bar{\lambda} x_0(\langle x_0^*, x \rangle). \)

Thus, \( Q(x) = \lambda x_0(\langle x_0^*, x \rangle) + \bar{\lambda} x_0(\langle x_0^*, x \rangle) = \lambda x_0(\langle x_0^*, x \rangle) + \bar{\lambda} x_0(\langle x_0^*, x \rangle). \) Hence, \( Q = \lambda x_0x_0^* + \bar{\lambda} x_0x_0^* + \bar{\lambda} x_0x_0^*. \)

By Propositions 13.22 and 13.43, we have \( Q = Q^* = \lambda x_0x_0^* + \bar{\lambda} x_0x_0^* + \bar{\lambda} x_0x_0^* = \lambda x_0x_0^* + \bar{\lambda} x_0x_0^* + \bar{\lambda} x_0x_0^*. \) Hence, we have \( Q = \lambda x_0x_0^* + \bar{\lambda} x_0x_0^*. \)

"Sufficiency" Let \( Q \) admit an eigenvector \( x_0 \) associated with the eigenvalue \( \lambda > 0 \). Then, \( \langle x_0, Qx_0 \rangle = \langle x_0, \lambda x_0 \rangle = \lambda \| x_0 \|^2 > 0. \) Hence, \( Q \neq \vartheta_{B(H,X)} \).

This proves the first statement.

(ii) By (i), \( Q = \lambda x_0x_0^* + Q_2 \), where \( Q_2 = \bar{P}_1Q\bar{P}_1 \). Denote \( \lambda_2 := \lambda, x_1 := x_0, Q_1 := Q, \) and \( \bar{P}_{1,1} := \bar{P}_1 \). Now, we will repeat (i) for the operator \( Q_2 \).

Clearly, \( Q_2 \in S_{\text{psd}}X \). We claim that

Claim 13.50.1 Any eigenvector \( x_2 \in X \) with \( \| x_2 \| = 1 \) of \( Q_2 \) associated with an eigenvalue \( \lambda_2 > 0 \) must be an eigenvector of \( Q_1 \) associated with the eigenvalue \( \lambda_2 \), and \( x_2 \perp x_1 \).

Proof of claim: \( Q_2x_2 = \lambda_2x_2 \). Since \( Q_2 = \bar{P}_1Q\bar{P}_1 \). Then, \( \lambda_2 \langle x_1, x_2 \rangle = \langle x_1, \lambda_2 x_2 \rangle = \langle x_1, \bar{P}_1Q\bar{P}_1x_2 \rangle = \langle \bar{P}_1x_1, Q\bar{P}_1x_2 \rangle = \langle \bar{P}_1(x_1), Q\bar{P}_1x_2 \rangle = 0 \), where the first equality follows from Definition 13.1 and the fact that \( \lambda_2 \in \Re \); 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 \( \bar{P}_1(x_0) = \vartheta_x \). Since, \( \lambda_2 > 0 \), then \( x_1 \perp x_2 \). Then, \( x_2 \in M^\perp \). It is then easy to check that \( Q(x_2) = Q_1(x_2) = \lambda x_0(\langle x_0^*, x_2 \rangle) + Q_2(x_2) = \lambda x_0(\langle x_2, x_0 \rangle + \lambda x_2 = \lambda x_0(\langle x_2, x_0 \rangle + \lambda x_2 = \lambda x_2, \) 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 (i) until \( Q_{n+1} = \vartheta_{B(X,X)} \) or indefinitely. After any step \( k \in \mathbb{N} \) (with \( k < n \)), we assume that we have found orthonormal sequence \( (x_i)_{i=1}^k \subseteq X \), \( x_i \) is an eigenvector of \( Q_i \) (and \( Q \)) associated with the eigenvalue \( \lambda_i \in (0, \infty) \subset \mathbb{R} \), \( i = 1, \ldots, k \), and \( Q = \sum_{j=1}^i \lambda_j x_j x_j^* + Q_{j+1}, \quad i = 1, \ldots, k \). At step \( k + 1 \), since \( Q_{k+1} \neq \vartheta_{B(X,X)} \), then \( \|Q_{k+1}\|_{B(X,Y)} =: \lambda_{k+1} > 0 \), by (i), there exists an eigenvector \( x_{k+1} \in X \) with \( \|x_{k+1}\| = 1 \) of \( Q_{k+1} \) associated with the eigenvalue \( \lambda_{k+1} \). Furthermore, \( Q_{k+1} = \lambda_{k+1} x_{k+1} x_{k+1}^* + P_{1,k+1} + P_{k+1} \). Now, by Claim 13.50.1, \( x_{k+1} \) is an eigenvector of \( Q_k \) associated with the eigenvalue \( \lambda_{k+1} \), and \( x_{k+1} \perp x_k \). Recursively, by Claim 13.50.1, \( x_{k+1} \) 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 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 \( Q_{i+1} =: \bar{P}_i Q_i \bar{P}_i \). By Definition 13.40, \( Q_{i+1} \in \mathcal{S}_{\text{psd}} X \). Then, \( Q \geq \sum_{j=1}^i \lambda_j x_j x_j^* , \forall i \in \mathbb{N} \). Hence, \( Q \geq \sum_{j=1}^\infty \lambda_j x_j x_j^* \). This completes the proof of (ii).

(iii) In the latter case of (ii), we have an orthonormal sequence \( (x_i)_{i=1}^\infty \). Each \( x_i \) is an eigenvector of \( Q \) associated with the eigenvalue \( \lambda_i > 0 \). Clearly, \( \lambda_1 \geq \lambda_2 \geq \cdots \). Suppose that \( \lim_{i \in \mathbb{N}} \lambda_i > 0 \). Then, \( Q x_i = \lambda_i x_i \). 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, \( \lim_{i \in \mathbb{N}} \lambda_i = 0 \). If the sequence \( (x_i)_{i=1}^\infty \) is complete, then \( \forall x \in X \), we have \( x = \sum_{i=1}^\infty x_i \langle x, x_i \rangle = \sum_{i=1}^\infty x_i \langle x_i, x \rangle \). We will first show that the right-hand-side of the equation we are to prove is a well-defined bounded linear operator (which will be denoted by \( \tilde{Q} \)). Since \( \lambda_i \)'s are eigenvalues of \( Q \), by Proposition 13.47, \( |\lambda_i| \leq \|Q\|_{B(X,X)} \). By Bessel's Inequality, we have \( \|x\|^2 \geq \sum_{i=1}^\infty \langle x, x_i \rangle^2 \). Then, \( \|\tilde{Q}(x)\|^2 \leq \|Q\|^2_{B(X,X)} \|x\|^2 \). Hence, \( \tilde{Q} : X \to X \), is linear, and has norm bounded by \( \|Q\|_{B(X,X)} \). Hence, it is a well-defined linear operator.

Then \( \tilde{Q}(x) = \lim_{k \in \mathbb{N}} Q(\sum_{i=1}^k x_i \langle x, x_i \rangle) = \lim_{k \in \mathbb{N}} (\sum_{i=1}^k \lambda_i x_i x_i^* + Q_{k+1})(\sum_{i=1}^k x_i \langle x, x_i \rangle) = \lim_{k \in \mathbb{N}} \left( \sum_{i=1}^k \lambda_i \langle x, x_i \rangle x_i + Q_{k+1}(\sum_{i=1}^k \lambda_i \langle x, x_i \rangle x_i) = \lim_{k \in \mathbb{N}} \tilde{Q}(\sum_{i=1}^k x_i \langle x, x_i \rangle) = \tilde{Q}(x) \right) \), where the first equality follows from the fact that \( Q \) is continuous and Proposition 3.66; the second equality follows from (ii); the third equality follows from \( (x_i)_{i=1}^k \) is orthonormal; the fourth equality follows from the expression for \( Q_{k+1} = \bar{P}_{1,k} \cdots \bar{P}_{1,1} Q_{k+1} \bar{P}_{1,1} \cdots \bar{P}_{1,k} \), which implies that \( Q_{k+1}(x_i) = \vartheta_X, i = 1, \ldots, k \); the fifth equality follows from the expression for \( Q \); and the last equality follows from Proposition 3.66 and the fact that \( \tilde{Q} \) is continuous. Hence, \( Q = \tilde{Q} \) by the arbitrariness of \( x \).
This completes the proof of the theorem. \qed
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 $\mathcal{Y}$ be a separable Banach space over $\mathbb{K}$. A $\mathcal{Y}$-valued random variable is a $\mathcal{B}$-measurable function $x : \Omega \to \mathcal{Y}$. When $\mathcal{Y} = \mathbb{R}$, 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 $E(x)$.

Lemma 14.2 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 $\mu_1(E_2 \setminus E_1) = \mu_1(E_2) - \mu_1(E_1) = \mu_2(E_2) - \mu_2(E_1) = \mu_2(E_2 \setminus E_1)$, where the first and last equalities follow 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 $A_i \in \mathcal{D}, \forall i \in \mathbb{N}$. Then, $\mu_1(\bigcup_{i=1}^\infty E_i) = \mu_1(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu_1(A_i) = \sum_{i=1}^\infty \mu_2(A_i) = \mu_2(\bigcup_{i=1}^\infty A_i) = \mu_2(\bigcup_{i=1}^\infty E_i)$. Hence, $\bigcup_{i=1}^\infty E_i \in \mathcal{D}$. The preceding discussion implies that $\mathcal{D}$ is a monotone class on $\Omega$.

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

Definition 14.3 Let $(\Omega, \mathcal{B})$ be a measurable space, $\mathcal{Y}$ be a topological space, $\mathcal{F} \subseteq \mathcal{B}$, and $x : \Omega \to \mathcal{Y}$ be $\mathcal{B}$-measurable. We will denote the $\sigma$-algebra
generated by $\mathcal{F}$ by $\sigma(\mathcal{F}) \subseteq \mathcal{B}$; and denote the $\sigma$-algebra generated by $\{E \in \mathcal{B} \mid E = x_{\text{inv}}(O), O \in \mathcal{O}_x \}$ by $\sigma(x) \subseteq \mathcal{B}$. Clearly, $\sigma(x)$ is the smallest $\sigma$-algebra on which $x$ is measurable.

**Proposition 14.4** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, $n \in \mathbb{N}$, $F_1, \ldots, F_n \in \mathcal{B}$. Then,

$$P(\bigcup_{i=1}^{n} F_i) = \sum_{l=1}^{n} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq n} P(\bigcap_{j=1}^{l} F_{i_j}) \quad (14.1)$$

**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

$$P(\bigcup_{i=1}^{k+1} F_i) = P(\bigcup_{i=1}^{k} F_i) + P(F_{k+1}) - P(\bigcup_{i=1}^{k} (F_i \cap F_{k+1}))$$

$$= \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} F_{i_j}) + P(F_{k+1})$$

$$- \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} (F_{i_j} \cap F_{k+1}))$$

$$= \sum_{l=1}^{k} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} F_{i_j}) + P(F_{k+1})$$

$$+ \sum_{l=1}^{k} (-1)^{l} \sum_{1 \leq i_1 < \cdots < i_l \leq k} P(\bigcap_{j=1}^{l} (F_{i_j} \cap F_{k+1}))$$

$$= \sum_{l=1}^{k+1} (-1)^{l-1} \sum_{1 \leq i_1 < \cdots < i_l \leq k+1} P(\bigcap_{j=1}^{l} F_{i_j})$$

This completes the induction process and therefore the proof of the proposition.

**Lemma 14.5 (Fatou’s Lemma)** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space, and $(E_n)_{n=1}^{\infty} \subseteq \mathcal{B}$. Define $F_1 := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m = \{\omega \in \Omega \mid \omega \in E_i \text{ for infinitely many numbers of } i \in \mathbb{N}\}$, and $F_2 :=$
\[ \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} E_m = \{ \omega \in \Omega \mid \omega \in E_i, \text{ eventually for all } i \in \mathbb{N} \text{ with } i \geq n_\omega \} \]. 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** \( \forall \omega \in F_2, \exists n_\omega \in \mathbb{N}, \omega \in \bigcap_{m=n}^{\infty} E_m, \forall n \in \mathbb{N} \) 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 \( \chi_{F_1, \Omega} = \liminf_{n \in \mathbb{N}} \chi_{E_n, \Omega} \leq \limsup_{n \in \mathbb{N}} \chi_{E_n, \Omega} = \chi_{F_1, \Omega} \).

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 \( G_n := \bigcup_{m=n}^{\infty} E_m, \forall n \in \mathbb{N} \). Let \( G_n \leq G_n = G_n, \forall n \in \mathbb{N} \). By Proposition 11.5, \( P(F_1) = P(\bigcap_{n=1}^{\infty} G_n) = \limsup_{n \in \mathbb{N}} P(G_n) \leq \limsup_{n \in \mathbb{N}} P(E_n) \).

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

**Definition 14.6** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( \Lambda \) be a set, and \( \mathcal{B}_\lambda \subseteq \mathcal{B} \) be a \( \sigma \)-algebra on \( \Omega \), \( \forall \lambda \in \Lambda \). We will say that the collection of \( \sigma \)-algebras \( \left( \mathcal{B}_\lambda \right)_{\lambda \in \Lambda} \) 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 14.7 (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) \( P(F_1) = \lim_{n \in \mathbb{N}} P(\bigcup_{m=n}^{\infty} E_m) \leq \limsup_{n \in \mathbb{N}} \sum_{m=n}^{\infty} P(E_m) = 0 \), where the first equality follows from the proof of Lemma 11.5; 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 \( P(\bigcap_{m=n}^{\infty} (\Omega \setminus E_m)) = \lim_{k \in \mathbb{N}} P(\bigcap_{m=n}^{\infty} (\Omega \setminus E_m)) = \lim_{k \in \mathbb{N}} \prod_{m=n}^{k} P(\Omega \setminus E_m) = \lim_{k \in \mathbb{N}} \prod_{m=n}^{k} (1 - p_m) = \prod_{m=n}^{\infty} (1 - p_m), \forall n \in \mathbb{N} \), where the first equality follows from Proposition 11.5; the second equality follows from the independence assumption; and the third equality follows from Proposition 11.6; and the last equality follows from the assumption.

This completes the proof of the lemma. \( \square \)
Definition 14.8 Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( \mathcal{Y} \) be a separable Banach space over \( \mathbb{K} \), \( x \in L_1(\Omega, \mathcal{Y}) \), \( \mathcal{B} \subseteq \mathcal{B} \) be a \( \sigma \)-algebra on \( \Omega \), and \( \bar{P} := P|_{\mathcal{B}} \). By Proposition 11.13, \( \Omega := (\Omega, \mathcal{B}, \bar{P}) \) is a probability measure space. Define \( \nu : \mathcal{B} \to \mathcal{Y} \) by \( \nu(E) = \int_E x \, dP \), \( \forall E \in \mathcal{B} \). By Proposition 11.116, \( \nu \) is a finite \( \mathcal{Y} \)-valued measure on \( (\Omega, \mathcal{B}) \) with \( P \circ \nu(E) = \int_E P \circ x \, dP \), \( \forall E \in \mathcal{B} \). By Definition 11.166 and Proposition 11.116, \( x = \frac{d\nu}{dP} \) 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 \( P \circ \nu \leq (P \circ \nu)|_{\mathcal{B}} \). Then, \( (\Omega, \mathcal{B}, \nu) \) is a finite \( \mathcal{Y} \)-valued measure space. \( \forall E \in \mathcal{B} \) with \( P(E) = 0 \), we have \( E \in \mathcal{B} \), \( P(E) = 0 \), and \( 0 \leq P \circ \nu(E) \leq P \circ \nu(E) = \int_E P \circ x \, dP = 0 \). Hence, \( P \circ \nu \ll \bar{P} \). Let there exists \( f : \Omega \to \mathcal{Y} \) that is the Radon-Nikodym derivative of \( \nu \) with respect to \( \bar{P} \), where \( f \) is \( \mathcal{B} \)-measurable. The function \( f \) is unique as described in Proposition 11.167, \( f = \frac{d\nu}{d\bar{P}} \) a.e. in \( \Omega \), and \( f \in L_1(\Omega, \mathcal{Y}) \). Then, \( \{ f \} \in L_1(\Omega, \mathcal{Y}) \) is said to the conditional expectation of \( x \) given \( \mathcal{B} \) and is denoted by \( E(x|\mathcal{B}) \). \( f \in E(x|\mathcal{B}) \) is a \( \mathcal{Y} \)-valued random variable and is said to be a version of the conditional expectation. Then,

\[
\int_{E} x \, d\bar{P} = \nu(E) = \int_{E} f \, d\bar{P} = \int_{E} f \, dP; \quad \forall E \in \mathcal{B}(14.2a)
\]

\[
\int_{E} P \circ f \, d\bar{P} = P \circ \nu(E) = \int_{E} P \circ f \, d\bar{P} \leq P \circ \nu(E) = \int_{E} P \circ x \, d\bar{P}
\]

\[
< \infty; \quad \forall E \in \mathcal{B}
\]

(14.2b)

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 \( E(x|\mathcal{B}) \in L_1(\Omega, \mathcal{Y}) \) exists by Radon-Nikodym Theorem 11.171.

Proposition 14.9 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 \( \Omega \), \( \bar{\Omega} := (\Omega, \mathcal{B}, \bar{P} := P|_{\mathcal{B}}) \), and \( f \in L_1(\Omega, \mathcal{Y}) \). Assume that \( f \) satisfies (14.2a), then \( E(x|\mathcal{B}) \) exists and \( f \in E(x|\mathcal{B}) \).

Proof Let \( \nu \) and \( \bar{\nu} \) be as defined in Definition 14.8. By Proposition 11.116, we may define a the \( \mathcal{Y} \)-valued measure \( \bar{\nu} \) with kernel \( f \) on \( (\Omega, \mathcal{B}) \). \( \bar{\nu} \) is finite since \( f \in L_1(\Omega, \mathcal{Y}) \). By (14.2a), \( \bar{\nu}(E) = \nu(E), \forall E \in \mathcal{B} \). Therefore, \( \bar{\nu} = \nu \). By Definition 11.166 and Proposition 11.116, \( f = \frac{d\nu}{d\bar{P}} \) a.e. in \( \Omega \). Hence, \( E(x|\mathcal{B}) \) exists and \( f \in E(x|\mathcal{B}) \). This completes the proof of the proposition. \( \square \)

Proposition 14.10 Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( \mathcal{Y} \) be a separable Banach space over \( \mathbb{K} \), \( Z \) be a separable Banach space over \( \mathbb{K} \),
\( \mathcal{W} \subseteq \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \) be a separable subspace, \( \hat{\mathcal{B}} \subseteq \mathcal{B} \) and \( \hat{\mathcal{B}} \subseteq \mathcal{B} \) be independent \( \sigma \)-algebras on \( \Omega \), \( \hat{\Omega} := (\Omega, \hat{\mathcal{B}}, \hat{\mathcal{P}} := \mathcal{P}|_{\hat{\mathcal{B}}}) \) and \( \hat{\Omega} := (\Omega, \hat{\mathcal{B}}, \hat{\mathcal{P}} := \mathcal{P}|_{\hat{\mathcal{B}}}) \) be probability measure spaces, \( x \in L_1(\hat{\Omega}, \mathcal{Y}) \) be \( \hat{\mathcal{B}} \)-measurable, \( A \in L_1(\hat{\Omega}, \mathcal{W}) \) be \( \hat{\mathcal{B}} \)-measurable. Assume that \( g := \| A(\cdot) \|_{B(\mathcal{Y}, \mathcal{Z})} \| x(\cdot) \|_{\Omega} : \Omega \rightarrow [0, \infty) \subseteq \mathbb{R} \) is integrable over \( \Omega \). Then, \( Ax \in L_1(\Omega, \mathcal{Z}) \) 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 \( \varnothing \in \mathcal{B}(\mathcal{Y}, \mathcal{Z}) \), and \( E_1, \ldots, E_n \in \hat{\mathcal{B}} \) are nonempty and pairwise disjoint. Let \( x \) admit the canonical representation \( x = \sum_{j=1}^{m} x_j \chi_{Y_j, \Omega} \), where \( m \in \mathbb{Z}_+ \), \( x_1, \ldots, x_m \in \mathcal{Y} \) are distinct and none equal to \( \varnothing \in \mathcal{Y} \) and \( F_1, \ldots, F_m \in \hat{\mathcal{B}} \) are nonempty and pairwise disjoint. Then, clearly, \( Ax \in L_1(\Omega, \mathcal{Z}) \) is \( \hat{\mathcal{B}} \)-measurable. This leads to

\[
E(Ax) = \int_{\Omega} \sum_{i=1}^{n} \sum_{j=1}^{m} A_i x_j \chi_{E_i \cap Y_j, \Omega} \, d\mathcal{P} = \sum_{i=1}^{n} \sum_{j=1}^{m} A_i x_j \int_{\Omega} \chi_{E_i \cap Y_j, \Omega} \, d\mathcal{P} = \sum_{i=1}^{n} \sum_{j=1}^{m} A_i x_j P(E_i)P(F_j) = \sum_{i=1}^{n} A_i P(E_i) \sum_{j=1}^{m} x_j P(F_j) = E(A)E(x)
\]

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 \( A \in L_1(\hat{\Omega}, \mathcal{W}) \), by Proposition 11.66, there exists a sequence of simple functions \( \phi_n \) \( \| \phi_n(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \leq \| A(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \), \( \forall \omega \in \hat{\Omega} \), \( \forall n \in \mathbb{N} \). Since \( x \in L_1(\hat{\Omega}, \mathcal{Y}) \), again by Proposition 11.66, there exists a sequence of simple functions \( \psi_n \) \( \| \psi_n(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \leq \| x(\omega) \|_{\mathcal{Y}} \), \( \forall \omega \in \hat{\Omega} \), \( \forall n \in \mathbb{N} \). By Propositions 7.65, 11.39, and 11.38, \( Ax \) is \( \hat{\mathcal{B}} \)-measurable and \( \phi_n \psi_n \) is \( \mathcal{B} \)-measurable, \( \forall n \in \mathbb{N} \). By the assumption, \( Ax \in L_1(\Omega, \mathcal{Z}) \). Clearly, \( \lim_{n \in \mathbb{N}} \phi_n = A \) a.e. in \( \hat{\Omega} \) and \( \lim_{n \in \mathbb{N}} \psi_n = x \) a.e. in \( \hat{\Omega} \) and \( \| \phi_n(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \leq \| \phi_n(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \| \psi_n(\omega) \|_{\mathcal{Y}} \leq \| \phi_n(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \| \psi_n(\omega) \|_{\mathcal{Y}} \leq \| A(\omega) \|_{B(\mathcal{Y}, \mathcal{Z})} \| x(\omega) \|_{\mathcal{Y}} = g(\omega) \). Then, we have

\[
E(Ax) = \int_{\Omega} Ax \, d\mathcal{P} = \lim_{n \in \mathbb{N}} \int_{\Omega} \phi_n \psi_n \, d\mathcal{P} = \lim_{n \in \mathbb{N}} E(\phi_n \psi_n) = \lim_{n \in \mathbb{N}} E(\phi_n)E(\psi_n) = \lim_{n \in \mathbb{N}} \int_{\Omega} \phi_n \, d\mathcal{P} \int_{\Omega} \psi_n \, d\mathcal{P}
\]

where \( \mathcal{P} \) is the probability measure.
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.

\[ \]
third equality follows from the fact that \( E_1, E_2 \in G \); and the fourth equality follows from the fact that \( E_1 \subseteq E_2 \). Hence, by the arbitrariness of \( E \in \mathcal{B} \), we have \( E_2 \setminus E_1 \in G \). (iii) Fix any \( (E_i)_{i=1}^{\infty} \subseteq G \) with \( E_i \subseteq E_{i+1} \), \( \forall i \in \mathbb{N} \). For each \( \forall \hat{E} \in \mathcal{B} \), we have \( P((\bigcup_{i=2}^{\infty} E_i) \cap \hat{E}) = P((\bigcup_{i=2}^{\infty} (E_i \setminus E_{i-1}) \cap \hat{E}) \cup (E_1 \cap \hat{E})) = \sum_{i=2}^{\infty} P((E_i \setminus E_{i-1}) \cap \hat{E}) + P(E_1 \cap \hat{E}) = \sum_{i=2}^{\infty} P(E_i \setminus E_{i-1})P(\hat{E}) + P(E_1)P(\hat{E}) = P(\hat{E})(\sum_{i=1}^{\infty} P(E_i \setminus E_{i-1}) + P(E_1)) = P(\hat{E})P(\bigcup_{i=1}^{\infty} E_i) \), where the first two equalities follow from Proposition 2.5; the third equality follows from the fact that \( \hat{E} \) is \( \sigma \)-algebra on \( \Omega \); the fourth equality follows from countable additivity of measure; the fourth equality follows from (iii); and the last equality follows from Proposition 2.5; the third equality follows from countable additivity of measure. By the arbitrariness of \( \hat{E} \in \mathcal{B} \), we have \( \bigcup_{i=1}^{\infty} E_i \in G \). This proves that \( G \) is a monotone class on \( \Omega \).

This proves that \( \forall \hat{E} \in \mathcal{B} \) and \( \forall \hat{E} \in \mathcal{B} \), we have \( P(\hat{E} \cap \hat{E}) = P(\hat{E})P(\hat{E}) \).

Hence, \( \mathcal{B} \) and \( \mathcal{B} \) are independent \( \sigma \)-algebras.

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

**Proposition 14.12** Let \( \Omega := (\Omega, \mathcal{B}, P) \) be a probability measure space, \( Y \) be a separable Banach space over \( \mathbb{K} \), \( x \in L_1(\Omega, Y) \), \( (x_n)_{n=1}^{\infty} \subseteq L_1(\Omega, Y) \), \( \mathcal{B} \subseteq \mathcal{B} \) be a \( \sigma \)-algebra on \( \Omega \), \( \Omega := (\Omega, \mathcal{B}, \hat{P} := P|_G) \), \( f \in E(x|\mathcal{B}) \), 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 \( [x] = E(x|\mathcal{B}) \in L_1(\Omega, Y) \).

(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 \( F \subseteq Y \) be nonempty closed convex set, and \( x : \Omega \to F \). Then, \( f(\omega) \in F \) a.e. \( \omega \in \Omega \).

(e) (Monotone Convergence) Let \( Y = \mathbb{R} \), \( 0 \leq x_n(\omega) \leq x_{n+1}(\omega) \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{N} \), and \( \lim_{n \to \infty} x_n(\omega) = x(\omega) \) a.e. \( \omega \in \Omega \). Then \( f_n \leq f_{n+1} \) a.e. in \( \Omega \) and \( \lim_{n \to \infty} f_n = \hat{f} \) a.e. in \( \Omega \).

(f) (Fatou) Let \( Y = \mathbb{R} \), \( 0 \leq x_n(\omega) < \infty \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{N} \), and \( \bar{f} \in E(\liminf_{n \to \infty} x_n|\mathcal{B}) \). Then \( \bar{f} \leq \liminf_{n \to \infty} f_n \) 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) \( \lim_{n \to \infty} x_n = \bar{x} \) a.e. in \( \Omega \), where \( \bar{x} : \Omega \to Y \) is \( \mathcal{B} \)-measurable;

(ii) \( P \circ x_n(\omega) \leq g_n(\omega) \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{N} \);

(iii) \( \lim_{n \to \infty} g_n = g \) a.e. in \( \Omega \), and \( E(g) = \lim_{n \to \infty} E(g_n) < \infty \).
Then, \( \bar{x} \in \hat{L}_1(\Omega, \mathcal{Y}) \), \( E(\bar{x} | \hat{\mathcal{B}}) \) exists, \( \lim_{n \to \infty} f_n = \tilde{f} \) a.e. in \( \hat{\Omega} \), \( \lim_{n \to \infty} \int_{\Omega} P \circ f_n dP = \int_{\Omega} \tilde{f} dP < \infty \), and \( \lim_{n \to \infty} \int_{\Omega} P \circ (f_n - \tilde{f}) dP = 0 \), where \( \tilde{f} \in E(\bar{x} | \hat{\mathcal{B}}) \).

(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 \( \hat{\mathcal{B}} \)-measurable, and \( \| A(\cdot) \|_{\mathcal{B}(\mathcal{Y}, \mathcal{Z})} \| x(\cdot) \|_y : \Omega \to [0, \infty) \subseteq \mathbb{R} \) be integrable over \( \Omega \). Then, \( E(Ax | \hat{\mathcal{B}}) \) exists and \( Af \in E(Ax | \hat{\mathcal{B}}) \).

(i) (Jensen’s Inequality) Let \( \mathcal{K} = \mathbb{R} \), \( F \subseteq \mathcal{Y} \) 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 \( \tilde{f} \in E(c \circ x | \hat{\mathcal{B}}) \). Then \( c \circ \tilde{f} \leq \tilde{f} \) a.e. in \( \hat{\Omega} \).

(j) (Tower Property) Let \( \hat{\mathcal{B}} \subseteq \hat{\mathcal{B}} \) be a \( \sigma \)-algebra on \( \Omega \). Then \( E(f | \hat{\mathcal{B}}) = E(f | \bar{\mathcal{B}}) \), whenever one of them exists.

(k) (Role of Independence) Let \( \hat{\mathcal{B}} \subseteq \hat{\mathcal{B}} \) be a \( \sigma \)-algebra on \( \Omega \), \( \hat{\mathcal{B}} \) and \( \sigma(\hat{\mathcal{B}} \cup \sigma(x)) \) are independent. Then \( f \in E(x | \sigma(\hat{\mathcal{B}} \cup \hat{\mathcal{B}})) \).

Proof

(a) \( E(f) = \int_{\Omega} f dP = \int_{\Omega} \tilde{f} d\hat{P} = \int_{\Omega} x d\hat{P} = E(x) \), where the first equality follows from the definition of \( E \); the second equality follows from \( f \) is \( \hat{\mathcal{B}} \)-measurable and Proposition 11.72; the third equality follows from the definition of \( E(x | \hat{\mathcal{B}}) \) and the fact \( \Omega \in \hat{\mathcal{B}} \); 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 14.8. \( \forall E \in \hat{\mathcal{B}} \), we have \( P \circ \hat{\nu}(E) = \int_{\hat{\mathcal{B}}} P \circ \hat{\nu} x d\hat{P} = \int_{\hat{\mathcal{B}}} P \circ x dP < \infty \), where the first equality follows from Proposition 11.116; and the second equality follows from the fact that \( P \circ x \) is \( \hat{\mathcal{B}} \)-measurable and Proposition 11.72. Furthermore, we have \( \nu(E) = \int_{\hat{\mathcal{B}}} x d\hat{P} = \int_{\hat{\mathcal{B}}} x d\hat{P} = \hat{\nu}(E) \), where the second equality follows from the fact that \( x \) is \( \hat{\mathcal{B}} \)-measurable and Proposition 11.72. By Proposition 14.9, we have \( x \in E(x | \hat{\mathcal{B}}) \).

(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 \( \forall E \in \hat{\mathcal{B}} \) with \( \hat{P}(E) > 0 \), \( \frac{1}{\hat{P}(E)} \int_{\hat{E}} f d\hat{P} = \frac{1}{P(F)} \int_{E} f dP \in F \), where the equality follows from \( f \in E(x | \hat{\mathcal{B}}) \). Hence, \( \frac{1}{\hat{P}(E)} \int_{\hat{E}} f d\hat{P} \in F, \forall E \in \hat{\mathcal{B}} \) with \( \hat{P}(E) > 0 \). Suppose that the result does not hold. Then, we have \( E := \{ \omega \in \Omega \mid f(\omega) \notin F \} \in \hat{\mathcal{B}} \) and \( \hat{P}(E) > 0 \). 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 \( \mathcal{Y}, \overline{B_n} \subseteq \mathcal{Y} \setminus F, \forall n \in \mathbb{N} \), and \( \mathcal{Y} \setminus F = \bigcup_{n=1}^{\infty} B_n \). Let \( E_n := f_{\text{inv}}(B_n) \in \hat{\mathcal{B}} \). Then, \( \hat{P}(E) = \hat{P}(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \hat{P}(E_n) \). Since \( \hat{P}(E) > 0 \), we have \( \exists n_0 \in \mathbb{N} \) such that \( \hat{P}(E_{n_0}) > 0 \). 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). \( \frac{1}{P(E_{n_0})} \int_{E_{n_0}} f \, dP \in B_{n_0} \) since \( B_{n_0} \) a nonempty convex set. This contradicts our earlier conclusion that \( \frac{1}{P(E_{n_0})} \int_{E_{n_0}} f \, dP \in F \subseteq \mathbb{Y} \setminus B_{n_0} \). Hence, the hypothesis does not hold. Hence, (d) holds.

(c) By (c), we have \( f_{n_1} - f_n \in E(x_{n_1} - x_n[B]) \), \( \forall n \in \mathbb{N} \). since \( x_{n_1} - x_n \geq 0 \), \( \forall \omega \in \Omega \), \( \forall n \in \mathbb{N} \), then, by (d), \( f_{n_1} - f_n \geq 0 \) a.e. in \( \tilde{\Omega} \), \( \forall n \in \mathbb{N} \). Then, \( f_n \leq f_{n_1} \) a.e. in \( \tilde{\Omega} \). \( \int_{\tilde{\Omega}} \mathcal{P} \circ (f - f_n) \, dP = \int_{\tilde{\Omega}} \mathcal{P} \circ (f - f) \, dP \leq \int_{\tilde{\Omega}} \mathcal{P} \circ (x_n - x) \, dP \), where the first equality follows from Proposition 11.72; and the inequality follows from (14.2b). By Monotone Convergence Theorem 11.81, \( \int_{\tilde{\Omega}} x \, dP = \lim_{n \in \mathbb{N}} \int_{\Omega} x_n \, dP \). By Definition 14.8, \( f_n, f \in L_1(\tilde{\Omega}, \mathbb{R}) \), \( \forall n \in \mathbb{N} \). Then, we have \( \lim_{n \in \mathbb{N}} f_n = f \) in \( L_1(\tilde{\Omega}, \mathbb{R}) \). Hence, \( f = \lim_{n \in \mathbb{N}} f_n \) a.e. in \( \tilde{\Omega} \).

(f) \( \forall n \in \mathbb{N} \), define \( \hat{x}_n : \Omega \to \mathbb{R} \) by \( \hat{x}_n(\omega) := \inf_{m \in \mathbb{N}, n \leq m} x_m(\omega) \in [0, \infty) \subset \mathbb{R}, \forall \omega \in \Omega \). Then, \( \hat{x}_n(\omega) \leq \hat{x}_{n+1}(\omega), \forall \omega \in \Omega \), and \( \hat{x}_n \) is integrable over \( \Omega \), \( \forall n \in \mathbb{N} \). By monotonicity, we have \( \lim_{n \in \mathbb{N}} \hat{x}_n = \liminf_{n \in \mathbb{N}} x_n =: \hat{x} : \Omega \to [0, \infty) \subset \mathbb{R} \). Since \( \tilde{f} \in E(\hat{x} | \hat{B}) \) exists, then \( [\hat{x}] \in L_1(\tilde{\Omega}, \mathbb{R}) \). Thus, we may have \( \hat{x} \in [\hat{x}] \) where \( \hat{x} \in L_1(\tilde{\Omega}, \mathbb{R}) \). This leads to \( \liminf_{n \in \mathbb{N}} x_n = \hat{x} \) a.e. in \( \tilde{\Omega} \). By (e), we have \( f = \lim_{n \in \mathbb{N}} f_n \) a.e. in \( \tilde{\Omega} \), where \( \tilde{f}_n \in E(\hat{x}_n | \hat{B}) \), \( \forall n \in \mathbb{N} \). By (e), we have \( \tilde{f}_n \leq \tilde{f}_m \) a.e. in \( \tilde{\Omega} \), \( \forall n, m \in \mathbb{N} \) with \( n \leq m \). Hence, \( \tilde{f} \leq \liminf_{n \in \mathbb{N}} f_n \) a.e. in \( \tilde{\Omega} \).

(g) By Lebesgue Dominated Convergence Theorem 11.91, \( \tilde{x} \) is absolutely integrable over \( \tilde{\Omega} \), \( 0 \leq \int_{\tilde{\Omega}} \mathcal{P} \circ \tilde{x} \, dP = \lim_{n \in \mathbb{N}} \int_{\tilde{\Omega}} \mathcal{P} \circ x_n \, dP \leq \liminf_{n \in \mathbb{N}} E(g_n) = E(g) < \infty \), and \( E(\hat{x}) = \liminf_{n \in \mathbb{N}} E(x_n) \in \mathbb{Y} \). This implies that \( \hat{x} \in L_1(\tilde{\Omega}, \mathbb{Y}) \) and \( x_n - \hat{x} \in L_1(\tilde{\Omega}, \mathbb{Y}) \). Note that \( \mathcal{P} \circ (x_n - \hat{x}) \leq g_n + g \) a.e. in \( \tilde{\Omega} \), \( \lim_{n \in \mathbb{N}} (x_n - \hat{x}) = 0 \) a.e. in \( \tilde{\Omega} \). Then, by Lebesgue Dominated Convergence Theorem 11.91, we have \( \lim_{n \in \mathbb{N}} \int_{\tilde{\Omega}} \mathcal{P} \circ (x_n - \hat{x}) \, dP = 0 \). This yields that \( \liminf_{n \in \mathbb{N}} x_n = \hat{x} \) in \( L_1(\tilde{\Omega}, \mathbb{Y}) \), and \( ((x_n)_{n=1}^\infty) \subseteq L_1(\tilde{\Omega}, \mathbb{Y}) \) is a Cauchy sequence. By (c), \( f_n \in E((x_n - \hat{x}) | \hat{B}) \). By (14.2b), \( \int_{\tilde{\Omega}} \mathcal{P} \circ (f_n - f_m) \, dP \leq \int_{\tilde{\Omega}} \mathcal{P} \circ (x_n - x_m) \, dP \), \( \forall n, m \in \mathbb{N} \). Then, \( ((f_n)_{n=1}^\infty) \subseteq L_1(\tilde{\Omega}, \mathbb{Y}) \) is a Cauchy sequence. By Example 11.179, \( L_1(\tilde{\Omega}, \mathbb{Y}) \) is a Banach space. Then, \( \exists \hat{f} \in L_1(\tilde{\Omega}, \mathbb{Y}) \) such that \( \lim_{n \in \mathbb{N}} f_n = \hat{f} \) in \( L_1(\tilde{\Omega}, \mathbb{Y}) \). Hence, \( f_n = \hat{f} \) a.e. in \( \tilde{\Omega} \) and \( \lim_{n \in \mathbb{N}} \int_{\tilde{\Omega}} \mathcal{P} \circ (f_n - \hat{f}) \, dP = 0 \). By the continuity of norm, we have \( \lim_{n \in \mathbb{N}} \int_{\tilde{\Omega}} \mathcal{P} \circ f_n \, dP = \int_{\tilde{\Omega}} \mathcal{P} \circ \hat{f} \, dP \).

Now, we will show that \( \hat{f} \in E(\hat{x} | \hat{B}) \). \( \forall E \in \hat{B} \), we have \( \int_E \hat{f} \, dP = \int_{\tilde{E}} \hat{f} \, dP = \lim_{n \in \mathbb{N}} \int_{\tilde{E}} f_n \, dP = \lim_{n \in \mathbb{N}} \int_{\tilde{E}} x_n \, dP = \int_{\tilde{E}} \tilde{x} \, dP \in \mathbb{Y} \), where the first equality follows from the fact that \( \hat{f} \) is \( \hat{B} \)-measurable and Proposition 11.72; the second equality follows from Lebesgue Dominated Convergence Theorem 11.91; the third equality follows from \( f_n \in E(x_n | \hat{B}) \); and the fourth equality follows from Lebesgue Dominated Convergence Theorem 11.91. By Proposition 14.9, we have \( f \in E(\hat{x} | \hat{B}) \).

(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 \( \vartheta_{B(y,z)} \), \( E_i \)'s are pairwise disjoint and nonempty. Since \( A \) is \( \mathcal{B} \)-measurable, we have \( E_i \in \mathcal{B} \), \( i = 1, \ldots, m \).

By Propositions 7.65, 11.39, and 11.38, \( Af \) is \( \mathcal{B} \) measurable, and \( Ax \) is \( \mathcal{B} \)-measurable. By \( A \) being a simple function, we have \( Af \in L_1(\Omega, \mathcal{B}, \hat{P}) \).

By the assumption, \( Ax \in L_1(\Omega, \mathcal{B}) \). \( \forall E \in \mathcal{B}, \int_E Af \, d\hat{P} = \int_E Af \, d\hat{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i f \, d\hat{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i \chi_{E_i, \Omega} \, d\hat{P} = \sum_{i=1}^{m} \int_{E \cap E_i} A_i \chi_{E_i, \Omega} \, d\hat{P} = \int_E Ax \, d\hat{P}, \) 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 \( f \in \mathcal{E}(x|\mathcal{B}) \); and the sixth through eighth equalities follow from Proposition 11.92. By Proposition 14.9, we have \( \mathcal{E}(Ax|\mathcal{B}) \) exists and \( Af \in \mathcal{E}(Ax|\mathcal{B}) \).

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 \( \hat{\Omega} \) and \( \| \phi_n(\omega) \|_{\mathcal{B}(y,z)} \leq \| A(\omega) \|_{\mathcal{B}(y,z)} \), \( \forall \omega \in \hat{\Omega}, \forall n \in \mathbb{N} \). 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 \( g : \hat{\Omega} \to [0, \infty) \subset \mathbb{R} \) be given by \( g(\omega) = \| A(\omega) \|_{\mathcal{B}(y,z)} \| x(\omega) \|_y \), \( \forall \omega \in \hat{\Omega} \), and \( g_n = g, \forall n \in \mathbb{N} \). Then, by \( (g) \), we have \( \mathcal{E}(Ax|\mathcal{B}) \) exists and \( \hat{f} = \lim_{n \to \infty} \phi_n f = Af \) a.e. in \( \hat{\Omega} \), where \( \hat{f} \in \mathcal{E}(Ax|\mathcal{B}) \).

(i) By Proposition 8.25, \( c \) is lower semicontinuous. By Definition 3.14, \( \forall \alpha \in \mathbb{R}, c_{\inf}((\alpha, +\infty)) = \{ y \in \mathbb{R} \mid c(y) > \alpha \} \in \mathcal{O}_F \subseteq \mathcal{B}(y,z) \), where \( \mathcal{O}_F \) is the subset topology of \( (\mathbb{Y}, \mathcal{O}_y) \) on the set \( F \). By Proposition 11.35, \( c \) is \( \mathcal{B}(y,z) \)-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. \( \omega \in \hat{\Omega} \). Thus, \( c \circ f(\omega) \in \mathbb{R} \) a.e. \( \omega \in \hat{\Omega} \). Fix any \( \omega \in f_{\text{mea}}(F) \). By Proposition 8.32, we have \( c(f(\omega)) = \sup_{y \in F_{\text{con}}}(\langle y, f(\omega) \rangle) - c_{\text{con}}(y) \), where \( c_{\text{con}} : F_{\text{con}} \to \mathbb{R} \) is the conjugate functional to \( c \). Then, \( \exists \gamma \in F_{\text{con}} \) such that \( c(f(\omega)) - \epsilon \leq \langle y, f(\omega) \rangle - c_{\text{con}}(y) \leq \langle y, f(\omega) \rangle - \langle y, f(\omega) \rangle + c(y) = \langle y, f(\omega) \rangle + c(y), \forall y \in F, \) 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 \( \hat{\Omega} \). Thus, by \( (d), (h), \) and \( (c) \) and taking conditional expectation of the preceding equation given \( \mathcal{B}, c(f(\omega)) - \epsilon \leq \langle y, f(\omega) - x(\omega) \rangle + \hat{f}(\omega) \) a.e. \( \omega \in \hat{\Omega} \), which implies that \( c(f(\omega)) - \epsilon \leq \hat{f}(\omega) \) a.e. \( \omega \in \hat{\Omega} \). By the arbitrariness of \( \epsilon \), we have \( c \circ f \leq \hat{f} \) a.e. in \( \hat{\Omega} \).

(j) Let \( \mathcal{B} := (\Omega, \mathcal{B}, \hat{P} := P|\mathcal{B}) \). Assume that \( \mathcal{E}(f|\mathcal{B}) \) exists and \( \hat{f} \in \mathcal{E}(f|\mathcal{B}) \). Clearly, \( \hat{f} \in L_1(\Omega, \mathcal{B}, \hat{P}) \). \( \forall E \in \mathcal{B}, \) we have \( \int_E \hat{f} \, d\hat{P} = \int_E f \, d\hat{P} = \int_E x \, d\hat{P}, \) where the first equality follows from \( \hat{f} \in \mathcal{E}(f|\mathcal{B}) \); and the second equality follows from \( f \in \mathcal{E}(x|\mathcal{B}) \). By the arbitrariness of \( E \) and Proposition 14.9, we have \( \hat{f} \in \mathcal{E}(x|\mathcal{B}) \). Hence, we have \( \mathcal{E}(x|\mathcal{B}) = \mathcal{E}(f|\mathcal{B}) = \mathcal{E}(\mathcal{E}(x|\mathcal{B})|\mathcal{B}) \).
On the other hand, assume $E(x|\tilde{B})$ exists and $\tilde{f} \in E(x|\tilde{B})$. Clearly, $\tilde{f} \in L_1(\Omega, Y)$. $\forall E \in \tilde{B}$, we have $\int_{E} \tilde{f} \, dP = \int_{E} x \, dP = \int_{E} f \, dP$, where the first equality follows from $\tilde{f} \in E(x|\tilde{B})$; and the second equality follows from $f \in E(x|\tilde{B})$. By the arbitrariness of $E$ and Proposition 14.9, we have $\tilde{f} \in E(f|\tilde{B})$. Hence, we have $E(x|\tilde{B}) = E(f|\tilde{B}) = E(E(x|\tilde{B})|\tilde{B})$. This establishes (j).

(k) $\forall E \in \tilde{B}$ and $\forall \tilde{E} \in \tilde{B}$, we have $\int_{E \cap \tilde{E}} x \, dP = \int_{\Omega} x \, dP = \int_{\tilde{E}} x \, dP = \int_{E} f \, dP = \int_{E} \tilde{f} \, dP = \int_{E} x \, dP$, where the first equality follows from Proposition 11.92; the second equality follows from the independence of $\tilde{B}$ and $\sigma(\tilde{B} \cup \sigma(x))$ and Proposition 14.10; the fourth equality follows from $f \in E(x|\tilde{B})$; and the fifth equality follows from Proposition 14.10. By assumption, $x \in L_1(\Omega, Y)$ and $f \in L_1(\Omega, Y)$. Then, by Proposition 11.116, we may define a finite $Y$-valued measure $\nu$ on $(\Omega, B)$ by $\nu(E) := \int_{E} x \, dP$, $\forall E \in \tilde{B}$: and we may define a finite $Y$-valued measure $\tilde{\nu}$ on $E(x|\tilde{B} \cup \tilde{B})$ by $\tilde{\nu}(E) := \int_{\tilde{E}} x \, dP$, $\forall E \in \sigma(\tilde{B} \cup \tilde{B})$. By Proposition 11.107 and Definition 14.1, $\tilde{\nu} := \nu|_{\sigma(\tilde{B} \cup \tilde{B})}$ is a finite $Y$-valued measure on $\sigma(\tilde{B} \cup \tilde{B})$. We have shown in the preceding $\tilde{\nu}(E) = \tilde{\nu}(E)$, $\forall E \in \mathcal{F}$, where $\mathcal{F} := \{ E \in \tilde{B} \mid \exists \tilde{E} \in \tilde{B}, \exists E \in \tilde{B} \neq \tilde{E} \}$. It is straightforward to show that $\mathcal{F}$ is a $\pi$-system on $\Omega$ and $\mathcal{F} \subseteq \tilde{B} \cup \tilde{B}$. Therefore, $\tilde{\nu} = \nu$ by Lemma 14.2. Note that $f$ is $\sigma(\tilde{B} \cup \tilde{B})$-measurable. Then, by Proposition 14.9, $f \in E(x|\sigma(\tilde{B} \cup \tilde{B}))$.

This completes the proof of the proposition.

Proposition 14.13 Let $X_i := (\Omega_i, B_i, P_i)$ be a probability measure space, $i = 1, 2$, $X := X_1 \times X_2 =: (\Omega, B, P)$ be the product probability measure space, $Y$ be a finite-dimensional Banach space over $K$, $x : \Omega_1 \times \Omega_2 \to Y$ be $B$-measurable, $x \in L_1(\Omega, Y)$, and $\exists M \in [0, \infty) \subset \mathbb{R}$, such that $\|x(\omega_1, \omega_2)\|_Y \leq M$, $\forall (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 \subseteq \Omega$. Define $g : \Omega \to Y$ by $g(\omega_1, \omega_2) := \int_{\Omega_1} x(\omega_1, \omega_2) \, dP(\omega_1, \omega_2)$, $\forall (\omega_1, \omega_2) \in \Omega$. Then, the following statements hold.

(i) The collection of $\sigma$-algebras $(\tilde{B}_1, \tilde{B}_2)$ is independent, where $\tilde{B}_i := \{ B_1 \times B_2 \in B \mid B_j = \Omega_j, j = 1, 2 \text{ with } i \neq j \}$.

(ii) There exists $g_1 : \Omega_1 \to Y$, such that $g(\omega_1, \omega_2) = g_1(\omega_1)$, $\forall (\omega_1, \omega_2) \in \Omega$, $g_1 \in L_1(X_1, Y)$, and $g \in E(x|\tilde{B}_1)$

Proof

(i) This holds by Proposition 12.21.

(ii) Clearly, $g(\omega_1, \omega_2) = g_1(\omega_1)$, $\forall (\omega_1, \omega_2) \in \Omega$, where $g_1(\omega_1) := \int_{\Omega_1} x(\omega_1, \omega_2) \, dP(\omega_1, \omega_2) = \int_{\Omega_1} \int_{\Omega_2} x(\omega_1, \omega_2) \, dP_1(\omega_1) \, dP_2(\omega_2) = \int_{\Omega_1} \int_{\Omega_2} x(\omega_1, \omega_2) \, dP_1(\omega_1) \, dP_2(\omega_2) = \int_{\Omega_1} x(\omega_1, \omega_2) \, dP_1(\omega_1) \, dP_2(\omega_2) \subseteq \mathbb{R}$,
from Proposition 11.92. By Fubini’s Theorem 12.31, \(g_1\) is \(B_1\)-measurable and \(g_1 \in L_1(X, Y)\). Hence, \(g\) is \(B_1\)-measurable and \(g \in L_1(X, Y)\). \(\forall E \in \mathcal{B}_1\), we have \(E = E_1 \times \Omega_2\), where \(E_1 \in \mathcal{B}_1\). Then,

\[
\int_E x \, dP = \int_\Omega x \chi_{E, \Omega} \, dP = \int_{E_1} \int_{\Omega_2} x(\omega_1, \omega_2) \chi_{E_1, \Omega_1}(\omega_1) \, dP_2(\omega_2) \, dP_1(\omega_1) = \int_{E_1} \chi_{E_1, \Omega_1}(\omega_1) \int_{\Omega_2} x(\omega_1, \omega_2) \, dP_2(\omega_2) \, dP_1(\omega_1) = \int_{E_1} \chi_{E_1, \Omega_1}(\omega_1) \, dP_1(\omega_1) = \int_{E_1} g_1(\omega_1) \, dP_1(\omega_1) = \int_{E_1} g_1 \, dP_1 = \int_{E_1} g \, dP
\]

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 second equality follows from Fubini’s Theorem 12.31. By the arbitrariness of \(E\) and Proposition 14.9, we have \(g \in E(x|\mathcal{B}_1)\).

This completes the proof of the proposition.

**Theorem 14.14 (Fundamental Theorem on Modeling)** Let \(\Lambda\) be a set, \(\mathcal{X}_\alpha := (X_\alpha, B_\alpha, P_\alpha)\) be a probability measure space, \(\forall \alpha \in \Lambda\), \(\mathcal{F} := \{\prod_{\alpha \in \Lambda} B_\alpha \subseteq X := \prod_{\alpha \in \Lambda} X_\alpha \mid B_\alpha \in B_\alpha, \forall \alpha \in \Lambda\}\) except a finite number of \(\alpha\)'s, and \(\mu : \mathcal{F} \to [0, 1] \subset \mathbb{R}\) be defined by \(\mu(\prod_{\alpha \in \Lambda} B_\alpha) = \prod_{\alpha \in \Lambda} P_\alpha(B_\alpha) = \prod_{\alpha \in \Lambda, B_\alpha \subseteq X_\alpha} P_\alpha(B_\alpha), \forall \prod_{\alpha \in \Lambda} B_\alpha \in \mathcal{F}\). Then, \(\mathcal{F}\) is a semialgebra on \(X\) and there exists a unique probability measure space \((X, \mathcal{B} := \sigma(\mathcal{F}), P)\) such that \(P|_\mathcal{F} = \mu\). Furthermore, the collection of \(\sigma\)-algebras \((\mathcal{B}_\alpha)_{\alpha \in \Lambda}\) is independent, where \(\mathcal{B}_\alpha := \{\prod_{\alpha \in \Lambda} B_\alpha \subseteq \mathcal{F} \mid B_\lambda = X_\lambda, \forall \lambda \in \Lambda\) 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 semialgebra on \(X\). Clearly, \(\emptyset, X \in \mathcal{F}\). \(\forall B := \prod_{\alpha \in \Lambda} B_\alpha, C := \prod_{\alpha \in \Lambda} C_\alpha \in \mathcal{F}, B \cap C = \prod_{\alpha \in \Lambda} (B_\alpha \cap C_\alpha)\). Clearly, \(B_\alpha \cap C_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda\), since \(B_\alpha, C_\alpha \in \mathcal{B}_\alpha\). Furthermore, \(B_\alpha \cap C_\alpha = X_\alpha, \forall \alpha \in \Lambda\) except a finite number of \(\alpha\)'s, since \(B_\alpha = X_\alpha, \forall \alpha \in \Lambda\) except for a finite number of \(\alpha\)'s and \(C_\alpha = X_\alpha, \forall \alpha \in \Lambda\) except a finite number of \(\alpha\)'s. Hence, \(B \cap C \in \mathcal{F}\). Fix any \(B := \prod_{\alpha \in \Lambda} B_\alpha \in \mathcal{F}\). Then, \(B_\alpha \in \mathcal{B}_\alpha, \forall \alpha \in \Lambda\), and \(B_\alpha = X_\alpha, \forall \alpha \in \Lambda \setminus \Lambda_B\), where \(\Lambda_B = \{\alpha_1, \ldots, \alpha_n\} \subseteq \Lambda, n \in \mathbb{Z}_+\). Then, \(X \setminus B = \bigcup_{j=1}^n \bigcup_{1 \leq i_1 < i_2 < \cdots < i_j \leq n} C_j, i_1, \ldots, i_j, \alpha\) and \(C_j, i_1, \ldots, i_j, \alpha\) = \(\left\{\begin{array}{ll}
X_\alpha \setminus B_\alpha & \alpha \in \{\alpha_1, \ldots, \alpha_j\} \\
B_\alpha & \alpha \in \Lambda \setminus \{\alpha_1, \ldots, \alpha_j\} \end{array}\right\}
\). Clearly, \(C_j, i_1, \ldots, i_j \in \mathcal{F}\) and \(C_j, i_1, \ldots, i_j\)'s are pairwise disjoint. Hence, \(\mathcal{F}\) is a semialgebra on \(X\).

We will next show that \(\mu\) satisfies the assumptions of Proposition 11.32.

(i) Fix any \(B \in \mathcal{F}\), any \(n \in \mathbb{Z}_+\), and any pairwise disjoint \((C_i)_{i=1}^n \subseteq \mathcal{F}\) with
$B = \bigcup_{i=1}^{n} C_i$. Let $B := \prod_{\alpha \in \Lambda} B_\alpha$, where $B_\alpha \in \mathcal{B}_\alpha$, $\forall \alpha \in \Lambda$, and $B_\alpha = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_B$, with $\Lambda_B \subseteq \Lambda$ being a finite set. $\forall i \in \{1, \ldots, n\}$, let $C_i := \prod_{\alpha \in \Lambda} C_{i,\alpha}$, where $C_{i,\alpha} \in \mathcal{B}_\alpha$, $\forall \alpha \in \Lambda$, and $C_{i,\alpha} = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_i$ with $\Lambda_i \subseteq \Lambda$ being a finite set. Then, let $\Lambda := \Lambda_B \cup (\bigcup_{i=1}^{n} \Lambda_i)$, which is a finite subset of $\Lambda$. Then, $B_\alpha = X_\alpha$, $\forall i \in \{1, \ldots, n\}$, $\forall \alpha \in \Lambda \setminus \Lambda$. Then, $\mu(B) = \prod_{\alpha \in \Lambda} \mathbb{P}_\alpha(B_\alpha) = \sum_{i=1}^{n} \prod_{\alpha \in \Lambda} \mathbb{P}_\alpha(C_{i,\alpha}) = \sum_{i=1}^{n} \mu(C_i)$, 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 $B := \prod_{\alpha \in \Lambda} B_\alpha$, where $B_\alpha \in \mathcal{B}_\alpha$, $\forall \alpha \in \Lambda$, and $B_\alpha = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_B$ with $\Lambda_B \subseteq \Lambda$ being a finite set. $\forall i \in \mathbb{N}$, let $C_i := \prod_{\alpha \in \Lambda} C_{i,\alpha}$, where $C_{i,\alpha} \in \mathcal{B}_\alpha$, $\forall \alpha \in \Lambda$, and $C_{i,\alpha} = X_\alpha$, $\forall \alpha \in \Lambda \setminus \Lambda_i$ with $\Lambda_i \subseteq \Lambda$ being a finite set. Then, $\hat{\Lambda} := \Lambda_B \cup (\bigcup_{i=1}^{\infty} \Lambda_i)$, which is a countable subset of $\Lambda$. Then, $B_\alpha = X_\alpha$, $\forall i \in \mathbb{N}$, $\forall \alpha \in \Lambda \setminus \hat{\Lambda}$.

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, $\exists \epsilon > 0 \in (0, \infty) \subset \mathbb{R}$ such that $\mu(B) > \epsilon + \sum_{i=1}^{\infty} \mu(C_i)$, $\forall n \in \mathbb{N}$. Let $\hat{\Lambda}_n := \hat{\Lambda} \cup (\bigcup_{i=1}^{n} \Lambda_i)$, $\forall n \in \mathbb{N}$, which is a finite subset of $\Lambda$. Let the elements of $\Lambda := \{\omega_1, \ldots, \omega_r(1), \ldots, \omega_r(n-1)+1, \ldots, \omega_r(n)\}$, where the $r(\cdot)$’s are defined such that $\hat{\Lambda}_n = \{\omega_1, \ldots, \omega_r(1), \ldots, \omega_r(n-1)+1, \ldots, \omega_r(n)\}$, $\forall n \in \mathbb{N}$.

Fix any $n \in \mathbb{N}$. By Proposition 12.21, we have $\hat{\mathcal{Y}}_{n,1} := \prod_{\alpha \in \hat{\Lambda}_n} X_\alpha := (Y_{n,1}, \hat{\mathcal{B}}_{n,1}, \hat{\mathbb{P}}_{n,1})$ is a probability measure space, on which the collection of $\sigma$-algebras $\{\hat{\mathcal{B}}_{n,1} \alpha \}_{\alpha \in \Lambda}$ is independent, where $\hat{\mathcal{B}}_{n,1} := \bigcup_{\alpha \in \hat{\Lambda}_n} \hat{B}_{\alpha} \subseteq \hat{\mathcal{B}}_{n,1}$ $\hat{B}_{\alpha} = X_\alpha$, $\forall \alpha \in \hat{\Lambda}_n \setminus \{\alpha\}$, $\forall \alpha \in \hat{\Lambda}_n$. Let $\hat{B}_{\alpha} := \prod_{\alpha \in \hat{\Lambda}_n} B_\alpha$ and $\hat{C}_{n,1} := \prod_{\alpha \in \Lambda_n} C_{n,\alpha}$, $\forall i \in \{1, \ldots, n\}$. Clearly, $\hat{B}_{n,1} \supseteq \bigcup_{i=1}^{n} \hat{C}_{n,1}$, and the sets in the union are pairwise disjoint, and all sets involved are in $\hat{B}_{n,1}$. Then, we have $\hat{P}_{n,1}(\hat{B}_{n,1}) = \mu(B) > \epsilon_0 + \sum_{i=1}^{n} \mu(C_i) = \epsilon_0 + \sum_{i=1}^{n} \hat{P}_{n,1}(\hat{C}_{n,1})$, where the equalities follows from the fact that $\Lambda_B \subseteq \Lambda_n$ and $\Lambda_i \subseteq \Lambda_n$, $\forall i \in \{1, \ldots, n\}$. Define, $\hat{H}_{n,1} := \hat{B}_{n,1} \setminus (\bigcup_{i=1}^{n} \hat{C}_{n,1})$. Then, $\hat{P}_{n,1}(\hat{H}_{n,1}) > \epsilon_0$. Define $H_{n,1} := B \setminus (\bigcup_{i=1}^{n} C_i)$. Clearly, $H_{n,1} \supseteq \hat{H}_{n,1}$ $\supseteq H_{n+1}$, and $\cap_{n=1}^{\infty} H_{n+1} = \emptyset$.

By Radon-Nikodym Theorem 11.169 and Definition 14.8, $E_{n,1}(\chi_{\hat{H}_{n,1}, Y_{n,1}} | \hat{B}_{n,1}, \alpha_1)$ exists, where $E_{n,1}$ is with respect to measure $\hat{P}_{n,1}$. Define $g_{n,1} : Y_{n,1} = \prod_{\alpha \in \hat{\Lambda}_n} X_\alpha \to [0, 1] \subset \mathbb{R}$ by, $\forall (\omega_{r(1), \ldots, \omega_r(n)}) \in Y_{n,1}$, $g_{n,1}(\hat{\omega}_{r(1), \ldots, \hat{\omega}_{r(n)})} = \int_{Y_{n,1}} \chi_{\hat{H}_{n,1}, Y_{n,1}}(\omega_{r(1), \ldots, \omega_r(n)}) d\hat{P}_{n,1}(\omega_{r(1), \ldots, \omega_r(n)})$.

By Proposition 14.13, $g_{n,1} \in \hat{E}_{n,1}(\chi_{\hat{H}_{n,1}, Y_{n,1}} | \hat{B}_{n,1}, \alpha_1)$, $g_{n,1}(\hat{\omega}_{r(1), \ldots, \hat{\omega}_{r(n)})} = \hat{g}_{n,1}(\hat{\omega}_{r(1), \ldots, \hat{\omega}_{r(n)})}, \forall (\hat{\omega}_{r(1), \ldots, \hat{\omega}_{r(n)})} \in Y_{n,1}$, and $\hat{g}_{n,1} : X_\alpha \to [0, 1] \subset \mathbb{R}$ is $\mathcal{B}_\alpha$-measurable. Since $\hat{P}_{n,1}(\hat{H}_{n,1}) > \epsilon_0$, then $E_{\alpha_1}(\hat{g}_{n,1}) = E_{n,1}(\hat{E}_{n,1}(\chi_{\hat{H}_{n,1}, Y_{n,1}} | \hat{B}_{n,1}, \alpha_1)) = E_{n,1}(\chi_{\hat{H}_{n,1}, Y_{n,1}} = \hat{P}_{n,1}(\hat{H}_{n,1}) > \epsilon_0$, where $E_{\alpha_1}$ is with respect
to $P_{\alpha_1}$; and the second equality follows from (a) of Proposition 14.12. This leads to the following inequalities

$$
\epsilon_0 < E_{\alpha_1}(\bar{g}_{n,1}) \leq P_{\alpha_1} \left( \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \bar{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \} \right) + 2^{-1}\epsilon_0 P_{\alpha_1} \left( \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \bar{g}_{n,1}(\omega_{\alpha_1}) \leq 2^{-1}\epsilon_0 \} \right)

\leq P_{\alpha_1} \left( \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \bar{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \} \right) + 2^{-1}\epsilon_0
$$

Thus, we have $P_{\alpha_1} \left( \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \bar{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \} \right) > 2^{-1}\epsilon_0$.

Now, since $H_{n,1} \supseteq H_{n+1,1}$, by the definition of $\bar{g}_{n,1}$, Proposition 11.92, and Fubini’s Theorem 12.31, we have $0 \leq \bar{g}_{n+1,1}(\omega_{\alpha_1}) \leq \bar{g}_{n,1}(\omega_{\alpha_1})$, $\forall \omega_{\alpha_1} \in X_{\alpha_1}$, (which requires us to work in the product measure space $\mathcal{Y}_{n+1,1}$ that includes both $H_{n,1} \times \prod_{r=n+1}^{(n+1)} X_{\alpha_1}$ and $H_{n+1,1}$). Therefore, we have $P_{\alpha_1}(\Gamma_{n,1}) > 2^{-1}\epsilon_0$, where $\Gamma_{n,1} := \{ \omega_{\alpha_1} \in X_{\alpha_1} \mid \bar{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0 \}$. Furthermore, $\Gamma_{n,1} \supseteq \Gamma_{n+1,1}$. By Proposition 11.5, we have $P_{\alpha_1}(\bigcap_{n=1}^{\infty} \Gamma_{n,1}) = \lim_{n \in \mathbb{N}} P_{\alpha}(\Gamma_{n,1}) \geq 2^{-1}\epsilon_0 > 0$. Then, $\exists \omega_{\alpha_1} \in X_{\alpha_1}$ such that $\omega_{\alpha_1} \in \bigcap_{n=1}^{\infty} \Gamma_{n,1}$. This implies that $\bar{g}_{n,1}(\omega_{\alpha_1}) > 2^{-1}\epsilon_0$, $\forall n \in \mathbb{N}$. By the definition of $\bar{g}_{n,1}$, we have $\int_{Y_{n,1}} \chi_{\mathcal{H}_{n,1},\mathcal{Y}_{n,1}}(\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,1}(\omega_{\alpha_1},\ldots,\omega_{\alpha_r(n)}) > 2^{-1}\epsilon_0$, $\forall n \in \mathbb{N}$.

Fix any $n \in \mathbb{N}$. By Proposition 12.21, we have $\mathcal{Y}_{n,2} := \prod_{\alpha \in \Lambda_n \setminus \{ \alpha_1 \}} X_{\alpha}$

$\overset{\text{=} \text{:}}{=} (Y_{n,2}, \hat{B}_{n,2}, \hat{P}_{n,2})$ is a probability measure space, on which the collection of $\sigma$-algebras $(\hat{B}_{n,2})_{\alpha \in \Lambda_n \setminus \{ \alpha_1 \}}$ is independent, where $\hat{B}_{\alpha,n,2} := \left\{ \prod_{\alpha \in \Lambda_n \setminus \{ \alpha_1 \}} B_{\alpha} \mid B_{\alpha} = X_{\alpha}, \forall \alpha \in \Lambda_n \setminus \{ \alpha_1 \} \right\}$. Let $\hat{B}_{n,2} := \bigcup_{\alpha \neq \alpha_1} \hat{B}_{\alpha,n,2}$ and $\hat{C}_{n,2} := \bigcup_{\alpha \neq \alpha_1} \hat{C}_{\alpha,n,2}$ and the sets in the union are pairwise disjoint, and all sets involved are in $\hat{B}_{n,2}$, by Proposition 12.25. Then, we have

$$
\hat{P}_{n,2}(\hat{B}_{n,2}) = \int_{Y_{n,2}} \chi_{\hat{B}_{n,2},Y_{n,2}}(\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,2}(\omega_{\alpha_1},\ldots,\omega_{\alpha_r(n)})

= \int_{Y_{n,1}} \chi_{\bar{B}_{n,1},Y_{n,1}}(\bar{\omega}_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,1}(\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)})

= \int_{Y_{n,1}} \chi_{\mathcal{H}_{n,1},\mathcal{Y}_{n,1}} + \sum_{i=1}^{n} \chi_{\bar{C}_{i,n,1},Y_{n,1}}(\bar{\omega}_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,1}(\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)})

= \int_{Y_{n,1}} \sum_{i=1}^{n} \chi_{\bar{C}_{i,n,1},Y_{n,1}}(\bar{\omega}_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,1}(\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)})

+ \int_{Y_{n,1}} \chi_{\mathcal{H}_{n,1},\mathcal{Y}_{n,1}}(\bar{\omega}_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)}) d\hat{P}_{n,1}(\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_r(n)})
$$
where the first equality follows from Proposition 11.75; the second equality follows from the definition of \( \bar{\mathcal{C}}_{i,n} \), \( \mathcal{X}_\alpha \) is a probability measure space, and Fubini’s Theorem 12.31; the third equality follows from the fact that \( \bar{\mathcal{C}}_{i,n} = \mathcal{H}_{n,1} \cup \bigcup_{i=1}^n \mathcal{C}_{i,n,1} \) 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 \( \hat{\mathcal{C}}_{i,n,2} \)’s and Fubini’s Theorem 12.31; and the last equality follows from Proposition 11.75. Thus, we have

\[
P_{n,2}(\mathcal{H}_{n,2}) := P_{n,2}(\bar{\mathcal{C}}_{i,n,2} \setminus (\bigcup_{i=1}^n \mathcal{C}_{i,n,1})) > 2^{1-\epsilon_0}.
\]

Define \( \mathcal{H}_{n,2} := \{ (\hat{\omega}_\alpha)_{\alpha \in \Lambda \setminus \{\alpha_1\}} \in \prod_{\alpha \in \Lambda \setminus \{\alpha_1\}} \mathcal{X}_\alpha \mid (\hat{\omega}_\alpha)_{\alpha \in \Lambda} \in \mathcal{H}_{n,1}, \text{ where } \hat{\omega}_\alpha = \omega_\alpha, \forall \alpha \in \Lambda \setminus \{\alpha_1\}, \hat{\omega}_{\alpha_1} = \bar{\omega}_{\alpha_1} \}. \)

Clearly, \( \mathcal{H}_{n,2} = \bar{H}_{n,2} \times \prod_{\alpha \in \Lambda \setminus \{\alpha_1\}} \mathcal{X}_\alpha, \mathcal{H}_{n,2} \supseteq \mathcal{H}_{n,1+1,2}, \) and \( \bigcap_{n=1}^\infty \mathcal{H}_{n,2} = \emptyset. \) Obviously, \( \bar{H}_{n,2} \subseteq \bar{\mathcal{C}}_{i,n,2}. \)

By Radon-Nikodym Theorem 11.169 and Definition 14.8, \( \bar{E}_{n,2}(X_{\mathcal{H}_{n,2},Y_{\mathcal{B}_{n,2}}},B_{n,2}) \) exists, where \( \bar{E}_{n,2} \) is with respect to measure \( \bar{P}_{n,2}. \)

Define \( g_{n,2} : Y_{n,2} = \prod_{\alpha \in \Lambda \setminus \{\alpha_1\}} \mathcal{X}_\alpha \to [0,1] \subset \mathbb{R} \) by

\[
g_{n,2}(\tilde{\omega}_{\alpha_2}, \ldots, \tilde{\omega}_{\alpha_{r(n)}}) = \int_{\mathcal{Y}_{n,2}} \chi_{\bar{H}_{n,2},Y_{\mathcal{B}_{n,2}}}(\tilde{\omega}_{\alpha_2}, \ldots, \tilde{\omega}_{\alpha_{r(n)}}) d\bar{P}_{n,2}(\omega_\alpha, \ldots, \omega_{\alpha_{r(n)}})
\]

\[\forall (\tilde{\omega}_{\alpha_2}, \ldots, \tilde{\omega}_{\alpha_{r(n)}}) \in Y_{n,2}.\] By Proposition 14.13, \( g_{n,2} \subseteq \bar{E}_{n,2}(X_{\mathcal{H}_{n,2},Y_{\mathcal{B}_{n,2}}},B_{n,2}), g_{n,2}(\tilde{\omega}_{\alpha_2}, \ldots, \tilde{\omega}_{\alpha_{r(n)}}) = \bar{g}_{n,2}(\tilde{\omega}_{\alpha_2}), \forall (\tilde{\omega}_{\alpha_2}, \ldots, \tilde{\omega}_{\alpha_{r(n)}}) \in Y_{n,2}, \) and \( \bar{g}_{n,2} : X_{\mathcal{B}_{n,2}} \to [0,1] \subset \mathbb{R} \) is \( \mathcal{B}_{\alpha_2} \)-measurable. Since \( \bar{P}_{n,2}(\mathcal{H}_{n,2}) > 2^{-1}\epsilon_0, \) then \( E_{\alpha_2}(\bar{g}_{n,2}) = \bar{E}_{n,2}(\bar{E}_{n,2}(X_{\mathcal{H}_{n,2},Y_{\mathcal{B}_{n,2}}},B_{n,2}),B_{n,2}) = E_{n,2}(X_{\mathcal{H}_{n,2},Y_{\mathcal{B}_{n,2}}}) = P_{n,2}(\mathcal{H}_{n,2}) > 2^{-1}\epsilon_0, \) where \( E_{\alpha_2} \) is with respect to \( P_{n,2} ; \) and the second equality follows from (a) of Proposition 14.12. This leads to the following inequalities

\[
2^{-1}\epsilon_0 < E_{\alpha_2}(\bar{g}_{n,2}) \leq P_{\alpha_2}(\{ \omega_{\alpha_2} \in X_{\mathcal{B}_{n,2}} \mid \bar{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2}\epsilon_0 \})
\]

\[+ 2^{-2}\epsilon_0 P_{\alpha_2}(\{ \omega_{\alpha_2} \in X_{\mathcal{B}_{n,2}} \mid \bar{g}_{n,2}(\omega_{\alpha_2}) \leq 2^{-2}\epsilon_0 \})
\]
Thus, we have $P_{\alpha_2}(\{\omega_{\alpha_2} \in X_{\alpha_2} \mid \tilde{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2}\varepsilon_0\}) + 2^{-2}\varepsilon_0$

Now, since $H_{n,2} \supseteq H_{n+1,2}$, by the definition of $\tilde{g}_{n,2}$, Proposition 11.92, and Fubini’s Theorem 12.31, we have $0 \leq \tilde{g}_{n+1,2}(\omega_{\alpha_2}) \leq \tilde{g}_{n,2}(\omega_{\alpha_2})$, $\forall \omega_{\alpha_2} \in X_{\alpha_2}$, (which requires us to work in the product measure space $\mathcal{Y}_{n+1,2}$ that includes both $H_{n,2} \times \prod_{i=r(n)+1}^{(n+1)} X_{\alpha_i}$ and $H_{n+1,2}$). Therefore, we have $P_{\alpha_2}(\Gamma_{n,2}) > 2^{-2}\varepsilon_0$, where $\Gamma_{n,2} := \{\omega_{\alpha_2} \in X_{\alpha_2} \mid \tilde{g}_{n,2}(\omega_{\alpha_2}) > 2^{-2}\varepsilon_0\}$. Furthermore, $\Gamma_{n,2} \supseteq \Gamma_{n+1,2}$. By Proposition 11.5, we have $P_{\alpha_2}(\bigcap_{n=1}^{\infty} \Gamma_{n,2}) = \lim_{n \in \mathbb{N}} P_{\alpha_2}(\Gamma_{n,2}) \geq 2^{-2}\varepsilon_0 > 0$. Then, $\exists \omega_{\alpha_2} \in X_{\alpha_2}$ such that $\omega_{\alpha_2} \in \bigcap_{n=1}^{\infty} \Gamma_{n,2}$.

Recursively, assume we have completed index $l \in \mathbb{N}$. Fix any $n \in \mathbb{N}$. By Proposition 12.21, we have $Y_{n,l+1} := \prod_{\alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}} X_{\alpha} =: (Y_{n,l+1}, \check{B}_{n,l+1}, \check{P}_{n,l+1})$ is a probability measure space, on which the collection of $\sigma$-algebras $\{\check{B}_{\alpha,n,l+1}\}_{\alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}}$ is independent, where

$\check{B}_{\alpha,n,l+1} := \{\prod_{\alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}} B_{\alpha} \subset \check{B}_{n,l+1} \mid \check{B}_{\alpha} = X_{\alpha}, \forall \alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}\}$.

Let $\check{B}_{n,l+1} := \{\prod_{\alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}} B_{\alpha} \subset \check{B}_{n,l+1} \mid \check{B}_{\alpha} = X_{\alpha}, \forall \alpha \in \check{\Lambda}_n \setminus \{\alpha_1, \ldots, \alpha_l\}\}$.

$\check{B}_{n,l+1} \supseteq \bigcup_{j=1}^{n} \check{C}_{n,l+1,j}$ and the sets in the union are pairwise disjoint, and all sets involved are in $\check{B}_{n,l+1}$, by Proposition 12.25. Then, we have

$\check{P}_{n,l+1}(\check{B}_{n,l+1})$
where the first equality follows from Proposition 11.75; the second equality follows from the definition of \( \tilde{B}_{n,l+1} \), \( \alpha_n \) is a probability measure space, and Fubini’s Theorem 12.31; the third equality follows from the fact that \( \tilde{B}_{n,l} = \hat{H}_{n,l} \cup (\bigcup_{i=1}^{n} \tilde{C}_{i,n,l}) \) 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 \( \tilde{C}_{i,n,l+1} \)’s and Fubini’s Theorem 12.31; and the last equality follows from Proposition 11.75. Thus, we have \( P_{n,l+1}(H_{n,l+1}) := P_{n,l+1}(B_{n,l+1} \setminus (\bigcup_{i=1}^{n} \tilde{C}_{i,n,l})) > 2^{-l} \epsilon_0 \).

Define \( H_{n,l+1} := \{ (\omega_\alpha)_{\alpha \in \Lambda} \setminus \{ a_1, \ldots, a_l \} : \omega_n \in \Lambda_n \setminus \{ \alpha_1, \ldots, \alpha_l \}, \omega_{a_j} = \omega_{\alpha_j}, j = 1, \ldots, l \} \). Clearly, \( \hat{H}_{n,l+1} = \hat{H}_{n,l+1} \times \prod_{\alpha \in \Lambda_n \setminus \{ \alpha_1, \ldots, \alpha_l \}} X_\alpha, H_{n,l+1} \supseteq H_{n,l+1} \), and \( \bigcap_{l=1}^{\infty} H_{n,l+1} = \emptyset \). Obviously, \( \hat{H}_{n,l+1} \subseteq \tilde{B}_{n,l+1} \).

By Radon-Nikodym Theorem 11.169 and Definition 14.8, \( \hat{E}_{n,l+1}(X_{H_{n,l+1}, Y_{n,l+1}} | \tilde{B}_{a_1, a_2, \ldots, a_l}) \) exists, where \( \hat{E}_{n,l+1} \) is with respect to measure \( P_{n,l+1} \). Define \( g_{n,l+1} : Y_{n,l+1} = \prod_{\alpha \in \Lambda_n \setminus \{ \alpha_1, \ldots, \alpha_l \}} X_\alpha \to [0, 1] \subset \mathbb{R} \) by, \( \forall (\omega_{a_1, \ldots, \omega_{a_{(l)}}}) \in Y_{n,l+1} \),

\[
g_{n,l+1}(\omega_{a_1, \ldots, \omega_{a_{(l)}}}) = \int_{Y_{n,l+1}} \chi_{\hat{H}_{n,l+1}, Y_{n,l+1}}(\tilde{\omega}_{a_1, \omega_{a_2}, \ldots, \omega_{a_{(l)}}}) \ dP_{n,l+1}(\omega_{a_1, \ldots, \omega_{a_{(l)}}})
\]

By Proposition 14.13, \( g_{n,l+1} \in \hat{E}_{n,l+1}(X_{H_{n,l+1}, Y_{n,l+1}} | \tilde{B}_{a_1, a_2, \ldots, a_l}) \), \( g_{n,l+1}(\omega_{a_1, \ldots, \omega_{a_{(l)}}}) = g_{n,l+1}(\tilde{\omega}_{a_1}), \forall (\tilde{\omega}_{a_1, \ldots, \omega_{a_{(l)}}}) \in Y_{n,l+1} \), and \( g_{n,l+1} : X_{a_{(l)}} \to [0, 1] \subset \mathbb{R} \) is \( B_{a_1, a_2, \ldots, a_l} \)-measurable. Since \( P_{n,l+1}(H_{n,l+1}) > 2^{-l} \epsilon_0 \), then \( E_{a_1, a_2, \ldots, a_l} = \hat{E}_{n,l+1}(E_{n,l+1}(X_{H_{n,l+1}, Y_{n,l+1}} | \tilde{B}_{a_1, a_2, \ldots, a_l})) = E_{n,l+1}(X_{\hat{H}_{n,l+1}, Y_{n,l+1}}) = P_{n,l+1}(H_{n,l+1}) > 2^{-l} \epsilon_0 \), where \( E_{a_1, a_2, \ldots, a_l} \) is with
CHAPTER 14. PROBABILITY THEORY

respect to $P_{\alpha+1}$; and the second equality follows from (a) of Proposition 14.12. This leads to the following inequalities

$$2^{-l_0} < E_{\alpha+1}(g_{n,l+1})$$

$$\leq P_{\alpha+1}(\{\omega_{\alpha+1} \in X_{\alpha+1} \mid g_{n,l+1}(\omega_{\alpha+1}) > 2^{-l_0}\})$$

$$+ 2^{-l_0} P_{\alpha+1}(\{\omega_{\alpha+1} \in X_{\alpha+1} \mid g_{n,l+1}(\omega_{\alpha+1}) \leq 2^{-l_0}\})$$

$$\leq P_{\alpha+1}(\{\omega_{\alpha+1} \in X_{\alpha+1} \mid g_{n,l+1}(\omega_{\alpha+1}) > 2^{-l_0}\}) + 2^{-l_0}$$

Thus, we have $P_{\alpha+1}(\{\omega_{\alpha+1} \in X_{\alpha+1} \mid g_{n,l+1}(\omega_{\alpha+1}) > 2^{-l_0}\}) > 2^{-l_0}$.

Now, since $H_{n,l+1} \supseteq H_{n+1,l+1}$, by the definition of $g_{n,l+1}$, Proposition 11.92, and Fubini's Theorem 12.31, we have $0 < g_{n,l+1}(\omega_{\alpha+1}) \leq g_{n,l+1}(\omega_{\alpha+1})$, $\forall \omega_{\alpha+1} \in X_{\alpha+1}$, (which requires us to work in the product measure space $X_{\alpha+1}$ 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 $P_{\alpha+1}(\Gamma_{n,l+1}) > 2^{-l_0}$, where $\Gamma_{n,l+1} := \{\omega_{\alpha+1} \in X_{\alpha+1} \mid g_{n,l+1}(\omega_{\alpha+1}) > 2^{-l_0}\}$. Furthermore, $\Gamma_{n,l+1} \supseteq \Gamma_{n+1,l+1}$. By Proposition 11.5, we have $P_{\alpha+1}(\bigcap_{n=1}^{\infty} \Gamma_{n,l+1}) = \lim_{n \to \infty} P_{\alpha+1}(\Gamma_{n,l+1}) \geq 2^{-l_0} > 0$. Then, $\exists \omega_{\alpha+1} \in X_{\alpha+1}$ such that $\omega_{\alpha+1} \in \bigcap_{n=1}^{\infty} \Gamma_{n,l+1}$. This implies that $g_{n,l+1}(\omega_{\alpha+1}) > 2^{-l_0}$, $\forall n \in \mathbb{N}$. By the definition of $g_{n,l+1}$, we have

$$\int_{\gamma_{\alpha+1}} \chi_{H_{n+1,l+1}}(\omega_{\alpha+1}, \omega_{\alpha+1}, \ldots, \omega_{\alpha(n)}) dP_{\alpha+1}(\omega_{\alpha+1}, \ldots, \omega_{\alpha(n)}) > 2^{-l_0}, \forall n \in \mathbb{N}.$$  

Thus, we may obtain $\omega_{\alpha_1} \in \bigcap_{n=1}^{\infty} \Gamma_{n,l+1} \subseteq X_{\alpha_1}, \forall i \in \mathbb{N}$ with $\alpha_1 \in \hat{\Lambda}$. By Axiom of Choice, there exists a $\hat{\omega} \in X = \prod_{\alpha \in \Lambda} X_{\alpha}$ such that $\pi_{\alpha}(\hat{\omega}) = \omega_{\alpha_1}, \forall \alpha_1 \in \hat{\Lambda}$. Fix any $n \in \mathbb{N}$. $\omega_{\alpha(n)} \in \bigcap_{n=1}^{\infty} \Gamma_{m,r(n)}$. This implies that $\omega_{\alpha(n)} \in \Gamma_{n,r(n)} \Rightarrow \omega_{\alpha(n)} \in \bigcap_{n=1}^{\infty} \Gamma_{n,r(n)}$. Then, we have $0 < g_{n,r(n)}(\omega_{\alpha(n)}) = \int_{\gamma_{\alpha(n)}} \chi_{H_{n,r(n)}}(\omega_{\alpha(n)}) dP_{n,r(n)}(\omega_{\alpha(n)}) = \chi_{H_{n,r(n)}}(\omega_{\alpha(n)}) = 1$, where the last equality follows since it is the value of an indicator function. Hence, $\omega_{\alpha(n)} \in \hat{H}_{n,r(n)}$. By the definition of $\hat{H}_{n,r(n)}$, we have $(\omega_{\alpha_1})_{1 \leq \alpha_1 \leq \hat{\Lambda}_n} \in \hat{H}_{n,1}$. This yields $\hat{\omega} \in H_{n,1}$. By the arbitrariness of $n$, we have $\hat{\omega} \in \bigcap_{n=1}^{\infty} H_{n,1} = \emptyset$. 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 11.32 holds.
14.2 Gaussian Random Variables and Vectors

**Definition 14.15** Let $\Omega := (\Omega, \mathcal{B}, P)$ 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 $L_x : \mathcal{B}_\mathbb{B}(\mathbb{R}^m) \to [0, 1] \subset \mathbb{R}$ by $L_x(E) := P(x_{\text{inv}}^{-1}(E)), \forall E \in \mathcal{B}_\mathbb{B}(\mathbb{R}^m)$. Then, $L_x$ is a probability measure on the measurable space $(\mathbb{R}^m, \mathcal{B}_\mathbb{B}(\mathbb{R}^m))$ and is said to be the law of $x$. Since $L_x$ is a finite measure, then we may define $F : \mathbb{R}^m \to [0, 1] \subset \mathbb{R}$ by $F(z) = L_x(\{ \bar{x} \in \mathbb{R}^m \mid \bar{x} \preceq z \}), \forall z \in \mathbb{R}^m$. Then, by Proposition 12.51, $F$ is a cumulative distribution function of $L_x$.

**Example 14.16** Let $\mathcal{I} := (((0, 1), \cdot, \cdot), \mathcal{B}, \mu)$ be the finite metric measure subspace of $\mathbb{R}$. Let $x : (0, 1) \to [0, 1] \subset \mathbb{R}$ be a random variable given by $x(\omega) = \omega, \forall \omega \in \mathcal{I}$. Then, any cumulative distribution function $F : \mathbb{R} \to [0, 1] \subset \mathbb{R}$ with $F$ being of bounded variation, $\lim_{z \to -\infty} F(z) = 0$, $\lim_{z \to \infty} F(z) = 1$, and $T_F = 1$, we seek a function $h : (0, 1) \to \mathbb{R}$ such that the random variable $y := h \circ x : \mathcal{I} \to \mathbb{R}$ admits $F$ as a cumulative distribution function of $L_y$, which is the probability measure on the measurable space $(\mathbb{R}, \mathcal{B}_\mathbb{B}(\mathbb{R}))$. Since $T_F = 1$ and $\lim_{z \to -\infty} F(z) = 0$ and $\lim_{z \to \infty} F(z) = 1$, 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 $h(\alpha) = \inf \{ z \in \mathbb{R} \mid F(z) \geq \alpha \} \in \mathbb{R}$, $\forall \alpha \in \mathcal{I}$. Then, it is straightforward to prove that $\mu(\{ \omega \in \mathcal{I} \mid y(\omega) \leq z \}) = \mu(\{ \omega \in \mathcal{I} \mid h(x(\omega)) \leq z \}) = \mu(\{ \omega \in \mathcal{I} \mid \omega \leq F(z) \}) = F(z)$. 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 $F : \mathbb{R} \to [0, 1] \subset \mathbb{R}$ satisfying the stated assumptions, there exists a probability measure space $\mathcal{I}$ and a random variable $y : \mathcal{I} \to \mathbb{R}$ such that $y$ admits $F$ as a cumulative distribution function for its law $L_y$.

**Definition 14.17** Let $\Omega := (\Omega, \mathcal{B}, P)$ 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 $K \in \mathbb{R}^{m \times m}$, where $K \in \mathcal{S}^+_m$, if $E(x) = \bar{x}$ and $E((x - \bar{x})(x - \bar{x})') = K$ and a cumulative distribution function for the law $L_x$ of $x$ is

$$F(z) = \lim_{k \to -\infty} \frac{1}{\sqrt{2\pi} \det(K)^{1/2}} \exp(-\frac{1}{2} (x - \bar{x})' K^{-1} (x - \bar{x})) d\mu_{\mathbb{B}^m}(x)$$

$\forall z \in \mathbb{R}^m$. 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 14.18** Let $m \in \mathbb{N}$, $\bar{x} \in \mathbb{R}^m$ and $K \in \mathbb{R}^{m \times m}$, where $K \in \mathcal{S}^+_m$. By Spectral Theory Theorem 13.50, there exists a unitary matrix $V \in \mathbb{R}^{m \times m}$ and a diagonal matrix $D = \text{block diagonal} \left( \sigma_1^2, \ldots, \sigma_m^2 \right) \in \mathcal{S}^+_m$...
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 an $\mathbb{R}^m$-valued Gaussian random variable with mean $\bar{y} := V\bar{x} =: (\bar{y}_1, \ldots, \bar{y}_m)$ and covariance $VK' = D$. Then, by Example 14.16 and Fundamental Theorem on Modeling 14.14, 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 $\sigma^2_i, i = 1, \ldots, m$. $y_i \sim N(\bar{y}_i, \sigma^2_i)$. Then, by independence of $y_1, \ldots, y_m$, we have the $\mathbb{R}^m$-valued random variable $y := (y_1, \ldots, y_m)$ admits a cumulative distribution function

$$F_y(z) = \prod_{i=1}^m \lim_{k \to -\infty} \int_{(-k, z]} \frac{1}{\sqrt{2\pi}\sigma_i} \exp(-\frac{1}{2}(y_i - \bar{y}_i)'\sigma_i^{-2}(y_i - \bar{y}_i)) \, d\mu_i(y_i)$$

$$= \lim_{k \to -\infty} \int_{\Omega \times (-k, z]} \frac{1}{\sqrt{2\pi}^m} \det(D)^{1/2} \exp(-\frac{1}{2}(y - \bar{y})'D^{-1}(y - \bar{y})) \, d\mu_m(y)$$

Thus, $x = V'y \sim N(\bar{x}, K)$ as desired. \hfill \diamond 

14.3 Law of Large Numbers

**Theorem 14.19 (Strong Law of Large Numbers)** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure, $(X_n)_{n=1}^\infty$ be a sequence of independent random variables, $X_n : \Omega \to \mathbb{R}$, with $E(X_n) = a \in \mathbb{R}$ and $E((X_n - a)^2) \leq b \in \mathbb{R}_+$, $\forall n \in \mathbb{N}$. Then, $\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n X_i = a \, \text{a.e. in } \Omega$.

**Proof** Let $S_n := \frac{1}{n} \sum_{i=1}^n X_i$, which is a random variable. Then, $E(S_n) = a, E((S_n - a)^2) = \frac{1}{n^2} E(\sum_{i=1}^n (X_i - a))^2 = \frac{1}{n^2} \sum_{i=1}^n E((X_i - a)(X_j - a)) = \frac{1}{n^2} \sum_{i=1}^n (X_i - a)^2 \leq \frac{1}{n^2} \sum_{i=1}^n (X_i - a)^2 \to 0, \text{ as } n \to \infty$, where the first two equalities follow from Proposition 11.92; and the third equality follows from the assumption that $(X_n)_{n=1}^\infty$ is independent. Then, $\lim_{n \to \infty} E((S_n - a)^2) = 0$. Now, $S_n : \Omega \to \mathbb{R}$. Then, the previous equality can be viewed as $\lim_{n \to \infty} \|S_n - a\|_{L^2(\Omega, \mathcal{B}, P)} = 0$. Hence, we have $\lim_{n \to \infty} S_n(\omega) = a \, \text{a.e. in } \Omega$. This completes the proof of the theorem.

14.4 Martingale Theory

**Definition 14.20** Let $\Omega := (\Omega, \mathcal{B}, P)$ be a probability measure space. $(\mathcal{B}_n)_{n=0}^\infty$ is said to be a filtration on $\Omega$, if $\mathcal{B}_0 \subseteq \mathcal{B}_1 \subseteq \cdots \subseteq \mathcal{B}$ are sub-$\sigma$-algebras of $\mathcal{B}$ on $\Omega$. Then, the quadruple $(\Omega, \mathcal{B}, (\mathcal{B}_n)_{n=0}^\infty, P)$ is said to be a filtered probability measure space. We will define $\mathcal{B}_\infty := \sigma(\bigcup_{n=0}^\infty \mathcal{B}_n) \subseteq \mathcal{B}$.

Let $\mathcal{Y}$ be a normed linear space over $\mathbb{K}_1$. $(\mathcal{B}_n)_{n=0}^\infty$ is said to be the natural filtration of a $\mathcal{Y}$-valued stochastic process $(W_n)_{n=0}^\infty$, where $W_n : \Omega \to \mathcal{Y}$,
\( \forall n \in \mathbb{Z}_+ \), are \( \mathcal{Y} \)-valued random variables, if \( \mathcal{B}_n = \sigma(W_0,W_1,\ldots,W_n) \), \( \forall n \in \mathbb{Z}_+ \).

Let \( \mathcal{Z} \) be a normed linear space over \( \mathbb{K}_2 \). Another \( \mathcal{Z} \)-valued stochastic process \( (X_n)_{n=0}^{\infty} \) is said to be adapted to the filtration \( (\mathcal{B}_n)_{n=0}^{\infty} \) if \( X_n \) is \( \mathcal{B}_n \)-measurable.

**Definition 14.21** 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 stochastic process. Then, the stochastic process \( (X_n)_{n=0}^{\infty} \) is said to be a Martingale if

\( (i) \) \( (X_n)_{n=0}^{\infty} \) is an adapted process (that is, it is adapted to the filtration \( (\mathcal{B}_n)_{n=0}^{\infty} \));

\( (ii) \) \( X_n \in L_1(\Omega,\mathbb{R}), \forall n \in \mathbb{Z}_+; \)

\( (iii) \) \( X_{n-1} \in E(X_n|\mathcal{B}_{n-1}), \forall n \in \mathbb{N}. \)

The stochastic process \( (X_n)_{n=0}^{\infty} \) is said to be a super Martingale if it satisfies (i) and (ii) and \( f_{n-1}(\omega) \leq X_{n-1}(\omega) \) a.e. \( \omega \in \Omega \), \( \forall f_{n-1} \in E(X_n|\mathcal{B}_{n-1}), \forall n \in \mathbb{N}. \)

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. \( \omega \in \Omega \), \( \forall f_{n-1} \in E(X_n|\mathcal{B}_{n-1}), \forall n \in \mathbb{N}. \)

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 14.22** 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 (or a Martingale), and \( (C_n)_{n=1}^{\infty} \) be such that

\( (i) \) \( C_n : \Omega \rightarrow [0,\alpha] \subset \mathbb{R}, \forall n \in \mathbb{N}, \) where \( \alpha \in \mathbb{R}_+; \)

\( (ii) \) \( C_n \) is \( \mathcal{B}_{n-1} \)-measurable, \( \forall n \in \mathbb{N}; \)

Then, the stochastic process \( (Y_n)_{n=0}^{\infty} \) defined by \( Y_0 = X_0 \) and \( Y_n = X_0 + \sum_{i=1}^{n-1} C_n(X_n - X_{n-1}), \forall n \in \mathbb{N}, \) is a super Martingale (or a Martingale).

**Proof** Since, \( \forall n \in \mathbb{N}, X_n \) is \( \mathcal{B}_n \)-measurable and \( C_n \) is \( \mathcal{B}_{n-1} \)-measurable. Then, \( C_n(X_n - X_{n-1}) \) is \( \mathcal{B}_n \)-measurable by Propositions 11.38 and 11.39 and Definition 14.21. Hence, \( Y_n \) is \( \mathcal{B}_n \)-measurable, \( \forall n \in \mathbb{Z}_+. \)

Clearly, \( |Y_n| = \left| \sum_{i=0}^{n-1} (C_{i+1} - C_i)X_i + C_nX_n \right| \leq \sum_{i=0}^{n-1} 2\alpha |X_i| + \alpha |X_n|, \forall n \in \mathbb{Z}_+. \) Then, by Definition 14.21, \( X_n \in L_1(\Omega,\mathbb{R}), \forall n \in \mathbb{Z}_+, \) implies that \( Y_n \in L_1(\Omega,\mathbb{R}) \) by Proposition 11.83.

Note that \( E(Y_n - Y_{n-1}|\mathcal{B}_{n-1}) = E(C_n(X_n - X_{n-1})|\mathcal{B}_{n-1}) = C_nE(X_n - X_{n-1}|\mathcal{B}_{n-1}) = C_nE(X_n|\mathcal{B}_{n-1}) - X_{n-1}, \forall n \in \mathbb{N}, \) where the first equality follows from the definition of \( Y_n \); the second and third equalities follow from
(h) and (c) of Proposition 14.12, respectively. 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 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. \(\square\)

This result implies that one cannot beat the system in a gambling scenario by varying his bet each time after observing the outcome of the past dice rolls.

**Example 14.24** Let \(\Omega := (\Omega, \mathcal{B}, (B_n)_{n=0}^\infty, P)\) be a filtered probability measure space, \(Y\) be a \(\mathcal{B}\)-valued adapted stochastic process, and \(B \in \mathcal{B}_Y(Y)\). Then, \(T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}\) defined by \(T(\omega) := \inf\{n \in \mathbb{Z}_+ \mid X_n(\omega) \in B\}\) is a stopping time. This is because \(\{\omega \in \Omega \mid T(\omega) \leq n\} = \bigcup_{i=0}^n \{\omega \in \Omega \mid X_i(\omega) \in B\} \in \mathcal{B}_n\).

**Definition 14.25** Let \(\Omega := (\Omega, \mathcal{B}, P)\) be a probability measure space, \(\mathcal{B} \subseteq \mathcal{B}\) be a sub-\(\sigma\)-algebra on \(\Omega\), and \(\Omega := (\Omega, \mathcal{B}, P := P_{|\mathcal{B}})\). \(\forall F \in \mathcal{B}\), the condition probability of the event \(F\) happens given \(\mathcal{B}\) is denoted by \(P(F|\mathcal{B})\), which is defined to be a version of \(E(\chi_{F,\Omega}|\mathcal{B}) \in L_1(\Omega, [0, 1])\).

**Proposition 14.26** Let \(\Omega := (\Omega, \mathcal{B}, P)\) be a probability measure space, \(\mathcal{B} \subseteq \mathcal{B}\) be a sub-\(\sigma\)-algebra on \(\Omega\), and \(\Omega := (\Omega, \mathcal{B}, P := P_{|\mathcal{B}})\). \(\forall F \in \mathcal{B}, \forall E \in \mathcal{B}, 0 = P(\emptyset) = \int_E \chi_{\emptyset, \Omega} \, dP = \int_E f \, dP\), 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 14.8. 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 \(E = \bigcup_{i=1}^\infty E_i\), \(\mu(E) = P(E|\mathcal{B}) = E(\chi_{E, \Omega}|\mathcal{B}) = E(\sum_{i=1}^\infty \chi_{E_i, \Omega}|\mathcal{B}) = \sum_{i=1}^\infty E(\chi_{E_i, \Omega}|\mathcal{B}) = \sum_{i=1}^\infty \mu(E_i)\), where the first two equalities follow from Definition 14.25; the fourth equality follows from (e) of Proposition 14.12; and the last equality follows from Definition 14.25. Furthermore, \(\infty > \|\mu(E)\|_{L_1(\Omega, [0, 1])} = \int_{\Omega} |\mu(E)(\omega)| \, dP(\omega) = \int_{\Omega} \mu(E)(\omega) \, dP(\omega) = \int_{\Omega} \sum_{i=1}^\infty \mu(E_i)(\omega) \, dP(\omega) = \sum_{i=1}^\infty \int_{\Omega} \mu(E_i)(\omega) \, dP(\omega)\).
14.4. MARTINGALE THEORY

$= \sum_{i=1}^{\infty} \|\mu(E_i)\|_{L_1(\Omega,[0,1])}$, where the inequality follows from Definition 14.25; 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, $\forall E \in \mathcal{B}$,

$$
P \circ \mu(E) = \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \|\mu(E_i)\|_{L_1(\Omega,[0,1])}
$$

$$
= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \int_{\Omega} |\mu(E_i)(\omega)| d\hat{P}(\omega)
$$

$$
= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \sum_{i=1}^{n} \int_{\Omega} \mu(E_i)(\omega) d\hat{P}(\omega)
$$

$$
= \sup_{n \in \mathbb{Z}_+, (E_i)_{i=1}^{n} \subseteq \mathcal{B}, E = \bigcup_{i=1}^{n} E_i, E_i \cap E_j = \emptyset, \forall 1 \leq i < j \leq n} \int_{\Omega} \mu(E)(\omega) d\hat{P}(\omega)
$$

$$
= \int_{\Omega} \mu(E)(\omega) d\hat{P}(\omega) = \int_{\Omega} \chi_{E,\omega} d\mathbb{P} = \mathbb{P}(E)
$$

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 14.8. Hence, $\mathbb{P} = P \circ \mu$.

Since $\mathbb{P}$ 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

**Theorem 14.27** 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 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).

**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 $\mathcal{B}_{n-1}$-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 14.22. Then, the result follows immediately from Theorem 14.22. \qed
Theorem 14.28 (Doob’s Optional Stopping) Let \( \Omega := (\Omega, \mathcal{B}, (\mathcal{B}_n)^\infty_{n=0}, P) \) be a filtered probability measure space, \((X_n)^\infty_{n=0}\) be a super Martingale adapted to the filtration \((\mathcal{B}_n)^\infty_{n=0}\), \(T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}\) be a stopping time. Then, \(X_T : \Omega \to \mathbb{R}\) defined by \(X_T(\omega) := X_{T(\omega)}(\omega), \forall \omega \in \Omega\), is absolutely integrable over \(\Omega\) and \(E(X_T) \leq E(X_0)\) if any of the following conditions hold.

1. \(X_n(\omega) \geq 0, \forall \omega \in \Omega, \forall n \in \mathbb{N}\), and \(T(\omega) < \infty\) a.e. \(\omega \in \Omega\).
2. \(\exists N \in \mathbb{N}\) such that \(T(\omega) \leq N, \forall \omega \in \Omega\).
3. \(\exists M \in [0, \infty) \subset \mathbb{R}\) such that \(|X_n(\omega)| \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{Z}_+, \) and \(T(\omega) < \infty\) a.e. \(\omega \in \Omega\).\(^{1}\)
4. \(E(T) < \infty\) and \(\exists M \in [0, \infty) \subset \mathbb{R}\) such that \(|X_n(\omega) - X_{n-1}(\omega)| \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{N}\).

If \((X_n)^\infty_{n=0}\) is a Martingale adapted to the filtration \((\mathcal{B}_n)^\infty_{n=0}\), 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. Let (i) hold. Then, the stopped stochastic process \((Y_n)^\infty_{n=0}\) defined in Theorem 14.27 satisfies \(Y_n(\omega) \geq 0, \forall \omega \in \Omega, \forall n \in \mathbb{Z}_+.\) By Theorem 14.27, \((Y_n)^\infty_{n=0}\) is a super Martingale. By (i), \(T(\omega) < \infty\) a.e. \(\omega \in \Omega\). Then, \(\lim_{n \to \infty} \mathbb{E}[Y_n(\omega)] = X_T(\omega)\) a.e. \(\omega \in \Omega\). Clearly, \(X_T(\omega) \geq 0, \forall \omega \in \Omega\). By Fatou’s Lemma 11.80, we have \(\mathbb{E}(X_T) = \int_\Omega X_T(\omega) d\mathbb{P}(\omega) \leq \lim \inf_{n \to \infty} \int_\Omega Y_n(\omega) d\mathbb{P}(\omega) = \lim \inf_{n \to \infty} \mathbb{E}(Y_n) = \lim \inf_{n \to \infty} \mathbb{E}(\mathbb{E}(Y_n|\mathcal{B}_n)) \leq \lim \inf_{n \to \infty} \mathbb{E}(Y_{n-1}) \leq \cdots \leq \mathbb{E}(Y_0) = \mathbb{E}(X_0),\) where the first equality follows from Definition 14.1; the second equality follows from Definition 14.1; the third equality follows from Proposition 14.12; the second inequality follows from Theorem 14.27; and the last equality follows from the definition of \((Y_n)^\infty_{n=0}\) in Theorem 14.27. Hence, the result holds.

Let (ii) hold. Then, the stopped stochastic process \((Y_n)^\infty_{n=0}\) satisfies that \(Y_n(\omega) = X_T(\omega), \forall \omega \in \Omega\). Hence, by Theorem 14.27, \(X_T\) is absolutely integrable over \(\Omega\) and \(\mathbb{E}(X_T) = \mathbb{E}(Y_n) \leq \mathbb{E}(Y_{n-1}) \leq \cdots \leq \mathbb{E}(Y_0) = \mathbb{E}(X_0)\). Hence, the result holds.

Let (iii) hold. Then, \((Y_n)^\infty_{n=0}\) satisfies that \(|Y_n(\omega)| \leq M, \forall \omega \in \Omega, \forall n \in \mathbb{Z}_+, \) and \(\lim_{n \to \infty} \mathbb{E}(Y_n(\omega)) = X_T(\omega)\) a.e. \(\omega \in \Omega\). By Lebesgue Dominated Convergence Theorem 11.91, \(X_T\) is absolutely integrable over \(\Omega\) and \(\mathbb{E}(X_T) = \lim_{n \to \infty} \mathbb{E}(Y_n) \leq \mathbb{E}(Y_0) = \mathbb{E}(X_0)\). Hence, the result holds.

Let (iv) hold. Then, the stopped stochastic process \((Y_n)^\infty_{n=0}\) defined in Theorem 14.27 satisfies \(Y_n = X_0 + \sum_{i=1}^n C_n(X_n - X_{n-1}), \forall n \in \mathbb{N},\) where \((C_n)^\infty_{n=1}\) is as defined in the proof of Theorem 14.27. Then, \(|Y_n(\omega)| \leq |X_0(\omega)| + \sum_{i=1}^n \chi_{\{T(\omega) \geq n\}, \Omega} |X_n(\omega) - X_{n-1}(\omega)| \leq |X_0(\omega)| + \sum_{i=1}^n M \chi_{\{T(\omega) \geq n\}, \Omega} \leq |X_0(\omega)| + MT(\omega), \forall \omega \in \mathbb{Z}_+, \) where the first inequality follows from the definition of \((C_n)^\infty_{n=1}\); and the second inequality follows...
from (iii). Then, $E(|Y_n|) \leq E(|X_0|) + M \int_{\Omega} T(\omega) \, dP(\omega) = E(|X_0|) + M E(T) < \infty$, $\forall n \in \mathbb{Z}_+$, 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, $E(X_T) = \lim_{n \in \mathbb{Z}_+} E(Y_n) \leq E(Y_0) = E(X_0)$, where the inequality follows from Theorem 14.27. Hence, the result holds.

Now, consider the second paragraph of the theorem statement. Let $(X_n)_{n=0}^\infty$ is a Martingale. Clearly, it is a super Martingale. Let (ii), (iii), or (iv) holds. By the super Martingale case, we have $E(X_T) \leq E(X_0)$. Clearly, $(- X_n)_{n=0}^\infty$ is also a super Martingale. By the super Martingale case, we have $E(-X_T) \leq E(-X_0)$. Then, we have $E(X_T) = E(X_0)$. Hence, the result holds.

This completes the proof of the theorem. 

Lemma 14.29 Let $\Omega := (\Omega, B, (B_n)_{n=0}^\infty, P)$ be a filtered probability measure space, and $T : \Omega \rightarrow \mathbb{Z}_+ \cup \{\infty\}$ be a stopping time. Assume that $\exists N \in \mathbb{N}$ and $\exists \epsilon \in (0, \infty) \subset \mathbb{R}$, $\forall n \in \mathbb{N}$, $P(\{\omega \in \Omega \mid T(\omega) \leq n + N\} | B_n) > \epsilon$. Then, $E(T) < \infty$.

Proof

14.5 Central Limit Theorem

Theorem 14.30 (Central Limit Theorem) Let $\Omega := (\Omega, B, P)$ be a probability measure, $(X_n)_{n=1}^\infty$ be a sequence of independent random variables, $X_n : \Omega \rightarrow \mathbb{R}$, with $E(X_n) = 0$ and $E(X_n^2) = b \in \mathbb{R}_+$, $\forall n \in \mathbb{N}$. Then, the random variable $X := \lim_{n \in \mathbb{N}} \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i / \sqrt{b} \sim N(0, 1)$.

Proof


Chapter 15

Numerical Methods

15.1 Newton’s Method

Let $X$ and $Y$ be Banach spaces and $F : X \rightarrow Y$. To solve the equation $F(x) = \vartheta_Y$, we may use Newton’s method. Assume that $F$ is twice Fréchet differentiable, we set $x_{n+1} = x_n - (F'(x_n))^{-1}F(x_n)$, $\forall n \in \mathbb{N}$. Then, when $x_1$ is sufficiently close to a solution, the sequence $(x_n)_{n=1}^{\infty}$ converges to a solution $x_{opt}$ with $F(x_{opt}) = \vartheta_Y$. This result is formalized in the following proposition.

Proposition 15.1 Let $X$ and $Y$ be Banach spaces over $\mathbb{K}$, $D_1 \subseteq X$, $F : D_1 \rightarrow Y$, and $x_1 \in X$. 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) $\text{dom} \left( F^{(2)} \right) \supseteq \mathcal{B}_X \left( x_1, \frac{2h_1}{\beta_1 - h_1} \right) =: D \subseteq \mathcal{B}_X \left( x_1, 2\eta_1 \right)$ and, $\forall x \in D$, $\|F^{(2)}(x)\| \leq K$;

(iii) $F^{(1)}(x_1)$ is bijective, $\left\| (F^{(1)}(x_1))^{-1} \right\| \leq \beta_1$, and $\left\| (F^{(1)}(x_1))^{-1} \cdot F(x_1) \right\| \leq \eta_1$;

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 $F(x_{opt}) = \vartheta_Y$.

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 $\left\| (F^{(1)}(x_n))^{-1} \right\| \leq \beta_n := \frac{\beta_{n-1}}{1 - h_{n-1}} \in [0, \infty) \subset \mathbb{R}$;

(b) $\left\| (F^{(1)}(x_n))^{-1} F(x_n) \right\| \leq \eta_n := \frac{h_{n-1} \eta_{n-1}}{\beta_{n-1}} \in [0, \infty) \subset \mathbb{R}$;

715
(c) $x_{n+1} \in B_{g} \left( x_1, \frac{2-2h_1}{2-3h_1} \left( 1 - \left( \frac{h_1}{2(1-h_1)} \right)^n \right) \eta_1 \right)$;

(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 $\left\| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \right\| \leq \left\| F^{(2)}(t_0 x_{k+1} + (1 - t_0) x_0) - t_0 x_0 \right\| x_{k+1} - x_k \leq K \eta_k$.

Then, $\left\| (F^{(1)}(x_k))^{-1} \right\| \leq \beta_k \eta_k K = h_k < 1$. By Proposition 9.55, $F^{(1)}(x_{k+1})$ is bijective and

\[
\left\| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1} \right\|
\leq \frac{\left\| (F^{(1)}(x_k))^{-1} \right\|^2 \left\| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \right\|}{1 - \left\| (F^{(1)}(x_k))^{-1} \right\| \left\| F^{(1)}(x_{k+1}) - F^{(1)}(x_k) \right\|}
\leq \frac{\beta_k^2 K \eta_k}{1 - \beta_k \eta_k K} = \frac{h_k \beta_k}{1 - h_k}
\]

This leads to $\left\| (F^{(1)}(x_{k+1}))^{-1} \right\| \leq \left\| (F^{(1)}(x_k))^{-1} \right\| + \left\| (F^{(1)}(x_{k+1}))^{-1} - (F^{(1)}(x_k))^{-1} \right\| \leq \beta_k + h_k \beta_k/(1 - h_k) = \beta_k/(1 - h_k) = \beta_{k+1}$. 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 $\left\| F^{(1)}(x_{k+1}) \right\| \leq \frac{1}{2} \left\| F^{(2)}(t_1 x_{k+1} + (1 - t_1) x_k) \right\| x_{k+1} - x_k \| \leq K \eta_k^2 / 2$.

Then, $\left\| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \right\| \leq \beta_k K \eta_k^2 / (2 - 2h_k) = \frac{h_k \eta_k}{2(1-h_k)} \eta_{k+1}$. Hence, (b) holds.

Note that $\left\| x_{k+2} - x_{k+1} \right\| = \left\| (F^{(1)}(x_{k+1}))^{-1} F(x_{k+1}) \right\| \leq \eta_{k+1} = \frac{h_k}{2(1-h_k)} \eta_k \leq \left( \frac{h_1}{2(1-h_1)} \right)^k \eta_1$. Then, $\| x_{k+2} - x_1 \| \leq \| x_{k+2} - x_{k+1} \| + \| x_{k+1} - x_1 \| \leq \left( \frac{h_k}{2(1-h_k)} \right)^k \eta_1 + \frac{2-2h_1}{2-3h_1} \left( 1 - \left( \frac{h_1}{2(1-h_1)} \right)^k \right) \eta_1 = \frac{2-2h_1}{2-3h_1} \left( 1 - \left( \frac{h_1}{2(1-h_1)} \right)^k \right) \eta_1$. Hence, (c) holds.

Note that $h_{k+1} = \beta_k \eta_k K = \frac{h_k^2}{2(1-h_k)}$. 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. \qed

Then, the sequence $(x_n)_{n=1}^{\infty} \subseteq D$ is well-defined. $\forall n \in \mathbb{N}$, $\| x_{n+2} - x_{n+1} \| = \left\| (F^{(1)}(x_{n+1}))^{-1} F(x_{n+1}) \right\| \leq \eta_{n+1} \leq \left( \frac{h_1}{2(1-h_1)} \right)^n \eta_1$. 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 $F(x_0) = F^{(1)}(x_0)(x_n - x_{n+1})$, $\forall n \in \mathbb{N}$. By Proposition 9.7, we have
15.1. NEWTON’S METHOD

$F$ and $F^{(1)}$ are continuous at $x_{\text{opt}}$. Then, $F(x_{\text{opt}}) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} F^{(1)}(x_n)(x_n - x_{n+1}) = F^{(1)}(x_{\text{opt}})(x_{\text{opt}} - x_{\text{opt}}) = \vartheta y$, 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 $x_{\text{opt}}$, 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 $\exists \beta_1, \eta_1, K, M_1, M_2, M_3 \in [0, \infty) \subseteq \mathbb{R}$ such that

(i) $h_1 := \beta_1 \eta_1 K \leq \frac{1}{2}$;

(ii) $\text{dom} \left( F^{(3)} \right) \supseteq D_2 \supseteq \mathcal{B}_X \left( x_1, \frac{2-2h_1}{2-\beta_1} \eta_1 \right) =: D$, where $D_2 \subseteq D_1$ is an open set in $X$, $\forall x \in D_2$, $F^{(1)}(x)$ is bijective, and, $\forall x \in D$, $\| F^{(2)}(x) \| \leq K$, $\| (F^{(1)}(x))^{-1} \| \leq M_1$, $\| F^{(3)}(x) \| \leq M_2$, and $\| F(x) \| \leq M_3$;

(iii) $\| (F^{(1)}(x_1))^{-1} \| \leq \beta_1$ and $\| (F^{(1)}(x_1))^{-1} F(x_1) \| \leq \eta_1$.

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 $x_{\text{opt}} \in D$ with $F(x_{\text{opt}}) = \vartheta y$. Furthermore, $\forall n \in \mathbb{N}$, $\| x_{n+1} - x_{\text{opt}} \| \leq c \| x_n - x_{\text{opt}} \|^2$, where $c := (M_1 K + 2M_2 K^2 M_3 + M_3^2 M_2 M_4)/2$.

**Proof** By Proposition 15.1, $(x_n)_{n=1}^{\infty} \subseteq D$ is well defined and converges to $x_{\text{opt}} \in D$ with $F(x_{\text{opt}}) = \vartheta y$. 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 $T^{(1)}(x(h_1) = h_1 - (F^{(1)}(x))(h_1) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x)) = (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x))$, $\forall x \in D_2$, $\forall h_1 \in X$. Then, $T^{(1)}(x_{\text{opt}})(h_1) = \vartheta y$, $\forall h_1 \in X$, since $F(x_{\text{opt}}) = \vartheta y$. This implies that $T^{(1)}(x_{\text{opt}}) = \vartheta y_B(x, X)$. The second order derivative of $T$ is given by, $\forall x \in D_2$, $\forall h_1, h_2 \in X$,

$$T^{(2)}(x)(h_1)(h_2) = - (F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1} F(x)) - (F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)) + (F^{(1)}(x))^{-1} F^{(2)}(x)(h_1) - (F^{(1)}(x))^{-1} F^{(2)}(x)(h_2) = (F^{(1)}(x))^{-1} F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1} F^{(2)}(x)(h_1))$$
\[-(F^{(1)}(x))^{-1}F^{(2)}(x)(h_1)((F^{(1)}(x))^{-1}F^{(2)}(x)(h_2)((F^{(1)}(x))^{-1}F(x)))
\+(F^{(1)}(x))^{-1}F^{(2)}(x)(h_1)(h_2)\]

Then, \(\forall x \in D, \|T^{(2)}(x)\| = \sup_{h_1, h_2 \in \mathbb{X}, \|h_1\| \leq 1, \|h_2\| \leq 1} \|T^{(2)}(x)(h_2)(h_1)\| \leq 2M_1^2K^2M_4 + M_2^2M_2M_4 + M_1K = 2c.\) \(\forall n \in \mathbb{N},\) by Taylor’s Theorem 9.48, \(\exists t_0 \in (0, 1) \subset \mathbb{R}\) such that \(\|x_{n+1} - x_{\text{opt}}\| = \|T(x_n) - T(x_{\text{opt}})\| = \|T(x_n) - T(x_{\text{opt}}) - T^{(1)}(x_{\text{opt}})(x_n - x_{\text{opt}})\| \leq \frac{1}{2} \|T^{(2)}(t_0x_n + (1 - t_0)x_{\text{opt}})\| \|x_n - x_{\text{opt}}\|^2 \leq c \|x_n - x_{\text{opt}}\|^2.\) 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° \( n = 0, \forall m \in \mathbb{Z}_+, \) we have LHS = m! = RHS. This case is proved.

2° Assume the result holds for \( n = k \in \mathbb{Z}_+. \)

3° 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, $\forall n \geq n_0$, 
\[ nx_{n+1} \leq (n-1)x_n - (a-1)x_n \]
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 
\[ nx_{n+1} \geq (n-a)x_n \geq (n-1)x_n, \forall n \geq n_0 \]
Then, $nx_{n+1} \geq n_0x_{n_0+1} =: c \in (0, \infty) \subset \mathbb{R}$, $\forall n \geq n_0$. 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. \qed
Bibliography


URL: http://www.loc.gov/catdir/description/dover031/95051359.html


Index

Doob’s Optional Stopping Theorem, 712

Absolute continuity of measure, 473

Absolute integrability, 405
  Banach space valued measure space, 438

Adjoint operator, 203

Admissible deviation, 253

Alaoglu Theorem, 212

Alexandroff One-Point Compactification, 120

Algebra, 168
  Boolean, 23
    Generated by, 23
  Alignment, 194

Almost everywhere, 370
  Banach space valued measure, 419
  Banach space valued pre-measure, 415

Analytic radius, 313

Antisymmetric, 20

Ascoli-Arzelá Theorem, 111

Baire, 86
  Category Theorem, 87

Ball
  Closed, 73
  Open, 73

Banach measure space, 365
  Banach space valued, 514

Banach space, 158
  Banach space valued pre-measure, 414

  On algebra, 538

Basis
  Topological space, 41
  Vector space, 145

Bernstein Approximation Theorem, 528

Bernstein function, 527

Bessel’s Inequality, 662

Bijective, 21

Bolzano-Weierstrass property, 104

Borel measure, 594

Borel sets, 361

Borel-Cantelli Lemma, 351, 691

Borel-Lebesgue Theorem, 108

Bound
  Greatest lower, 25
  Least upper, 25
  Lower, 25
  Upper, 25

Bounded
  Function, 153
  Linear operator, 175
  Set, 153

Bounded Convergence Theorem, 391

Bourbaki, 112

Carathéodory Extension Theorem, 356, 534

Cartesian metric, 80

Cartesian product, 33
  Normed linear space, 156
  Vector space, 136

Category
  First, 48
  Second, 48
INDEX

Everywhere, 48
Cauchy
   Net, 87
   Sequence, 76
      In measure, 377
Cauchy-Schwarz Inequality, 150,
      499, 647
Central Limit Theorem, 713
Chain Rule, 261
Change of Variable, 543, 636
Choice, 26
Closed Graph Theorem, 201, 202
Commutative
   Ring, 133
Compact, 97, 163
      Countably, 103
      Locally, 114
      Sequentially, 104
      \( \sigma \)-, 123
Complete, 76
   Banach space valued measure
      space, 419
   Banach space valued pre-measure
      space, 415
   Measure space, 351
   Metric measure space, 365
      Banach space valued, 513
   Sequence, 663
Completely regular, 58
Completion of a Banach space valued measure space, 421
Completion of a Banach space valued pre-measure space, 416
Completion of a measure space, 352
Component, 51
Composition, 21
Concave functional, 225
Conditional expectation, 692
Conditional probability, 710
Cone, 143
   Conjugate, 243
   Negative, 243
   Positive, 243, 244
Conic segment, 143
Conjugate
   Group, 132
Connected, 50
      Arcwise, 54
      Locally, 52
      Simply, 54
Continuity, 38
      Absolute, 595
      One-sided, 571
      Semi-, 40
Continuum, 102
Contraction index, 288
Contraction mapping, 288
Contraction Mapping Theorem, 289
Convergence in measure in \( X \), 376
      Banach space valued measure,
         419
Convex
   Functional, 225
   Local, 214
   Mapping, 244
Convex combination, 142
Convex hull, 141
Coset, 165
Countability
   First axiom of, 41
   Second axiom of, 41
Countable, 23
Covering, 42
Cube, 93
Hilbert, 93
Cumulative distribution function,
   576
      With origin, 587
Curve, 54
   Closed, 54
Dense, 48
   Nowhere, 48
Derivative
   Directional, 254
   Fréchet, 254
   Higher-order, 267
   Partial, 255
Higher-order, 273
Diameter, 612
Dimension, 145
Co-, 669
Dini’s Lemma, 106
Direct sum, 655
Directed system, 59
Distance, 75
Domain, 21
Dual functional, 247
Dual space, 179
Second, 192
Egoroff’s Theorem, 375
Eidelheit Separation Theorem, 221
Eigenvalue, 681
Eigenvector, 681
Empty set, 19
Epigraph, 225
Equicontinuity, 80
Equivalence, 20
Equivalent
  Metric spaces, 75
  Normed linear spaces, 163
Essential supremum, 495
Exhaustion, 123
Expectation, 689
Exponential of linear operator, 324
Extensionality, 19
Extreme point, 326
Extreme subset, 326
Fσ, 518
Fatou’s Lemma, 394, 467
  For sets, 690
Fenchel Duality Theorem, 236
Field, 134
  Sub-, 134
Finite, 23
  Banach space valued measure, 419
  Banach space valued pre-measure, 415
  Locally, 124
  Banach space valued measure, 576
Topological measure space, 365
  Measure, 351
  Star, 126
Finite intersection property, 97
Fourier series, 666
Fubini Theorem, 560, 562, 570
Function, 21
  Analytic, 313
  Linear, 138
  Measurable, 368
  Of bounded variation, 575
Functional, 138
  Sublinear, 184
Fundamental Theorem of Calculus, 623, 627
Gδ, 518
Generation process, 427
Gram determinant, 660
Gram matrix, 660
Gram-Schmidt Procedure, 659
Graph, 21
Greatest element, 25
Group, 131
  Abelian, 132
  Semi, 132
  Abelian, 132
Hahn Decomposition Theorem, 478
Hahn-Banach Theorem
  Extension form, 185
    Complex, 187
    Geometric form, 220
    Simple version, 188
Hausdorff
  Maximal Principle, 26
  Space, 47
Heine-Borel, 99
Hermitian adjoint, 656
Hermitian operator, 656
Hilbert space, 649
Hölder’s Inequality, 149, 497
Homeomorphical isoméasure, 539
Homeomorphism, 40
Homomorphism, 132
Ring, 133
Vector space, 138
Homotopic, 54
Hyperplane, 217
Supporting, 221

Ideal, 133
Identity element, 133
Image, 21
Implicit Function Theorem, 295, 297
Analytic, 323
Increment on semi-open rectangle, 575
Indicator function, 379
Induced measure, 540
Infinity, 19
Injective, 21
Injective Mapping Theorem, 290
Inner product, 647
Integral
Finite Banach space valued measure space, 429
Finite measure space, 384
Infinite Banach space valued measure space, 429
Infinite measure space, 385
Nonnegative extended real-valued functions, 394
Integration by Parts, 633
Integration system, 383
Interval, 359
Inverse Function Theorem, 293
Analytic, 320
Global, 304, 305
Inverse image, 21
Invert, 300
Invertible, 21
Isomorphism, 132
Isometry, 75
Isometrical, 157
Ring, 133

Vector space, 138
Iterated Limit Theorem, 94
Jensen’s Inequality, 413
Joint Limit Theorem, 93
Jordan Decomposition Theorem, 479, 480
Kernel
For measure, 424
Krein-Milman, 327
Kuhn-Tucker Theorem, 339
Lp(\mathcal{X}, \mathcal{Y}), 494, 497
Lagrange multiplier, 249, 334, 335
Lattice, 168
Law, 707
Law of Large Numbers
Strong, 708
Lp(\mathcal{X}, \mathcal{Y}), 494, 497
Least element, 24
Lebesgue Decomposition Theorem, 491
Lebesgue Dominated Convergence Theorem, 402, 405, 437, 440, 467, 468
Lebesgue measure, 360, 594
Lebesgue number, 107
Lebesgue outer measure, 361, 594
Lebesgue point, 614
Lebesgue-Stieltjes integral, 595
Legendre polynomials, 663
Limit, 59, 63
Inferior, 67, 68
Joint, 71
Superior, 67, 68
Linear combination, 140
Linear variety, 140
Generated by, 140
Closed, 156
Linearly independent, 144
Vectors, 144
Lipschitz continuity, 602
m-dimensional Lebesgue measure space, 594
Maximal, 24
Mazur’s Theorem, 220
Meager, 48
Co-, 48
Non-, 48
Mean Value Theorem, 51, 263, 264
Measurable
Set, 349
Space, 349
With respect to an outer measure, 355
Measure, 349
Banach space valued, 419
Induced outer, 356
Outer, 355
Product, 547
Space, 349
Vector, 463
Measure on algebra, 356
Banach space valued, 533
Metric, 73
Pseudo-, 89
Metric measure space, 365
Banach space valued, 513
Metric space, 73
Product, 80
Metrizable, 92
Minimal, 24
Minkowski functional, 218
Minkowski’s Inequality, 150, 494
Monotone class, 544
Monotone Class Lemma, 544
Monotone Convergence Theorem, 395
Mutually singular measures, 473
Norm
Pseudo, 167
Normal
Subgroup, 132
Normal equation, 661
Normal space, 47
Normed linear measure space, 365
Banach space valued, 514
Null space, 138
Open mapping, 199
Open Mapping Theorem, 200, 291
Operator
Affine, 138
Compact, 684
Linear, 138
Spectral theory, 684
Pseudoinverse, 678
Order
Group, 132
Ordering
Partial, 24
Total, 24
Orthogonal, 194, 649
Orthogonal complement, 194
Pre-, 194
Orthonormal, 658
Outer product of tensors, 312
Pairing, 19
Paracompact, 125
Parallelogram Law, 648
Partition of Unity, 118
π-system, 544
Point
Accumulation, 35
At infinity, 120
Beginning, 54
Boundary, 35
Cluster, 59
End, 54
Exterior, 35
Interior, 35
Relative, 156
Of closure, 35
Negative definite, 326, 672
Negative semi-definite, 326, 672
Net, 59
Joint, 71
Sub-, 62
Newton’s Method, 715
INDEX

Point of minimum, 325
Point of relative extremum, 325
Point of relative minimum, 325
Positive definite, 325, 672
Positive linear functional, 520
Positive semi-definite, 326, 672
Power set, 19
Pre-Hilbert space, 647
Primal functional, 245
Principal of a region, 573
Probability measure space, 689
Filtered, 708
Projection, 33, 43
Projection operator, 677
Projection Theorem, 652, 668
Proper, 121
Countably, 121
Pythagorean Theorem, 649
Quotient, 20
Quotient space, 89, 165–167
Raabe’s Test, 719
Radon-Nikodym derivative, 482
Radon-Nikodym Theorem, 486, 488
Random variable, 689
Gaussian, 707
Range, 21
Range space, 138
Rectangle in \(\mathbb{R}^m\), 572, 573
Refinement, 100
Reflexive, 20
Banach space, 192
Region, 573
Regular point, 334
Regular space, 47
Regularity, 19
Relation, 20
Replacement, 19
Residual, 48
Restriction, 21
Riemann-Lebesgue Theorem, 644
Riesz Representation Theorem, 507, 521, 531, 532, 641, 642
Riesz-Fréchet Theorem, 653
Ring, 133
Sub-, 133
Semialgebra, 366
Separable, 48
Separation, 50
Set
Boundary, 35
Closed, 35
Closure, 35
Relative, 37
Complement, 35
Convex, 141
Exterior, 35
Interior, 35
Negative, 477
Open, 35
Relative, 156
Positive, 477
\(\sigma\)-algebra, 23
Generated by, 23
Independence, 691
\(\sigma\)-finite
Banach space valued measure, 419
Measure, 351
Simple function, 379
Banach space valued measure space, 431
Stationary point, 335
Stochastic process, 708
Adapted, 709
Martingale, 709
Sub Martingale, 709
Super Martingale, 709
Stone-Čech compactification, 128
Stopping time, 710
Subordinate, 117
Subset, 20
Subspace, 137
Banach space valued measure space, 423
Banach space valued pre-measure space, 418
Generated by, 140
Measure space, 354
Proper, 137
Topological measure space, 366
Support
Of a convex set, 222
Of a function, 117, 520
Support functional, 223
Surjective, 21
Surjective Mapping Theorem, 290
Symmetric, 20
Taylor Series, 313
Taylor’s Theorem, 285
Tensor, 311
Transpose, 311
Tietze’s Extension Theorem, 56
Tonelli Theorem, 559
Topological measure space, 365
Banach space valued, 513
Product, 564
Topological space, 35
Topology, 35
Generated by, 38
Natural, 73
Of pointwise convergence, 57
Product, 42
Stronger, 38
Subset, 37
Weak, 57, 206
Weak∗, 209
Weaker, 38
Total variation, 414, 419, 533, 538
Of function, 575
Totally bounded, 79
Transitive, 20
Tychonoff
Space, 47
Theorem, 112

Uniform
Boundedness Principle, 50
Cauchy sequence, 80
Continuity, 78
Convergence, 80
Equivalent

Metric spaces, 78
Homeomorphism, 78
Union, 19
Urysohn Metrization Theorem, 93
Urysohn’s Lemma, 55
Vector space, 134
Vitali, 613
Vitali’s Lemma, 613, 634
Well-ordering, 27
Principle, 27
Zero-element, 132
Zorn’s Lemma, 27