# Notes on Real Analysis

Kai-Jyun Wang*

Fall 2024

## Foreword

I took Real Analysis in the fall of 2024 with Professor Tien. This is my note on the course. I tried to include all the proofs and details that has or has not been covered in the class, in order to make this note as self-contained as possible. Some of the proofs might be taken from somewhere and some might be wrong. The following topics are covered in the lecture: measure theory, Lebesgue integration, Banach space and Hilbert space.

Some funny things happened in the class. The professor had taught so fast that we had already reached the Banach space before our first midterm. Every student was wondering if the professor forgot that this is actually a one-year course. Time comes to the eleventh week, the professor walked into the class and said, "Few days ago, someone told me that we actually have two semesters for real analysis, and I didn't know that before!" It turns out that our concern was right. The professor then said, "But that is also a good thing, because we can learn more advanced topics in the second semester, like the harmonic analysis, the Fourier analysis. . . "

*National Taiwan University, Department of Economics.

# Contents


**1 Lebesgue Measure and Integration** **3**

1.1 Lebesgue Measure . . . . . . . . . . . . 3

1.2 Lebesgue Integration . . . . . . . . . . . . 7

1.3 Differentiation . . . . . . . . . . . . 13

1.4 Radon-Nikodym Theorem . . . . . . . . . . . . 16

**2 Banach Space** **22**

2.1 Banach Space and Bounded Linear Functional . . . . . . . . . . . . 22

2.2 $\ell^p$ Space . . . . . . . . . . . . 24

2.3 $\mathscr{L}^p$ Space . . . . . . . . . . . . 27

2.4 Weak Convergence . . . . . . . . . . . . 40

2.5 Open Mapping Theorem and Closed Graph Theorem . . . . . . . . . . . . 48

**3 Hilbert Space** **52**

3.1 Cauchy-Schwarz Inequality . . . . . . . . . . . . 52

3.2 Bounded Linear Functional on $\mathscr{H}$ . . . . . . . . . . . . 57

# 1. Lebesgue Measure and Integration

## 1.1. Lebesgue Measure

**Definition 1.1**
*The **length** of an open interval $(a,b) = I$ is $b-a$ in the extended sense, denoted by $\ell(I)$.*

**Remark**
*We define $(a,a) = \varnothing$.*

**Definition 1.2**
*The **Lebesgue outer measure** (or in brief, **outer measure**) of a set $E \subset \mathbb{R}$ is*

$$\mu^*(E) = \inf\left\{\sum_n \ell(I_n) \,\middle|\, I_n \text{ are countable open intervals covering } E\right\}.$$

**Proposition 1.3**

(a) *Countable sets are of outer measure zero.*

(b) *If $A \subset B$, then $\mu^*(A) \leq \mu^*(B)$.*

(c) *For $x \in \mathbb{R}$ and $A \subset \mathbb{R}$, $\mu^*(A+x) = \mu^*(A)$.*

(d) *For countable $A_n \subset \mathbb{R}$, $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$.*

*Proof.* For (a), let $x_n$ denumerate a countable set $A$. Then consider

$$I_n = (x_n - 2^{-n}\epsilon, x_n + 2^{-n}\epsilon)$$

for $n \in \mathbb{N}$. Then $A \subset \bigcup_n I_n$ and $\mu^*(A) \leq \sum_n 2 \cdot 2^{-n}\epsilon = 2\epsilon$. Since $\epsilon$ is arbitrary, $\mu^*(A) = 0$.

For (b), note that any cover of $B$ must cover $A$. The result follows.

For (c), note that the translations of open intervals preserve their lengths.

For (d), let $\left\{I_j^n\right\}$ cover $A_n$ for each $n$ such that $\sum_j \ell(I_j^n) < \mu^*(A_n) + 2^{-n}\epsilon$. Then we have that $\bigcup_n \bigcup_j I_j^n$ covers $\bigcup_n A_n$ and

$$\sum_n \sum_j \ell(I_j^n) < \sum_n \mu^*(A_n) + 2^{-n}\epsilon = \epsilon + \sum_n \mu^*(A_n).$$

Since $\epsilon$ is arbitrary, it follows that $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$. ■

**Definition 1.4**
*A family of sets $\mathscr{M}$ is called a $\sigma$-**algebra** if*

(a) *$\varnothing \in \mathscr{M}$.*

(b) *$A \in \mathscr{M}$ implies $A^c \in \mathscr{M}$.*

(c) *For countably many $A_n \in \mathscr{M}$ we have $\bigcup_n A_n \in \mathscr{M}$.*

*The space $(X,\mathscr{M})$ is called a* ***measurable space*** *and the sets in $\mathscr{M}$ are called* ***measurable sets****.*

**Proposition 1.5**

*$\mathscr{M}$ is a $\sigma$-algebra if and only if the following hold:*

(a) $X \in \mathscr{M}$.

(b) $A,B \in \mathscr{M}$ *implies* $A \bigcap B, A \bigcup B, A-B \in \mathscr{M}$.

(c) *For countably many* $A_n \in \mathscr{M}$ *we have* $\bigcap_n A_n \in \mathscr{M}$.

*Proof.* Omitted. ∎

**Proposition 1.6**

*Let $\mathscr{F}$ be a family of sets in $X$. Then there exists a unique smallest $\sigma$-algebra containing $\mathscr{F}$.*

*Proof.* Let $\mathscr{M}$ be the intersection of all $\sigma$-algebras containing $\mathscr{F}$. Since $\mathscr{P}(X)$ must be such a $\sigma$-algebra, $\mathscr{M}$ is non-empty. Now we verify that $\mathscr{M}$ is a $\sigma$-algebra. First, $\varnothing \in \mathscr{M}$ since $\varnothing$ is in every $\sigma$-algebra. Second, if $A \in \mathscr{F}$ then $A$ must belong to every $\sigma$-algebra containing $\mathscr{F}$ and so does $A^c$. Hence $A^c \in \mathscr{M}$. The closure under countable unions follows from a similar argument. We conclude that $\mathscr{M}$ is the desired $\sigma$-algebra. ∎

**Definition 1.7**

*For a family of sets $\mathscr{F}$, we denote the smallest $\sigma$-algebra containing $\mathscr{F}$ by $\sigma(\mathscr{F})$.*

**Definition 1.8**

*Let $\mathscr{T}$ be the family of all open sets. The* ***Borel $\sigma$-algebra*** *is defined as $\mathscr{B} = \sigma(\mathscr{T})$. The sets in $\mathscr{B}$ are called* ***Borel sets****.*

**Definition 1.9**

*A set $E$ is called* ***Lebesgue measurable*** *if for $\epsilon > 0$, there exists an open set $V$ such that $E \subset V$ and $\mu^*(V-E) \leq \epsilon$.*

**Remark**

*The Lebesgue measurable sets form a $\sigma$-algebra.*

**Remark**

*The Borel sets are Lebesgue measurable.*

**Remark**

*Not all subsets in $\mathbb{R}$ are Lebesgue measurable. Consider the Vitali set. For a Lebesgue measurable set that is not Borel, consider the preimage of a Vitali set of Cantor-Lebesgue function.*

**Definition 1.10**

*A function $f:(X,\mathscr{M}) \to (\mathbb{R},\mathscr{B})$ is called* ***$\mathscr{M}$-measurable*** *if $f^{-1}(B) \in \mathscr{M}$ for all $B \in \mathscr{B}$.*

**Proposition 1.11**
*Let $f : X \to Y$ and $A$ be an index set. Then*

(a) $f^{-1}(B^c) = f^{-1}(B)^c$.

(b) $f^{-1}(\bigcup_{a\in A} B_a) = \bigcup_{a\in A} f^{-1}(B_a)$.

(c) $f^{-1}(\bigcap_{a\in A} B_a) = \bigcap_{a\in A} f^{-1}(B_a)$

*Proof.* Omitted. ■

**Proposition 1.12**
$f : (X, \mathscr{M}) \to (\mathbb{R}, \mathscr{B})$ *is* $\mathscr{M}$*-measurable if* $f^{-1}((a,\infty)) \in \mathscr{M}$.

*Proof.* Observe that $\{A \subset \mathbb{R} \mid f^{-1}(A) \in \mathscr{F}\}$ is a $\sigma$-algebra. By assumption, $[a,b]$, $(a,b]$, $[a,b)$ and $(a,b)$ are in this $\sigma$-algebra for $a,b \in \overline{\mathbb{R}}$. ■

**Proposition 1.13**
$f_n$ *are measurable. Then* $\sup_n f_n$, $\inf_n f_n$, $\limsup f_n$ *and* $\liminf f_n$ *are measurable.*

*Proof.* Note that $\{\sup_n f_n > a\} = \bigcup_n \{f_n > a\}$ and $\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\}$ are measurable. $\limsup_n f_n = \inf_k \sup_{n\geq k} f_n$ and $\liminf_n f_n = \sup_k \inf_{n\geq k} f_n$ are measurable as well. ■

**Remark**
$\lim_n f_n = \limsup_n f_n = \liminf_n f_n$ *is measurable.*

**Definition 1.14**
*Let* $(X, \mathscr{M})$ *be a measurable space. A* ***measure*** *on* $X$ *is a function* $\mu : \mathscr{M} \to [0,\infty]$ *satisfying*

(a) $\mu(\varnothing) = 0$.

(b) $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$ *for disjoint* $A_n$.

*The triple* $(X, \mathscr{M}, \mu)$ *is called a* ***measure space***.

**Proposition 1.15**
*Let* $(X, \mathscr{M}, \mu)$ *be a measure space and* $A, B \in \mathscr{M}$. *Then*

(a) $A \subset B$ *implies* $\mu(A) \leq \mu(B)$.

(b) $\mu(A - B) = \mu(A) - \mu(B)$ *if* $B \subset A$ *and* $\mu(B) < \infty$.

*Proof.* Omitted. ■

**Proposition 1.16**
*Let* $(X, \mathscr{M}, \mu)$ *be a measure space and* $E_n$ *be a sequence of measurable sets. Then*

(a) *If* $E_n \nearrow E$, *then* $\mu(E_n) \to \mu(E)$ *as* $n \to \infty$.

(b) *If* $E_n \searrow E$ *and* $\mu(E_1) < \infty$, *then* $\mu(E_n) \to \mu(E)$ *as* $n \to \infty$.

*Proof.* Suppose $\mu(E_n) < \infty$ for all $n$. Consider $S_n = E_n - E_{n-1}$ with $E_0 = \varnothing$. Then $S_n$ are disjoint and $\bigcup_n S_n = E$. Then

$$\mu(E) = \mu(\bigcup_n S_n) = \sum_n \mu(S_n) = \sum_n \mu(E_n) - \mu(E_{n-1}) = \lim_n \mu(E_n).$$

If $\mu(E_n) = \infty$ for some $n$, then $\mu(E) = \infty$ and the result follows.

For the second part, note that $E_1 - E_n \nearrow E_1 - E$. Then

$$\mu(E_1) - \mu(E_n) = \mu(E_1 - E_n) \to \mu(E_1 - E) = \mu(E_1) - \mu(E).$$

Rearranging gives the desired result. ■

**Theorem 1.17** (Egorov)
*Let $E$ be a measurable set with $\mu(E) < \infty$ and $f_n : E \to \mathbb{R}$ are measurable functions. If $f_n \to f$ a.e. on $E$, then for all $\epsilon > 0$, there exists a closed set $A_\epsilon \subset E$ such that $\mu(E - A_\epsilon) < \epsilon$ and $f_n \to f$ uniformly on $A_\epsilon$.*

*Proof.* Consider the case where $f_n \to f$ everywhere on $E$ since $\{x \in E \mid f_n(x) \not\to f(x)\}$ is of measure zero. For each $n, k \in \mathbb{N}$, let $E_k^n = \left\{x \in E \,\middle|\, |f_j(x) - f(x)| < 1/n \text{ for all } j > k\right\}$. Then fix $n$ and note that $E_k^n \nearrow E$ as $k \to \infty$. By proposition 1.16, there exists $k_n$ such that $\mu(E - E_{k_n}^n) < 2^{-n}$. Then we have $|f_j(x) - f(x)| < 1/n$ for every $j > k_n$ and $x \in E_{k_n}^n$. Choose $N$ such that $\sum_{n \ge N} 2^{-n} < \epsilon/2$ and let $\hat{A}_\epsilon = \bigcap_{n \ge N} E_{k_n}^n$. Then $\mu(E - \hat{A}_\epsilon) \le \sum_{n \ge N} \mu(E - E_{k_n}^n) < \epsilon/2$. Also, for any $\delta > 0$, we may pick $n > N$ with $1/n < \delta$ and for $x \in \hat{A}_\epsilon$, $|f_j(x) - f(x)| < \delta$ whenever $j > k_n$. Hence $f_n \to f$ uniformly on $\hat{A}_\epsilon$. We may further find a closed $A_\epsilon \subset \hat{A}_\epsilon$ such that $\mu(\hat{A}_\epsilon - A_\epsilon) < \epsilon/2$. Then $A_\epsilon$ is the desired set. ■

**Definition 1.18**
*A function $s : X \to Y$ is called **simple** if it only takes finitely many values.*

**Lemma 1.19**
*$f : E \to [0, \infty]$ is measurable. Then there exists a sequence of simple functions $s_n \nearrow f$; furthermore, if $f$ is bounded, then $s_n \to f$ uniformly.*

*Proof.* Consider $s_n = \sum_{k=0}^{n2^n - 1} k 2^{-n} \chi_{f^{-1}([k2^{-n}, (k+1)2^{-n}))} + n \chi_{f^{-1}([n, \infty])}$. Then $s_n$ are simple and $s_n \nearrow f$. If $f$ is bounded, then $f^{-1}([n, \infty]) = \varnothing$ for some $n$ large enough and $s_n \to f$ uniformly. ■

**Theorem 1.20** (Lusin)
*Let $E \subset \mathbb{R}$ be a set of finite measure and $f : E \to \mathbb{R}$ be a measurable, finite-valued function. Then for all $\epsilon > 0$, there exists a closed set $F_\epsilon \subset E$ such that $\mu(E - F_\epsilon) < \epsilon$ and $f|_{F_\epsilon}$ is continuous.*

*Proof.* First we may partition $E$ into $E = \bigcup_{i \in \mathbb{N}} E_i$ where $E_i = E \bigcap [-i, i]$. We first prove the result for simple functions. Let $f = \sum_{j=1}^N c_j \chi_{A_j}$ be a simple function with the stated properties. Then for each $j$, we may find a closed set $F_j \subset A_j$ such that $\mu(A_j - F_j) < \epsilon/N$. Now since $E_i$ are bounded, $F_j \bigcap E_i$ are compact and hence $f$ being constant on each $F_j \bigcap E_i$ is continuous. Note that $F_\epsilon = \bigcup_{i,j=1}^N F_j \bigcap E_i$ satisfies the desired properties. Next, for a general measurable function $f$, we may find a sequence of simple functions $s_n \nearrow f$ by lemma 1.19. Now by Egorov's theorem, we may find a closed set $F_\epsilon \subset E$ such that $\mu(E - F_\epsilon) < \epsilon$ and $s_n \to f$ uniformly on $F_\epsilon$. Since $s_n$ are continuous on $F_\epsilon$, $f$ is continuous on $F_\epsilon$. ■

**Remark**
*By Tietze's extension theorem, $f$ can be extended to a continuous function on all of $\mathbb{R}$.*

**Proposition 1.21**
*$E$ is Lebesgue measurable if and only if $\mu(E \triangle B) = 0$ for some Borel set $B$.*

*Proof.* Suppose $E$ is Lebesgue measurable. Then for each $n$, there exists an open set $V_n$ such that $E \subset V_n$ and $\mu(V_n - E) < 1/n$. Let $B = \bigcap_n V_n$. Then $B$ is a Borel set and $\mu(E \triangle B) = 0$. Conversely, if $\mu(E \triangle B) = 0$ for some Borel set $B$, since $B$ is measurable, there exists an open $V \supset B$ such that $\mu(V - B) < \epsilon$. Then $B = (E \bigcap B) \bigcup (B - E)$ and since the later set has outer measure zero, $E \bigcap B$ is measurable. And since $E - B$ is outer measure zero, $E \bigcap B = E$ is measurable. ■

**Proposition 1.22**
*If $f$ is Lebesgue measurable, then there exists a Borel measurable function $g$ such that $f = g$ a.e.*

*Proof.* Let $s_k \nearrow f$ be a sequence of simple functions with $s_k = \sum_{i=1}^{n_k} c_i \chi_{E_i}$ where $E_i$ are measurable. Then for each $E_i$ we may find a Borel set $B_i \subset E_i$ such that $\mu(E_i - B_i) = 0$ by the previous proposition. Then $t_k = \sum_{i=1}^{n_k} c_i \chi_{B_i}$ is a Borel measurable function. Let $g = \lim_{k \to \infty} t_k$. Then $g$ is Borel measurable and $f = g$ a.e. since $\mu(E_i - B_i) = 0$ for countably many $i$. ■

## 1.2. Lebesgue Integration

**Definition 1.23**
*For a simple function $s = \sum_{i=1}^n c_i \chi_{E_i}$, its* ***Lebesgue integral*** *is defined as*

$$\int s \, d\mu = \sum_{i=1}^n c_i \mu(E_i).$$

**Definition 1.24**
*For a non-negative measurable function $f$, its* ***Lebesgue integral*** *is defined as*

$$\int f \, d\mu = \sup \left\{ \int s \, d\mu \,\middle|\, s \text{ is simple and } 0 \le s \le f \right\}.$$

**Definition 1.25**
*For a measurable function $f : X \to [-\infty, \infty]$, its **Lebesgue integral** is defined as*

$$\int f\,d\mu = \int f^+ d\mu - \int f^- d\mu,$$

*where $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$ provided that*

$$\int |f|\,d\mu = \int f^+ d\mu + \int f^- d\mu < \infty.$$

*In such a case, we say that $f$ is **integrable**.*

**Proposition 1.26**
*For $f, g$ integrable and $c \in \mathbb{R}$,*

*(a)* $\int cf + g\,d\mu = c\int f\,d\mu + \int g\,d\mu$.

*(b) If $f \le g$ a.e., then $\int f\,d\mu \le \int g\,d\mu$.*

*Proof.* Omitted. ■

**Theorem 1.27** (Lebesgue Monotone Convergence Theorem)
*Let $f_n : X \to [0, \infty]$ be a sequence of measurable functions with $f_n \nearrow f$ a.e. Then*

$$\int f\,d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

*Proof.* By the monotonicity we have

$$\int f_n d\mu \le \int f\,d\mu$$

for all $n$ and hence

$$\lim_{n\to\infty} \int f_n d\mu \le \int f\,d\mu.$$

To obtain the reverse inequality, note that for any $c \in (0, 1)$, there exists $N$ such that $f_n \ge cf$ a.e. for all $n \ge N$. Then

$$\int f_n d\mu \ge c \int f\,d\mu$$

for all $n \ge N$. Letting $n \to \infty$,

$$\lim_{n\to\infty} \int f_n d\mu \ge c \int f\,d\mu.$$

Taking $c \to 1^-$ then

$$\lim_{n\to\infty} \int f_n d\mu \ge \int f\,d\mu \implies \lim_{n\to\infty} \int f_n d\mu = \int f\,d\mu.$$

■

**Remark**
*As a consequence,*

$$\int \sum_n f_n d\mu = \sum_n \int f_n d\mu.$$

**Theorem 1.28** (Bounded Covergence Theorem)
*Suppose $\mu(X) < \infty$. Let $f_n : X \to \mathbb{R}_+$ be a sequence measurable functions such that $f_n \le M$ a.e. for some $M \in \mathbb{R}$. If $f_n \to f$ a.e., then $f$ is integrable and*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

*Proof.* For any $\epsilon > 0$, by Egorov's theorem, there exists $F \subset X$ such that $\mu(X - F) < \epsilon$ and $f_n \to f$ uniformly on $F$. Then there exists $N$ such that $|f_n - f| < \epsilon$ on $F$ for all $n \ge N$. We have

$$\begin{aligned}
\left| \int f_n d\mu - \int f d\mu \right| &\le \int_X |f_n - f| d\mu \\
&= \int_F |f_n - f| d\mu + \int_{X-F} |f_n - f| d\mu \\
&\le \epsilon\mu(F) + 2M\mu(X - F) = \epsilon(\mu(F) + 2M\epsilon).
\end{aligned}$$

Since $\mu(X) < \infty$ and $\epsilon$ is arbitrary, we may conclude that

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

■

**Lemma 1.29** (Fatou)
*$f_n : X \to [0,\infty]$ are measurable. Then*

$$\int \liminf_n f_n d\mu \le \liminf_n \int f_n d\mu.$$

*Proof.* Let $g_n = \inf_{k\ge n} f_k$. Then $g_n \nearrow g = \liminf_n f_n$. By LMCT,

$$\int g_n d\mu \to \int g d\mu = \int \liminf_n f_n d\mu.$$

Note that $f_n \ge g_n$ and thus $\int f_n d\mu \ge \int g_n d\mu$. Hence

$$\liminf_n \int f_n d\mu \ge \liminf_n \int g_n d\mu = \int g d\mu = \int \liminf_n f_n d\mu.$$

■

**Theorem 1.30** (Lebesgue Dominated Convergence Theorem)
*Let $f_n : X \to [-\infty,\infty]$ be a sequence of measurable functions such that $f_n \to f$ a.e. and*

*there exists an integrable function $g$ such that $|f_n| \le g$ a.e. for all $n$. Then*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

*Proof.* Since $|f_n| \le g$ a.e., $|f| \le g$ a.e. Now $|f_n - f| \le |f_n| + |f| \le 2g$ a.e. Let $h_n = 2g - |f_n - f| \ge 0$ a.e. By Fatou's lemma,

$$\begin{aligned}\int 2g d\mu &= \int \liminf_n h_n d\mu \le \liminf_n \int h_n d\mu = \liminf_n \int 2g - |f_n - f| d\mu \\ &= \int 2g d\mu - \limsup_n \int |f_n - f| d\mu.\end{aligned}$$

It follows that

$$0 \le \liminf_n \int |f_n - f| d\mu \le \limsup_n \int |f_n - f| d\mu \le 0.$$

Hence

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0.$$

By the triangle inequality,

$$\left| \int f d\mu - \int f_n d\mu \right| \le \int |f - f_n| d\mu \to 0.$$

So

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

■

**Remark**

*If* supp$(f)$ *has finite measure and $f$ is bounded, then*

$$\int f = \inf_{s \ge f} \int s d\mu,$$

*where $s$ is simple.*

**Definition 1.31**

*$\mathscr{L}^1 = \{f : X \to \mathbb{R} \mid f$ is integrable$\}$ with the norm $\|f\|_{\mathscr{L}^1} = \int |f| d\mu$ is called the $\mathscr{L}^1$* ***space****.*

**Remark**

*The elements in $\mathscr{L}^1$ are in fact equivalence classes of functions that are equal a.e.*

**Proposition 1.32**

*Let $f \in \mathscr{L}^1$ be a nonegative function. Then for every $\epsilon > 0$, there is some $\delta > 0$ such that for any measurable $E$ with $\mu(E) \le \delta$,*

$$\int_E f d\mu \le \epsilon.$$

*Proof.* Let $E_n = \{x \in X \mid f(x) > n\}$. Then by Lebesgue dominated convergence theorem, since $f\chi_{E_n} \le f$,

$$\int_{E_n} f\,d\mu \to 0.$$

For any $\epsilon > 0$, there exists $n$ such that

$$\int_{E_n} f\,d\mu \le \frac{\epsilon}{2}.$$

Pick $\delta \le \epsilon/(2n)$. Then for any measurable $E$ with $\mu(E) \le \delta$,

$$\int_E f\,d\mu = \int_{E\cap E_n} f\,d\mu + \int_{E\cap E_n^c} f\,d\mu \le \int_{E_n} f\,d\mu + n\mu(E) \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

since $f \le n$ on $E_n^c$. This completes the proof. ■

**Theorem 1.33** (Lebesgue-Vitali)
*$f : X \to \mathbb{R}$ is Riemann integrable if and only if the discontinuity set of $f$ has Lebesgue measure zero. Furthermore, if $f$ is Riemann integrable, then the Riemann integral and the Lebesgue integral agrees.*

*Proof.* Define the oscillation of $f$ at $x$ as

$$\mathrm{osc}(f,x) = \inf_{U: x\in U} \mathrm{diam}(f(U)),$$

where $U$ is open.

We first claim that $f$ is continuous at $x$ if and only if $\mathrm{osc}(f,x) = 0$. Indeed, if $f$ is continuous at $x$, then $\forall\,\epsilon > 0$, $\exists\,\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $y \in B_\delta(x)$. Then $\mathrm{diam}(f(B_\delta(x))) \le 2\epsilon$. Since $\epsilon$ is arbitrary, $\mathrm{osc}(f,x) = 0$. Conversely, if $\mathrm{osc}(f,x) = 0$, then $\forall\,\epsilon > 0$, $\exists$ open $U$ containing $x$ such that $\mathrm{diam}(f(U)) < \epsilon$. This implies that $|f(x) - f(y)| < \epsilon$ for all $y \in U$ and hence $f$ is continuous at $x$.

Next, let $D_\epsilon$ collect all points $x$ such that $\mathrm{osc}(f,x) \ge \epsilon > 0$. We claim that $D_\epsilon$ is closed. For any convergent sequence $x_k \in D_\epsilon$, let $x_k \to x$. For any open $U$ containing $x$, $\exists N$ such that $x_k \in U$ for all $k \ge N$. Then $\exists$ an open neighborhood of $x_N$, $U'$, such that $U' \subset U$ and $\mathrm{diam}\big(f(U')\big) \ge \epsilon$. Hence $\mathrm{osc}(f,x) \ge \epsilon$ and $x \in D_\epsilon$, showing that $D_\epsilon$ is closed. Observe that $D = \bigcup_{n=1}^\infty D_{1/n}$.

Now suppose that $f$ is Riemann integrable. Then for any $\epsilon > 0$, $\exists \mathscr{P}$ such that $\mathrm{U}(f,\mathscr{P}) - \mathrm{L}(f,\mathscr{P}) < \frac{1}{n}$ and $\|\mathscr{P}\| < \frac{1}{n}$. Then

$$\begin{aligned}&\sum_{\substack{Q\in\mathscr{P},\\ Q\cap D_{\frac{1}{n}}\neq\varnothing}} (\sup_Q f - \inf_Q f)|Q| + \sum_{\substack{Q\in\mathscr{P},\\ Q\cap D_{\frac{1}{n}}=\varnothing}} (\sup_Q f - \inf_Q f)|Q| \\ &= \sum_{Q\in\mathscr{P}} (\sup_Q f - \inf_Q f)|Q| = \mathrm{U}(f,\mathscr{P}) - \mathrm{L}(f,\mathscr{P}) < \epsilon.\end{aligned}$$

Note that $\sup_Q f - \inf_Q f = \operatorname{diam}(f(Q))$. This gives that $2M\mu^*(D_{\frac{1}{n}}) < \epsilon$ for every $n$. Since $\epsilon$ is arbitrary, we conclude that $\mu^*(D_{\frac{1}{n}}) = 0$ for each $n$. Thus $D$ is an union of sets of measure zero and hence also has measure zero.

For the converse, suppose that $m(D) = 0$. Then $D_\epsilon$ also has measure zero. Let $\mathscr{P}$ be a partition on $E$ with $\|\mathscr{P}\| < \delta$ for some $\delta > 0$, which will be determined later. Then

$$\begin{aligned} \mathrm{U}(f,\mathscr{P}) - \mathrm{L}(f,\mathscr{P}) &= \sum_{Q\in\mathscr{P}} (\sup_Q f - \inf_Q f)|Q| \\ &= \sum_{\substack{Q\in\mathscr{P},\\ Q\bigcap D_\epsilon=\varnothing}} (\sup_Q f - \inf_Q f)|Q| + \sum_{\substack{Q\in\mathscr{P},\\ Q\bigcap D_\epsilon\neq\varnothing}} (\sup_Q f - \inf_Q f)|Q| \end{aligned}$$

For the first term, $\sup_Q f - \inf_Q f < \epsilon$ for $\|\mathscr{P}\| < \delta_1$ for some $\delta_1 > 0$. And thus the first term is bounded by $\epsilon m(E)$. For the second term, $\sup_Q f - \inf_Q f < 2M$ and since $D_\epsilon$ has measure zero, $\exists Q_k$ cubic cover of $D_\epsilon$ such that $\sum_k |Q_k| < \epsilon$. Now if $\operatorname{diam}(Q) < \delta_2$ for some $\delta_2 > 0$, then those $Q$ intersecting $D_\epsilon$ nonempty are subset of $\bigcup_k Q_k$. Thus the second term is bounded by $2M\epsilon$. Choosing $\delta = \min\{\delta_1, \delta_2\}$ yields that

$$\mathrm{U}(f,\mathscr{P}) - \mathrm{L}(f,\mathscr{P}) < \epsilon m(E) + 2M\epsilon$$

whenever $\|\mathscr{P}\| < \delta$. Since $\epsilon$ is arbitrary, $f$ is Riemann integrable. ∎

**Proposition 1.34**

*(a) Step functions are dense in $\mathscr{L}^1$.*

*(b) Continuous functions with compact support are dense in $\mathscr{L}^1$.*

*Proof.* Let $f \in \mathscr{L}^1$. By lemma 1.19, we already know that simple functions are dense in $\mathscr{L}^1$. It now remains to show that step functions can approximate simple functions. Since simple functions are linear combinations of finitely many characteristic functions, it suffices to show that characteristic functions can be approximated by step functions. Now for any measurable $E$, there is a family of almost disjoint cubes $Q_i$ such that $\mu(E \triangle \cup_{i=1}^M Q_i) \leq 2\epsilon$, and thus we may set the step function to be $\phi = \sum_{i=1}^M \chi_{Q_i}$, with $\left\|\chi_E - \phi\right\|_{\mathscr{L}^1} \leq 2\epsilon$.

For the second part, let it now suffices to show that continuous functions with compact support can approximate characteristic functions of a rectangle, say $[a,b]$. Then set

$$g(x) = \begin{cases} 0 & x \leq a-\epsilon, \\ \frac{x-a+\epsilon}{\epsilon} & a-\epsilon \leq x \leq a, \\ 1 & a \leq x \leq b, \\ 1 - \frac{x-b}{\epsilon} & b \leq x \leq b+\epsilon, \\ 0 & x \geq b+\epsilon. \end{cases}$$

Then $g$ is continuous with compact support and $\left\|\chi_{[a,b]} - g\right\|_{\mathscr{L}^1} \leq \epsilon/2 + \epsilon/2 = \epsilon$. ∎

## 1.3. Differentiation

**Definition 1.35**

*Let $f \in \mathscr{L}^1(\mathbb{R}^d)$. The **Hardy-Littlewood maximal function** is defined as*

$$f^*(x) = \sup_{B:x\in B} \frac{1}{\mu(B)} \int_B |f(y)|\,dy,$$

*where the supremum is taken over all balls containing $x$.*

**Proposition 1.36**

*$f^*$ is measurable.*

*Proof.* Let $E_\alpha = \{x \mid f^*(x) > \alpha\}$. We claim that it is an open set. Indeed, if $p \in E_\alpha$, there exists a ball $B$ containing $p$ such that

$$\frac{1}{\mu(B)} \int_B |f(y)|\,dy > \alpha.$$

Now any $x$ close enough to $p$ will be contained in $B$ and hence in $E_\alpha$. Thus $E_\alpha$ is open. Hence $f^*$ is measurable. ∎

**Lemma 1.37**

*[Vitali Covering Lemma] Suppose $\{B_1, \ldots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^d$. Then there exists a disjoint subcollection $\left\{B_{i_1}, \ldots, B_{i_k}\right\}$ such that*

$$\mu\left(\bigcup_{j=1}^N B_j\right) \leq 3^d \sum_{j=1}^k \mu(B_{i_j}).$$

*Proof.* First we make an observation that if $B$ and $B'$ are balls intersecting with, say, the radius of $B$ is greater than the radius of $B'$, then $B'$ is contained in the ball $\tilde{B}$ that is concentric with $B$ but with 3 times the radius.

The construction of the subcollection is proceeded as follows. First, pick a ball $B_{i_1}$ with the largest radius. Then remove all balls intersecting with $\tilde{B}_{i_1}$, the ball concentric with $B_{i_1}$ but with 3 times the radius. Among the remaining balls, we repeat the process and pick $B_{i_2}$. The process terminates when no more balls can be picked, after at most $N$ steps and we obtain a disjoint subcollection of balls $\left\{B_{i_1}, \ldots, B_{i_k}\right\}$.

Lastly, we verify the inequality. By the construction, we know that $\cup_{j=1}^N B_j \subset \cup_{j=1}^k \tilde{B}_{i_j}$ and thus

$$\mu\left(\bigcup_{j=1}^N B_j\right) \leq \mu\left(\bigcup_{j=1}^k \tilde{B}_{i_j}\right) \leq \sum_{j=1}^k \mu(\tilde{B}_{i_j}) = \sum_{j=1}^k 3^d \mu(B_{i_j}).$$

∎

**Theorem 1.38** (Weak-Type Inequality)
*Let $f \in \mathscr{L}^1(\mathbb{R}^d)$. Then for all $\alpha > 0$,*

$$\mu\left(\left\{x \in \mathbb{R}^d \,\middle|\, f^*(x) > \alpha\right\}\right) \leq \frac{A}{\alpha} \|f\|_{\mathscr{L}^1(\mathbb{R}^d)},$$

*where $A = 3^d$.*

*Proof.* Let $E_\alpha = \{x \mid f^*(x) > \alpha\}$. For each $x \in E_\alpha$ there exists a ball $B_x$ containing $x$ such that

$$\frac{1}{\mu(B_x)} \int_{B_x} |f(y)|\,dy > \alpha \quad \Rightarrow \quad \mu(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)|\,dy.$$

Now for any fixed compact $K \subset E_\alpha$, $K$ is covered by $\bigcup_{x \in E_\alpha} B_x$, and hence there exists a finite subcover $\{B_1, \ldots, B_N\}$ of $K$. By the Vitali covering lemma, there exists a disjoint subcollection $\{B_{i_1}, \ldots, B_{i_k}\}$ with

$$\mu\left(\bigcup_{j=1}^N B_j\right) \leq 3^d \sum_{j=1}^k \mu(B_{i_j}).$$

As a result,

$$\begin{aligned}
\mu(K) \leq \mu\left(\bigcup_{j=1}^N B_j\right) &\leq 3^d \sum_{j=1}^k \mu(B_{i_j}) \leq \frac{3^d}{\alpha} \sum_{j=1}^k \int_{B_{i_j}} |f(y)|\,dy \\
&\leq \frac{3^d}{\alpha} \int_{\cup_{j=1}^k B_{i_j}} |f(y)|\,dy \leq \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)|\,dy.
\end{aligned}$$

Since the inequality holds for all compact subset $K$ of $E_\alpha$, the proof is complete. ∎

**Remark**
*Note that $\{x \mid f^*(x) = \infty\} \subset \{x \mid f^*(x) > \alpha\}$ for every $\alpha > 0$. Taking $\alpha \to \infty$ yields*

$$\mu\left(\left\{x \,\middle|\, f^*(x) = \infty\right\}\right) = 0.$$

*Hence $f^*(x) < \infty$ a.e.*

**Theorem 1.39** (Lebesgue Differentiation Theorem)
*Let $f \in \mathscr{L}^1(\mathbb{R}^d)$. Then for almost every $x \in \mathbb{R}^d$,*

$$\lim_{m(B) \to 0, x \in B} \frac{1}{m(B)} \int_B f(y)dy = f(x).$$

*Proof.* Since continuous functions are dense in $\mathscr{L}^1$, we may find a continuous $g$ such that $\|f - g\|_{\mathscr{L}^1} < \epsilon$. For such $g$, by the continuity, there exists a ball such that $|g(y) - g(x)| < \epsilon$

for all $x, y \in B$. Thus

$$\begin{aligned}\left|\frac{1}{m(B)}\int_B f(y)dy - f(x)\right| &= \left|\frac{1}{m(B)}\int_B (f(y)-g(y))dy + \frac{1}{m(B)}\int_B g(y)-g(x)dy + g(x)-f(x)\right| \\ &\le \frac{1}{m(B)}\int_B |(f(y)-g(y))|\,dy + \frac{1}{m(B)}\int_B |g(y)-g(x)|\,dy + |g(x)-f(x)| \\ &\le (f-g)^*(x) + \epsilon + |g(x)-f(x)|.\end{aligned}$$

Since $\epsilon$ can be arbitrary small, we have

$$\left|\frac{1}{m(B)}\int_B f(y)dy - f(x)\right| \le (f-g)^*(x) + |g(x)-f(x)|.$$

Now we let

$$E_\alpha = \left\{x \;\middle|\; \limsup_{m(B)\to 0, x\in B}\left|\frac{1}{m(B)}\int_B f(y)dy - f(x)\right| > 2\alpha\right\}.$$

We claim that $E_\alpha$ has measure zero. Set

$$F_\alpha = \left\{x \;\middle|\; (f-g)^*(x) > \alpha\right\} \quad \text{and} \quad G_\alpha = \{x \mid |g(x)-f(x)| > \alpha\}.$$

Then we have $E_\alpha \subset F_\alpha \cup G_\alpha$. By the weak-type inequality and Tchebyshev's inequality,

$$\mu(F_\alpha) \le \frac{A}{\alpha}\|f-g\|_{\mathscr{L}^1} < \frac{A}{\alpha}\epsilon \quad \text{and} \quad \mu(G_\alpha) \le \frac{1}{\alpha}\|f-g\|_{\mathscr{L}^1} < \frac{1}{\alpha}\epsilon.$$

Thus $\mu(E_\alpha) \le \mu(F_\alpha \cup G_\alpha) < \frac{A+1}{\alpha}\epsilon$. Since $\epsilon$ is arbitrary, we have $\mu(E_\alpha) = 0$ and the proof is complete. ■

**Remark**
*For $f \in \mathscr{L}^1(\mathbb{R})$, and $F(x) = \int_{-\infty}^x f(y)dy$, we have $F'(x) = f(x)$ a.e. Indeed,*

$$\begin{aligned}\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| &= \frac{1}{h}\left|\int_x^{x+h} f(y)-f(x)dy\right| \le \frac{1}{h}\int_x^{x+h} |f(y)-f(x)|\,dy \\ &\le \frac{1}{h}\int_{x-h}^{x+h} |f(y)-f(x)|\,dy \le 2\frac{1}{2h}\int_{x-h}^{x+h} |f(y)-f(x)|\,dy \to 0\end{aligned}$$

*as $h \to 0$ a.e. $x$.*

**Remark**
*In fact, the requirement that $f \in \mathscr{L}^1$ can be relaxed to $f \in \mathscr{L}^1_{loc}$, which is defined as the set of all locally integrable functions, i.e., $f\chi_B \in \mathscr{L}^1$ for all finite balls $B$ since the proof only requires $B$ to be a ball near $x$.*

## 1.4. Radon-Nikodym Theorem

**Definition 1.40**
*Let* $(X,\mathscr{A})$ *be a measurable space. A* ***signed measure*** *is a function* $\mu:\mathscr{A}\to[-\infty,\infty]$ *such that* $\mu(\varnothing)=0$ *and for any countable disjoint collection* $\{A_i\}_{i\in\mathbb{N}}$,

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right)=\sum_{i\in\mathbb{N}}\mu(A_i).$$

**Remark**
*The range of* $\mu$ *can only include one of* $\pm\infty$.

**Definition 1.41**
*Let* $(X,\mathscr{A},\mu)$ *be a measure space.* $\mu$ *is called* $\sigma$***-finite*** *if* $X$ *can be covered by countably many* $A_n\in\mathscr{A}$ *such that* $\mu(A_n)<\infty$ *for all* $n$. *In this case, we also call* $X$ $\sigma$***-finite***.

**Definition 1.42**
*Let* $\nu,\lambda$ *be two measures defined on a measurable space.* $\nu$ *is said to be* ***absolutely continuous*** *with respect to* $\lambda$ *if* $\lambda(A)=0$ *implies that* $\nu(A)=0$ *for all measurable* $A$, *denoted as* $\nu\ll\lambda$.

**Example**
*Let*

$$\nu(A)=\int_A f\,d\lambda$$

*where* $f\geq 0$ *is measurable. Then* $\lambda(A)=0$ *implies* $\nu(A)=0$. $\nu\ll\lambda$.

**Definition 1.43**
*Let* $\nu,\lambda$ *be two measures defined on a measurable space.* $\nu$ *is said to be* ***singular*** *with respect to* $\lambda$ *if there exists a measurable set* $A$ *such that* $\lambda(A)=0$ *and* $\nu(A^c)=0$, *denoted as* $\nu\perp\lambda$.

**Example**
*Let* $\lambda$ *be the Lebesgue measure on* $[0,1]$ *and*

$$\nu(A)=\sum_i c_i\delta_{q_i}(A),\quad \textit{with}\quad \sum_i c_i<\infty,\quad \delta_{q_i}(A)=\mathbf{1}\{q_i\in A\},$$

*where* $q_i$ *enumerates the rationals in* $[0,1]$ *and* $\mathbf{1}$ *is the indicator function. Then* $\nu\perp\lambda$.

**Definition 1.44**
$\nu$ *and* $\lambda$ *are said to be* ***equivalent*** *if* $\nu\ll\lambda$ *and* $\lambda\ll\nu$.

**Definition 1.45**
*Let* $(X,\mathscr{A},\mu)$ *be a measure space. A set* $P\in\mathscr{A}$ *is said to be* ***positive*** *if* $\mu(A)\geq 0$ *for all measurable* $A\subset P$; *a set* $N\in\mathscr{A}$ *is said to be* ***negative*** *if* $\mu(A)\leq 0$ *for all measurable* $A\subset N$.

**Theorem 1.46** (Hahn Decomposition)
*Let $\mu$ be a signed measure on a measurable space $(X,\mathscr{A})$. Then $X$ can be partitioned into a positive set $P$ and a negative set $N$. Furthermore, if $P'$, $N'$ form another such partition, then $P \triangle P'$ and $N \triangle N'$ are measure zero.*

*Proof.* We may consider the case where $\mu(A) \neq -\infty$ for all $A \in \mathscr{A}$. The other case is similar. We first claim that every measurable set $A$ contains a postive set $P$ such that $\mu(P) \geq \mu(A)$.

To prove the claim, we first show that for every $\epsilon > 0$, there exists $A_\epsilon \subset A$ such that $\mu(A_\epsilon) \geq \mu(A)$ and $B \subset A_\epsilon$ implies $\mu(B) > -\epsilon$. Otherwise, we can pick a sequence of set $B_k$ inductively, such that $B_1 \subset A, \dots, B_k \subset A - (B_1 \cup \cdots \cup B_{k-1}), \dots$ with $\mu(B_k) \leq -\epsilon$. Put $B = \cup_k B_k$. Since $B_k$ are disjoint, $\mu(B) = -\infty$. Also, $\mu(A-B) = \mu(A) - \mu(B) = \infty$, contradicting to the remark that $\mu$ cannot take both $\pm\infty$. Now choose $\epsilon_n \to 0$ and let $P = \cap_n A_{\epsilon_n}$. $A_{\epsilon_n} \searrow P$ and then $\mu(A_{\epsilon_n}) \to \mu(P)$ by proposition 1.16. Thus $\mu(P) \geq \mu(A)$.

Next, let $s = \sup\{\mu(A) \mid A \in \mathscr{A}\}$. There is a sequence $P_n$ such that $\mu(P_n) \to s$. Note that $s \geq 0$ since $\varnothing \in \mathscr{A}$. By the claim, we may assume that $P_n$ are positive. Putting $P = \cup_n P_n$, we have $\mu(P) = s$ and $P$ is positive. Now let $N = X - P$. $N$ is negative; otherwise if $E \subset N$ and $\mu(E) > 0$, then $\mu(P \cup E) = \mu(P) + \mu(E) > s$, which contradicts to the definition of $s$.

Finally, suppose $P'$ and $N'$ are another such partition. Then $P \cap N'$ and $N \cap P'$ are both negative and positive, implying that they are measure zero. $\mu(P \triangle P') = \mu(P \cap N') + \mu(N \cap P') = 0$. This furnishes the proof. ∎

**Corollary 1.47** (Hahn-Jordan Decomposition)
*If $\nu$ is a signed measure on a measurable space $(X,\mathscr{A})$, then there exists a unique pair of positive measures $\nu^+$ and $\nu^-$ such that $\nu = \nu^+ - \nu^-$.*

*Proof.* By the Hahn decomposition, $X$ can be partitioned into a positive set $P$ and a negative set $N$. Define $\nu^+(A) = \nu(A \cap P)$ and $\nu^-(A) = -\nu(A \cap N)$. Then $\nu^+$ and $\nu^-$ are positive measures and $\nu = \nu^+ - \nu^-$. The uniqueness follows from the uniqueness of the Hahn decomposition. ∎

**Theorem 1.48** (Radon-Nikodym)
*Let $(X,\mathscr{A})$ be a measurable space and $\nu$, $\lambda$ are $\sigma$-finite measures on $(X,\mathscr{A})$. If $\nu \ll \lambda$, then there exists an $\mathscr{A}$-measurable function $f : X \to [0,\infty)$ such that for every $A \in \mathscr{A}$,*

$$\nu(A) = \int_A f \, d\lambda.$$

*Furthermore, if $f$ and $f'$ are two such functions, then $f = f'$ a.e.*

*Proof.* We first consider the case where $\nu$ and $\lambda$ are finite. Let

$$F = \left\{ f : X \to [0,\infty] \,\middle|\, \int_A f \, d\lambda \leq \nu(A) \text{ for all } A \in \mathscr{A} \right\}.$$

$F \neq \emptyset$ since $f = 0$ is in $F$. Now let $f_1, f_2 \in F$ and $A \in \mathscr{A}$ and define

$$A_1 = \{x \in A \mid f_1(x) > f_2(x)\}, \quad A_2 = \{x \in A \mid f_1(x) \le f_2(x)\}.$$

Then

$$\int_A \max\{f_1, f_2\} d\lambda = \int_{A_1} f_1 d\lambda + \int_{A_2} f_2 d\lambda \le \nu(A_1) + \nu(A_2) = \nu(A).$$

Thus $\max\{f_1, f_2\} \in F$. Next, for any sequence of functions $f_n \in F$ such that

$$\lim_{n\to\infty} \int_X f_n d\lambda = \sup_{f\in F} \int_X f d\lambda,$$

we may assume that $f_n \nearrow$ by replacing $f_n$ with the maximum among $f_1, \ldots, f_n$. Let $g$ be the pointwise limit of $f_n$. By Lebesgue's monotone convergence theorem,

$$\int_A g d\lambda = \lim_{n\to\infty} \int_A f_n d\lambda \le \nu(A),$$

so $g \in F$. Also, by construction,

$$\int_X g d\lambda = \sup_{f\in F} \int_X f d\lambda.$$

Now define

$$\nu_0(A) = \nu(A) - \int_A g d\lambda.$$

Since $g \in F$, $\nu_0$ is a nonnegative measure. To prove the equality, we need to show that $\nu_0(A) = 0$ for all $A \in \mathscr{A}$. Suppose $\nu_0 > 0$. Then there exists $\epsilon > 0$ such that $\nu_0(X) > \epsilon\lambda(X)$. By the Hahn decomposition theorem, we can find a positive set $P$ such that $\nu_0(A) \ge \epsilon\lambda(A)$ for each $A \subset P$. Thus

$$\nu(A) = \int_A g d\lambda + \nu_0(A) \ge \int_A g d\lambda + \nu_0(P \cap A) \ge \int_A g d\lambda + \epsilon\lambda(P\cap A) = \int_A (g + \epsilon\chi_P) d\lambda.$$

Note that $\lambda(P) > 0$, for otherwise $\lambda(P) = 0$ and $\nu_0(P) \le \nu(P) = 0 \implies \nu(P) = 0$ by the absolute continuity and hence

$$\nu_0(X) - \epsilon\lambda(X) = (\nu_0 - \epsilon\lambda)(N) \le 0,$$

posing a contradiction. Meanwhile,

$$\int_X (g + \epsilon\chi_P) d\lambda \le \nu(X) < \infty \implies g + \epsilon\chi_P \in F,$$

and

$$\int_X (g + \epsilon\chi_P) d\lambda > \int_X g d\lambda = \sup_{f\in F} \int_X f d\lambda.$$

This violates the definition of the supremum. Thus $\nu_0 = 0$ and we obtain that

$$\nu(A) = \int_A g d\lambda.$$

Finally, if we define

$$f(x) = \begin{cases} g(x) & \text{if } g(x) < \infty, \\ 0 & \text{if } g(x) = \infty, \end{cases}$$

since $g$ is $\lambda$-integrable, $f = g$ $\lambda$-a.e. and $f$ is the desired function.

For the uniqueness, suppose $f$ and $f'$ are two such functions. Then

$$\nu(A) = \int_A f d\lambda = \int_A f' d\lambda \implies \int_A (f - f') d\lambda = 0$$

for every $A$. In particular, letting $A = \{x \in X \mid f(x) \le f'(x)\}$ or $A = \{x \in X \mid f(x) \ge f'(x)\}$ gives

$$\int_X (f - f')^+ d\lambda = \int_X (f - f')^- d\lambda = 0.$$

Thus $f = f'$ $\lambda$-a.e.

For the general case where $\nu$ and $\lambda$ are $\sigma$-finite, we can write $X = \cup_n X_n$ such that $\lambda(X_n) < \infty$ and $X_n$ are disjoint. For each $n$ we can find $f_n$ such that

$$\nu(A) = \int_A f_n d\lambda.$$

for every $\mathscr{A}$-measurable $A \subset X_n$. Let $f = \sum_n f_n \chi_{X_n}$.

$$\int_A f d\lambda = \sum_n \int_{A \cap X_n} f_n d\lambda = \sum_n \nu(A \cap X_n) = \nu(A),$$

for every $A \in \mathscr{A}$. The uniqueness follows from the uniqueness of $f_n$. ∎

**Remark**

*The function $f$ can be chosen in $\mathscr{L}^1(X, \lambda)$ if $\nu$ is finite.*

**Definition 1.49**

*The function $f$ in the Radon-Nikodym theorem is called the* ***Radon-Nikodym derivative*** *of $\nu$ with respect to $\lambda$, denoted as $f = \frac{d\nu}{d\lambda}$.*

**Proposition 1.50**

*Let $\nu$, $\mu$ and $\lambda$ be $\sigma$-finite measures defined on measurable space $(X, \mathscr{A})$. If $\nu \ll \lambda$ and $\mu \ll \lambda$, then*

*(a) $\frac{d(\nu+\mu)}{d\lambda} = \frac{d\nu}{d\lambda} + \frac{d\mu}{d\lambda}$ $\lambda$-a.e.*

*(b) If $\nu \ll \mu \ll \lambda$, then $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ $\lambda$-a.e.*

*(c) If $\nu$ and $\mu$ are equivalent, then $\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1}$ $\mu$-a.e.*

*(d) If $g$ is $\nu$-integrable, then*

$$\int_X g\,d\nu = \int_X g\frac{d\nu}{d\lambda}d\lambda.$$

*Proof.* For (a), note that $\nu+\mu \ll \lambda$ as well. Let $f = \frac{d\nu}{d\lambda}$ and $g = \frac{d\mu}{d\lambda}$. Then

$$\int_A (f+g)d\lambda = \int_A f\,d\lambda + \int_A g\,d\lambda = \nu(A)+\mu(A) = (\nu+\mu)(A) = \int_A \frac{d(\nu+\mu)}{d\lambda}d\lambda \quad \text{for all } A \in \mathscr{A}.$$

Thus $\frac{d\nu}{d\lambda} + \frac{d\mu}{d\lambda} = f+g = \frac{d(\nu+\mu)}{d\lambda}$ $\lambda$-a.e.

Next, we jump to (d). We start by considering the case where $g = \chi_A$ with $A \in \mathscr{A}$. By the Radon-Nikodym theorem,

$$\int_X g\,d\nu = \int_X \chi_A\,d\nu = \nu(A) = \int_A \frac{d\nu}{d\lambda}d\lambda = \int_X \chi_A\frac{d\nu}{d\lambda}d\lambda = \int_X g\frac{d\nu}{d\lambda}d\lambda.$$

By linearity, the result holds for simple functions. For a nonnegative $g \in \mathscr{L}^1(\nu)$, we can find a sequence of simple functions $g_n \nearrow g$ so that

$$\int_X g\,d\nu = \lim_{n\to\infty}\int_X g_n\,d\nu = \lim_{n\to\infty}\int_X g_n\frac{d\nu}{d\lambda}d\lambda = \int_X g\frac{d\nu}{d\lambda}d\lambda$$

by Lebesgue's monotone convergence theorem. For general $g \in \mathscr{L}^1(\nu)$, we can write $g = g^+ - g^-$ and apply the result to $g^+$ and $g^-$.

$$\int_X g\,d\nu = \int_X g^+d\nu - \int_X g^-d\nu = \int_X g^+\frac{d\nu}{d\lambda}d\lambda - \int_X g^-\frac{d\nu}{d\lambda}d\lambda = \int_X g\,d\nu.$$

With (d) established, we can now prove (b). By the Radon-Nikodym theorem,

$$\int_A \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}d\lambda = \int_A \frac{d\nu}{d\mu}d\mu = \int_A d\nu = \nu(A) = \int_A \frac{d\nu}{d\lambda}d\lambda.$$

Finally, for (c), letting $\lambda = \nu$ and applying (b) gives $1 = \frac{d\nu}{d\nu} = \frac{d\nu}{d\mu}\frac{d\mu}{d\nu}$ $\nu$-a.e. and thus $\mu$-a.e. by the equivalence of $\nu$ and $\mu$. Hence $\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1}$ $\mu$-a.e. ■

**Theorem 1.51** (Lebesgue Decomposition)
*Let $\nu,\lambda$ be two $\sigma$-finite measures defined on a measurable space $(X,\mathscr{A})$. Then $\nu$ can be decomposed uniquely into $\nu = \nu_a + \nu_s$ where $\nu_a \ll \lambda$ and $\nu_s \perp \lambda$.*

*Proof.* We first assume that $\nu,\lambda$ are finite measures. Let $\mu = \nu+\lambda$. Then clearly $\lambda \ll \mu$ and $\mu$ is $\sigma$-finite. By the Radon-Nikodym theorem, there exists a Radon-Nikodym derivative $f$ such that

$$\lambda(A) = \int_A f\,d\mu.$$

Denote $\{x \in X \mid f(x) = 0\}$ by $E$. Define

$$\nu_a(A) = \nu(A \cap E^c), \quad \nu_s(A) = \nu(A \cap E)$$

for each $A \in \mathscr{A}$. Then clearly $\nu_a(A) + \nu_s(A) = \nu(A \cap E^c) + \nu(A \cap E) = \nu(A)$ for all $A \in \mathscr{A}$. Also, suppose $\lambda(A) = 0$. Then by proposition 1.50,

$$0 = \lambda(A) = \int_A f\, d\mu = \int_A f\, d\lambda + \int_A f\, d\nu = \int_A f\, d\nu.$$

Hence $f(x) = 0$ $\nu$-a.e. on $A$. This implies that $\nu(A) = \nu(A \cap E)$ and thus $\nu_a(A) = \nu(A \cap E^c) = \nu(A) - \nu(A \cap E) = 0$, so $\nu_a \ll \lambda$. Also, since $\lambda(E) = 0$ and $\nu_s(E^c) = \nu(\varnothing) = 0$, $\nu_s \perp \lambda$. For the uniqueness, suppose $\nu = \nu_a + \nu_s = \nu'_a + \nu'_s$ both satisfy the conditions. Since $\nu_a \ll \lambda$ and $\nu'_a \ll \lambda$, by the uniqueness of the Radon-Nikodym derivative, $\nu_a = \nu'_a$ and hence $\nu_s = \nu'_s$ as well.

Finally, for the general case where $\nu, \lambda$ are $\sigma$-finite, write $X = \cup_n X_n$ where $\lambda(X_n) < \infty$ and $X_n$ are disjoint. For each $n$ we can find the corresponding decomposition $\nu_a^n$ and $\nu_s^n$. Let $\nu_a = \sum_n \nu_a^n$ and $\nu_s = \sum_n \nu_s^n$. Then $\nu_a \ll \lambda$ and $\nu_s \perp \lambda$. The uniqueness follows from the uniqueness of the decompositions in each $X_n$. This establishes the proof. ■

**Corollary 1.52**
*Let $\nu$ be a signed measure and $\lambda$ be a measure defined on a measurable space $(X, \mathscr{A})$. Suppose both $\nu$ and $\lambda$ are finite and $\nu \ll \lambda$. Then there exists a unique $f \in \mathscr{L}^1(X, \lambda)$ such that*

$$\nu(A) = \int_A f\, d\lambda.$$

*Proof.* By Hahn decomposition, there exists a positive set $P$ and a negative set $N$ such that $P \cup N = X$. Define

$$\nu_P(A) = \nu(A \cap P), \quad \nu_N(A) = -\nu(A \cap N).$$

Then clearly $\nu_P - \nu_N = \nu$ and $|\nu| = \nu_P + \nu_N$. Note that $\nu_P$ and $\nu_N$ are both positive measures. Also, by assumption, if $\lambda(A) = 0$ then $\nu(A) = 0$ and hence so are $\nu_P$ and $\nu_N$. Thus $\nu_P \ll \lambda$ and $\nu_N \ll \lambda$. By the Radon-Nikodym theorem, there exists $f_P, f_N \in \mathscr{L}^1(X, \lambda)$ such that

$$\nu_P(A) = \int_A f_P d\lambda, \quad \nu_N(A) = \int_A f_N d\lambda.$$

Hence

$$\nu(A) = \nu_P(A) - \nu_N(A) = \int_A f_P d\lambda - \int_A f_N d\lambda = \int_A (f_P - f_N) d\lambda.$$

By setting $f = f_P - f_N$, we obtain the desired function. Uniqueness follows from the uniqueness of the Radon-Nikodym derivative. ■

# 2. Banach Space

## 2.1. Banach Space and Bounded Linear Functional

**Definition 2.1**
*A space $X$ is called a **Banach space** if it is a complete normed vector space.*

**Remark**
*$\mathscr{L}^1$ is a Banach space with the norm*

$$\|f\|_{\mathscr{L}^1} = \int |f|\, d\mu.$$

*We treat $f = g$ a.e. as the same element in $\mathscr{L}^1$.*

**Definition 2.2**
*Let $V, W$ be vector spaces. A map $T : V \to W$ is **linear** if for every $c \in \mathbb{R}$, $f, g \in V$, $T(cf+g) = cT(f) + T(g)$.*

**Definition 2.3**
*A linear map $T : V \to W$ has **operator norm** defined by*

$$\|T\| = \sup_{\|f\|_V = 1} \|T(f)\|_W.$$

*$T$ is **bounded** if $\|T\| < \infty$. We denote the set of all bounded linear operators from $V$ to $W$ by $B(V, W)$.*

**Proposition 2.4**
*Suppose $W$ is a Bnach space. Then $B(V, W)$ is a Banach space with the operator norm.*

*Proof.* It suffices to show that $B(V, W)$ is complete. Let $\{T_i\} \subset B(V, W)$ be a Cauchy sequence. Then for $f \in V$,

$$\left\|T_i(f) - T_j(f)\right\|_W \le \left\|T_i - T_j\right\| \|f\|_V.$$

Hence $\{T_i(f)\}$ is a Cauchy sequence in $W$. By the completeness of $W$, we may define $Tf$ as the limit of $T_i(f)$ as $i \to \infty$. Now,

$$\|Tf\| \le \sup_i \|T_i(f)\| \le \sup_i \|T_i\| \|f\|.$$

Since Cauchy sequences are bounded, $\|Tf\| < \infty$ for all $f \in V$ and $T \in B(V, W)$. It remains to show that $T_i$ converges to $T$ in the operator norm. For any $f \in V$, pick $N$ such that $\left\|T_i(f) - T_j(f)\right\| \le \epsilon$ for all $i, j \ge N$. Then for fixed $i$,

$$\left\|(T_i - T_j)f\right\| \le \left\|T_i - T_j\right\| \|f\| \le \epsilon \|f\|$$

for every $f \in V$ and $j \geq N$. Hence $\|T_i - T\| \leq \epsilon$ for all $i \geq N$ and the proof is complete. ■

**Definition 2.5**
*$T$ is continuous if $f_i \to f$ in $V$ implies that $T(f_i) \to T(f)$ in $W$.*

**Proposition 2.6**
*Suppose $T : V \to W$ is linear. Then $T$ is continuous if and only if $T$ is bounded.*

*Proof.* Suppose $T$ is not bounded. Then there exists $f_i \in V$ with $\|f_i\| \leq 1$ for all $i$ and $\|Tf_i\| \to \infty$. Thus

$$\frac{f_i}{\|Tf_i\|} \to 0, \quad \text{but} \quad T\left(\frac{f_i}{\|Tf_i\|}\right) = \frac{Tf_i}{\|Tf_i\|} \not\to 0 \quad \text{as} \quad \frac{\|Tf_i\|}{\|Tf_i\|} = 1.$$

Hence $T$ is not continuous.

Conversely, suppose $T$ is bounded. Let $f_i \to f$ in $V$. Then

$$\|Tf_i - Tf\| = \|T(f_i - f)\| \leq \|T\| \|f_i - f\| \to 0.$$

Hence $T$ is continuous. ■

**Definition 2.7**
*A **linear functional** $T$ is a linear map $T : V \to \mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $C$ is the scalar field of $V$.*

**Definition 2.8**
*Let $V$, $W$ be vector spaces. $T : V \to W$ is linear. The **kernel** of $T$ is defined as*

$$\ker(T) = \{f \in V \mid T(f) = 0\}.$$

**Proposition 2.9**
*Let $X$ be a normed vector space and $T \in X'$. Then*

- *(a)* $\ker(T)$ *is a closed subspace of $X$.*
- *(b) If $T \neq 0$, there exists $x \in X$ such that $T(x) \neq 0$. Then for any $y \in X$, there exists $c \in \mathbb{R}$ and $z \in \ker(T)$ such that $y = cx + z$.*

*Proof.* For (a), let $x, y \in \ker(T)$ and $c \in \mathbb{R}$.

$$T(cx + y) = cT(x) + T(y) = 0. \implies cx + y \in \ker(T).$$

Also, let $x_i \to x$ in $X$. Then since $T$ is continuous,

$$T(x) = \lim_{n\to\infty} T(x_n) = 0. \implies x \in \ker(T).$$

Hence $\ker(T)$ is a closed subspace of $X$.

For the rest part, fix $x \in X$ and $f(x) \neq 0$. For each $y \in X$, let $\alpha = T(y)/T(x)$ and $z = y - T(y)x/T(x)$. Then

$$\alpha x + z = \frac{T(y)}{T(x)}x + y - \frac{T(y)}{T(x)}x = y - \frac{T(y)}{T(x)}x = y.$$

■

**Definition 2.10**
*The **dual space** of $V$ is defined as $V' = B(V, \mathbb{F})$, where $\mathbb{F} = \mathbb{R}$ or $C$.*

**Remark**
*The dual space is a Banach space.*

**Remark**
*$T : X \to Y$ is bounded and linear. Then*

$$\|T\| = \inf\{c \in [0,\infty) \mid \|Tx\|_Y \le c\,\|x\|_X \text{ for all } x \in X\}.$$

**Example**
*Let $X = C([0,1])$ with the supremum norm and $Y = \mathbb{R}$ with the usual norm. For $g \in X$, $g(t) \neq 0$ on $[0,1]$, define $Tg : X \to \mathbb{R}$ by*

$$Tg(f) = \int_0^1 f(t)g(t)dt.$$

*Now for $\|f\|_\infty \le 1$,*

$$\begin{aligned} |Tg(f)| &= \left|\int_0^1 f(t)g(t)dt\right| \le \int_0^1 |f(t)g(t)|\,dt \le \int_0^1 |g(t)| \sup_{[0,1]} |f(t)|\,dt \\ &= \|f\|_\infty \int_0^1 |g(t)|\,dt \le \int_0^1 |g(t)|\,dt. \end{aligned}$$

*Take $f = g/|g|$,*

$$|Tgf| = \left|\int_0^1 \frac{g^2(t)}{|g(t)|}dt\right| = \int_0^1 |g(t)|\,dt. \implies \|Tg\| = \int_0^1 |g(t)|\,dt.$$

**Example**
*Consider $X = Y = C([0,1])$ with the supremum norm. Define $T : C^1([0,1]) \to Y$ by $Tf = f'$. Then consider the sequnce $f_n(x) = e^{-n(x-1/2)^2}$, $f_n'(x) = e^{-n(x-1/2)^2}(-2n(x-1/2))$. Hence $\|Tf_n\|/\|f_n\| = \sqrt{2n}e^{-1/2} \to \infty$ as $n \to \infty$. Thus $T$ is not bounded.*

## 2.2. $\ell^p$ Space

**Definition 2.11**
*$\ell^p = \{\{x_i\}_{i \in \mathscr{I}} \mid \|x\|_p < \infty\}$, where $\mathscr{I}$ is an countable index set and $\|x\|_p = (\sum_i |x_i|^p)^{1/p}$, $1 \le$*

$p < \infty$, *is called the* $\ell^p$ ***space***. *For* $p = \infty$, *the norm is defined as* $\|x\|_\infty = \sup_i |x_i|$.

**Definition 2.12**
$f : X \to Y$ *is called a* ***homomophism*** *if it preserves the algebraic structure. In particular, for* $X, Y$ *being vector spaces,* $f$ *is a homomorphism if* $f(cx + y) = cf(x) + f(y)$.

**Definition 2.13**
$f : X \to Y$ *is called an* ***isomorphism*** *if it is a bijective homomorphism.*

**Definition 2.14**
$f : X \to Y$ *is called an* ***isometry*** *if* $\|f(x)\|_Y = \|x\|_X$ *for all* $x \in X$.

**Example**
*A rightward shift operator* $S_R : \ell^p(\mathbb{N}) \to \ell^p(\mathbb{N})$ *is not an isomorphism, but* $S_R : \ell^p(\mathbb{Z}) \to \ell^p(\mathbb{Z})$ *is.*

**Lemma 2.15** (Young's Inequality)
*Let* $1 < p, p' < \infty$ *with* $\frac{1}{p} + \frac{1}{p'} = 1$. *Then for all* $a, b \geq 0$,

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

*Furthermore, the equality holds if and only if* $a^p = b^{p'}$.

*Proof.* If $a = 0$ or $b = 0$, the inequality is trivial. Suppose $a, b > 0$. Let $t = 1/p$ and we can write

$$\log(ab) = \log(a) + \log(b) = t\log(a^p) + (1-t)\log(b^{p'}) \leq \log\left(ta^p + (1-t)b^{p'}\right)$$

by the concavity of logarithm and Jensen's inequality. Exponentiating both sides yields the desired inequality. The equality holds if and only if $a^p = b^{p'}$ by the Jensen's inequality. ∎

**Theorem 2.16** (Hölder's Inequality in $\ell^p$)
*Let* $1 \leq p, p' \leq \infty$ *with* $\frac{1}{p} + \frac{1}{p'} = 1$. *Then for all* $f \in \ell^p$ *and* $g \in \ell^{p'}$,

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}.$$

*Moreover, the equality holds if and only if* $f = cg$ *for some constant* $c$.

*Proof.* If one of $f$ or $g$ is zero, the inequality is trivial. If $p = 1$ and $p' = \infty$, $|f_i g_i| \leq \|g\|_\infty |f_i|$. Summing over $i$ yields the desired inequality. For the case $p = \infty$ and $p' = 1$ the proof is similar. Now suppose $1 < p < \infty$ and $1 < p' < \infty$. Without loss of generality, we may assume that $\|f\|_p = \|g\|_{p'} = 1$. By Young's inequality,

$$|f_i g_i| \leq \frac{|f_i|^p}{p} + \frac{|g_i|^{p'}}{p'}.$$

Thus

$$\|fg\|_1 = \sum_i |f_i g_i| \le \sum_i \frac{|f_i|^p}{p} + \sum_i \frac{|g_i|^{p'}}{p'} = \frac{1}{p}\|f\|_p^p + \frac{1}{p'}\|g\|_{p'}^{p'} = 1.$$

Hence we obtain the desired inequality. The equality holds if and only if $|f_i|^p = |g_i|^{p'}$ for all $i$ by the Young's inequality. In general, the equality holds if and only if $f = cg$ for some constant $c$ after scaling the both sides of the inequality by $c$. ∎

**Remark**
*We call $p'$ the **conjugate exponent** of $p$ for $1/p + 1/p' = 1$.*

**Theorem 2.17** (Minkowski's Inequality in $\ell^p$)
*Let $1 \le p \le \infty$. Then for all $f, g \in \ell^p$,*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

*Proof.* If $p = 1$, the inequality comes from the triangle inequality. For $1 < p < \infty$,

$$\begin{aligned}
\|f+g\|_p^p &= \sum_i |f_i + g_i||f_i + g_i|^{p-1} \\
&\le \sum_i |f_i||f_i + g_i|^{p-1} + \sum_i |g_i||f_i + g_i|^{p-1} \\
&\le \|f\|_p \left(\sum_i |f_i + g_i|^{(p-1)p'}\right)^{1/p'} + \|g\|_p \left(\sum_i |f_i + g_i|^{(p-1)p'}\right)^{1/p'} \\
&= \|f\|_p \|f+g\|_p^{p/p'} + \|g\|_p \|f+g\|_p^{p/p'}
\end{aligned}$$

by the Hölder's inequality. Rearranging the inequality yields

$$\|f+g\|_p = \|f+g\|_p^{p-p/p'} \le \|f\|_p + \|g\|_p.$$

For $p = \infty$,

$$\|f+g\|_\infty = \sup_i |f_i + g_i| \le \sup_i |f_i| + \sup_i |g_i| = \|f\|_\infty + \|g\|_\infty.$$

The proof is complete. ∎

**Remark**
*The Minkowski's inequality is exactly the triangle inequality in $\ell^p$ spaces. We can thus confirm that $\ell^p$ norms are indeed norms.*

**Theorem 2.18** (Dualities of $\ell^p$ Spaces)
*Let $1 < p < \infty$. Then $(\ell^p)' \cong \ell^{p'}$, where $p'$ is the conjugate exponent of $p$.*

*Proof.* We need to prove that there exists an isometric isomorphism $\psi : \ell^{p'} \to (\ell^p)'$ such that $\psi g f = \sum_i f_i g_i$ for all $g \in \ell^{p'}$ and $f \in \ell^p$. We show that $\psi$ is well-defined, linear, bounded, bijective, and isometric.

First, we show that $\psi$ is well-defined. For $f \in \ell^p$ and $g \in \ell^{p'}$,

$$\left|\psi g f\right| \le \sum_i |f_i g_i| \le \|f\|_p \|g\|_{p'} < \infty$$

by the Hölder's inequality. Thus $\psi g \in (\ell^p)'$ is well-defined.

Next, $\psi$ is linear since for $g_1, g_2 \in \ell^{p'}$ and $c \in \mathbb{R}$,

$$\psi(cg_1 + g_2)(f) = \sum_i f_i(cg_{1i} + g_{2i}) = c\sum_i f_i g_{1i} + \sum_i f_i g_{2i} = c\psi g_1(f) + \psi g_2(f)$$

for all $f \in \ell^p$. Hence $\psi(cg_1 + g_2) = c\psi g_1 + \psi g_2$.

Now, to show that $\psi$ is bounded,

$$\begin{aligned}\left\|\psi g\right\| &= \sup\left\{\left|\psi g f\right| \,\middle|\, \|f\|_p = 1\right\} = \sup\left\{\left|\sum_i f_i g_i\right| \,\middle|\, \|f\|_p = 1\right\} \\ &\le \sup_{\|f\|_p=1}\left\{\|g\|_{p'}\right\} \le \|g\|_{p'}.\end{aligned}$$

We see that $\left\|\psi\right\| \le 1$. Next, let $h \in (\ell^p)'$ and define $g$ by $g_i = h(e_i)$. Then

$$\|g\|_{p'} = \left(\sum_i |g_i|^{p'}\right)^{1/p'} = \left(\sum_i |h(e_i)|^{p'}\right)^{1/p'} \le \left(\sum_i \|h\|^{p'}\right)^{1/p'} = \|h\|.$$

Then $g \in \ell^{p'}$. Furthermore, for such $g$,

$$\psi g(f) = \sum_i f_i g_i = \sum_i f_i h(e_i) = h\left(\sum_i f_i e_i\right) = h(f)$$

for every $f \in \ell^p$. Hence $\psi$ is surjective and $\left\|\psi g\right\| = \|h\|$. The isometry of $\psi$ is immediate from that

$$\left\|\psi g\right\| \le \|g\|_{p'} \le \|h\| = \left\|\psi g\right\|.$$

Finally, $\psi$ is injective since otherwise there exists $g \neq 0$ such that $\psi g = 0$. Then $\|g\|_{p'} = 0$ by the isometry of $\psi$, which implies that $g = 0$, a contradiction. We conclude that $\psi$ is an isometric isomorphism and the proof is complete. ∎

## 2.3. $\mathscr{L}^p$ Space

**Definition 2.19**

*Let $(X, \mathscr{A}, \mu)$ be a measure space and $1 \le p < \infty$. The space $\mathscr{L}^p(X)$ consists of all equivalence classes of measurable functions $f : X \to \mathbb{R}$ such that*

$$\|f\|_{\mathscr{L}^p} = \left(\int_X |f|^p \, d\mu\right)^{1/p} < \infty,$$

*where $f \sim g$ if $f = g$ a.e. and the norm is defined on a representative of the equivalence class.*

**Definition 2.20**
*$f : X \to \mathbb{R}$ is measurable. The* ***essential supremum*** *of $f$ on $X$ is defined as*

$$\operatorname{ess\,sup}_X f = \inf\left\{\sup_X g \,\middle|\, g = f\ \mu\text{-a.e.}\right\} = \inf\left\{c \in \mathbb{R} \,\middle|\, \mu(\{x \mid f(x) > c\}) = 0\right\}.$$

*We called $f$* ***essentially bounded*** *if $\operatorname{ess\,sup}_X f < \infty$. The space $\mathscr{L}^\infty(X)$ consists of all equivalence classes of essentially bounded measurable functions with the norm*

$$\|f\|_{\mathscr{L}^\infty} = \operatorname{ess\,sup}_X |f|.$$

**Theorem 2.21** (Hölder's Inequality in $\mathscr{L}^p$)
*Let $1 \le p, p' \le \infty$ with $\frac{1}{p} + \frac{1}{p'} = 1$. Then for all $f \in \mathscr{L}^p$ and $g \in \mathscr{L}^{p'}$,*

$$\|fg\|_1 \le \|f\|_p \, \|g\|_{p'}.$$

*Moreover, the equality holds if and only if $f = cg$ for some constant $c$.*

*Proof.* For the case $p = 1$ and $p' = \infty$, notice that

$$|fg| \le |f| \operatorname{ess\,sup}|g| \implies \|fg\|_1 = \int |fg| \, d\mu \le \int |f| \operatorname{ess\,sup}|g| \, d\mu = \|f\|_1 \, \|g\|_\infty.$$

For the case $p = \infty$ and $p' = 1$, the proof is similar. Now suppose $1 < p < \infty$ and $1 < p' < \infty$. If one of $f$ or $g$ is zero, the inequality is trivial. Without loss of generality, we may assume that $\|f\|_p = \|g\|_{p'} = 1$. By the Young's inequality,

$$|fg| \le \frac{|f|^p}{p} + \frac{|g|^{p'}}{p'}.$$

Integrating both sides yields

$$\|fg\|_1 = \int |fg| \, d\mu \le \int \frac{|f|^p}{p} d\mu + \int \frac{|g|^{p'}}{p'} d\mu = \frac{1}{p} + \frac{1}{p'} = 1.$$

Hence we obtain the desired inequality. The equality holds if and only if $|f|^p = |g|^{p'}$ a.e. by the Young's inequality. In general, the equality holds if and only if $f = cg$ a.e. for some constant $c$ after scaling the both sides of the inequality by $c$. ∎

**Theorem 2.22** (Minkowski's Inequality in $\mathscr{L}^p$)
*Let $1 \le p \le \infty$. Then for all $f, g \in \mathscr{L}^p$,*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

*Proof.* If $p = 1$, the inequality comes from the triangle inequality. For $1 < p < \infty$,

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p \, d\mu = \int |f+g|\,|f+g|^{p-1} \, d\mu \\
&\leq \int |f|\,|f+g|^{p-1} \, d\mu + \int |g|\,|f+g|^{p-1} \, d\mu \\
&\leq \|f\|_p \left( \int |f+g|^{(p-1)p'} \, d\mu \right)^{1/p'} + \|g\|_p \left( \int |f+g|^{(p-1)p'} \, d\mu \right)^{1/p'} \\
&= \|f\|_p \, \|f+g\|_p^{p/p'} + \|g\|_p \, \|f+g\|_p^{p/p'}.
\end{aligned}$$

Rearranging the inequality yields

$$\|f+g\|_p = \|f+g\|_p^{p-p/p'} \leq \|f\|_p + \|g\|_p.$$

For $p = \infty$,

$$\|f+g\|_\infty = \operatorname{ess\,sup} |f+g| \leq \operatorname{ess\,sup} |f| + \operatorname{ess\,sup} |g| = \|f\|_\infty + \|g\|_\infty.$$

The proof is complete. ■

**Theorem 2.23**
$1 \leq p \leq \infty$. *Simple functions are dense in* $\mathscr{L}^p$.

*Proof.* For $p < \infty$, consider $f \geq 0$ and $f \in \mathscr{L}^1$. There exists a sequence of simple functions $f_n \nearrow f$ a.e. Note that $|f - f_n|^p \leq |f|^p \in \mathscr{L}^1$. By Lebesgue's dominated convergence theorem, $\|f_n - f\|_p \to 0$ as $n \to \infty$. For $p = \infty$, pick an $f$ in the $f$-equivalent class such that $f$ is bounded. Then since the approximation of simple functions can be done uniformly, the result follows. ■

**Remark**
*A simple function* $s = \sum_{i=1}^n c_i \chi_{A_i} \in \mathscr{L}^p$ *must have* $\mu(A_i) < \infty$ *for every* $i$ *such that* $c_i > 0$. *Since contiuous functions can approximate simple functions, they are dense in* $\mathscr{L}^p$ *as well.*

**Lemma 2.24**
$1 \leq p < \infty$. $g_k \in \mathscr{L}^p$ *and* $\sum_k \|g_k\|_p < \infty$. *Then there exists* $f \in \mathscr{L}^p$ *such that* $\sum_k g_k = f$ *pointwise a.e. and in* $\mathscr{L}^p$.

*Proof.* Define $h_n$ and $h$ by $h_n = \sum_{k=1}^n |g_k|$ and $h = \sum_k |g_k|$. Then $h_n \nearrow h$. By Lebesgue's monotone convergence theorem,

$$\lim_{n\to\infty} \int h_n^p d\mu = \int h^p \, d\mu.$$

By Minkowski's inequality,

$$\left(\int h_n^p d\mu\right)^{1/p} = \left(\int \left(\sum_{k=1}^n |g_k|\right)^p d\mu\right)^{1/p} \le \sum_{k=1}^n \left(\int |g_k|^p \, d\mu\right)^{1/p} \le \sum_{k=1}^n \|g_k\|_p < \infty$$

for every $n$, so $h \in \mathscr{L}^p$ and $\|h\|_p \le M$ for some $M$ bounding $\sum_k \|g_k\|_p$. Now since $\sum_k g_k$ converges absolutely to some $f$ pointwisely a.e. and $|f| \le h$,

$$\left|f - \sum_{k=1}^n g_k\right|^p \le \left(|f| + \sum_{k=1}^n |g_k|\right)^p \le (2h)^p \in \mathscr{L}^1.$$

By Lebesgue's dominated convergence theorem, $\left\|f - \sum_{k=1}^n g_k\right\|_p \to 0$ as $n \to \infty$. Thus the proof is complete. ■

**Theorem 2.25** (Riesz-Fischer)
*$\mathscr{L}^p$ spaces are complete.*

*Proof.* First, we focus on the case where $1 \le p < \infty$. Let $f_k$ be a Cauchy sequence in $\mathscr{L}^p$. Take a subsequence $f_{k_j}$ such that $\left\|f_{k_{j+1}} - f_{k_j}\right\| \le 2^{-j}$. Let $g_j = f_{k_{j+1}} - f_{k_j} \in \mathscr{L}^p$ and we have $\sum_j \left\|g_j\right\|_p < \infty$. By the lemma 2.24, there exists $f \in \mathscr{L}^p$ such that $f = \sum_j g_j$ a.e. and

$$\lim_{j\to\infty} f_{k_j} = \lim_{j\to\infty} f_{k_1} + \sum_{i=1}^{j-1} g_i = f_{k_1} + f \in \mathscr{L}^p.$$

Since $f_k$ is Cauchy and a subsequence converges, the original sequence $f_k$ converges to $f_{k_1} + f \in \mathscr{L}^p$ as well. We now consider the case where $p = \infty$. Let $f_k$ be a Cauchy sequence in $\mathscr{L}^\infty$. Then for almost every $x$, $\{f_k(x)\}$ is a Cauchy sequence in $\mathbb{R}$. Thus we can define $f(x)$ as the limit of $f_k(x)$ as $k \to \infty$. On the set where $f_k(x)$ does not converge, we let $f(x)$ be zero. Then $f \in \mathscr{L}^\infty$ since $\{f_k\}$ is Cauchy and has an uniform bound except on a measure zero set. Also, for any $\epsilon > 0$, we can find $N$ such that $\left\|f_k - f_j\right\|_\infty < \epsilon$ for all $k, j \ge N$. Hence $\|f_k - f\|_\infty < \epsilon$ for all $k \ge N$. Thus $f_k \to f$ in $\mathscr{L}^\infty$. We conclude that $\mathscr{L}^p$ spaces are complete. ■

**Proposition 2.26**
$1 \le p < \infty$. $1/p + 1/p' = 1$. *Let $g \in \mathscr{L}^{p'}(X, \mu)$. Then the mapping $Tg : \mathscr{L}^p(X, \mu) \to \mathbb{R}$ defined by*

$$Tg(f) = \int_X f g d\mu$$

*is a bounded linear functional. Furthermore, $\|Tg\|_{\mathscr{L}^p \to \mathbb{R}} = \|g\|_{p'}$.*

*Proof.* We start by checking that $Tg$ is well-defined. For $f \in \mathscr{L}^p$,

$$|Tg(f)| = \left|\int f g d\mu\right| \le \int |fg| \, d\mu \le \|f\|_p \|g\|_{p'}$$

by Hölder's inequality. Thus $Tg(f) \in \mathbb{R}$. Also, we obtain that $\|Tg\|_{\mathscr{L}^p \to \mathbb{R}} \le \|g\|_{p'}$. For the linearity, let $c \in \mathbb{R}$ and $f_1, f_2 \in \mathscr{L}^p$.

$$Tg(cf_1 + f_2) = \int (cf_1 + f_2) g \, d\mu = c \int f_1 g \, d\mu + \int f_2 g \, d\mu = cTg(f_1) + Tg(f_2).$$

Lastly, to furnish the isometry, let $g \neq 0$ and define

$$f = \operatorname{sgn}(g) \left( \frac{|g|}{\|g\|_{p'}} \right)^{p'/p} \implies \int |f|^p \, d\mu = \int \left( \frac{|g|}{\|g\|_{p'}} \right)^{p'} d\mu < \infty.$$

Then $f \in \mathscr{L}^p$ and $\|f\|_p = 1$. Also,

$$Tg(f) = \int \operatorname{sgn}(g) \left( \frac{|g|}{\|g\|_{p'}} \right)^{p'/p} g \, d\mu = \|g\|_{p'}.$$

It follows that $\|Tg\|_{\mathscr{L}^p \to \mathbb{R}} = \|g\|_{p'}$. ■

**Theorem 2.27** (Riesz Representation)
*Let $(X, \mathscr{A} \mu)$ be a $\sigma$-finite measure space and $1 \le p < \infty$. Then the mapping $T : \mathscr{L}^{p'}(X, \mu) \to (\mathscr{L}^p(X, \mu))'$ defined by $Tg \in \mathscr{L}^p(X, \mu)$,*

$$Tg(f) = \int f g \, d\mu,$$

*is an isometric isomorphism.*

*Proof.* By proposition 2.26, $Tg$ is a bounded linear functional. Besides, let $c \in \mathbb{R}$ and $g_1, g_2 \in \mathscr{L}^{p'}$,

$$T(cg_1 + g_2)(f) = \int (cg_1 + g_2) f \, d\mu = c \int g_1 f \, d\mu + \int g_2 f \, d\mu = cTg_1(f) + Tg_2(f) = (cTg_1 + Tg_2)(f)$$

for all $f \in \mathscr{L}^p$. Thus $T$ is linear. It remains to show that $T$ is a bijection. We first verify that $T$ is surjective.

Consider the case where $p > 1$ and $\mu(X) < \infty$. Let $h \in (\mathscr{L}^p)'$. Define $\nu : \mathscr{A} \to \mathbb{R}$ by $\nu(A) = h(\chi_A)$. We claim that $\nu$ is a finite measure and $\nu \ll \mu$. Since

$$|\nu(A)| = \left| h(\chi(A)) \right| \le \|h\|_{\mathscr{L}^p \to \mathbb{R}} \left\| \chi_A \right\|_p = \|h\|_{\mathscr{L}^p \to \mathbb{R}} \mu(A)^{1/p},$$

we see that $\nu$ is finite since so is $\mu$. Also, if $\mu(A) = 0$, then $|\nu(A)| = 0$ and hence $\nu(A) = 0$. Thus $\nu \ll \mu$. For finite additivity, let $A_1, A_2 \in \mathscr{A}$ be disjoint.

$$\nu(A_1 \cup A_2) = h(\chi_{A_1 \cup A_2}) = h(\chi_{A_1} + \chi_{A_2}) = h(\chi_{A_1}) + h(\chi_{A_2}) = \nu(A_1) + \nu(A_2).$$

To show the $\sigma$-additivity, let $A_j \in \mathscr{A}$ be countably many disjoint sets. Put $A = \cup_j A_j$,

$A = B_n + C_n$ where $B_n = \cup_{j=1}^n A_j$ and $C_n = \cup_{j=n+1}^\infty A_j$. Then since $B_n \cap C_n = \varnothing$,

$$\nu(A) = \nu(B_n + C_n) = \nu(B_n) + \nu(C_n) = \sum_{j=1}^n \nu(A_j) + \nu(C_n)$$

for all $n$. Since $\mu(X) < \infty$, $\sum_j \mu(A_j) < \infty$ and $\mu(C_n) \to 0$ as $n \to \infty$. Thus

$$|\nu(C_n)| = |h(C_n)| \le \|h\|_{\mathscr{L}^p \to \mathbb{R}} \mu(C_n)^{1/p} \to \infty.$$

We conclude that $\nu(A) = \sum_j \nu(A_j)$ and $\nu$ is a measure.

Next, since $\nu \ll \lambda$, by the Radon-Nikodym theorem, there exists a unique $g \in \mathscr{L}^1(X,\mu)$ such that

$$h(\chi_A) = \nu(A) = \int_A g d\mu = \int_X \chi_A g d\mu = Tg(\chi_A).$$

for arbitrary $A \in \mathscr{A}$. Extend by linearity to $p$-integrable simple functions, say $s = \sum_{i=1}^n c_i \chi_{A_i}$.

$$h(s) = \sum_{i=1}^n c_i h(\chi_{A_i}) = \sum_{i=1}^n c_i \int_X \chi_{A_i} g d\mu = \int_X \sum_{i=1}^n c_i \chi_{A_i} g d\mu = \int_X s g d\mu = Tg(s).$$

For a general $f \in \mathscr{L}^p$, by separating $f = f^+ - f^-$ if necessary, we may assume that $f \ge 0$. By lemma 1.19, there exists a sequence of simple functions $s_n \nearrow f$. Then by Lebesgue's monotone convergence theorem, $\|f - s_n\|_p \to 0$. Since $h$ is a bounded linear functional, it is continuous, and hence $h(s_n) \to h(f)$ as $n \to \infty$. We obtain that

$$h(f) = \lim_{n\to\infty} h(s_n) = \lim_{n\to\infty} \int_X s_n g d\mu = \int_X f g d\mu = Tg(f)$$

for all $f \in \mathscr{L}^p$. Thus $Tg = h$. It remains to check that $g \in \mathscr{L}^{p'}$. Let

$$f_n = \begin{cases} |g|^{p'-1} \operatorname{sgn}(g) & \text{if } |g(x)|^{p'-1} \le n, \\ n \operatorname{sgn}(g) & \text{otherwise.} \end{cases}$$

Then $f_n \in \mathscr{L}^p$ and $f_n g \nearrow |g|^{p'}$.

$$|Tg(f_n)| = \left| \int f_n g d\mu \right| \le \|Tg\|_{\mathscr{L}^p \to \mathbb{R}} \|f_n\|_p.$$

Also, $f_n g = |f_n||g| \ge |f_n||f_n|^{1/(p'-1)} = |f_n|^p$ and

$$\|f_n\|_p^p = \int |f_n|^p d\mu \le \int f_n g d\mu \le \|Tg\|_{\mathscr{L}^p \to \mathbb{R}} \|f_n\|_p.$$

As a result,

$$\|g\|_{p'}^{p'} = \int |g|^{p'} d\mu = \lim_{n\to\infty} \int f_n g d\mu \le \|Tg\|_{\mathscr{L}^p \to \mathbb{R}} \|f_n\|_p < \infty.$$

Hence $g \in \mathscr{L}^{p'}$ and $T$ is indeed surjective. Furthermore, such $g$ is unique by the uniqueness of the Radon-Nikodym derivative. We also conclude that $T$ is injective.

For the case where $p = 1$ and $\mu(X) < \infty$, $p' = \infty$. We consider the same mapping $T$ with $Tg(f) = \int fgd\mu$. We claim that $g \in \mathscr{L}^\infty$. Suppose $g \notin \mathscr{L}^\infty$. Then for every $K$, the set $A_K = \{x \in X \mid |g(x)| > K\}$ has positive measure. Define $f_K = \operatorname{sgn}(g)\chi_{A_K}/\mu(A_K)$. Note that $\|f_K\|_1 = 1$. If $g \geq 0$, then

$$|Tg(f_K)| = \int f_K g d\mu > K$$

for all $K$. But $Tg$ is a bounded linear functional, which is a contradiction. Thus $g \in \mathscr{L}^\infty$.

Finally, we prove the case where $X$ is $\sigma$-finite. Write $X = \cup_n X_n$ where $\mu(X_n) < \infty$ and $X_n \subset X_{n+1}$. For every $f \in \mathscr{L}^p(X_k, \mu)$, consider $\hat{f} \in \mathscr{L}^p(X, \mu)$ defined by $\hat{f} = f$ on $X_k$ and $\hat{f} = 0$ on $X - X_k$. Then $\|f\|_{\mathscr{L}^p(X_k)} = \|f\|_{\mathscr{L}^p(X)}$. Let $h \in (\mathscr{L}^p(X))'$ and consider $h_k \in (\mathscr{L}^p(X_k))'$ by $h_k(f) = h(\hat{f})$. Then $\|h_k\| \leq \|h\|$. By the previous result, we can find a unique $g_k \in \mathscr{L}^{p'}(X_k, \mu)$ such that

$$h_k(f) = \int fg_k d\mu, \|g_k\|_{\mathscr{L}^{p'}(X_k)} \leq \|h_k\| \leq \|h\|.$$

Since $X_n \subset X_{n+1}$, for $f \in \mathscr{L}^p(X_k)$, we have $h_k(f) = h(\hat{f}) = h_{k+1}(f)$ and $g_k = g_{k+1}$ $\mu$-a.e. in $X_k$. Define $g = g_k$ on $X_k$ with $\|g\|_{\mathscr{L}^{p'}(X)} \leq \|h\|$. Let $f \in \mathscr{L}^p(X, \mu)$. Hölder's inequality implies that $fg \in \mathscr{L}^1(X, \mu)$ and

$$h(f\chi_{X_k}) = h_k(f) = \int f\chi_{X_k} g_k d\mu$$

Since $f\chi_{X_k} \leq |f|$, $f\chi_k \to f \in \mathscr{L}^p(X, \mu)$ by Lebesgue's dominated convergence theorem. Also,

$$h_k(f) = \int f\chi_{X_k} g_k d\mu \to \int fgd\mu = Tg(f)$$

by Lebesgue's dominated convergence theorem. Thus $T$ is indeed the desired isometric isomorphism. ∎

**Remark**

*$\mathscr{L}^\infty \not\cong \mathscr{L}^1$. Consider $C^\infty([-1,1])$, a subspace of $\mathscr{L}^\infty$. Define a linear functional $\delta : \mathbb{C}^\infty([-1,1]) \to \mathbb{R}$ by $\delta(f) = f(0)$. Clearly $\delta \in (\mathscr{L}^\infty)'$. Now suppose there exists $g \in \mathscr{L}^1$ such that $\delta(f) = \int_{-1}^1 fgdx$. Let $f = \chi_A$ where $A = [-1,1] - [-\epsilon, \epsilon]$ for some small $\epsilon > 0$. Then $f \in \mathscr{L}^\infty$ and by definition,*

$$0 = f(0) = \delta(f) = \int_{-1}^1 fgdx = \int_A gdx.$$

*Thus $g = 0$ a.e. and $\delta = 0$, a contradiction.*

**Definition 2.28**

*$M(X)$ is a space consisting of all finite signed measures. For $\nu \in M(X)$, the total varia-*

*tion norm of $\nu$ is defined by $\|\nu\| = \nu^+(X) + \nu^-(X)$, where $\nu^+$ and $\nu^-$ are the Hahn-Jordan decompositions of $\nu$.*

**Proposition 2.29**

*$M(X)$ with the total variation norm forms a Banach space.*

*Proof.* Clearly, $M(X)$ forms a vector space. We check that $\|\cdot\|$ is indeed a norm. For $\nu \in M(X)$, clearly $\|\nu\| \geq 0$. If $\|\nu\| = 0$, then $\nu^+(X) = \nu^-(X) = 0$, $\nu^+(A)$ and $\nu^-(A)$ are zero for all $A \in \mathscr{A}$, and hence $\nu = 0$. Conversely, if $\nu = 0$, then so are $\nu^+$ and $\nu^-$ and hence $\|\nu\| = 0$. For $c \in \mathbb{R}$,

$$\|c\nu\| = |c|\,\nu^+(X) + |c|\,\nu^-(X) = |c|(\nu^+(X) + \nu^-(X)) = |c|\,\|\nu\|.$$

Lastly, let $\nu, \mu \in M(X)$. Notice that $(\nu+\mu)^+ \leq \nu^+ + \mu^+$ and $(\nu+\mu)^- \leq \nu^- + \mu^-$. Thus

$$\|\nu + \mu\| = (\nu+\mu)^+(X) + (\nu+\mu)^-(X) \leq \nu^+(X) + \mu^+(X) + \nu^-(X) + \mu^-(X) = \|\nu\| + \|\mu\|,$$

proving that $\|\cdot\|$ is indeed a norm.

For the completeness, let $\nu_n$ be a Cauchy sequence in $M(X)$. We define a measure $\nu$ by $\nu(A) = \lim_{n\to\infty} \nu_n(A)$ for all $A \in \mathscr{A}$. We claim that the limit exists and $\nu$ is indeed a finite signed measure. Since the sequence is Cauchy, for every $\epsilon > 0$, there exists $N$ such that

$$(\nu_m - \nu_n)^+(X) + (\nu_m - \nu_n)^-(X) = \|\nu_m - \nu_n\| \leq \epsilon$$

for all $m, n \geq N$. Since both $(\nu_m - \nu_n)^+$ and $(\nu_m - \nu_n)^-$ are positive measures, we have

$$(\nu_m - \nu_n)^+(A) \leq (\nu_m - \nu_n)^+(X) \leq \epsilon, \quad \text{and} \quad (\nu_m - \nu_n)^-(A) \leq (\nu_m - \nu_n)^-(X) \leq \epsilon$$

for every $A \in \mathscr{A}$. Thus

$$|\nu_m(A) - \nu_n(A)| = \left|(\nu_m - \nu_n)^+(A) - (\nu_m - \nu_n)^-(A)\right| \leq \epsilon.$$

It follows that for any fixed $A \in \mathscr{A}$, $\nu_n(A)$ is a Cauchy sequence in $\mathbb{R}$ and hence the limit exists. Also, taking $A = X$, we see that $\nu(X)$ is finite. To show that $\nu$ is a measure, first note that $\nu(\varnothing) = 0$. For finite additivity, let $A_1, A_2 \in \mathscr{A}$ be disjoint. Then

$$\nu(A_1 \cup A_2) = \lim_{n\to\infty} \nu_n(A_1 \cup A_2) = \lim_{n\to\infty} \nu_n(A_1) + \nu_n(A_2) = \nu(A_1) + \nu(A_2).$$

For the $\sigma$-additivity, let $A_n \in \mathscr{A}$ be countably many disjoint sets. Put $A = \cup_n A_n$, $A = B_n \cup C_n$ where $B_n = \cup_{j=1}^n A_j$ and $C_n = \cup_{j=n+1}^\infty A_j$. Since $\nu(X) < \infty$, $\sum_j \nu(A_j) < \infty$ and hence $\nu(C_n) \to 0$ as $n \to \infty$. Thus

$$\nu(A) = \nu(B_n) + \nu(C_n) = \sum_{j=1}^n \nu(A_j) + \nu(C_n)$$

for every $n$ and by letting $n \to \infty$, we obtain $\nu(A) = \sum_j \nu(A_j)$. Finally, fix $n$ and let $m \to \infty$,

$$\|\nu - \nu_n\| = \lim_{m\to\infty} \|\nu_m - \nu_n\| = \lim_{m\to\infty} |\nu_m(X) - \nu_n(X)| = |\nu(X) - \nu_n(X)| \le \epsilon$$

for all $n \ge N$. Thus $\nu_n \to \nu$ in norm and $M(X)$ is complete. ■

**Definition 2.30**
*Let $f : [a,b] \to \mathbb{R}$. The **variation** of $f$ is defined by*

$$V_{\mathscr{P}}(f) = \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|,$$

*where $\mathscr{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of $[a,b]$. The **total variation** of $f$ on $[a,b]$ is defined by*

$$V(f) = \sup_{\mathscr{P}} V_{\mathscr{P}}(f).$$

**Definition 2.31**
*The **bounded variation space** $BV([a,b])$ consists of all functions $f : [a,b] \to \mathbb{R}$ such that $V(f) < \infty$. For $f \in BV([a,b])$, the **total variation norm** is defined by $\|f\|_{TV} = |f(a)| + V(f)$.*

**Proposition 2.32**
*$BV([a,b])$ with the total variation norm forms a Banach space.*

*Proof.* It clearly forms a vector space. We check that $\|\cdot\|_{TV}$ is indeed a norm. First, clearly $\|f\|_{TV} \ge 0$. If $\|f\|_{TV} = 0$, then $f(a) = 0$ and $f(t) = f(t')$ for all $t, t' \in [a,b]$. Hence $f = 0$; if $f = 0$, then $V(f) = 0$ and $f(a) = 0$ and $\|f\|_{TV} = 0$. Next, for $c \in \mathbb{R}$,

$$\|cf\|_{TV} = |cf(a)| + \sum_{i=0}^{n-1} |cf(t_{i+1}) - cf(t_i)| = |c| \left( |f(a)| + \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| \right) = |c| \, \|f\|_{TV}.$$

Lastly, let $f, g \in BV([a,b])$. Then

$$\begin{aligned}
\|f+g\|_{TV} &= \sup_{\mathscr{P}} |(f+g)(a)| + \sum_{i=0}^{n-1} |(f+g)(t_{i+1}) - (f+g)(t_i)| \\
&\le \sup_{\mathscr{P}} |f(a)| + |g(a)| + \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| + \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)| \\
&\le \sup_{\mathscr{P}} |f(a)| + \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| + \sup_{\mathscr{P}} |g(a)| + \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)| = \|f\|_{TV} + \|g\|_{TV}.
\end{aligned}$$

Thus $\|\cdot\|_{TV}$ is indeed a norm.

For the completeness, let $f_n$ be a Cauchy sequence in $BV([a,b])$. For $\epsilon > 0$, there exists $N$ such that $\|f_m - f_n\|_{TV} < \epsilon$ for all $m, n \ge N$. Given any $x \in [a,b]$, consider the partition

$\mathscr{P} = \{a < x < b\}$.

$$\begin{aligned}|f_m(x) - f_n(x)| &= |f_m(x) - f_m(a) + f_m(a) - f_n(a) + f_n(a) - f_n(x)| \\ &\le |((f_m(x) - f_n(x))) - (f_m(a) - f_n(a))| + |f_m(a) - f_n(a)| \\ &\le V(f_m - f_n) + |f_m(a) - f_n(a)| = \epsilon.\end{aligned}$$

Thus $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$ and hence converges pointwisely to, say $f(x)$. Furthermore, observe that the choice of $N$ does not depend on $x$, and thus the convergence is uniform. We claim that $f \in BV([a,b])$. Indeed, for any partition $\mathscr{P} = \{a = t_0 < \cdots < t_n = b\}$,

$$\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| \le \sum_{i=0}^{n-1} |f(t_{i+1}) - f_N(t_{i+1})| + \sum_{i=0}^{n-1} |f(t_i) - f_N(t_i)| + V(f_N).$$

Since the convergence is uniform, we can choose $N$ such that $|f(t) - f_N(t)| \le \epsilon/(2n)$. Thus

$$\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| \le \epsilon + V(f_N).$$

Since $f_N$ is of bounded variation, we see that $f \in BV([a,b])$ as well. Lastly, to show that $\|f - f_n\|_{TV} \to 0$, first note that by definition we have $|f_n(a) - f(a)| \to 0$. It remains to show that $V(f_n - f) \to 0$. For any $\epsilon > 0$, there exists $N$ such that $V_{\mathscr{P}}(f_m - f_n) < \epsilon$ for all $m, n \ge N$ and some partition $\mathscr{P}$. Taking $m \to \infty$, we obtain $V_{\mathscr{P}}(f - f_n) < \epsilon$ for all $n \ge N$. Since the partition is arbitrary, we have $V(f - f_n) < \epsilon$ for all $n \ge N$. Thus $f_n \to f$ in $BV([a,b])$ and $BV([a,b])$ is complete. ∎

**Theorem 2.33**
*$M([a,b])$ is isometrically isomorphic to $BV([a,b])$.*

*Proof.* We define the mapping $\phi : M([a,b]) \to BV([a,b])$ by

$$\rho(t) = \phi\nu(t) = \nu([a,t]).$$

First, we show that $\rho \in BV([a,b])$. For any partition $\mathscr{P} = \{a = t_0 < \cdots < t_n = b\}$,

$$\begin{aligned}\sum_{i=0}^{n-1} \left|\rho(t_{i+1}) - \rho(t_i)\right| + \left|\rho(a)\right| &= \sum_{i=0}^{n-1} |\nu([a, t_{i+1}]) - \nu([a, t_i])| + |\nu(\{a\})| \\ &= \sum_{i=0}^{n-1} |\nu((t_i, t_{i+1}])| + |\nu(\{a\})| \\ &= \sum_{i=0}^{n-1} |\nu|((t_i, t_{i+1}]) + |\nu|(\{a\}) = |\nu|([a,b]) = \|\nu\|.\end{aligned}$$

Since $\nu$ is a finite signed measure, $\rho \in BV([a,b])$. Furthermore, taking supremum over all partitions, we obtain that $\left\|\rho\right\|_{TV} = \|\nu\|$. It remains to show that $\phi$ is an isomorphism. Suppose $\nu, \mu \in M([a,b])$ and $\phi\nu = \phi\mu$. Then $\nu([a,t]) = \mu([a,t])$ for all $t \in [a,b]$. Since

$[a,t]$ generates the Borel $\sigma$-algebra on $[a,b]$, we have $\nu = \mu$. Thus $\phi$ is injective. For surjectivity, let $\rho \in BV([a,b])$. Consider the signed measure $\nu$ defined by $\nu([a,t]) = \rho(t)$ and $\nu(\varnothing) = 0$. Then $\nu$ is a finite signed measure and $\phi\nu = \rho$. The proof is complete. ■

**Lemma 2.34**
*Let $X$ be a normed vector space and $M \subset X$ be a proper subspace. Suppose $S : M \to \mathbb{R}$ is a bounded linear functional. Then for every $x \in X \setminus M$, there exists a linear $U : M' \to \mathbb{R}$ such that $\|U\|_{M' \to \mathbb{R}} = \|S\|_{M \to \mathbb{R}}$, where $M' = M + \mathbb{R}x$.*

*Proof.* Clearly $M'$ is a subspace; furthermore, $M' = M \bigoplus \mathbb{R}x$ since if $v = w + cx = w' + c'x$ for some $w, w' \in M$ and $c, c' \in \mathbb{R}$, then $(c - c')x = w - w' \in M$. Since $x \notin M$, this implies that $c = c'$, $w = w'$ and hence the representation is unique.

Now we can define $U$ on $M'$ by $U(w + cx) = Sw + c\lambda$ for any $w + cx \in M'$ and some $\lambda \in \mathbb{R}$ to be determined. To make $U$ have the same norm as $U$, we need to find $\lambda$ such that $|Sw + c\lambda| \le \|S\| \, \|w + cx\|$ holds for all $w \in M$ and $c \in \mathbb{R}$. Clearly if $c = 0$, the inequality is already satisfied. For $c \neq 0$, by deviding both sides by $|c|$, we see that the condition is equivalent to $|Sw + \lambda| \le \|S\| \, \|w + x\|$ for all $w \in M$. Now for any $w, v \in M$,

$$Sw - Sv = S(w - v) \le |S(w - v)| \le \|S\| \, \|w - v\| = \|S\| \, \|w + x - (v + x)\| \le \|S\| \, (\|w + x\| + \|v + x\|).$$

Thus

$$Sw - \|S\| \, \|w + x\| \le Sv + \|S\| \, \|v + x\| .$$

Fix $v$ and taking supremum over all $w \in M$ on the left,

$$\sup_{w \in M} Sw - \|S\| \, \|w + x\| \le Sv + \|S\| \, \|v + x\| .$$

Taking infimum over all $v \in M$ on the right,

$$\sup_{w \in M} Sw - \|S\| \, \|w + x\| \le \inf_{v \in M} Sv + \|S\| \, \|v + x\| .$$

Hence there exists $\lambda \in \mathbb{R}$ such that

$$S(w) - \|S\| \, \|w + x\| \le -\lambda \le S(w) + \|S\| \, \|w + x\|$$

for all $w \in M$. Picking this $\lambda$, we see that

$$|Sw + \lambda| \le \|S\| \, \|w + x\|$$

as desired. Thus $U$ is a bounded linear functional on $M'$ with $\|U\|_{M' \to \mathbb{R}} = \|S\|_{M \to \mathbb{R}}$. Also, on $M$, $U = S$ and hence $U$ is an extension of $S$. ■

**Theorem 2.35** (Hahn-Banach)
*Let $X$ be a normed vector space and $M \subset X$ be a subspace. Suppose $S : M \to \mathbb{R}$ is a bounded*

*linear functional on $M$. Then there exists a bounded linear functional $T: X \to \mathbb{R}$ such that $T|_M = S$ and $\|T\|_{X\to\mathbb{R}} = \|S\|_{M\to\mathbb{R}}$.*

*Proof.* The proof relies on Zorn's lemma.[1] We start by constructing a partial order space. Let $(P, \preceq)$ be a partial order space with

$$P = \{(U,Y) \mid M \subset Y \subset X, Y \text{ is a subspace of } X, U \text{ is a bounded extension of } S \text{ on } V\}$$

and the partial order: $(U_1,Y_1) \preceq (U_2,Y_2)$ if $Y_1 \subset Y_2$ and $U_2$ is a bounded extension of $U_1$ on $Y_2$. Clearly the pair indeed forms a partial order space. We now check the assumptions of Zorn's lemma. Let $C = \{(U_\alpha, Y_\alpha) \mid \alpha \in A\}$ with an arbitrary index set $A$ be a chain in $P$. Put $Y = \cup_{\alpha\in A} Y_\alpha$. We claim that $Y$ is a subspace of $X$. Indeed, for $y_1, y_2 \in Y$ and $c_1, c_2 \in \mathbb{R}$, there exist $\alpha_1, \alpha_2 \in A$ such that $y_1 \in Y_{\alpha_1}$ and $y_2 \in Y_{\alpha_2}$. Since $Y$ is a chain, one of them is a subspace of the other, say $Y_{\alpha_1}$ is a subspace of $Y_{\alpha_2}$. Then $y_1, y_2 \in Y_{\alpha_2}$ and hence $c_1y_1 + c_2y_2 \in Y_2 \subset Y$. Thus $Y$ is a subspace.

Next we need to define a bounded linear functional $U$ on $Y$ so that $U$ is a bounded extension of $S$ on $Y$. For $y \in Y$, we can find an $\alpha \in A$ such that $y \in Y_\alpha$ and set $U(y) = U_\alpha(y)$. Such $U$ is well-defined since if $\alpha_1$ and $\alpha_2$ are two indices satisfying $y \in Y_{\alpha_1} \cap Y_{\alpha_2}$, then $U_{\alpha_1}(y) = U_{\alpha_2}(y)$ since one of them is an extension of the other. Also, $U$ is linear since $U_\alpha$ is linear for every $\alpha \in A$. Lastly, $U$ is a bounded extension of $U_\alpha$ on $Y$ for any $\alpha \in A$ because every $U_{\alpha'}$ with $(U_\alpha, Y_\alpha) \preceq (U_{\alpha'}, Y_{\alpha'})$ is a bounded extension of $U_\alpha$. We conclude that $(U,Y) \in P$ is an upper bound of $C$.

By Zorn's lemma, there exists a maximal element $(T,Z) \in P$. We claim that $Z = X$. Suppose $Z \subsetneq X$. Then there exists $x \in X \setminus Z$ and also a bounded extension $T'$ of $T$ on $Z + \mathbb{R}x \supsetneq Z$ by lemma 2.34. But then $(T', Z+\mathbb{R}x) \in P$ and $(T,Z) \preceq (T', Z+\mathbb{R}x)$, contradicting the maximality of $(T,Z)$. Thus $Z = X$ and $T$ is a bounded extension of $S$ on $X$. ■

**Theorem 2.36** (Riesz Representation of $C([a,b])$)
*$C([a,b])' \cong BV([a,b]) \cong M([a,b])$ isometrically.*

*Proof.* In theorem 2.33, we have shown that $M([a,b]) \cong BV([a,b])$. We are going to show this by constructing an isometric isomorphism between $C([a,b])'$ and $BV([a,b])$.

Let $X = C([a,b])$ and $\ell \in X'$. $\ell : X \to \mathbb{R}$ is a bounded linear functional. We need to find a $\nu \in M([a,b])$ such that

$$\ell(f) = \int_{[a,b]} f \, d\nu$$

for $f \in C([a,b])$. Let $Y = B([a,b]) = \{f : [a,b] \to \mathbb{R} \mid f \text{ is bounded}\}$. By Hahn-Banach theorem, there exists a bounded linear extension $L : Y \to \mathbb{R}$ of $\ell$. Now if $f = \chi_{[a,t]} \in Y$, then

$$L(f) = \int_{[a,b]} \chi_{[a,t]} d\nu = \nu([a,t]) = \rho(t).$$

[1] Zorn's lemma states that if every chain in a partially ordered set has an upper bound, then the set has a maximal element. It is a direct consequence of the axiom of choice.

We claim that $\rho \in BV([a,b])$. For any partition $\mathscr{P} = \{a = t_0 < \cdots < t_n = b\}$,

$$
\begin{aligned}
V_{\mathscr{P}}(\rho) &= \sum_{i=0}^{n-1} \left|\rho(t_{i+1}) - \rho(t_i)\right| = \sum_{i=0}^{n-1} \left|L(\chi_{[a,t_{i+1}]}) - L(\chi_{[a,t_i]})\right| \\
&= \sum_{i=0}^{n-1} L(\chi_{(t_i,t_{i+1}]}) s_i = L\left(\sum_{i=0}^{n-1} \chi_{(t_i,t_{i+1}]} s_i\right) \le \|L\| \left\|\sum_{i=0}^{n-1} \chi_{(t_i,t_{i+1}]} s_i\right\|_\infty \le \|L\|
\end{aligned}
$$

by letting $s_i = \operatorname{sgn}(\rho(t_{i+1}) - \rho(t_i))$. Thus $\rho \in BV([a,b])$ and $\|\rho\|_{TV} \le \|L\| = \|\ell\|$. To extend to $f \in C([a,b])$ so that

$$
\ell(f) = L(f) = \int_{[a,b]} f \, d\nu,
$$

we first note that by our established result, $f = \chi_{[a,t]} \in Y$ holds. By linearity so does simple functions. For $f \in C([a,b])$, consider

$$
h_{\mathscr{P}}(t) = f(a) + \sum_{i=0}^{n-1} f(t_i) \chi_{(t_i,t_{i+1}]}(t).
$$

Since $L$ is continuous and $h_{\mathscr{P}} \to f$ uniformly as $\|\mathscr{P}\| \to 0$, we have

$$
L(f) = \lim_{\|\mathscr{P}\| \to 0} L(h_{\mathscr{P}}) = \int_a^b f \, d\rho.
$$

$L$ is an extension of $\ell$ and hence

$$
\ell(f) = \int_a^b f \, d\rho = f(a)\rho(a) + \int_a^b f \, d\rho.
$$

Finally, we claim that $\|\ell\| \le \|\rho\|_{TV} \le \|L\| = \|\ell\|$. Take $f \in X$.

$$
|\ell(f)| = \left|\int_a^b f \, d\rho\right| \le \|f\|_\infty \|\rho\|_{TV} \le \|f\|_\infty \|L\| = \|\ell\| \|f\|_\infty .
$$

Hence $\|\ell\| \le \|\rho\|_{TV} \le \|L\| = \|\ell\|$. It follows that the mapping $\ell \mapsto \rho$ is isometric. Conversely, if $\rho \in BV([a,b])$, define

$$
\ell_\rho(f) = f(a)\rho(a) + \int_a^b f \, d\rho.
$$

We need to check that $\ell_\rho$ is linear and $\|\rho\|_{TV} \le \|\ell\| \le \|\rho\|_{TV}$. $\ell_\rho$ has an extension $L_\rho : Y \to \mathbb{R}$. Define $\lambda(t) = L_\rho(\chi_{[a,t]})$. Then $\|\rho\|_{TV} = \|\lambda\| \le \|L_\rho\| = \|\ell_\rho\|$. ■

**Remark**

*If $\ell \in C([a,b])'$, there exists $\rho \in BV([a,b])$ such that*

$$
\ell(f) = \int_a^b f \, d\rho;
$$

*if $\rho \in BV([a,b])$,*

$$\ell_\rho(f) = f(a)\rho(a) + \int_a^b f \, d\rho$$

*and $\|\ell_\rho\| = \|\rho\|_{TV}$.*

## 2.4. Weak Convergence

**Definition 2.37**

*Let $(X, \|\cdot\|)$ be a normed space. A sequence $\{x_n\}$ in $X$ is said to* ***converge weakly*** *to $x \in X$, denoted by $x_n \rightharpoonup x$, if for every $L \in X'$, $L(x_n) \to L(x)$ as $n \to \infty$.*

**Remark**

*Strong convergence implies weak convergence. If $x_n \to x$,*

$$|L(x_n) - L(x)| = |L(x_n - x)| \leq \|L\| \, \|x_n - x\| \to 0$$

*as $n \to \infty$. Thus $x_n \rightharpoonup x$. However, the converse is not true in general.*

**Example**

*Consider $\ell^2$. Note that $(\ell^2)' \cong \ell^2$. For all $L \in (\ell^2)'$, there exists $y \in \ell^2$ such that $L(x) = \sum_{n=1}^{\infty} x_n y_n$. Let $x_n = e^n$ be the sequence with $1$ at the $n$-th position and $0$ elsewhere. Then $x_n \rightharpoonup 0$ since for every $L \in (\ell^2)'$,*

$$L(x_n) = \sum_i e_i^n y_i = y_n \to 0$$

*for $y \in \ell^2$. However, $\|x_n\|_{\ell^2} = 1$ for every $n$ and thus $x_n \not\to 0$.*

**Example**

*Consider $X = C([0,1])$ with the supremum norm. Let*

$$x_n(t) = \begin{cases} nt & \text{if } 0 \leq t \leq 1/n, \\ 2 - nt & \text{if } 1/n \leq t \leq 2/n, \\ 0 & \text{if } 2/n \leq t \leq 1. \end{cases}$$

*Then $\|x_n\|_\infty = 1$ and thus $x_n \not\to 0$. Instead, we have $x_n \rightharpoonup 0$. Assume not, then we can find $T \in X'$ and a subsequence $\{x_{n_k}\}$ such that $\left|T(x_{n_k})\right| \geq \delta > 0$. For simplicity, we consider the case $T(x_{n_k}) \geq \delta$, but the other case is similar. Since $T \in X'$, $\left|T(x_{n_k})\right| \leq \|T\|_{X \to \mathbb{R}} \left\|x_{n_k}\right\|_\infty$. Let $y_K = \sum_{k=1}^{K} x_{n_k}$. Then $T(y_K) = \sum_{k=1}^{K} T(x_{n_k}) \geq K\delta$ and $T(y_K) \leq \|T\|_{X \to \mathbb{R}} \|y_K\|_\infty$. This implies that $y_K$ cannot be bounded. Now consider $x_{n_k}$ with $n_{k+1} \geq 2n_k$. For $t \in [0, 1/n_K]$, $x_{n_k}(t) = n_k t$.*

$$y_K(t) = \sum_{k=1}^{K} n_k t \leq \sum_{k=1}^{K} n_k / n_K \leq 1 + \sum_{k=1}^{K} 2^{K-k} \leq 1 + \sum_k 2^{-k} = 2.$$

*For $t \in [1/n_K, 1/n_{K-1}]$,*

$$y_K(t) = \sum_{k=1}^{K} x_{n_k}(t) \le 1 + \sum_{k=1}^{K-1} n_k t \le 1 + \frac{1}{n_{K-1}} \sum_{k=1}^{K-1} n_k \le 1 + 1 + \sum_k 2^{-k} = 3.$$

*On $[1/n_K, 1/n_{K-1}]$, we have $\|y_K\| \le 3$. Thus $\delta K \le \|T\|_{X\to\mathbb{R}} \|y_K\|_\infty \le 3\|T\|_{X\to\mathbb{R}}$, which is impossible for sufficiently large $K$. Hence $x_n \rightharpoonup 0$.*

**Proposition 2.38**
*$(X, \|\cdot\|_X)$ is a normed space and $x_n \in X$. If $\|x_n\|_X \le C$ for all $n \in \mathbb{N}$ and $L(x_n) \to L(x)$ for all $L \in A \subset X'$, where $A$ is dense in $X'$, then $x_n \rightharpoonup x$ in $X$.*

*Proof.* Let $\epsilon > 0$ be given. $A$ is dense in $X'$. For $T \in X'$, there is an $L \in A$ such that $\|T - L\|_{X'\to\mathbb{R}} \le \epsilon$. Also, there exists $N$ such that $|L(x_n) - L(x)| \le \epsilon$ for all $n \ge N$. Then

$$\begin{aligned}|T(x_n) - T(x)| &\le |T(x_n) - L(x_n)| + |L(x_n) - L(x)| + |L(x) - T(x)| \\ &\le \|T - L\|_{X'\to\mathbb{R}}(\|x_n\|_X + \|x\|_X) + |L(x_n) - L(x)| \le 2C\epsilon + \epsilon\end{aligned}$$

for all $n \le N$. Since $\epsilon$ is arbitrary, $x_n \rightharpoonup x$. ■

**Definition 2.39**
*A space $X$ is called a **Baire space** if for any sequence of open dense subsets $\{E_n\}$, $\cap_n E_n$ is dense in $X$.*

**Theorem 2.40** (Baire Category Theorem)
*A complete metric space is a Baire space.*

*Proof.* Let $X$ be a complete metric space and $\{E_n\}$ be a sequence of open dense subsets in $X$. Put $E = \cap_n E_n$. We want to show that any nonempty open set $G \subset X$ intersects $E$.

$E_1$ is dense in $X$ so $G \cap E_1$ is nonempty. Then there exists $x_1 \in E_1 \cap G$. Note that $E_1 \cap G$ is open; there exists $1 > \delta_1 > 0$ such that $B_{\delta_1}(x_1) \subset E_1 \cap G$. By shrinking $\delta_1$, we can have $\overline{B_{\delta_1}(x_1)} \subset E_1 \cap G$. Now since $E_2$ is dense in $X$, there exists $x_2 \in E_2 \cap B_{\delta_1}(x_1)$ and also a $1/2 > \delta_2 > 0$ such that $\overline{B_{\delta_2}(x_2)} \subset E_2 \cap B_{\delta_1}(x_1)$. Continue this process, we obtain a sequence $\{x_n\}$ and $\delta_n \le 1/n$ such that $\overline{B_{\delta_n}(x_n)} \subset E_n \cap B_{\delta_{n-1}}(x_{n-1})$.

For every $m, n \ge N$, we have $x_n \in B_{\delta_n}(x_n) \subset \cdots \subset B_{\delta_N}(x_N)$ and $x_m \in B_{\delta_m}(x_m) \subset \cdots \subset B_{\delta_N}(x_N)$ by construction. Hence $d(x_n, x_m) \le 2\delta_N \le 2/N$ and $\{x_n\}$ is a Cauchy sequence. Since $X$ is complete, $\{x_n\}$ converges to some $x \in X$. We claim that $x \in E \cap G$. Clearly $x \in G$. By construction $x_m \in \overline{B_{\delta_n}(x_n)}$ for all $m \ge n$. Thus $x \in B_{\delta_n}(x_m) \subset E_N$ for $m \ge n \ge N$. We see that $x \in \cap_n E_n$. Notice that $G$ is arbitrary, so $E$ is dense in $X$, proving that $X$ is a Baire space. ■

**Theorem 2.41** (Uniform Boundedness Principle I)
*$X$ is a complete metric space. $f_\alpha : X \to \mathbb{R}$ is continuous for every $\alpha \in A$, where $A$ is an index*

*set. If for every $x \in X$, there exists $M(x) < \infty$ such that*

$$\sup_{\alpha \in A} |f_\alpha(x)| \le M(x),$$

*then there exists an open $G$ and a constant $C < \infty$ such that*

$$\sup_{\alpha \in A} |f_\alpha(x)| \le C$$

*for all $x \in G$.*

*Proof.* By Baire Category Theorem, $X$ is a Baire space. For each $n$, let

$$F_n = \left\{ x \in X \,\middle|\, \sup_{\alpha \in A} |f_\alpha(x)| \le n \right\}.$$

We claim that $F_n$ is closed and $X = \cup_n F_n$. Indeed, set $x_k \to x \in X$, where $x_k \in F_n$ for all $k$. For any $\alpha \in A$, $|f_\alpha(x_k)| \le n$ for all $k$ and by continuity of $f_\alpha$,

$$|f_\alpha(x)| = \lim_{k \to \infty} |f_\alpha(x_k)| \le n.$$

Hence $x \in F_n$ and $F_n$ is closed. Next, for any $x \in X$, take $N \ge M(x)$. Then $x \in F_N \subset \cup_n F_n$. This shows that $X = \cup_n F_n$.

Finally, observe that $F_n$ cannot have empty interiors for all $n$. Otherwise, $\varnothing = X^c = (\cup_n F_n)^c = \cap F_n^c \ne \varnothing$ since $F_n^c$ are open dense subsets of $X$, which is absurd. Hence there is some $n$ such that $F_n$ has nonempty interior, say $G \subset F_n$. Then $\sup_{\alpha \in A} |f_\alpha(x)| \le n$ for all $x \in G$ as desired. ∎

**Definition 2.42**
*A function $f : X \to \mathbb{R}$ is said to be **sub-additive** if $f(x+y) \le f(x) + f(y)$ for all $x, y \in X$.*

**Theorem 2.43** (Uniform Boundedness Principle II)
*$X$ is a Banach space. $\alpha \in A$ is an arbitrary index set. $f_\alpha : X \to \mathbb{R}$ are continuous, sub-additive and satisfy $f_\alpha(cx) = |c| f_\alpha(x)$ for all $x \in X$ and $c \in \mathbb{R}$. If for every $x \in X$, there exists $M(x) < \infty$ such that*

$$\sup_{\alpha \in A} |f_\alpha(x)| \le M(x),$$

*then there exists a constant $C < \infty$ such that*

$$\sup_{\alpha \in A} |f_\alpha(x)| \le C \|x\|_X$$

*for all $x \in X$.*

*Proof.* By theorem 2.41, there exists an open $G$ and a constant $C < \infty$ such that

$$\sup_{\alpha \in A} |f_\alpha(x)| \le C$$

for all $x \in G$. The proof will be complete if we can extend $G$ to $X$. Since $G$ is open, there exists $r > 0$ such that $B_r(z) \subset G$ for all $z \in G$. For any $x \in B_r(z)$, $\sup_{\alpha\in A} |f_\alpha(x)| \le C$ and hence $\sup_{\alpha\in A} |f_\alpha(z+y)| \le C$ for all $y \in B_r(0)$. Take $y$ with $\|y\| \le r/2$. Then

$$-2C \le f_\alpha(y+z) - f_\alpha(z) \le f_\alpha(y) \le f_\alpha(y+z) + f_\alpha(-z) = f_\alpha(y+z) + f_\alpha(z) \le 2C.$$

Hence $|f_\alpha(y)| \le 2C$ for all $y$ with $\|y\| \le r/2$. Take $x \in X$.

$$|f_\alpha(x)| = \left| f_\alpha\left( \frac{x}{\|x\|} \frac{r}{2} \frac{2}{r} \|x\| \right) \right| = \frac{2}{r} \|x\| \, |f_\alpha(y)| \le \frac{4C}{r} \|x\|.$$

Thus

$$\sup_{\alpha\in A} |f_\alpha(x)| \le \frac{4C}{r} \|x\|$$

for all $x \in X$. ■

**Corollary 2.44**

*$X$ is a Banach space. $L_\alpha \in X'$ and $\alpha \in A$. If for every $x \in X$, there exists $M(x) < \infty$ such that $\sup_{\alpha\in A} |L_\alpha(x)| \le M(x)$, then there exists a constant $C < \infty$ such that $\sup_{\alpha\in A} \|L_\alpha\| \le C$.*

*Proof.* Apply theorem 2.43 to $f_\alpha(x) = |L_\alpha(x)|$. First, $L_\alpha$ is linear and the sub-linearity follows from the triangle inequality. Next, $|L_\alpha(cx)| = |c|\,|L_\alpha(x)|$ for all $c \in \mathbb{R}$. Also, $L_\alpha \in X'$ implies that $f_\alpha$ is continuous. The conclusion follows from theorem 2.43. ■

**Corollary 2.45**

*$X$ is a normed space. $x_\alpha \in X$ for all $\alpha \in A$ with the property that for every $L \in X'$, there is $M(L) < \infty$ such that $\sup_\alpha |L(x_\alpha)| \le M(L)$ and $(X', \|\cdot\|_{X\to\mathbb{R}})$ is a Banach space. Then there exists $C < \infty$ such that $\|x_\alpha\|_X \le C$ for all $\alpha \in A$.*

*Proof.* Apply the theorem 2.43 to $f_\alpha(L) = |L(x_\alpha)|$. First, for $L, T \in X'$,

$$f_\alpha(L+T) = |L(x_\alpha) + T(x_\alpha)| \le |L(x_\alpha)| + |T(x_\alpha)| = f_\alpha(L) + f_\alpha(T).$$

Next, for $c \in \mathbb{R}$,

$$f_\alpha(cL) = |cL(x_\alpha)| = |c|\,|L(x_\alpha)| = |c|\, f_\alpha(L).$$

Finally, to verify that $f_\alpha$ is continuous, note that for $L_n \to L$ in $X'$,

$$|f_\alpha(L_n) - f_\alpha(L)| = |L_n(x_\alpha) - L(x_\alpha)| \le \|L_n - L\|_{X'\to\mathbb{R}} \|x_\alpha\|_X \to 0$$

for each $\alpha \in A$. The conclusion follows from theorem 2.43. ■

**Corollary 2.46**

*$X$ is a normed space and $x_n \in X$ with $x_n \rightharpoonup x$ in $X$. Then there exists $C < \infty$ such that $\|x_n\|_X \le C$ for all $n$.*

*Proof.* This is a direct consequence of corollary 2.45 with $A = \mathbb{N}$. ■

**Proposition 2.47**
*Let $f_n \in \mathscr{L}^p(X,\mu)$ and $1 \le p < \infty$. Then $f_n \rightharpoonup f \in \mathscr{L}^p$ if*

$$\lim_{n\to\infty} \int f_n g d\mu = \int f g d\mu$$

*for all $g \in \mathscr{L}^{p'}(X,\mu)$ and some $f$ in $\mathscr{L}^p$ where $p'$ is the conjugate exponent of $p$.*

*Proof.* By the assumption and Riesz representation theorem, for every $T \in (\mathscr{L}^p)'$, there exists a unique $g \in \mathscr{L}^{p'}$ such that

$$T(f_n) = \int f_n g d\mu \to \int f g d\mu = T(f).$$

Hence $f_n \rightharpoonup f$. ■

**Proposition 2.48**
*$f_n \in \mathscr{L}^p(X,\mu)$ and $1 \le p < \infty$. If $f_n \rightharpoonup f$ in $\mathscr{L}^p$, then $f_n$ is bounded and*

$$\|f_n\|_p \le \liminf_{n\to\infty} \|f_n\|_p.$$

*Proof.* Consider the function

$$g = \frac{|f|^{p/p'}}{\|f\|_p^{p/p'}}.$$

Note that

$$\|g\|_{p'}^{p'} = \int |g|^{p'} d\mu = \int \frac{|f|^p}{\|f\|_p^p} d\mu = 1.$$

Hence $g \in \mathscr{L}^{p'}$ with $\|g\|_{p'} = 1$. Also notice that $|g| = |f|^{p/p'}/\|f\|_p^{p/p'} = |f|^{p-1}/\|f\|_p^{p-1}$. By the weak convergence and Riesz representation theorem,

$$\|f\|_p = \int \frac{|f|^p}{\|f\|_p^{p-1}} d\mu = \int |fg| d\mu = \lim_{n\to\infty} \int |f_n g| d\mu \le \liminf_{n\to\infty} \|f_n\|_p \|g\|_{p'} = \liminf_{n\to\infty} \|f_n\|_p$$

by the Hölder inequality. Note that by corollary 2.46, $f_n$ is bounded uniformly in $n$. ■

**Proposition 2.49**
*$1 \le p < \infty$ and $1/p + 1/p' = 1$. Suppose $f_n \to f$ in $\mathscr{L}^p$ and $g_n \to g$ in $\mathscr{L}^{p'}$. Then*

$$\lim_{n\to\infty} \int f_n g_n d\mu = \int f g d\mu.$$

*Proof.* By the Hölder inequality,

$$\left|\int f_n g_n d\mu - \int f g d\mu\right| \le \left|\int f_n (g_n - g) d\mu\right| + \left|\int (f_n - f) g d\mu\right|$$
$$\le \|f_n\|_p \|g_n - g\|_{p'} + \|f_n - f\|_p \|g\|_{p'}.$$

Note that by proposition 2.48, $f_n$ converges to $f$ strongly and hence weakly. It follows that $\|f_n\|$ is bounded by some $C < \infty$. Since $g_n$ converges to $g$ and $f_n$ converges to $f$ in their respective norms, the right hand side of the inequality converges to 0 as $n \to \infty$. ■

**Remark**
*If we loosen the condition to $f_n \rightharpoonup f$ in $\mathscr{L}^p$ and $g_n \rightharpoonup g$ in $\mathscr{L}^{p'}$, then the conclusion fails.*

**Example**
*Suppose $p = p' = 2$ and $f_n(x) = \sqrt{2/\pi}\sin(nx)$ for $x \in [0,\pi]$. Then $f_n \in \mathscr{L}^2([0,\pi])$ and*

$$\int_0^\pi f_n^2 dx = \frac{2}{\pi}\int_0^\pi \sin^2(nx) dx = 1.$$

*To see that $f_n \rightharpoonup 0$, let $g \in \mathscr{L}^2([0,\pi])$. For every $\epsilon > 0$, there is a step function $\phi$ such that $\|g - \phi\|_2 < \epsilon$. Note that every step function is a finite linear combination of characteristic functions of intervals. Hence it suffices to show that $f_n \chi_I$ can be arbitrary small for $n$ sufficiently large. On every interval,*

$$\left|\int_I \sin(nx) dx\right| \le \int_0^{\pi/n} \sin(nx) dx = \frac{2}{n} \to 0$$

*as $n \to \infty$. Thus $f_n \rightharpoonup 0$ in $\mathscr{L}^2([0,\pi])$. However, $f_n$ does not converge to 0 strongly in $\mathscr{L}^2([0,\pi])$ since $\|f_n\|_2 = 1 \neq 0$ for all $n$.*

**Proposition 2.50**
*$1 \le p < \infty$. Let $f_n \in \mathscr{L}^p(X,\mu)$ be a bounded sequence of functions. Then $f_n \rightharpoonup f$ in $\mathscr{L}^p$ if and only if*

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A f d\mu$$

*for all $A \in \mathscr{A}$ when $p = 1$ and for $A$ with finite measure when $p > 1$.*

*Proof.*

$$\begin{aligned} f_n \rightharpoonup f &\iff \int f_n g d\mu \to \int f g d\mu \text{ for all } g \in \mathscr{L}^{p'} \\ &\iff \int_A f_n s d\mu \to \int_A f s d\mu \text{ for all simple } s \in \mathscr{L}^{p'} \\ &\iff \int_A f_n d\mu = \int f_n \chi_A d\mu \to \int f \chi_A d\mu = \int_A f d\mu \end{aligned}$$

for $A \in \mathscr{A}$ such that $\chi_A \in \mathscr{L}^{p'}$. If $p = 1$, then $A$ can be taken to be any $A \in \mathscr{A}$; if $p > 1$, then $A$ must have finite measure. ■

**Proposition 2.51**

*$1 < p < \infty$. Let $f_n \in \mathscr{L}^p(X, \mu)$ be a sequence with $\|f_n\|_p \le M$ and $f_n \to f$ pointwise a.e. Then $f_n \rightharpoonup f$ in $\mathscr{L}^p$.*

*Proof.* Since $\|f_n\|_p \le M$,

$$\int |f|^p \, d\mu = \int \liminf_{n\to\infty} |f_n|^p \, d\mu \le \liminf_{n\to\infty} \int |f_n|^p \, d\mu = M^p$$

by Fatou's lemma. Hence $f \in \mathscr{L}^p$. It remains to show that the convergence is weak. By proposition 2.50, it is equivalent to show that

$$\lim_{n\to\infty} \int_A f_n d\mu = \int_A f \, d\mu$$

for all $A \in \mathscr{A}$ with $\mu(A) < \infty$. Indeed, by Egorov's theorem, for every $\epsilon > 0$, there exists $F_\epsilon \subset A$ with $\mu(A - F_\epsilon) \le \epsilon$ and $f_n \to f$ uniformly on $F_\epsilon$. Furthermore, by proposition 1.32, we can choose $F_\epsilon$ so that

$$\int_{A-F_\epsilon} |f_n - f|^p \, d\mu \le \epsilon$$

since $f_n, f \in \mathscr{L}^p$ and so does $|f_n - f|^p$. Also, let $E = \{x \in A - F_\epsilon \mid |f_n - f| > 1\}$. Then for $n$ sufficiently large,

$$\begin{aligned}
\int_A |f_n - f| \, d\mu &\le \int_{F_\epsilon} |f_n - f| \, d\mu + \int_{A-F_\epsilon} |f_n - f| \, d\mu \\
&\le \int_A \epsilon \, d\mu + \int_{A-F_\epsilon-E} |f_n - f| \, d\mu + \int_E |f_n - f| \, d\mu \\
&\le \epsilon\mu(A) + \mu(A - F_\epsilon) + \int_{A-F_\epsilon} |f_n - f|^p \, d\mu \le \epsilon\mu(A) + \epsilon + \epsilon.
\end{aligned}$$

Hence $f_n \rightharpoonup f$. ■

**Remark**

*The proposition fails for $p = 1$. Consider $f_n = n\chi_{[0,1/n]}$. Then $\|f_n\|_1 = 1$ and $f_n \to 0$ pointwise a.e. However,*

$$\int_0^1 f_n(x) dx = 1 \neq 0 = \int_0^1 0 dx.$$

*Thus $f_n$ does not converge weakly to $0$ in $\mathscr{L}^1$.*

**Theorem 2.52** (Radon-Riesz)

*$1 < p < \infty$. Then $f_n \to f$ in $\mathscr{L}^p$ if and only if $\lim_{n\to\infty} \|f_n\|_p = \|f\|_p$ and $f_n \rightharpoonup f$ in $\mathscr{L}^p$.*

*Proof.* Suppose $f_n \to f$ in $\mathscr{L}^p$. Then the strong convergence immediately implies the weak convergence. Also, note that $\|f_n\|_p \le \|f_n - f\|_p + \|f\|_p$ and thus

$$\big|\|f_n\|_p - \|f\|_p\big| \le \|f_n - f\|_p \to 0$$

by the strong convergence. Conversely, suppose that $\|f_n\|_p \to \|f\|_p$ and $f_n \rightharpoonup f$ in $\mathscr{L}^p$.

Assume $p \ge 2$. For any $y \in \mathbb{R}$, notice that $|1+y|^p \ge 1 + py + c|y|^p$ for some $c \in (0,1)$. Let $E = \{x \in X \mid f(x) = 0\}$ and apply $y = (f_n - f)/f$ on $E^c$. Then on $E^c$,

$$\left|\frac{f_n}{f}\right|^p \ge 1 + p\left(\frac{f_n - f}{f}\right) + c\left|\frac{f_n - f}{f}\right|^p$$

Thus

$$|f_n|^p \ge |f|^p + p(f_n - f)|f|^{p-1}\operatorname{sgn}(f) + c|f_n - f|^p.$$

Rearranging the inequality and integrating both sides on $E^c$ gives

$$c\int_{E^c} |f_n - f|^p \, d\mu \le \int_{E^c} |f_n|^p - |f|^p \, d\mu - p\int_{E^c} |f|^{p-1}\operatorname{sgn}(f)(f_n - f) d\mu$$

Note that as shown in the proof of proposition 2.48, $|f|^{p-1}\operatorname{sgn}(f) \in \mathscr{L}^{p'}$. By the assumptions we see that

$$\int_{E^c} |f_n - f|^p \, d\mu \to 0$$

as $n \to \infty$. On $E$, we have $f = 0$ and

$$\int_E |f_n - f|^p \, d\mu = \int_E |f_n|^p \, d\mu \to 0$$

as $n \to \infty$. Hence $f_n \to f$ in $\mathscr{L}^p$.

Assume $1 < p < 2$. Then we have the same inequality for $|z| \ge 1$, i.e.,

$$|1+z|^p \ge 1 + p|z| + c|z|^p$$

Also, for $|z| \le 1$,

$$\frac{|1+z|^p - 1 - pz}{z^2}$$

is strictly positive. Now let $E_n = \{x \in X \mid |f_n(x) - f(x)| \le |f(x)|\}$. Then by applying the same argument above on $E_n^c$, we have

$$\int_{E_n^c} |f_n - f|^p \, d\mu \le \frac{1}{c}\int_{E_n^c} |f_n|^p - |f|^p \, d\mu - \frac{p}{c}\int_{E_n^c} |f|^{p-1}\operatorname{sgn}(f)(f_n - f) d\mu$$

as $n \to \infty$. On $E_n$, replacing $z$ by $(f_n - f)/f$,

$$\left|\frac{f_n}{f}\right|^p \ge 1 + p\frac{f_n - f}{f} + c'\left(\frac{f_n - f}{f}\right)^2 \implies |f_n|^p \ge |f|^p + p(f_n - f)|f|^{p-1}\operatorname{sgn}(f) + c'|f_n - f|^2|f|^{p-2}$$

for some $c' > 0$. Thus

$$\int_{E_n} |f_n - f|^2 |f|^{p-2}\, d\mu \le \frac{1}{c'} \int_{E_n} |f_n|^p - |f|^p\, d\mu - \frac{p}{c'} \int_{E_n} |f|^{p-1} \operatorname{sgn}(f)(f_n - f) d\mu.$$

Adding up the two inequalities, we have

$$\int_{E_n^c} |f_n - f|^p\, d\mu + \int_{E_n} |f_n - f|^2 |f|^{p-2}\, d\mu \to 0$$

as $n \to \infty$ by the assumptions. Note that on $E_n$, $|f| \ge |f_n - f|$ and

$$\begin{aligned}\int_{E_n} |f_n - f|^p\, d\mu \le \int_{E_n} |f_n - f||f|^{p-1}\, d\mu &\le \left(\int_{E_n} |f_n - f|^2 |f|^{p-2}\, d\mu\right)^{1/2} \left(\int_{E_n} |f|^p\, d\mu\right)^{1/2} \\ &\le \left(\int_{E_n} |f_n - f|^2 |f|^{p-2}\, d\mu\right)^{1/2} \|f\|_p^{p/2} \to 0.\end{aligned}$$

Hence $f_n \to f$ in $\mathscr{L}^p$. We conclude that $f_n \to f$ strongly in $\mathscr{L}^p$ if and only if $f_n \rightharpoonup f$ in $\mathscr{L}^p$ and $\|f_n\|_p \to \|f\|_p$. ■

**Remark**
*Radon-Riesz theorem fails for $p = 1$. Consider $f_n(x) = 1 + \sin(nx)$ on $X = [-\pi, \pi]$. Then for every $g \in \mathscr{L}^\infty$,*

$$\int (f_n - 1) g\, d\mu \le \int \sin(nx) g\, d\mu \to 0$$

*by the step function approximation argument. Also, $\|f_n\|_1 = 2\pi$ for all $n$ and hence converges to $\|1\|_1 = 2\pi$. However, $f_n$ does not converge to 1 in $\mathscr{L}^1$ since*

$$\int_{-\pi}^{\pi} |f_n - 1|\, d\mu = \int_{-\pi}^{\pi} |\sin(nx)|\, d\mu = 2n \int_0^{\frac{\pi}{2n}} \sin(nx) dx = 2$$

*for all $n$.*

## 2.5. Open Mapping Theorem and Closed Graph Theorem

**Proposition 2.53**
*If $X$ is a Baire space and $F_n$ is a sequence of closed sets in $X$ such that $\bigcup_{n=1}^\infty F_n = X$, then there exists some $n$ and a nonempty open set $G$ such that $G \subseteq F_n$.*

*Proof.* Let $G_n = F_n^c$ be open sets in $X$. Then $\bigcap_{n=1}^\infty G_n = \bigcap_{n=1}^\infty F_n^c = \left(\bigcup_{n=1}^\infty F_n\right)^c = \varnothing$. By the Baire category theorem, at least one of the $G_n$ is not dense in $X$. Thus there is some $x \in G^c$ and an open neighborhood $U$ of $x$ such that $U \bigcap G_n = \varnothing$. This implies $U \subseteq F_n$. ■

**Theorem 2.54** (Open Mapping Theorem)
*Let $X$ and $Y$ be Banach spaces and $T : X \to Y$ be a bounded surjective linear map. Then for any open set $U \subset X$, $T(U)$ is open in $Y$.*

*Proof.* We first claim that for any open ball $B$ centered at 0 in $X$, $\overline{T(B)}$ contains an open neighborhood of zero in $Y$. By the surjectivity, $Y \subset T(X) = T(\bigcup_n nB) = \bigcup_n T(nB) \subset \bigcup_n \overline{T(nB)}$. By proposition 2.53, there is some $n$ such that $\overline{T(nB)}$ contains an interior point, say $y$, and some open ball $B_r(y) \subset \overline{T(nB)}$. Then for every $z \in Y$ with $\|z\| < r$, $z - y \in B_r(-y) \subset \overline{T(-nB)} = \overline{T(nB)}$ and

$$z = y + (z - y) \in y + B_r(-y) \subset \overline{T(nB)} + \overline{T(nB)} \subset \overline{T(2nB)}.$$

Deviding $z$ by $2n$ gives that $z/2n \in \overline{T(B)}$ and $B_{r/2n}(0) \subset \overline{T(B)}$.

Next, let $B$ be an unit ball. To shorten the notation, denote $r/2n$ by $\delta$ and $B_{r/2n}(0)$ by $B_\delta$. Let $y \in B_\delta$ and $c_n > 0$ be a sequence. Since $B_\delta \subset \overline{T(B)}$, $\overline{B_\delta} \subset \overline{T(B)}$. Thus for every $z \in Y$ and $\epsilon > 0$, we can find some $x \in X$ such that $\|x\| < \delta^{-1}\|z\|$ and $z \in B_\epsilon(T(x))$. Now taking $z = y$ and $\epsilon = c_1$, we can find an $x_1$ such that $\|x_1\| < \delta^{-1}\|y\|$ and $\|y - Tx_1\| < c_1$. Similarly, we can take $z = y - Tx_1$ and $\epsilon = c_2$ to find an $x_2$ such $\|x_2\| < \delta^{-1}\|y - Tx_1\| < \delta^{-1}c_1$ and $\|y - Tx_1 - Tx_2\| < c_2$. Iductively, we find a sequence $\{x_n\}$ such that $\|x_n\| < \delta^{-1}c_{n-1}$ and $\left\|y - T\left(\sum_{k=1}^n x_k\right)\right\| < c_n$. Now we choose $c_n = 2^{-n}c$ for arbitrary $c > 0$. Then

$$\left\|\sum_{k=1}^n x_k\right\| \le \sum_{k=1}^n \|x_k\| \le \frac{\|y\|}{\delta} + \sum_{k=2}^n \frac{c_{k-1}}{\delta} \le \frac{\|y\|}{\delta} + \frac{c}{\delta}\sum_{k=1}^\infty 2^{-k} = \frac{\|y\|}{\delta} + \frac{c}{\delta}.$$

Hence $\sum_n x_n$ converges in $X$ to some $x$ with $\|x\| < 1$ by making $c$ arbitrarily small. Also,

$$\left\|y - T\left(\sum_{k=1}^n x_k\right)\right\| \le c_n = 2^{-n}c \to 0.$$

Thus $Tx = y$ and $y \in T(B)$, which implies $B_\delta \subset T(B)$.

Finally, let $U$ be an open set in $X$. Then for any $y \in T(U)$, there is some $x \in U$ such that $y = Tx$. Since $U$ is open, there is some $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. By the previous claim, there is some $s > 0$ such that $B_s(0) \subset T(B_1(0))$. Multiplying both sides by $\epsilon$ gives $B_{s\epsilon}(0) \subset T(B_\epsilon(0))$. Then

$$B_{s\epsilon}(y) = y + B_{s\epsilon}(0) \subset y + T(B_\epsilon(0)) = Tx + T(B_\epsilon(0)) = T(x + B_\epsilon(0)) = T(B_\epsilon(x)) \subset T(U).$$

Thus $T(U)$ is open. This completes the proof. ∎

**Theorem 2.55** (Bounded Inverse Theorem)
*Let $X$ and $Y$ be Banach spaces and $T : X \to Y$ be a bounded linear map. If $T$ is bijective, then $T^{-1}$ is bounded.*

*Proof.* By the open mapping theorem, there is $r > 0$ such that $B_r(0) \subset T(B_r(0))$. For any $y \in Y$ with $\|y\| = r/2$, there exists $x \in B_1(0)$ such that $y = Tx$. For $z \in Y$, write

$$z = \frac{rz}{2\|z\|}\frac{2}{r}\|z\|.$$

Then since $\left\| \frac{rz}{2\|z\|} \right\| = r/2$, there is some $x \in B_1(0)$ such that $\frac{rz}{2\|z\|} = Tx$. Thus $z = \frac{2}{r} \|z\| Tx$,

$$T^{-1}z = \frac{2}{r} \|z\| x \implies \left\| T^{-1}z \right\| \le \frac{2}{r} \|z\| \|x\| .$$

Note that $\|x\| \le 1$. We see that $\left\| T^{-1} \right\|$ is bounded by $2/r$. ■

**Remark**
*The completeness in the open mapping theorem is essential. For counterexample, consider $X$ as the space of all sequences with finitely many nonzero terms equipped with the supremum norm. Define $T : X \to X$ by*

$$T(x_1, x_2, \ldots) = \left( x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots \right).$$

*Note that $X$ is not complete since the sequence $x^{(n)} = (1, 1/2, \ldots, 1/n, 0, 0, \ldots)$ converges to $(1, 1/2, \ldots)$, which does not belong to $X$. In this case $T^{-1}$ exists but is not bounded.*

**Definition 2.56**
*$X, Y$ are Banach spaces. $T : X \to Y$ is a bounded linear map. The set*

$$\Gamma(T) = \{(x, Tx) \in X \times Y \mid x \in X\}$$

*is called the **graph** of $T$. We define the norm of $x$ on the graph by*

$$\|(x, Tx)\|_\Gamma = \|x\|_X + \|Tx\|_Y .$$

*Note that $(\Gamma(T), \|\cdot\|_\Gamma)$ forms a normed space.*

**Definition 2.57**
*A linear map $T : X \to Y$ is called **closed** if its graph is a closed, i.e., for any sequence $x_n \in X$, if $x_n \to x \in X$ and $Tx_n \to y \in Y$, then $Tx = y$ and $(x, y) \in \Gamma(T)$.*

**Remark**
*If $T$ is bounded, it is closed. To see this, note that if $x_n \to x \in X$, by the continuity we have $Tx_n \to Tx \in Y$.*

**Theorem 2.58** (Closed Graph Theorem)
*Let $X$ and $Y$ be Banach spaces and $T : X \to Y$ be a linear map. If $T$ is closed, then $T$ is bounded.*

*Proof.* Observe that $\Gamma(T)$ is a Banach space with the norm $\|\cdot\|$ on $\Gamma(T)$. This follows from the closedness of $T$. Now define $S : \Gamma(T) \to X$ by $S(x, Tx) = x$. We claim that $S$ is bounded, linear and bijective. For linearity, let $(x_1, Tx_1), (x_2, Tx_2) \in \Gamma(T)$ and $c \in \mathbb{R}$.

$$S(c(x_1, Tx_1) + (x_2, Tx_2)) = S(cx_1 + x_2, cTx_1 + Tx_2) = cx_1 + x_2 = cS(x_1, Tx_1) + S(x_2, Tx_2).$$

For boundedness,

$$\|S(x,Tx)\|_X = \|x\|_X \le \|x\|_X + \|Tx\|_Y = \|(x,Tx)\|_\Gamma .$$

Thus $\|S\| \le 1$. For bijectivity, notice that

$$S(x_1,Tx_1) = S(x_2,Tx_2) \implies x_1 = S(x_1,Tx_1) = S(x_2,Tx_2) = x_2.$$

and for any $x \in X$, $(x,Tx) \in \Gamma(T)$ and $S(x,Tx) = x$. Thus $S$ is bounded, linear and bijective.

By the bounded inverse theorem, $S^{-1} : X \to \Gamma(T)$ is bounded as well. For any $x \in X$,

$$\|Tx\|_Y = \|(x,Tx)\|_\Gamma - \|x\|_X = \left\|S^{-1}x\right\|_\Gamma - \|x\|_X \le C\,\|x\|_X - \|x\|_X = (C-1)\,\|x\|_X$$

for some constant $C < \infty$. Thus $T$ is bounded. ∎

**Definition 2.59**
*Suppose $X$ is a vector space with norms $\|\cdot\|_1$ and $\|\cdot\|_2$. The norms are said to be **compatible** if $x_n \to x$ in $\|\cdot\|_1$ and $x_n \to y$ in $\|\cdot\|_2$ implies $x = y$.*

**Definition 2.60**
*Let $X$ be a vector space with norms $\|\cdot\|_1$ and $\|\cdot\|_2$. The norms are said to be **equivalent** if there are constants $c_1, c_2 > 0$ such*

$$c_1\,\|x\|_1 \le \|x\|_2 \le c_2\,\|x\|_1$$

*for all $x \in X$.*

**Proposition 2.61**
*If $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are Banach spaces. Then the norms are equivalent.*

*Proof.* By the closed graph theorem, the identity map $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is a closed linear map and thus bounded. Suppose $x_n \to x$ in $\|\cdot\|_1$. Then $x_n = Ix_n \to Ix = x$ in $\|\cdot\|_2$ by the continuity of $I$. Since $I$ is bounded, $\|x\|_2 = \|Ix\|_2 \le c_1\,\|x\|_1$ for some $c_1 > 0$. Applying the same argument exchanging the roles of $\|\cdot\|_1$ and $\|\cdot\|_2$ gives $\|x\|_1 \le c_2\,\|x\|_2$ for some $c_2 > 0$. Hence

$$\frac{1}{c_2}\,\|x\|_1 \le \|x\|_2 \le c_1\,\|x\|_1$$

and the norms are equivalent. ∎

# 3. Hilbert Space

## 3.1. Cauchy-Schwarz Inequality

**Definition 3.1**

*Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ is an inner product on $X$ if it satisfies*

*(a) $\langle cx + y, z \rangle = c\langle x, z \rangle + \langle y, z \rangle$ for all $x, y, z \in X$ and $c \in \mathbb{F}$.*

*(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in X$.*

*(c) $\langle x, x \rangle > 0$ for all $x \neq 0$.*

**Remark**

*An inner product automatically induces a norm on $X$ by $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$.*

**Definition 3.2**

*A **Hilbert space** is a complete vector space with an inner product inducing a norm that makes it a Banach space. We denote the Hilbert space by $\mathscr{H}$.*

**Remark**

*If $X$ is a vector space with an inner product but not complete, then $X$ is called a **pre-Hilbert space**.*

**Proposition 3.3** (Cauchy-Schwarz Inequality)

*For all $x, y \in \mathscr{H}$,*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

*Furthermore, equality holds if and only if $x$ and $y$ are linearly dependent.*

*Proof.* If $\langle x, y \rangle = 0$, then the inequality is trivial. Otherwise, let $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 &\leq \left\langle t\frac{|\langle x, y \rangle|}{\langle x, y \rangle}x + y, t\frac{|\langle x, y \rangle|}{\langle x, y \rangle}x + y \right\rangle \\ &= t^2 \|x\|^2 + 2t\Re\left(\frac{|\langle x, y \rangle|}{\langle x, y \rangle}\langle x, y \rangle\right) + \|y\|^2 = t^2 \|x\|^2 + 2t|\langle x, y \rangle| + \|y\|^2. \end{aligned}$$

Hence

$$4|\langle x, y \rangle|^2 - 4\|x\|^2 \|y\|^2 \leq 0 \implies |\langle x, y \rangle| \leq \|x\| \|y\|.$$

Note that if the equality holds, then

$$t^2 \|x\|^2 + 2t|\langle x, y \rangle| + \|y\|^2 = t^2 \|x\|^2 + 2t\|x\| \|y\| + \|y\|^2 = (t\|x\| + \|y\|)^2 = 0$$

by taking $t = -\|y\|/\|x\|$. But this implies that

$$t\frac{|\langle x, y \rangle|}{\langle x, y \rangle}x + y = 0$$

and so $x$ and $y$ are linearly dependent. Conversely, suppose $cx = y$. Then $|\langle x, y\rangle| = |c|\,\|y\|^2 = \|x\|\,\|y\|$. ■

**Proposition 3.4** (Parallelogram Law)
*For all $x, y \in \mathscr{H}$,*

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2.$$

*Proof.* Note that

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \|x\|^2 + 2\Re(\langle x, y\rangle) + \|y\|^2,\\ \|x-y\|^2 &= \langle x-y, x-y\rangle = \|x\|^2 - 2\Re(\langle x, y\rangle) + \|y\|^2.\end{aligned}$$

Adding the two equations gives

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2.$$

■

**Proposition 3.5**
*For all $x \in \mathscr{H}$,*

$$\|x\| = \sup_{\|y\|=1} |\langle x, y\rangle|.$$

*Proof.* By the Cauchy-Schwarz inequality,

$$|\langle x, y\rangle| \le \|x\|\,\|y\| \implies \left|\left\langle x, \frac{y}{\|y\|}\right\rangle\right| \le \|x\| \implies \sup_{\|y\|=1} |\langle x, y\rangle| \le \|x\|.$$

Taking $y = x/\|x\|$ gives the equality and $\|y\| = 1$. ■

**Theorem 3.6** (Completion of Pre-Hilbert Space)
*Let $(X, \langle\cdot,\cdot\rangle)$ be a pre-Hilbert space. Then there exists a Hilbert space $\mathscr{H}$ such that $X$ is dense in $\mathscr{H}$ and $\langle\cdot,\cdot\rangle_*$ on $\mathscr{H}$ is an extension of $\langle\cdot,\cdot\rangle$.*

*Proof.* Define $\langle x, y\rangle_* = \lim_{n\to\infty} \langle x_n, y_n\rangle$ for Cauchy sequences $\{x_n\}, \{y_n\} \subset X$ and $x, y \in \overline{X}$. We first check that $\langle\cdot,\cdot\rangle_*$ is well-defined. Note that

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x_m, y_m\rangle| &\le |\langle x_n, y_n\rangle - \langle x_n, y_m\rangle| + |\langle x_n, y_m\rangle - \langle x_m, y_m\rangle|\\ &\le \|x_n\|\,\|y_n - y_m\| + \|x_n - x_m\|\,\|y_m\| \to 0\end{aligned}$$

as $n, m \to \infty$ by the Cauchy-Schwarz inequality. Since $\mathbb{F}$ is complete, the limit exists. To see that $\langle\cdot,\cdot\rangle_*$ is independent of the choice of sequences, suppose $\{x_n^1\}, \{y_n^1\}$ and $\{x_n^2\}, \{y_n^2\}$ are two pairs of Cauchy sequences converging to $x$ and $y$ respectively. Then

$$\begin{aligned}\left|\langle x_n^1, y_n^1\rangle - \langle x_n^2, y_n^2\rangle\right| &\le \left|\langle x_n^1, y_n^1\rangle - \langle x_n^1, y_n^2\rangle\right| + \left|\langle x_n^1, y_n^2\rangle - \langle x_n^2, y_n^2\rangle\right|\\ &\le \left\|x_n^1\right\|\,\left\|y_n^1 - y_n^2\right\| + \left\|x_n^1 - x_n^2\right\|\,\left\|y_n^2\right\| \to 0.\end{aligned}$$

Hence $\langle x, y\rangle_*$ is well-defined. We now show that $\langle\cdot,\cdot\rangle_*$ is indeed an inner product on $\overline{X}$. For the linearity in the first argument, let $x, y, z \in \overline{X}$, $\{x_n\}, \{y_n\}, \{z_n\} \subset X$ be Cauchy sequences converging to $x, y, z$ respectively and $c \in \mathbb{F}$. Then

$$\begin{aligned}\langle cx + y, z\rangle_* &= \lim_{n\to\infty} \langle cx_n + y_n, z_n\rangle = \lim_{n\to\infty} c\,\langle x_n, z_n\rangle + \langle y_n, z_n\rangle \\ &= c \lim_{n\to\infty} \langle x_n, z_n\rangle + \lim_{n\to\infty} \langle y_n, z_n\rangle = c\,\langle x, z\rangle_* + \langle y, z\rangle_* .\end{aligned}$$

For the conjugate symmetry, let $x, y \in \overline{X}$ and $\{x_n\}, \{y_n\} \subset X$ be Cauchy sequences converging to $x$ and $y$ respectively. Then

$$\overline{\langle x, y\rangle_*} = \overline{\lim_{n\to\infty} \langle x_n, y_n\rangle} = \lim_{n\to\infty} \overline{\langle x_n, y_n\rangle} = \lim_{n\to\infty} \langle y_n, x_n\rangle = \langle y, x\rangle_* .$$

For the positive definiteness, let $x \in \overline{X}$, $x \neq 0$ and $\{x_n\} \subset X$ be a Cauchy sequence converging to $x$. Then

$$\langle x, x\rangle_* = \lim_{n\to\infty} \langle x_n, x_n\rangle = \lim_{n\to\infty} \|x_n\|^2 > 0.$$

Hence $\langle\cdot,\cdot\rangle_*$ is an inner product on $\overline{X}$ and induces a norm on $\overline{X}$. Lastly, for every $x, y \in \overline{X}$, pick $x_n = x$ and $y_n = y$ to see that

$$\langle x, y\rangle_* = \lim_{n\to\infty} \langle x, y\rangle = \langle x, y\rangle ,$$

which shows that $\langle\cdot,\cdot\rangle_*$ is an extension of $\langle\cdot,\cdot\rangle$. We conclude that $\mathscr{H} = \overline{X}$ forms a Hilbert space. ■

**Example**
*Let $X = C([0,1])$ with the inner product*

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx.$$

*Then $X$ is a pre-Hilbert space. To see this, set $f_n(x) = x^n$.*

$$\|f_m - f_n\|^2 = \int_0^1 (x^m - x^n)^2 dx = \frac{1}{2m+1} + \frac{2}{m+n+1} + \frac{1}{2n+1} \to 0$$

*as $m, n \to \infty$. Hence $\{f_n\}$ is Cauchy in $X$. However, $f_n$ converges to*

$$f(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1, \end{cases}$$

*which is not in $X$. Hence $X$ is not complete. But by the proposition 1.34, $X$ is dense in $\mathscr{L}^2([0,1])$ and so $X$ can be completed to a Hilbert space $\mathscr{H} = \mathscr{L}^2([0,1])$, which is complete* *by Riesz-Fischer theorem.*

**Definition 3.7**
*A set $X$ is called* ***convex*** *if for all $x, y \in X$ and $t \in [0,1]$, $tx + (1-t)y \in X$.*

**Theorem 3.8**
*Let $K \subset \mathcal{H}$ be a closed convex set. For $x \in \mathcal{H}$, define the distance from $x$ to $K$ as*

$$d(x,K) = \inf_{y \in K} \|x - y\|.$$

*Then there exists a unique $z \in K$ such that $d(x,K) = \|x - z\|$.*

*Proof.* Let $\{y_n\} \subset K$ be a sequence such that $\|y_n - x\| \to d(x,K)$. We claim that $\{y_n\}$ is Cauchy. Let $\epsilon > 0$ be given. By the parallelogram law,

$$2\left(\|x - y_n\|^2 + \|x - y_m\|^2\right) = \|2x - y_n - y_m\|^2 + \|y_n - y_m\|^2$$

Rearranging gives

$$\begin{aligned}
\frac{1}{4}\|y_n - y_m\|^2 &= \frac{1}{2}\|x - y_n\|^2 + \frac{1}{2}\|x - y_m\|^2 - \left\|x - \frac{y_n + y_m}{2}\right\|^2 \\
&\le \frac{1}{2}(d(x,K)+\epsilon)^2 + \frac{1}{2}(d(x,K)+\epsilon)^2 - d(x,K)^2 \\
&= \epsilon^2 + 2\epsilon d(x,K)
\end{aligned}$$

for all $m, n \ge N$ for some $N \in \mathbb{N}$. The inequality follows from the fact that $(y_n + y_m)/2 \in K$ by the convexity of $K$. Since $\epsilon > 0$ is arbitrary, we conclude that $\{y_n\}$ is Cauchy. By the completeness of $\mathcal{H}$, $\{y_n\}$ converges to some $z \in \mathcal{H}$. Since $K$ is closed, $z \in K$. To see the uniqueness, suppose $z_1, z_2 \in K$ are such that $\|x - z_1\| = \|x - z_2\| = d(x,K)$. Then by the parallelogram law,

$$\begin{aligned}
4d(x,K)^2 = 2\|x - z_1\|^2 + 2\|x - z_2\|^2 &= \|z_1 - z_2\|^2 + \|2x - z_1 - z_2\|^2 \\
&= \|z_1 - z_2\|^2 + 4\left\|x - \frac{z_1 + z_2}{2}\right\|^2.
\end{aligned}$$

Hence

$$\|z_1 - z_2\|^2 = 4d(x,K)^2 - 4\left\|x - \frac{z_1 + z_2}{2}\right\|^2 \le 4d(x,K)^2 - 4d(x,K)^2 = 0$$

and so $z_1 = z_2$. ∎

**Definition 3.9**
*$Y \subset \mathcal{H}$ is a closed subspace. The* ***orthogonal complement*** *of $Y$, denoted by $Y^\perp$, is defined as*

$$Y^\perp = \{x \in \mathcal{H} \mid \langle x, y \rangle = 0 \text{ for all } y \in Y\}.$$

**Proposition 3.10**
*$Y \subset \mathcal{H}$ is a closed subspace. Then*

(a) *$Y^{\perp}$ is a closed subspace.*

(b) *$\mathscr{H} = Y \oplus Y^{\perp}$.*

(c) *$\left(Y^{\perp}\right)^{\perp} = Y$.*

*Proof.* For (a), we first check that $Y^{\perp}$ is a subspace. First note that $0 \in Y^{\perp}$. Also, if $x, z \in Y^{\perp}$ and $c \in \mathbb{F}$, then

$$\langle cx + z, y \rangle = c \langle x, y \rangle + \langle z, y \rangle = 0$$

for all $y \in Y$. Hence $cx + z \in Y^{\perp}$. This shows that $Y^{\perp}$ is a subspace. To see that $Y^{\perp}$ is closed, let $\{x_n\} \subset Y^{\perp}$ be a sequnce converging to $x$. Then for all $y \in Y$,

$$|\langle x, y \rangle - \langle x_n, y \rangle| = |\langle x - x_n, y \rangle| \leq \|x - x_n\| \, \|y\| \to 0$$

by the Cauchy-Schwarz inequality. Thus $\langle \cdot, y \rangle$ is a continuous functional on $\mathscr{H}$. Therefore,

$$\langle x, y \rangle = \lim_{n \to \infty} \langle x_n, y \rangle = 0$$

for all $y \in Y$ and so $x \in Y^{\perp}$. This shows that $Y^{\perp}$ is closed.

For (b), notice that $Y$ as a closed subspace is convex. By theorem 3.8, for any $u \in \mathscr{H}$, there exists a unique $y \in Y$ such that $\|u - y\| \leq \|u - y'\|$ for all $y' \in Y$. Let $z = u - y$. We claim that $z \in Y^{\perp}$. To see this, let $y' \in Y$ and $t \in \mathbb{R}$. Then

$$\begin{aligned}
\|z\|^2 = \|u - y\|^2 &\leq \left\|u - y - ty'\right\|^2 \\
&= \|u - y\|^2 - 2t\Re\left(\langle u - y, y' \rangle\right) + t^2 \left\|y'\right\|^2 \\
&= \|z\|^2 - 2t\Re\left(\langle z, y' \rangle\right) + t^2 \left\|y'\right\|^2.
\end{aligned}$$

Rearranging gives

$$2t\Re\left(\langle z, y' \rangle\right) - t^2 \left\|y'\right\|^2 \leq 0.$$

If $y' = 0$, we have $\langle z, y' \rangle = 0$; if $y' \neq 0$, then take $t = \Re\left(\langle z, y' \rangle / \left\|y'\right\|^2\right)$. Substituting this back gives

$$0 \geq 2\frac{\left(\Re\left(\langle z, y' \rangle\right)\right)^2}{\|y'\|^2} - \frac{\left(\Re\left(\langle z, y' \rangle\right)\right)^2}{\|y'\|^2} = \frac{\left(\Re\left(\langle z, y' \rangle\right)\right)^2}{\|y'\|^2}.$$

Hence $\Re\left(\langle z, y' \rangle\right) = 0$ for all $y' \in Y$. Similarly, replacing $t$ with $it$ gives $\Im\left(\langle z, y' \rangle\right) = 0$ for all $y' \in Y$. Therefore, $\langle z, y' \rangle = 0$ for all $y' \in Y$ and so $z \in Y^{\perp}$. Since our choice of $y$ is unique, we can write $u = y + z$ uniquely for $y \in Y$ and $z \in Y^{\perp}$. This shows that $\mathscr{H} = Y \oplus Y^{\perp}$.

For (c), note that we can apply (a) and (b) to $Y^{\perp}$ and obtain that $(Y^{\perp})^{\perp}$ is a closed subspace and $\mathscr{H} = Y \oplus Y^{\perp} = \left(Y^{\perp}\right)^{\perp} \oplus Y^{\perp}$. It follows that for every $u \in \mathscr{H}$, we can write $u = y + z = x + z$ for $x \in \left(Y^{\perp}\right)^{\perp}$, $y \in Y$ and $z \in Y^{\perp}$ by the uniqueness of decomposition. This implies that $y = x$ and hence $\left(Y^{\perp}\right)^{\perp} = Y$. ∎

## 3.2. Bounded Linear Functional on $\mathscr{H}$

**Proposition 3.11**

*$\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. $T : \mathscr{H} \to \mathbb{F}$ is a nonzero bounded linear functional.*

*(a)* $\mathscr{H} = \mathrm{span}(\{w\}) \oplus \ker(T)$ *for* $w \notin \ker(T)$.

*(b) If* $S, T$ *are bounded linear functionals and* $\ker(S) = \ker(T)$, *then there exists* $c \in \mathbb{F}$ *such that* $S = cT$.

*(c)* $\ker(T)$ *is closed.*

*Proof.* For (a), since $T$ is nonzero, there is some $w$ such that $Tw \neq 0$. For $x \in \mathscr{H}$, set $\alpha = Tx/Tw$ and $u = x - \alpha w$. Then $x = \alpha w + u$ and

$$Tu = Tx - \frac{Tx}{Tw} Tw = 0.$$

Hence $u \in \ker(T)$. Also, if $v \in \mathrm{span}(\{w\}) \bigcap \ker(T)$, $v = cw$ and $Tv = 0$. Then $cTw = Tv = 0$; $c = 0$ and thus $v = 0$. Therefore, $\mathrm{span}(\{w\}) \bigcap \ker(T) = \{0\}$ and $\mathscr{H} = \mathrm{span}(\{w\}) \oplus \ker(T)$. ■