Jie Xiao

# Introductory Functional Analysis

April 17, 2005

Memorial University of Newfoundland
Department of Mathematics & Statistics
St. John's, NL A1C 5S7, Canada

# Preface

This book is based on a one-semester course in the Introductory Functional Analysis the author offered at MUN in the winter of 2005 for both undergraduate and graduate students. The prerequisites of this book are deliberately modest, and it is assumed that this will be the student's first experience with abstract mathematical reasoning. Thus the exposition of this book is careful and detailed with an emphasis on simplicity, concreteness and abstraction, and the examples are all quite elementary, requiring at most some knowledge of elementary linear algebra and real analysis.

As can be seen from the table of Contents, the students who take and finish this course should have a good understanding of normed vector spaces, Banach spaces with fixed point theorems, linear operators, and four fundamental theorems – Hahn-Banach Theorem, Uniform Boundedness Principle, Open Mapping Theorem and Closed Graph Theorem, Hilbert spaces and their adjoint operators. In the process of learning this course, the students are strongly suggested to follow an important principle; that is, the best way to learn mathematics is to do mathematics. Moreover, the students are urged to acquire the habit of studying with paper and pencil in hand; in this way mathematics will become increasingly meaningful to them. At the end of each chapter there are some basic problems for the students. Answers to those problems can be found in the part of Solutions to Exercises. This part is followed by a list of references. These books also supplement the author's text since they were used while preparing the course. Last but not least, an index is attached.

The author would like to express his gratitude to Professor B. Watson of MUN who had made some valuable suggestions on a course outline. The author would also like to take this opportunity to thank his undergraduate students: J. Bishop, P. Deal, J. Martin, S. Maye, B. Rowe, M. White and graduate students: J. Howell, N. Kirby, L. Li, J. McGraw, X. Yang and O. Yasar, as well as his teaching assistant: F. Zhang, who participated in the course. Without their compliments, criticisms, comments, and careful scrutiny of the text at all stages from classroom notes to page proofs, the author would not have ended up with the typewriting of this textbook.

St. John's, 2005 *J. Xiao*

# Contents

# 1. Normed Linear Spaces

In this chapter, we discuss the definition of a linear space which is the same as thing as a vector space – simply an abstract version of the familiar vector spaces $\mathbf{R}^n$, $n \in \mathbf{N}$.We shall always use the former terminology in order to emphasize the linearity that permeates the subject of this course. Note that vector spaces have certain algebraic properties: vectors may be added, multiplied by scalars, and vector spaces have bases and subspaces. Linear maps between vector spaces may be described in terms of matrices. Using Euclidean norm or distance, vector spaces have other analytic properties: for instance, certain functions from $\mathbf{R}^n$ to $\mathbf{R}$ are continuous, differentiable, Riemann integrable and so on.

## 1.1 Linear Spaces, Subspaces and Independence

**Definition 1.1.1.** *A linear space over a field* $\mathbf{F}$ *is a set* $V$ *equipped with maps* $\oplus : V \times V \to V$ *and* $\odot : \mathbf{F} \times V \to V$ *with the properties below:*

(i) *(commutative for vector addition)* $x \oplus y = y \oplus x \ \forall x, y \in V$*;*

(ii) *(associative for vector addition)* $(x \oplus y) \oplus z = x \oplus (y \oplus z) \ \forall x, y, z \in V$*;*

(iii) *(a zero element for vector addition)* $\exists\, 0 \in V \ni x \oplus 0 = 0 \oplus x = x \ \forall x \in V$*;*

(iv) *(additive inverse for vector addition)* $\exists -x \in V \ni x \oplus (-x) = (-x) \oplus x = 0 \ \forall x \in V$*;*

(v) *(associative for scalar multiplication)* $\alpha \odot (\beta \odot x) = (\alpha\beta) \odot x \ \forall \alpha, \beta \in \mathbf{F}, x \in V$*;*

(vi) *(distributive for scalar multiplication over scalar addition)* $(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x) \ \forall \alpha, \beta \in \mathbf{F}, x \in V$*;*

(vii) *(distributive for scalar multiplication over vector addition)* $\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y) \ \forall \alpha \in \mathbf{F}, x, y \in V$*;*

(viii) *(a unit element for scalar multiplication)* $1 \odot x = x \ \forall x \in V$ *where* $1$ *is the multiplicative identity in* $\mathbf{F}$*.*

*Example 1.1.1.* (i) $V = \mathbf{R}^n$ is a linear space with the usual vector addition and scalar multiplication over $\mathbf{R}$.

(ii) Let $V$ be the set of all polynomials with coefficients in $\mathbf{R}$ of degree less than $n$. Then $V$ is a linear space with usual addition of polynomials and scalar multiplication over $\mathbf{R}$.

(iii) Let $V$ be $M_{m,n}(\mathbf{C})$ of complex-valued $m \times n$ matrices. Then $V$ is a linear space with usual addition of matrices and scalar multiplication over $\mathbf{C}$.

(iv) Let $\ell_\infty$ denote the set of infinite sequences $(x_j)$ that are bounded: $\sup_j |x_j| < \infty$. Then $\ell_\infty$ is a linear space over $\mathbf{R}$ with:

$$\sup_j |x_j + y_j| \le \sup_j |x_j| + \sup_j |y_j|; \quad \sup_j |\alpha x_j| = |\alpha| \sup_j |x_j| < \infty.$$

(v) Let $C(S)$ be the class of continuous functions $f : S \to \mathbf{R}$ with $(f \oplus g)(x) = f(x) + g(x)$ and $(\alpha \odot f)(x) = \alpha f(x)$. Here $S$ is any nonempty subset of $\mathbf{R}$. Then $C(S)$ is a linear space over $\mathbf{R}$.

(vi) Let $V$ be the class of Riemann-integrable functions $f : (0,1) \to \mathbf{R}$ for which $\int_0^1 |f|^2 < \infty$. Then $V$ is a linear space over $\mathbf{R}$ with usual addition and scalar multiplication. Indeed, by the Cauchy-Schwartz inequality:

$$\Big(\int_0^1 |fg|\Big)^2 \le \Big(\int_0^1 |f|^2\Big)\Big(\int_0^1 |g|^2\Big) \quad \forall f, g \in V,$$

we have

$$\begin{aligned}
\int_0^1 |f+g|^2 &\le \int_0^1 \big(|f|^2 + 2|f||g| + |g|^2\big) \\
&\le \int_0^1 |f|^2 + 2\Big(\int_0^1 |f|^2\Big)^{\frac{1}{2}}\Big(\int_0^1 |g|^2\Big)^{\frac{1}{2}} + \int_0^1 |g|^2 \\
&= \Big(\big(\int_0^1 |f|^2\big)^{\frac{1}{2}} + \big(\int_0^1 |g|^2\big)^{\frac{1}{2}}\Big)^2.
\end{aligned}$$

Meanwhile,

$$\int_0^1 |\alpha f|^2 = |\alpha|^2 \int_0^1 |f|^2.$$

(vii) Let $C^\infty[a,b]$ be the space of infinitely differentiable functions on $[a,b]$. Then it is linear space over $\mathbf{R}$ with usual addition and scalar multiplication.

(viii) Let $\Omega \subset \mathbf{R}^n$ be nonempty, and $C^k(\Omega)$ the space of $k$ times continuously differentiable functions. This means that if $\mathbf{a} = (a_1, ..., a_n) \in \mathbf{N}^n$ has $|\mathbf{a}| = a_1 + \cdots a_n \le k$, then the partial derivatives

$$D^{\mathbf{a}} = \frac{\partial^{|\mathbf{a}|} f}{\partial x_1^{a_1} ... \partial x_n^{a_n}}$$

exists and are continuous. Then it is linear space over $\mathbf{R}$ with usual addition and scalar multiplication.

*Note*: For convenience, we shall drop the special notation $\oplus$, $\odot$ for vector addition and scalar multiplication, and simply refer to $V$ as a linear space over $\mathbf{F}$. Moreover, if $\mathbf{F} = \mathbf{R}$ then we shall say that $V$ is a real linear space; whereas if $\mathbf{F} = \mathbf{C}$, then we shall say that $V$ is a complex linear space.

As in the linear algebra of finite-dimensional vector spaces, subsets of linear spaces that are themselves linear spaces are called linear subspaces. With this, we have

**Definition 1.1.2.** *Let $V$ be a linear space over the field $\mathbf{F}$. A subset $W \subset V$ is called a linear subspace of $V$ if $\alpha x + \beta y \in W$ for $\forall \alpha, \beta \in \mathbf{F}$ and $\forall x, y \in W$.*

*Example 1.1.2.* (i) The set of vectors in $\mathbf{R}^n$ of the form $(x_1, x_2, x_3, 0, ..., 0)$ forms a three-dimensional $(3D)$ linear subspace of $\mathbf{R}^n$.

(ii) The set of polynomials of degree $\leq r$ forms a linear subspace of the set of polynomials of degree $\leq n$ for any $r \leq n$.

(iii) The space $C^{k+1}(\Omega)$ is a linear subspace of $C^k(\Omega)$.

**Definition 1.1.3.** *Let $V$ be a linear space over $\mathbf{F}$. If $W = \{x_0 + c : x_0 \in S\}$ where $S$ is a linear subspace of $V$ and $c$ is a fixed element of $V$, then $W$ is called an affine subset of $V$.*

*Example 1.1.3.* (i) If $W = \{x = \underbrace{(x_1, x_2, 1, ..., 1)}_{n} \in \mathbf{R}^n\}$ then it is an affine subset of $\mathbf{R}^n$.

(ii) In Example 1.1.1 (iii), let $W$ be the set of matrices with certain blocks of $1'$s. Then $W$ is an affine subset of $M_{m,n}(\mathbf{C})$.

(iii) In $\mathbf{R}^3$ all lines and planes through the origin are subspaces, whereas lines and planes not passing through the origin are affine subsets.

**Definition 1.1.4.** *A hyperplane $W$ of a linear space $V$ is a maximal proper affine subset of $V$; that is, $W = \{x \in V : x = x_0 + c, x_0 \in S\}$ where $c \in V$ is fixed, and $S$ is a maximal linear subspace of $V$ in the sense that any other subspace of $V$ containing $S$ is either $S$ or $V$ itself.*

Note that a hyperplane $W$ is proper subset of $V$. So $W \neq V$ which implies $S \neq V$.

*Example 1.1.4.* Hyperplanes in $\mathbf{R}^2$ are lines; hyperplanes in $\mathbf{R}^3$ are planes.

A fundamental concept for linear spaces is that of dimension, but first we need a few more definitions.

Given a linear space $V$. Elements $x_1, x_2, ..., x_n \in V$ are linearly dependent provided that there are scalars $\alpha_1, \alpha_2, ..., \alpha_n$ (not all zero) such that $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$. If there is no such set of scalars, then they are linearly independent.

The linear span of the vectors $x_1, x_2, ..., x_n \in V$ is the linear subspace of $V$:

$$\operatorname{span}\{x_1, ..., x_n\} = \Big\{x = \sum_{j=1}^{n} \alpha_j x_j : \ \alpha_j \in \mathbf{F}\Big\}.$$

**Definition 1.1.5.** *If the linear space $V$ is equal to the space spanned by a linearly independent set of $n$ vectors, then $V$ is said to have dimension $n$. If there is no such set of vectors, then $V$ is infinite-dimensional. Furthermore, a linearly independent set of vectors that spans $V$ is called a basis for $V$.*

.

*Example 1.1.5.* (i) The space $\mathbf{R}^n$ has dimension $n$; the standard basis is given by the vectors $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0)$,..., $e_n = (0, 0, ..., 0, 1)$.

(ii) $\{1, t, t^2, ..., t^n\}$ is a basis of the linear space of polynomials of degree $\leq n$ which has dimension $n + 1$.

(iii) All linear spaces given in Example 1.1 (iv)-(viii) are infinite-dimensional.

## 1.2 Norms

A norm on a a vector space is a way of measuring the length of a vector and hence the distance between two vectors.

**Definition 1.2.1.** *A norm on a linear space $V$ over the field $\mathbf{F}$ is a non-negative function $\|\cdot\| : V \to \mathbf{R}$ with the following properties:*

(i) $\|x\| = 0$ *iff* $x = 0$*;*

(ii) $\|x + y\| \leq \|x\| + \|y\|$ *for* $\forall x, y \in V$*;*

(iii) $\|\alpha x\| = |\alpha|\|x\|$ *for* $\forall x \in V$ *and* $\forall \alpha \in \mathbf{F}$.

In the definition we are assuming that $|\cdot|$ denotes the usual absolute value. If $\|\cdot\|$ is a function with (ii) and (iii) only it is called a semi-norm.

**Definition 1.2.2.** *Let $V$ be a linear space over the field $\mathbf{F}$.*

(i) *The normed linear space $(V, \|\cdot\|)$ is $V$ with a norm $\|\cdot\|$(denoted by $\|\cdot\|_V$, sometimes);*

(ii) *A set $C \subset V$ is convex if for any two points $x, y \in C$, $tx + (1 - t)y \in C$ for $\forall t \in [0, 1]$;*

(iii) *A norm $\|\cdot\|$ is strictly convex if $\|x\| = \|y\| = 2^{-1}\|x + y\| = 1$ implies $x = y$.*

*Example 1.2.1.* (i) Let $V = \mathbf{R}^n$ with the usual Euclidean norm

$$\|x\|_2 = \Big(\sum_{j=1}^{n} |x_j|^2\Big)^{\frac{1}{2}}.$$

Note that the only difficulty is the triangle inequality–for this we use the Cauchy-Schwarz inequality:

$$\sum_{j=1}^{n} |x_j y_j| \leq \Big(\sum_{j=1}^{n} |x_j|^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{n} |y_j|^2\Big)^{\frac{1}{2}}.$$

(ii) There are many other norms on $\mathbf{R}^n$, called the $p$-norms. For $p \in [1, \infty)$ define

$$\|x\|_p = \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{\frac{1}{p}}.$$

Then $\|\cdot\|_p$ is a norm on $V$: it suffices to verify the triangle inequality which follows from the Minkowski's inequality

$$\Big(\sum_{j=1}^{n}|x_j+y_j|^p\Big)^{\frac{1}{p}} \le \Big(\sum_{j=1}^{n}|x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^{n}|x_j|^p\Big)^{\frac{1}{p}}.$$

In case of $p=\infty$, define

$$\|x\|_\infty = \sup_{1\le j\le n}|x_j|.$$

It is conventional to write $\ell_p^n$ for these spaces. Note that $\ell_p^n$ and $\ell_q^n$ have exactly the same elements.

(iii) Let $V=\ell_\infty$, the linear space of bounded infinite sequences. Define

$$\|x\|_p = \begin{cases} \Big(\sum_{j=1}^{\infty}|x_j|^p\Big)^{\frac{1}{p}}, \text{ if } \ 1\le p<\infty, \\ \sup_{1\le j<\infty}|x_j|, \quad \text{ if } \ p=\infty. \end{cases}$$

If we restrict attention to the linear subspace on which $\|\cdot\|_p$ is finite, then $\|\cdot\|_p$ is a norm: the somewhat difficult triangle inequality follows from the infinite version of Minkowski's inequality. This yields an infinite family of normed linear spaces:

$$\ell_p = \{x=(x_j) : \|x\|_p<\infty\}.$$

Notice that for $p\in[1,\infty)$ there is a strict inclusions: $\ell_p\subset\ell_\infty$ and $\ell_p\subset\ell_q$ and hence $\ell_p$ is a linear subspace of $\ell_q$ whenever $p<q$. Consequently, $\ell_p\ne\ell_q$ if $p\ne q$.

(iv) Let $V=C[a,b]$. Then it becomes a normed space with the $p$-norm below

$$\|f\|_p = \begin{cases} \Big(\int_a^b|f(t)|^p dt\Big)^{\frac{1}{p}}, \text{ if } \ 1\le p<\infty, \\ \sup_{a\le t\le b}|f(t)|, \quad \text{ if } \ p=\infty. \end{cases}$$

Note that the triangle inequality for $p\in[1,\infty)$ follows from the integral form of the Minkowski inequality:

$$\Big(\int_a^b|f+g|^p\Big)^{\frac{1}{p}} \le \Big(\int_a^b|f|^p\Big)^{\frac{1}{p}} + \Big(\int_a^b|g|^p\Big)^{\frac{1}{p}}.$$

(v) Let $V$ be the class of Riemann-integrable functions $f:(0,1)\to\mathbf{R}$ with

$$\|f\|_2 = \Big(\int_0^1|f|^2\Big)^{\frac{1}{2}}.$$

Then it is a normed linear space.

## 1.3 Isomorphism and Product of Normed Spaces

Recall from linear algebra that linear spaces $V$ and $W$ over the field $\mathbf{F}$ are algebraically isomorphic if there is a bijection $T:V\to W$ that is linear:

$$T(\alpha x+\beta y)=\alpha T(x)+\beta T(y), \quad \forall\alpha,\beta\in\mathbf{F};\forall x,y\in V.$$

**Definition 1.3.1.** (i) *A pair* $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ *of normed linear spaces are topologically isomorphic if there is a linear bijection* $T : X \to Y$ *with the property that there are two positive constants* $c_1$ *and* $c_2$ *with*

$$c_1\|x\|_X \le \|T(x)\|_Y \le c_2\|x\|_X.$$

*We shall usually denote topological isomorphism by* $X \cong Y$*. If* $c_1 = c_2 = 1$ *then* $T$ *is called an isometry and the normed spaces* $X$ *and* $Y$ *are called isometric.*

(ii) *Two norms* $\|\cdot\|_{(1)}$ *and* $\|\cdot\|_{(2)}$ *defined on* $X$ *are said to be equivalent if there are two constants* $c_1, c_2 > 0$ *such that*

$$c_1\|\cdot\|_{(1)} \le \|\cdot\|_{(2)} \le c_2\|\cdot\|_{(1)}.$$

*Example 1.3.1.* (i) Let $X$ be the set of all real polynomials of the form $f(t) = a + bt + \frac{c}{2}t^2$ with the norm $\|f\|_X = \max\{|a|, |b|, |c|\}$, and $Y$ the set of all three dimensional vectors $z = ai + bj + ck$ with the norm $\|z\|_Y = \sqrt{a^2 + b^2 + c^2}$. If $T : Y \to X$ is given by $T(z) = a + bt + \frac{c}{2}t^2$, then $T$ is linear, bijective and has

$$\|T(z)\|_X = \max\{a, b, c\} \le \sqrt{a^2 + b^2 + c^2} = \|z\|_Y; \ \ \frac{1}{\sqrt{3}}\|z\|_Y \le \|T(z)\|_X \le \|z\|_Y.$$

So the two spaces are topologically isomorphic.

(ii) The real linear spaces $(\mathbf{C}, |\cdot|)$ and $(\mathbf{R}^2, \|\cdot\|_2)$ are topologically isometric.

**Theorem 1.3.1.** (i) *Any two norms on a finite dimensional linear spaces are equivalent.*

(ii) *If* $X$ *and* $Y$ *are* $n$*-dimensional normed linear spaces over the field* $\mathbf{F}$ *then* $X$ *and* $Y$ *are topologically isomorphic.*

*Proof.* (i) Let $X$ be an $n$-dimensional linear space over $\mathbf{F}$ and let $\|\cdot\|_{(k)}$, $k = 1, 2$, be two norms on $X$. Choose a basis $x_1, ..., x_n$ for $X$ and define a third norm, $\|\cdot\|_{(3)}$, as follows: For each $x \in X$ there is a unique set of scalars $\alpha_1, ..., \alpha_n$ in $\mathbf{F}$ such that $x = \sum_{j=1}^n \alpha_j x_j$. Let

$$\|\cdot\|_{(3)} = \sum_{j=1}^n |\alpha_j|.$$

Suppose that each of the norms $\|\cdot\|_{(k)}$, $k = 1, 2$ is equivalent to $\|\cdot\|_{(3)}$. Then there are positive constants $m_1, M_1$ and $m_2, M_2$ such that

$$m_k\|\cdot\|_{(k)} \le \|\cdot\|_{(3)} \le M_k\|\cdot\|_{(k)}, \quad k = 1, 2.$$

It follows that

$$\frac{m_1}{M_2}\|\cdot\|_{(1)} \le \|\cdot\|_{(2)} \le \frac{M_1}{m_2}\|\cdot\|_{(1)},$$

implying the desired.

Now let $\|\cdot\|$ denote either $\|\cdot\|_{(1)}$ or $\|\cdot\|_{(2)}$. We shall show that $\|\cdot\|$ is equivalent to $\|\cdot\|_{(3)}$. If $x = \sum_{j=1}^n \alpha_j x_j$ then

$$\|x\| \leq \sum_{j=1}^{n} |\alpha_j| \|x_j\| \leq \big( \max_{1\leq j\leq n} \|x_j\|\big) \|x\|_{(3)}.$$

This implies one-side estimate. To prove that there is a constant $c > 0$ such that $\|\cdot\|_{(3)} \leq c\|\cdot\|$, i.e.,

$$\sum_{j=1}^{n} |\alpha_j| \leq c \Big\| \sum_{j=1}^{n} \alpha_j x_j \Big\|. \tag{1.1}$$

We may assume that $\|x\|_{(3)} = 1$ by dividing the last inequality by $\sum_{j=1}^{n} |\alpha_j|$. Now if (1.1) is not true then there would be a sequence $y_k = \sum_{j=1}^{n} \alpha_{kj} x_j$ with $\|y_k\|_{(3)} = \sum_{j=1}^{n} |\alpha_{kj}| = 1$ but $\lim_{k\to\infty} \|y_k\| = 0$. Note that $|\alpha_{kj}| \leq 1$ for $j = 1, ..., n$ and every $k$. So there is a subsequence of $(y_k)$, denoted itself also, such that $\lim_{k\to\infty} \alpha_{kj}$ exists and equals, say, $\alpha_j$ for $j = 1, 2, ..., n$ as well as

$$\lim_{k\to\infty} \|y_k - \sum_{j=1}^{n} \alpha_j x_j\| \leq \max_{1\leq j\leq n} \|x_j\| \lim_{k\to\infty} \sum_{j=1}^{n} |\alpha_{kj} - \alpha_j| = 0.$$

Thus $\sum_{j=1}^{n} |\alpha_j| = 1$ while $\sum_{j=1} \alpha_j x_j = 0$. This is impossible, since the $x_j$ are linearly independent. This completes the argument.

(ii) It suffice to show that $X$ is topologically isomorphic to $\mathbf{F}^n$ with the Euclidean norm. There is an isomorphism $T$ from $X$ onto $\mathbf{F}^n$. We can use this map to define a new norm on $X$ as follows: For each $x \in X$ let $|||x|||$ be the norm of $T(x)$ in $\mathbf{F}^n$. When $X$ is given this new norm $T$ becomes a topological isomorphism. However, (i) shows that $|||\cdot|||$ is equivalent to any norm $\|\cdot\|_X$ on $X$. Hence $T$ is a topological isomorphism from $(X, \|\cdot\|_X)$ onto $\mathbf{F}^n$.

If $Y$ is a subspace of a linear normed space $(X, \|\cdot\|_X)$ then $\|\cdot\|_X$ restricted to $Y$ makes $Y$ into a normed subspace.

*Example 1.3.2.* Let $Y$ be the space of infinite real sequences with only finitely many non-zero terms. Then $Y$ is a linear subspace of $\ell_p$, $1 \leq p \leq \infty$, so the $p$-norm makes $Y$ into a normed space.

**Definition 1.3.2.** *If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed linear spaces, then their product is defined by*

$$X \times Y = \{(x, y) : \ x \in X, \ y \in Y\}.$$

*It may be made into a normed space under any of the following norms:*
(i) $\|(x, y)\| = (\|x\|_X^p + \|y\|_Y^p)^{\frac{1}{p}}, \ p \in [1, \infty)$;
(ii) $\|(x, y)\| = \max\{\|x\|_X, \|y\|_Y\}$.

Of course, this does not exhaust all the possible combinations of the norms $\|x\|_X$ and $\|y\|_Y$, but these are the most commonly used ones. The extension to products of $n$ normed linear spaces is defined in a similar manner.

*Example 1.3.3.* If $n = m + k$ with $m > 0, k > 0, n < \infty$ then $\ell_p^n$ may be viewed as the product of normed linear spaces $\ell_p^m$ and $\ell_p^k$.

## 1.4 Continuous Mappings between Normed Spaces

We have seen continuous mappings between $\mathbf{R}$ and $\mathbf{R}$ in real analysis. To make this definition we used the distance function $|x-y|$ on $\mathbf{R}$: a function $f : \mathbf{R} \to \mathbf{R}$ is continuous at $a \in \mathbf{R}$ if

$$\forall \epsilon > 0, \exists \delta > 0 \quad \ni \quad |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Looking at this, we find that exactly the same definition can be made for mappings between linear normed spaces. Thus, on suitably defined spaces, questions like "is the mapping $f \longmapsto f'$ or $\int_0^x f$ continuous?" can be asked.

**Definition 1.4.1.** *A mapping $T : X \to Y$ between normed linear spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ is continuous at $a \in X$ if*

$$\forall \epsilon > 0, \exists \delta = \delta(\epsilon, a) > 0 \quad \ni \quad \|x-a\|_X < \delta \Rightarrow \|T(x) - T(a)\|_Y < \epsilon.$$

*If $f$ is continuous at every $a \in X$ then we simply say $T$ is continuous on $X$. Moreover, $T$ is called uniformly continuous if*

$$\forall \epsilon > 0, \exists \delta = \delta(\epsilon) > 0 \quad \ni \quad \|x-a\|_X < \delta \Rightarrow \|T(x) - T(a)\|_Y < \epsilon.$$

*Example 1.4.1.* (i) The mapping $x \longmapsto x^2$ from $(\mathbf{R}, |\cdot|)$ to itself is continuous but not uniformly continuous.

(ii) Let $T(x) = Ax$ be the non-trivial linear map from $\mathbf{R}^n$ to $\mathbf{R}^m$ (with Euclidean norms) defined by the $m \times n$ matrix $A = (a_{ij})$. Using the Cauchy-Schwarz inequality, we see that $f$ is uniformly continuous. In fact, fix $a \in \mathbf{R}^n$ and $b = Aa$. Then for any $x \in \mathbf{R}^n$ we have

$$\begin{aligned} \|Ax - Aa\|_2^2 &= \sum_{i=1}^m \Big| \sum_{j=1}^n a_{ij}(x_j - a_j) \Big|^2 \\ &\le \sum_{i=1}^m \Big( \sum_{j=1}^n |a_{ij}|^2 \Big) \Big( \sum_{j=1}^n |x_j - a_j|^2 \Big) \\ &= C^2 \|x-a\|_2^2 \end{aligned}$$

where

$$C^2 = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 > 0.$$

(iii) Let $X = C[-1,1]$ with the sup-norm. Define a map $T : X \to X$ by

$$T(u)(t) = 1 + \int_0^t \big( \sin u(s) + \tan s \big) ds.$$

The map $T$ is uniformly continuous on $X$. To see this, we calculate by the Mean Value Theorem for Derivatives

$$\begin{aligned}||T(u)-T(v)||_\infty &= \sup_{t\in[-1,1]} |T(u)(t)-T(v)(t)| \\ &= \sup_{t\in[-1,1]} \Big|\int_0^t \big(\sin u(s)-\sin v(s)\big)ds\Big| \\ &\le \sup_{t\in[-1,1]} \int_0^t \Big|\sin u(s)-\sin v(s)\Big|ds \\ &\le \sup_{t\in[-1,1]} \int_0^t \Big|u(s)-v(s)\Big|ds \\ &\le \|u-v\|_\infty.\end{aligned}$$

(iv) Let $X$ be the space of complex-valued square-integrable Riemann integrable functions on $[0,1]$ with 2-norm. Define a map $T : X \to X$ by

$$T(u)(t) = \int_0^t u^2(s)ds.$$

Then $T$ is continuous. To see this, we estimate by Hölder's inequality and Minkowski's inequality

$$\begin{aligned}|T(u)(t)-T(v)(t)| &= \Big|\int_0^t \big(u^2(s)-v^2(s)\big)ds\Big| \\ &\le \Big(\int_0^t \big(|u(s)+v(s)|\big)^2 ds\Big)^{\frac{1}{2}} \Big(\int_0^t \big(|u(s)-v(s)|\big)^2 ds\Big)^{\frac{1}{2}} \\ &\le (\|u\|_2+\|v\|_2)\|u-v\|_2,\end{aligned}$$

so that

$$\|T(u)-T(v)\|_2 \le \sup_{t\in[0,1]} |T(u)(t)-T(v)(t)| \le (\|u\|_2+\|v\|_2)\|u-v\|_2.$$

(v) The same map as in (iv) applied to square-integrable Riemann integrable functions on $[0,\infty)$ is not continuous. To see this, let $a, b > 0$ and define

$$u(t) = \begin{cases} a, & 0 \le t \le 2b^2, \\ ia, & 2b^2 < t \le 4b^2, \\ 0, & \text{otherwise} \end{cases}$$

Then $\|u-0\|_2 = 2ab$. On the other hand,

$$T(u)(t) = \begin{cases} a^2 t, & 0 \le t \le 2b^2, \\ 4b^2a^2 - a^2 t, & 2b^2 < t \le 4b^2, \\ 0, & \text{otherwise} \end{cases}$$

Then $\|T(u)-T(0)\|_2 = \frac{3}{\sqrt{3}}a^2b^3$. Now, given any $\delta > 0$ we may choose constants $a, b$ with $2ab < \delta$ but $\frac{4}{\sqrt{3}}a^2b^3 = 1$. This is, given any $\delta > 0$ there is a function $u$ with the property that $\|u-0\|_2 < \delta$ but $\|T(u)-T(0)\|_2 = 1$, showing that $T$ is not continuous.

## 1.5 Sequences and Completeness in Normed Spaces

Let $X = (X, \|\cdot\|_X)$ be a normed linear space. Just as for continuity, we can employ $\|\cdot\|_X$ to define convergence for sequences and series in $X$ using the corresponding notion for $\mathbf{R}$.

**Definition 1.5.1.** *A sequence $(x_j)$ in $X$ is said to converge to $a \in X$ if $\lim_{j\to\infty} \|x_j - a\|_X = 0$. Similarly, a series $\sum_j x_j$ converges if the sequence of partial sums $(s_k)$ defined by $s_k = \sum_{j=1}^k x_j$ is convergent in $X$.*

*Example 1.5.1.* (i) If $(x_j)$ is a sequence in $\mathbf{R}^n$, with $x_j = (x_j^{(1)}, ..., x_j^{(n)})$, then for $p \in [1,\infty]$, $\|x_j\|_p \to 0$ iff $x_j^{(k)} \to 0$ for each $k = 1, ..., n$.

(ii) For infinite-dimensional spaces, it is not enough to check convergence on each component using a basis. Given $p \in [1,\infty]$. Suppose $(x_j)$ is the sequence in $\ell_p$ given by $x_j = (0, 0, ..., 1, ...)$ where the 1 appears in the $j$th position. Then if we write $x_j = (x_j^{(1)}, x_j^{(2)}, ...)$ we certainly have $x_j^{(k)} \to 0$ as $j \to \infty$ for each $k$. However, we also have $\|x_j\|_p = 1$ for all $j$, so the sequence is certainly not converging to 0. As a matter of fact, it is not converging to anything.

**Theorem 1.5.1.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be normed linear spaces. A map $T : X \to Y$ is continuous at $a \in X$ iff $\lim_{j\to\infty} T(x_j) = T(a)$ in $Y$ for every sequence $(x_j)$ converging to $a$ in $X$.*

*Proof.* Replace $|\cdot|$ with $\|\cdot\|$ in the proof of this statement for functions from $\mathbf{R}$ to itself.

**Definition 1.5.2.** *A sequence $(x_j)$ is a Cauchy sequence in the normed linear space $X = (X, \|\cdot\|_X)$ if*

$$\forall \epsilon > 0, \exists N \ \ni \ n, m > N \Rightarrow \|x_n - x_m\|_X < \epsilon.$$

Clearly, a convergence sequence is a Cauchy sequence. We know that in the normed linear space $(\mathbf{R}, |\cdot|)$ the converse also holds, and it is a simple matter to verify that in $\mathbf{R}^n$ the converse holds too. In many reasonable infinite-dimensional normed linear spaces however there are Cauchy sequences that do not converge.

*Example 1.5.2.* Let $X = (C[0,1], \|\cdot\|_2)$. If $(f_j)$ in $X$ is defined by

$$f_j(t) = \begin{cases} 0, & 0 \le t \le \frac{1}{2} - \frac{1}{j} \\ \frac{jt}{2} - \frac{j}{4} + \frac{1}{2}, & \frac{1}{2} - \frac{1}{j} \le t \le \frac{1}{2} + \frac{1}{j} \\ 1, & \frac{1}{2} + \frac{1}{j} \le t \le 1 \end{cases}$$

Note that $(f_j)$ is Cauchy due to

$$\begin{aligned} \|f_n - f_m\|_2^2 &= \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} |f_n(t) - f_m(t)|^2 dt + \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{n}} |f_n(t) - f_m(t)|^2 dt \\ &\to 0 \quad \text{as} \quad m > n \to \infty. \end{aligned}$$

But nevertheless, it is not hard to verify $\lim_{j\to\infty}\|f_j - f\|_2 = 0$, where

$$f(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2} \\ \frac{1}{2}, & t = \frac{1}{2} \\ 1, & \frac{1}{2} < t \le 1 \end{cases}$$

is not continuous at $t = \frac{1}{2}$ and so not in $C[0,1]$. In other words, $(f_j)$ does not converge in $C[0,1]$ with respect to $\|\cdot\|_2$. In fact, if there is a function $g \in C[0,1]$ such that $\lim_{j\to\infty}\|f_j - g\|_2 = 0$, then the triangle inequality implies $\|f-g\|_2 = 0$. Now consider $g(\frac{1}{2})$. If $g(\frac{1}{2}) \neq f(\frac{1}{2})$, then $|f-g|$ must be positive on $(\frac{1}{2}-\delta, \frac{1}{2})$ for some $\delta > 0$, which contradicts $\|f-g\|_2 = 0$. If $g(\frac{1}{2}) = f(\frac{1}{2})$, then $|f-g|$ must be positive on $(\frac{1}{2}, \delta + \frac{1}{2})$ for some $\delta > 0$, again contradicting $\|f-g\|_2 = 0$.

**Definition 1.5.3.** *A normed linear space is said to be complete if all Cauchy sequences are convergent.*

*Example 1.5.3.* (i) If $C[0,1]$ is equipped with the sup-norm, then it is complete.
(ii) $(C[0,1], \|\cdot\|_2)$ is not complete.

## 1.6 Some Topology

There are some properties of subsets of normed linear spaces and other more general spaces that we use very often. Topology is a subject that begins by attaching names to these properties and then develops a shorthand for talking about such things.

**Definition 1.6.1.** *Let $X = (X, \|\cdot\|_X)$ be a normed linear space.*

(i) *A set $C \subset X$ is closed if whenever $(c_j)$ is a sequence in $C$ that is convergent in $X$, the limit $\lim_{j\to\infty} c_j$ also lies in $C$.*

(ii) *A set $U \subset X$ is open if for every $u \in U$ there exists $\epsilon > 0$ such that $\|x-u\|_X < \epsilon \Rightarrow x \in U$.*

(iii) *A set $S \subset X$ is bounded if there is an $M > 0$ with the property that $x \in S \Rightarrow \|x\|_X \le M$.*

(iv) *A set $S \subset X$ is connected if there do not exist open sets $A, B$ in $X$ with $S \subset A \cup B$, $S \cap A \neq \emptyset$, $S \cap B \neq \emptyset$ and $S \cap A \cap B = \emptyset$.*

(v) *The interior of $S \subset X$ is the set*

$$S^\circ = \{x \in X : \ \exists \delta > 0 \ \ni \ \|x-y\|_X < \delta \Rightarrow y \in S\}.$$

(vi) *The closure of $S \subset X$ is the set*

$$\bar{S} = \{x \in X : \ \forall \epsilon > 0 \ \exists \ s \in S \ \ni \ \|s-x\|_X < \epsilon\}.$$

Clearly, $S^\circ \subset S \subset \bar{S}$ for any $S \subset X$. Moreover, a map $f : X \to Y$ between two normed linear spaces is continuous iff for every open set $U \subset Y$, the pre-image $f^{-1}(U) \subset X$ is also open. But, there is a continuous map $f : \mathbf{R} \to \mathbf{R}$ such that $f(U)$ is not open even if $U$ is open.

Recall that the Bolzano-Weierstrass theorem: if $S \subset \mathbf{R}^n$ is bounded and closed, then a continuous function $f : S \to \mathbf{R}$ attains its maximum and minimum: $\exists x_1, x_2 \in S \ \ni \ f(x_1) = \sup_{x \in S} f(x), \ f(x_2) = \inf_{x \in S} f(x)$. This is really the same as the Heine-Borel theorem that reads: a subset of $\mathbf{R}^n$ is compact iff it is bounded and closed.

**Definition 1.6.2.** (i) *A subset $S$ of a normed linear space is compact iff every sequences $(s_j)$ in $S$ has a subsequence $(s_{j_k})$ that converges in $S$.*

(ii) *A subset $S$ of a normed linear space is relatively (conditionally) compact iff its closure $\bar{S}$ is compact.*

Below is a more general version of the Bolzano-Weierstrass theorem.

**Theorem 1.6.1.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be two normed linear spaces, and $f : X \to Y$ a continuous map. If $A \subset X$ is compact, then $f(A) \subset Y$ is compact too. Consequently, if $Y = (\mathbf{R}, |\cdot|)$ then $f$ attains its maximum and minimum on $A$.*

*Proof.* Assume that $(y_j)$ is a sequence in $f(A)$. For each $j \in \mathbf{N}$ let $x_j \in A$ satisfy $f(x_j) = y_j$. Since $A$ is compact, we conclude that $(x_j)$ has a subsequence $(x_{j_k})$ converging to $x_0 \in A$. Note that $f$ is continuous at $x_0$. So $y_{j_k} = f(x_{j_k}) \to f(x_0)$ in $Y$ as $k \to \infty$. Obviously, $f(x_0) \in f(A)$. Thus, $f(A)$ is compact. In case $Y = (\mathbf{R}, |\cdot|)$, we know $f(A)$ is bounded and closed in $\mathbf{R}$ and thus get the desired result.

Some standard sets are used so often that we give them special names.

**Definition 1.6.3.** *Let $X = (X, \|\cdot\|_X)$ be a normed linear space.*

(i) *The open ball of radius $r > 0$ and center $x_0$ is the set*

$$B_r(x_0) = \{x \in X : \|x - x_0\|_X < r\}.$$

(ii) *The closed ball of radius $r > 0$ and center $x_0$ is the set*

$$\overline{B_r(x_0)} = \{x \in X : \|x - x_0\|_X \le r\}.$$

(iii) *A subset $S \subset X$ is dense if every open ball in $X$ has non-empty intersection with $S$. $X$ is said to be separable if there is a countable set $S = \{x_1, x_2, ...\}$ that is dense in $X$.*

*Example 1.6.1.* (i) Open and closed balls in normed linear spaces are convex.

(ii) $\mathbf{R}^n$, equipped with the norm $\|\cdot\|_2$, is separable because $\mathbf{Q}^n$ is a countable set and is dense in $\mathbf{R}^n$.

(iii) $C[a, b]$ is separable under the sup-norm: Let $S$ be the set of piecewise linear functions of the form

$$f(t) = s_k + \frac{s_{k+1} - s_k}{t_{k+1} - t_k}(t - t_k), \quad t_k \le t \le t_{k+1},$$

where $a = t_0 < t_1 < \cdots < t_n = b$ is a partition of $[a,b]$, and the $s_k, t_k$ are rational (with the possible exceptions of $t_0 = a$ and $t_n = b$). The set $S$ is enumerable. Given $g \in C[a,b]$. For any $\epsilon > 0$. Then there is a $\delta > 0$ such that

$$|t-s| < \delta \Longrightarrow |g(t) - g(s)| < \frac{\epsilon}{4}.$$

Let $a = t_0 < t_1 < \cdots < t_n = b$ be a partition of $[a,b]$ with $t_1, ..., t_{n-1}$ rational and such that $\max_k (t_{k+1} - t_k) < \delta$. Let $s_0, ..., s_n$ be rational numbers such that $\max_k |g(t_k) - s_k| < \frac{\epsilon}{4}$. Then for $t_k \le t \le t_{k+1}$ we have

$$f(t) - g(t) = \frac{t_{k+1} - t}{t_{k+1} - t_k}(s_k - g(t)) + \frac{t - t_k}{t_{k+1} - t_k}(s_{k+1} - g(t)),$$

and hence

$$|f(t) - g(t)| \le |s_k - g(t_k)| + |g(t_k) - g(t)| + |s_{k+1} - g(t_{k+1})| + |g(t_{k+1}) - g(t)| < \epsilon.$$

Thus $S$ is dense in $C[a,b]$.

(iii) $(\ell_p, \|\cdot\|_p)$ is separable for $1 \le p < \infty$ but not for $p = \infty$. In fact, let $W$ be the set of all elements of $\ell_p$ of the form $x = (x_1, ..., x_j, ...)$, where all of the $x_j$ are rational and all but a finite number of them vanish. Clearly, $W$ is countable. If $p \in [1, \infty)$ then $W$ is dense in $\ell_p$. To see this, let $\epsilon > 0$ and $x \in \ell_p$ be given. Then take $N$ so large that $\sum_{k=N+1}^{\infty} |x_k|^p < \epsilon/2$. Now for each $k \le N$, there is a rational number $r_k$ such that

$$1 \le k \le N \Rightarrow |x_k - r_k|^p < \frac{\epsilon}{2N}.$$

Set $r = (r_1, ..., r_N, 0, 0, ...)$. Then $r \in W$ and

$$\|x - r\|_p^p = \sum_{k=1}^{N} |x_k - r_k|^p + \sum_{k=N+1}^{\infty} |x_k|^p < \epsilon.$$

This shows that $W$ is dense in $\ell_p$, $p \in [1, \infty)$. But, if $p = \infty$ then let $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, ..., x_j^{(n)}, ...)$, $n \in \mathbf{N}$, be any sequence of elements in $\ell_\infty$. Define $x = (x_1, ..., x_j, ...)$, where

$$x_j = \begin{cases} x_j^{(j)} + 1, & \text{if} \quad |x_j^{(j)}| \le 1 \\ 0, & \text{if} \quad |x_j^{(j)}| > 1. \end{cases}$$

Thus $x \in \ell_\infty$. Moreover,

$$\|x - x^{(n)}\|_\infty \ge |x_n - x_n^{(n)}| \ge 1, \quad n \in \mathbf{N}.$$

This shows that the sequence $(x^{(n)})$ cannot be dense in $\ell_\infty$.

As an application of the above concepts, quotients of normed linear spaces may be formed. Note that we need both the algebraic structure (subspace of a linear space) and a topological property (closed) to make it all work.

**Definition 1.6.4.** *Let $X$ be a normed linear space over $\mathbf{F}$ and $Y$ be a linear subspace of $X$. Then the linear space $X/Y$–the quotient or factor space is formed as follows. The elements of $X/Y$ are cosets of $Y$ – sets of the form $x + Y = \{x + y : \ y \in Y\}$ for $x \in X$. The set of cosets is a linear space under the operations:*

$$(x_1 + Y) \oplus (x_2 + Y) = (x_1 + x_2 + Y); \quad \lambda \odot (x + Y) = \lambda x + Y$$

*for any $x_1, \ x_2, \ x \in X$ and $\lambda \in \mathbf{F}$.*

It is worth remarking that the operations make sense precisely because $Y$ is itself a linear space, so for instance $Y + Y = \{y_1 + y_2 : \ y_1, y_2 \in Y\} = Y$ and $\lambda Y = \{\lambda y : \ y \in Y\} = Y$ for $\lambda \neq 0$. Moreover, $x_1 + Y$ and $x_2 + Y$ are equal iff as sets $x_1 + Y = x_2 + Y$; this is true iff $x_1 - x_2 \in Y$ which is denoted by $x_1 \sim x_2$ – $x_1$ is equivalent to $x_2$ with respect to $Y$.

*Example 1.6.2.* (i) Let $X = \mathbf{R}^3$, and let $Y$ be the subspace spanned by $(1, 1, 0)$. Then $X/Y$ is a two-dimensional real vector space: since $(1, 0, 1) + Y$ and $(0, 0, 1) + Y$ generate $X/Y$.

(ii) The linear space $Y$ of finitely supported sequences (of which all but a finite number of entries vanish) in $\ell_1$ is a linear subspace. The quotient space $\ell_1/Y$ is very hard to visualize: its elements are equivalence classes under the relation $(x_j) \sim (y_j)$ if $(x_j)$ and $(y_j)$ differ in finitely many positions.

(iii) The linear space $Y$ of $\ell_1$ sequences of the form $(0, ..., 0, x_{n+1}, ...)$ (first $n$ are zero) is a linear subspace of $\ell_1$. Here the quotient space $\ell_1/Y$ is quite reasonable: in fact it is isomorphic to $\mathbf{R}^n$.

(iv) Recall that for $p, q \in [1, \infty)$, $p < q \Rightarrow \ell_p \subset \ell_q$. This means that for any $p < q$ there is a linear quotient space $\ell_q/\ell_p$. These quotient spaces are very pathological.

(v) Let $X = C[0, 1]$ and $Y = \{f \in X : \ f(0) = 0\}$. Then $X/Y$ is isomorphic to $\mathbf{R}$.

(vi) $Y = C[0, 1]$ is a linear subspace of the space $X$ of square-Riemann-integrable functions on $[0, 1]$. The quotient $X/Y$ is again a linear space that is impossible to work with.

Evidently, these examples tell us that not all linear subspaces are equally good: (i), (iii) and (v) are quite reasonable, whereas (ii), (iv) and (vi) are examples of linear spaces unlike any we have seen. The reason is the following: the space $X/Y$ is guaranteed to be a normed space with a norm related to the original norm on $X$ only when the subspace $Y$ is itself closed. Notice that (i), (iii) and (v) are precisely the ones in which the subspace is closed.

**Theorem 1.6.2.** *If $X = (X, \| \cdot \|_X)$ is a normed linear space, and $Y$ is a closed subspace of $X$, then $X/Y$ is a normed space under the norm $\|x + Y\| = \inf_{z \in x+Y} \|z\|_X$.*

*Proof.* Note that $\|x + Y\| = \inf_{z \in x+Y} \|z\|_X = \inf_{y \in Y} \|x - y\|$.

Firstly, if $\|x+Y\| = 0$ then there is a sequence of vectors $y_j \in Y$ such that $\lim_{j\to\infty} \|x - y_j\| = 0$. Since $Y$ is closed, $x \in Y$ and then $x + Y = 0 + Y$. Conversely, $\|0 + Y\| = 0$.

Secondly, the homogeneity is clear:

$$\|\lambda(x+Y)\| = \inf_{z\in x+Y} \|\lambda z\|_X = |\lambda| \inf_{z\in x+Y} \|z\|_X = |\lambda|\|x+Y\|.$$

Finally, the triangle inequality:

$$\begin{aligned}\|(x_1+Y)+(x_2+Y)\| &= \inf_{z_1\in x_1+Y, z_2\in x_2+Y} \|z_1+z_2\|_X \\ &\le \inf_{z_1\in x_1+Y} \|z_1\|_X + \inf_{z_2\in x_2+Y} \|z_2\|_X \\ &= \|x_1+Y\| + \|x_2+Y\|\end{aligned}$$

*Example 1.6.3.* (i) If $X = (\mathbf{R}^2, \|\cdot\|_2)$ and $Y = (1,0)\mathbf{R}$, then $X/Y$ consists of lines in $X$ of the form $(s,t)+Y$. Note that each such line may be written uniquely in the form $(0,t)+Y$, and this choice minimizes the norm of the element of $X$ that represents the line.

(ii) The quotient space may be a little odd. For instance, let $c$ denote the space of all sequences $(x_j)$ with the property that $\lim_{j\to\infty} x_j$ exists. This is a closed subspace of $\ell_\infty$. What is the quotient $\ell_\infty/c$?

(iii) $\|x+Y\|$ defines a norm on $X/Y$ only if $Y$ is a closed subspace.

(iv) $\dim(X/Y)$–the dimension of $X/Y$ is called the codimension of $Y$. Moreover, a linear subspace of codimension one is called a hyperplane.

## Exercises

**1.1** Let $p \in (1,\infty)$ and $q = p/(p-1)$. Prove Hölder's inequality in the following forms. For vectors:

$$\sum_{j=1}^n |x_j y_j| \le \Big(\sum_{j=1}^n |x_j|^p\Big)^{\frac{1}{p}} \Big(\sum_{j=1}^n |y_j|^q\Big)^{\frac{1}{q}}.$$

For sequences $(x_j)$ and $(y_j)$ with $\sum_j |x_j|^p < \infty$ and $\sum_j |y_j|^q < \infty$:

$$\sum_j |x_j y_j| \le \Big(\sum_j |x_j|^p\Big)^{\frac{1}{p}} \Big(\sum_j |y_j|^q\Big)^{\frac{1}{q}}.$$

For Riemann-integrable functions $f, g$ with $\int_0^1 |f|^p < \infty$, $\int_0^1 |g|^q < \infty$:

$$\int_0^1 |fg| \le \Big(\int_0^1 |f|^p\Big)^{\frac{1}{p}} \Big(\int_0^1 |g|^q\Big)^{\frac{1}{q}}.$$

**1.2** Let $p \in [1, \infty)$. Show Minkowski's inequality in the following forms. For vectors:

$$\Big(\sum_{j=1}^{n} |x_j + y_j|^p\Big)^{\frac{1}{p}} \le \Big(\sum_{j=1}^{n} |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^{n} |y_j|^p\Big)^{\frac{1}{p}}.$$

For sequences $(x_j)$ and $(y_j)$ with $\sum_j |x_j|^p < \infty$ and $\sum_j |y_j|^p < \infty$:

$$\Big(\sum_{j} |x_j + y_j|^p\Big)^{\frac{1}{p}} \le \Big(\sum_{j} |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j} |y_j|^p\Big)^{\frac{1}{p}}.$$

For Riemann-integrable functions $f, g$ with $\int_0^1 |f|^p < \infty$, $\int_0^1 |g|^p < \infty$:

$$\Big(\int_0^1 |f + g|^p\Big)^{\frac{1}{p}} \le \Big(\int_0^1 |f|^p\Big)^{\frac{1}{p}} + \Big(\int_0^1 |g|^p\Big)^{\frac{1}{p}}.$$

**1.3** For each of the following linear spaces, determine the dimension:
(i) The set of vectors $x = (x_j)$ in $\mathbf{R}^n$ with $\sum_{j=1}^n x_j = 0$;
(ii) The set of continuous functions $f : [0, 1] \to \mathbf{R}$;
(iii) The set of polynomials on $[0, 1]$.

**1.4** Prove that if $X$ is an $n$-dimensional normed linear spaces over $\mathbf{R}$ then $X$ is topologically isomorphic to $\mathbf{R}^n$.

**1.5** Norms $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are said to be equivalent if the identity map $I : (X, \|\cdot\|_{(1)}) \to (X, \|\cdot\|_{(2)})$ is a topological isomorphism. Prove that on $C[0, 1]$ the sup-norm is not equivalent to any $p$-norm, $1 \le p < \infty$.

**1.6** Prove that any norm $\|\cdot\| : V \to \mathbf{R}$ on a linear space $V$ over the field $\mathbf{F}$ is continuous on $V$, but also vector addition and scalar multiplication are continuous whenever $X \times Y$ is equipped with the norm $\|\cdot\|_X + \|\cdot\|_Y$.

**1.7** Recall that $\ell_\infty$ is the space of all bounded infinite sequences $x = (x_j)$ of complex numbers with the sup-norm $\|x\|_\infty = \sup_{j \in \mathbf{N}} |x_j|$. Prove: if $\ell_0$ is the class of all infinite sequence of complex numbers which have only finitely many non-zero terms, then $\ell_0$ is not closed in $\ell_\infty$.

**1.8** Given a normed linear space $X = (X, \|\cdot\|_X)$. Prove the following statements:
(i) $S^\circ \subset S \subset \bar{S}$ for any $S \subset X$;
(ii) Suppose $Y = (Y, \|\cdot\|_Y)$ is another normed linear space. Then a map $f : X \to Y$ is continuous iff for every open set $U \subset Y$, the pre-image $f^{-1}(U) \subset X$ is also open. But, there is a continuous map $f : \mathbf{R} \to \mathbf{R}$ such that $f(U)$ is not open even if $U$ is open.
(iii) Any open and closed balls in $X$ are convex.

**1.9** Prove that if $X = (X, \|\cdot\|_X)$ is a normed linear space, and $Y$ is a subspace of $X$, then $X/Y$ is a normed space under the norm $\|x + Y\| = \inf_{z \in x+Y} \|z\|_X$ iff $Y$ is closed.

# 2. Banach Spaces

In this chapter, we consider complete spaces - trying to do functional analysis in non-complete spaces is a little like trying to do elementary analysis over the rationals. In particular, we discuss the contraction mappings on Banach spaces and their applications in differential and integral equations.

## 2.1 Definition

**Definition 2.1.1.** *A complete normed linear space is called a Banach space.*

*Example 2.1.1.* (i) Any finite-dimensional normed linear space $(\ell_p^n, 1 \le p \le \infty)$ is a Banach space.

(ii) $C[0,1]$ with sup-norm is a Banach space over $\mathbf{R}$. In fact, if $(f_j)$ is Cauchy in $C[0,1]$, then for any $\epsilon > 0$ there exists $N \in \mathbf{N}$ such that

$$m, n \ge N \Rightarrow |f_m(x) - f_n(x)| \le \|f_m - f_n\|_\infty < \epsilon \quad \text{for all} \quad x \in [0,1].$$

So $(f_j(x))$ is Cauchy sequence in $\mathbf{R}$. Therefore it converges to some real number $f(x)$ for each $x \in [0,1]$; this defines a new function $f$ such that $f_j \to f$ pointwise. It remains to be shown that $(f_j)$ converges to $f$ uniformly on $[0,1]$, and that $f \in C[0,1]$. But this follows readily from the last estimate when letting $m \to \infty$.

(iii) The sequence space $\ell_p$ is a Banach space. To see this, assume that $(x_j)$ is a Cauchy in $\ell_p$, and write $x_j = (x_j^{(1)}, x_j^{(2)}, \ldots)$. Recall that $\|\cdot\|_p \ge \|\cdot\|_\infty$ for all $p \in [1, \infty]$. So, given $\epsilon > 0$ we may find $N$ such that $m, n > N \Rightarrow \|x_n - x_m\|_p < \epsilon$ which in turn implies that $\|x_n - x_m\|_\infty < \epsilon$, so for each $k$, $|x_n^{(k)} - x_m^{(k)}| < \epsilon$. That is, if $(x_j)$ is Cauchy in $\ell_p$ then $(x_j^{(k)})$ is a Cauchy in $\mathbf{R}$. Since $\mathbf{R}$ is complete, we deduce that for each $k$ we have $x_j^{(k)} \to y^{(k)}$. Note that this does not imply by itself $x_j \to y$. However, if we know that $(x_j)$ is Cauchy, then it does. In fact, we prove this for $p < \infty$ but the $p = \infty$ case is similar. Fix $\epsilon > 0$, and use the Cauchy criterion to find $N$ such that $n, m > N$ implies that

$$\sum_{k=1}^{\infty} |x_n^{(k)} - x_m^{(k)}|^p < \epsilon.$$

Now fix $n$ and let $m \to \infty$ to see that

$$\sum_{k=1}^{\infty} |x_n^{(k)} - y^{(k)}|^p \le \epsilon.$$

This last inequality means that $\|x_n - y\|_p \le \epsilon^{\frac{1}{p}}$, showing that $x_n \to y = (y^{(1)}, y^{(2)}, ...) \in \ell_p$.

**Definition 2.1.2.** *For a sequence $(x_j)$ in a normed linear space $(X, \|\cdot\|)$, we say that the series $\sum_{j=1}^{\infty} x_j$ is absolutely convergent provided $\sum_{j=1}^{\infty} \|x_j\| < \infty$.*

**Theorem 2.1.1.** *Let $(X, \|\cdot\|)$ be a normed linear space. Then it is a Banach space if and only if the absolute convergence of $\sum_{j=1}^{\infty} x_j$ implies its convergence.*

*Proof.* On the one hand, suppose that $(X, \|\cdot\|)$ is a Banach space. Consider the sequence of partial sums $s_k = \sum_{j=1}^{k} x_j$. Since $\sum_{j=1}^{\infty} x_j$ is absolutely convergent, we conclude that

$$\|s_m - s_k\| \le \sum_{j=k+1}^{m} \|x_j\| \to 0 \quad \text{as} \quad m > k \to \infty.$$

It follows that $(s_m)$ is Cauchy; since $X$ is complete this sequence converges, so $\sum_{j=1}^{\infty} x_j$ converges.

On the other hand, assume that $(x_j)$ is a Cauchy sequence in $(X, \|\cdot\|)$. Then for each $k \in \mathbf{N}$ there is a $N_k \in \mathbf{N}$ such that

$$i, j \ge N_k \Rightarrow \|x_i - x_j\| < 2^{-k}.$$

Without loss of generality, we may assume that $N_{k+1} \ge N_k$. This yields that $(x_{N_k})$ is a subsequence of $(x_k)$. Set $y_1 = x_{N_1}$ and $y_k = x_{N_k} - x_{N_{k-1}}$ when $k \ge 2$. Note that

$$\sum_{k=1}^{l} \|y_k\| < \|y_1\| + \sum_{k=2}^{l} 2^{1-k} \le \|y_1\| + 1.$$

So $(\sum_{k=1}^{l} \|y_k\|)$ is non-decreasing and bounded, and hence is convergent. By hypothesis, $\sum_{k=1}^{\infty} y_k$ is convergent. Since $\sum_{k=1}^{n} y_k = x_{N_n}$, it follows that $(x_{n+K})$ is convergent in $(X, \|\cdot\|)$. This, together with the fact that $(x_k)$ is Cauchy, infers that $(x_k)$ is convergent. Therefore, $(X, \|\cdot\|)$ is a Banach space.

*Example 2.1.2.* Theorem 2.1.1 is clearly not true for general normed linear spaces. For example, if $f_j(x) = x^{j-1} - x^j$ for $j \in \mathbf{N}$, then $f_j \in C[0,1]$, $\sum_{j=1}^{\infty} \|f_j\|_2 < \infty$, and

$$\sum_{j=1}^{\infty} f_j(x) = \lim_{k\to\infty} (1 - x^k), \quad x \in [0,1],$$

is not an element of $C[0,1]$. This is, of course, due to the fact that $(C[0,1], \|\cdot\|_2)$ is not complete.

## 2.2 Contraction Mapping Theorem

In this section we prove the simplest of the many fixed-point theorems. Such theorems are useful for solving equations, and with the formalism of function spaces one uniform treatment may be given for equations like $x = \cos x$, and $\frac{dy}{dx} = x + \tan(xy)$, $y(0) = y_0$.

**Definition 2.2.1.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be two normed linear spaces. A map $T : X \to Y$ is called a contraction if there is $\alpha \in [0,1)$ such that*

$$\|T(x) - T(y)\|_Y \le \alpha \|x - y\|_X, \quad x, y \in X.$$

*Here $\alpha$ is called the contraction constant.*

*Example 2.2.1.* (i) Let $X = Y = C[0,1]$ be equipped with the sup-norm and set $T(f)(x) = \alpha \int_0^x f(t)dt$, $\alpha \ge 0$. Clearly, if $\alpha \in [0,1)$ then $T$ is a contraction since

$$\|T(f) - T(g)\|_\infty \le \alpha \|f - g\|_\infty.$$

(ii) If $f(x) = 3^{-1}(x + \sin x)$ then $f$ is a contraction mapping from $\mathbf{R}$ to itself.

A contraction mapping contracts or shrinks the distance between points by the factor $\alpha$. Clearly, any contraction map is uniformly continuous on $X$. A mapping $T : X \to X$ has a fixed point if $T(p) = p$ for some $p \in X$. We show the following Banach's theorem on contraction mapping and fixed point.

**Theorem 2.2.1.** *Let $X = (X, \|\cdot\|_X)$ be a Banach space. If $T : X \to X$ is a contraction mapping, then $T$ has a unique fixed point.*

*Proof.* Let $p_0 \in X$. Define $p_{k+1} = T(p_k)$ for $k \in \mathbf{N} \cup \{0\}$. One claims that $\{p_k\}$ is a Cauchy sequence in $X$. In fact,

$$\|p_2 - p_1\|_X = \|T(p_1) - T(p_0)\|_X \le \alpha \|p_1 - p_0\|_X,$$

and so

$$\|p_3 - p_2\|_X = \|T(p_2) - T(p_1)\|_X \le \alpha \|p_2 - p_1\|_X \le \alpha^2 \|p_1 - p_0\|_X.$$

Generally, one has that if $k > j$ then

$$\|p_k - p_j\|_X \le \sum_{i=j}^{k-1} \|p_{i+1} - p_i\|_X \le \sum_{i=j}^{k-1} \alpha^i \|p_1 - p_0\|_X \le \frac{\alpha^j}{1-\alpha} \|p_1 - p_0\|_X.$$

This, together with $\alpha \in [0,1)$, implies $\{p_k\}$ is Cauchy and hence it converges to a point $p \in X$: $\lim_{k\to\infty} p_k = p$ in $X$.

Since $T$ is uniformly continuous,

$$T(p) = \lim_{k\to\infty} T(p_k) = \lim_{k\to\infty} p_{k+1} = p.$$

That is to say, $p$ is a fixed point of $T$.

Regarding the uniqueness, suppose $q$ is also a fixed point of $T$. Then

$$\|p-q\|_X = \|T(p)-T(q)\|_X \le \alpha\|p-q\|_X$$

and hence $\|p-q\|_X = 0$ due to $\alpha \in [0,1)$. This implies $p = q$.

It is worth remarking that the above proof is constructive in the sense that the fixed point is the limit of the iterates given by

$$p_{k+1} = T(p_k)$$

where the initial point or initial guess $p_0$ is an arbitrary point in $X$. The previous estimate gives the rapidity of the convergence of $p_k \to p$:

$$\|p-p_j\|_X \le \frac{\alpha^j}{1-\alpha}\|T(p_0)-p_0\|_X.$$

**Corollary 2.2.1.** *If $S$ is a closed subset of the Banach space $X = (X, \|\cdot\|_X)$, and $T : S \to S$ is a contraction mapping, then $T$ has a unique fixed point in $S$.*

*Proof.* Simply notice that $S$ is itself complete (since it is a closed subset of a complete space), and the proof of the above theorem does not use the linear space structure of $X$.

**Corollary 2.2.2.** *Let $S$ be a closed subset of the Banach space $X = (X, \|\cdot\|_X)$. Suppose there is an $n \in \mathbf{N}$ such that the $n$-th composition*

$$T_n(p) = \underbrace{T(T(T(\cdots(T(p)))))}_{n}$$

*is a contraction mapping, then $T$ has a unique fixed point.*

*Proof.* Since $T_n$ is a contraction mapping, it has a unique fixed point $p$. If $\alpha$ is the contraction constant for $T_n$, then

$$\|T(p)-p\|_X = \|T(T_n(p))-T_n(p)\|_X = \|T_n(T(p))-T_n(p)\|_X \le \alpha\|T(p)-p\|_X$$

and hence $T(p) = p$.

If $q$ is another fixed point of $T$ then it is also a fixed point of $T_n$ since

$$T_n(q) = T_{n-1}(T(q)) = T_{n-1}(q) = \cdots = q.$$

By the uniqueness of the fixed point of $T_n$, it follows that $q = p$.

*Example 2.2.2.* (i) Each map defined in Example 2.2.1 has a unique fixed point.

(ii) Let $X = Y = C[a,b]$ be equipped with the sup-norm, $0 < b-a < \infty$ and

$$T(f)(x) = \int_a^x f(t)dt, \quad a \le x \le b.$$

It is easy to see

$$\|T(f) - T(g)\|_\infty \le (b-a)\|f-g\|_\infty.$$

If $b - a \ge 1$ then $T$ is not a contraction. However,

$$T_n(f)(x) = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f(t)dt, \quad n \in \mathbf{N}$$

with the estimate

$$\|T_n(f) - T_n(g)\|_\infty \le \frac{1}{n!}(b-a)^n \|f-g\|_\infty,$$

so that $T_n$ is a contraction for $n$ large enough.

(iii) If the contraction constant $\alpha = 1$ then the theorem fails. For example, let $X = Y = \mathbf{R}$ and $f(x) = \frac{\pi}{2} + x - \arctan x$, then

$$|f(x) - f(y)| = |f'(\zeta)||x-y| = \frac{\zeta^2}{1+\zeta^2}|x-y| < |x-y|$$

where $\zeta$ is a point lying between $x$ and $y$. However, there is no $x \in \mathbf{R}$ such that $f(x) = x$.

(iv) There is a discontinuous function $f$ such that its iteration becomes a contraction mapping. For instance, if

$$f(x) = \begin{cases} \frac{1}{4}, & \text{if} \quad x \in [0, 1/2], \\ \frac{1}{2}, & \text{if} \quad x \in (1/2, 1]. \end{cases}$$

then $f(f(x)) = 1/4$ for $x \in [0,1]$.

(v) A basic linear problem is the following: let $T : \mathbf{R}^n \to \mathbf{R}^n$ be the affine map defined by $T(x) = Ax + b$ where $A = (a_{ij})$ is an $n \times n$ matrix. Equivalently, $T(x) = y$ where $y_i = \sum_{j=1}^n a_{ij}x_j + b_i$ for $i = 1, 2, ..., n$. If $T$ is a contraction map, then we can use the above theorem to solve the equation $T(x) = x$. The conditions under which $T$ is a contraction depend on the choice of norm for $\mathbf{R}^n$. Three cases follow.

(1) $\|x\|_\infty = \max_i\{|x_i|\}$. In this case,

$$\|T(x) - T(y)\|_\infty = \max_i \Big|\sum_{j=1}^n a_{ij}(x_j - y_j)\Big| \le \Big(\max_i \sum_{j=1}^n |a_{ij}|\Big)\|x-y\|_\infty.$$

Thus the contraction condition is

$$\max_{1\le i\le n} \sum_{j=1}^n |a_{ij}| \le \alpha < 1. \tag{2.1}$$

(2) $\|x\|_1 = \sum_{i=1}^n |x_i|$. In this case,

$$\|T(x) - T(y)\|_1 = \sum_i \Big|\sum_{j=1}^n a_{ij}(x_j - y_j)\Big| \le \Big(\max_j \sum_{i=1}^n |a_{ij}|\Big)\|x-y\|_1.$$

Thus the contraction condition is

$$\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|\le\alpha<1. \tag{2.2}$$

(3) $\|x\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$. In this case,

$$\|T(x)-T(y)\|_2^2=\sum_{i=1}^{n}\Big(\sum_{j=1}^{n}a_{ij}(x_j-y_j)\Big)^2\le\Big(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\Big)^2\|x-y\|_2^2.$$

Thus the contraction condition is

$$\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\le\alpha<1. \tag{2.3}$$

It follows that if any one of these three conditions holds, then there exists a unique solution in $\mathbf{R}^n$ to the affine equation $Ax+b=x$. Moreover, the solution may be approximated using the iterative scheme $x_1=T(x_0), x_2=T(x_1),\ldots.$ Note that each of these three conditions is sufficient for the method to work, but none of them are necessary, since (2.1), (2.2) and (2.3) are not equivalent.

## 2.3 Applications to Differential and Integral Equations

As mentioned before, the most important applications of the contraction mapping method are to differential and integral Equations. The first result in this direction is due to Picard.

**Theorem 2.3.1.** *Let $D\subseteq\mathbf{R}^2$ be open and $(x_0,y_0)\in D$. Let $f:D\to\mathbf{R}$ be a continuous and satisfy a Lipschitz condition of the form:*

$$|f(x,y_1)-f(x,y_2)|\le L|y_1-y_2|,\quad (x,y_1),\ (x,y_2)\in D.$$

*Consider the Initial Value Problem (IVP):*

$$\frac{dy}{dx}=f(x,y(x)),\quad y(x_0)=y_0.$$

*Then there exists a $\delta>0$ such that the IVP has a unique solution in the interval $[x_0-\delta,x_0+\delta]$.*

*Proof.* Clearly, solving this IVP is equivalent to solving the integral equation

$$y(x)=y_0+\int_{x_0}^{x}f(t,y(t))dt.$$

So, it is our wish to show that the last equation has a unique solution.

Suppose $R \subset D$ is a closed rectangle centered at $(x_0, y_0)$. Then there is $M > 0$ such that

$$\max_{x,y \in R} |f(x,y)| \leq M.$$

Now choose $\delta > 0$ such that $L\delta < 1$ and

$$[x_0 - \delta, x_0 + \delta] \times [y_0 - M\delta, y_0 + M\delta] \subseteq R.$$

Set

$$S = \Big\{ \phi \in C[x_0 - \delta, x_0 + \delta] : \phi([x_0 - \delta, x_0 + \delta]) \subseteq [y_0 - M\delta, y_0 + M\delta] \Big\}.$$

Note that $S$ is a closed subset of $C[x_0 - \delta, x_0 + \delta]$. So $S$ is complete.

Define a mapping $T : S \to S$ by

$$T(\phi)(x) = y_0 + \int_{x_0}^{x} f(t, \phi(t))dt.$$

Obviously, $T(\phi)$ is continuously differentiable and

$$|T(\phi)(x) - y_0| \leq M|x - x_0| \leq M\delta.$$

This implies $T(\phi) \in S$. Now, our problem is equivalent to showing that $T$ has a unique fixed point in $S$. Thus, it suffices to verify that $T$ is a contraction. For this, if $x \in [x_0 - \delta, x_0 + \delta]$ and $\phi, \psi \in S$, then

$$|T(\phi)(x) - T(\psi)(x)| \leq \int_{x_0}^{x} |f(t, \phi(t)) - f(t, \psi(t))|dt \leq L|x - x_0| \|\phi - \psi\|_\infty,$$

and the result follows.

*Example 2.3.1.* (i) Note that any existence theorem for the IVP in the above must be local in nature. For example,

$$\frac{dy}{dx} = y^2, \quad y(1) = -1,$$

has the solution $y(x) = -1/x$, which is not defined at $x = 0$ even though $f(x,y) = y^2$ is continuous there.

(ii) If the Lipschitz condition is dropped, the IVP still has a solution, but the solution may fail to be unique. For instance the IVP:

$$\frac{dy}{dx} = y^{1/3}, \quad y(0) = 0,$$

has an infinite number of solutions

$$y_c(x) = \begin{cases} 0, & \text{if} \quad 0 \leq x \leq c, \\ \left(\frac{2(x-c)}{3}\right)^{3/2}, & \text{if} \quad c < x \leq 1. \end{cases}$$

where $c \in [0, 1]$.

The condition on the set $D$ used in the last theorem arise very often so it is useful to have a short description for them.

**Definition 2.3.1.** *A domain in a normed linear space $X$ is an open connected set.*

An example of a domain in $\mathbf{R}$ containing the point $a$ is an interval $(a-\delta, a+\delta)$ for some $\delta > 0$. Note that if $D$ is a domain in $(X, \|\cdot\|_X)$ and $a \in D$ then for some $r > 0$ the open ball $B_r(a) = \{x \in X : \|x - a\|_X < r\}$ lies in $D$.

Picard's theorem easily generalizes to systems of simultaneous differential equations.

**Theorem 2.3.2.** *Let $D \subset \mathbf{R}^{n+1}$ be a domain containing $(x_0, y_{01}, ..., y_{0n})$ and let $f_1, ..., f_n$ be continuous functions from $D$ to $\mathbf{R}$ each satisfying a Lipschitz condition*

$$|f_i(x, y_1, ..., y_n) - f_i(x, z_1, ..., z_n)| \leq L \max_{1 \leq i \leq n} |y_i - z_i|$$

*for $(x, y_1, ..., y_n), (x, z_1, ..., z_n) \in D$. Then there is an interval $(x_0 - \delta, x_0 + \delta)$ on which the system of simultaneous ordinary differential equations*

$$\frac{dy_i}{dx} = f_i(x, y_1, ..., y_n) \quad \text{for} \quad i = 1, ..., n$$

*has a unique solution $y_i = \phi_i(x)$, $i = 1, ..., n$ satisfying the initial conditions $\phi_i(x_0) = y_{0i}$, $i = 1, ..., n$.*

*Proof.* As in the proof of the last theorem, write the system in integral form

$$\phi_i(x) = y_{0i} + \int_{x_0}^{x} f_i(t, \phi_1(t), ..., \phi_n(t))dt, \ \ i = 1, ..., n.$$

Since each of $f_i$ is continuous on $D$, there is a bound $|f_i(x, y_1, ...y_n)| \leq M$ in some domain $D' \subset D$ with $(x_0, y_{01}, ..., y_{0n}) \in D'$. Choose $\delta > 0$ with the properties that $M\delta < 1$ and

$$|x - x_0| \leq \delta \quad \text{and} \quad \max_i |y_i - y_{0i}| \leq M\delta \Rightarrow (x, y_1, ..., y_n) \in D'.$$

Let now $S$ be the set of $n$-tuples $\phi = (\phi_1, ..., \phi_n)$ of continuous functions defined on the interval $[x_0 - \delta, x_0 + \delta]$ and such that $|\phi_i(x) - y_{0i}| \leq M\delta$ for all $i = 1, ..., n$. The set $S$ may be equipped with the norm

$$\|\phi\| = \max_{x,i} |\phi_i(x)|.$$

It is easy to check that $S$ is complete. The mapping $T$ defined by the set of integral operators

$$(T(\phi))_i(x) = y_{0i} + \int_{x_0}^{x} f_i(t, \phi_1(t), ..., \phi_n(t))dt, \ \ i = 1, ..., n.$$

for $|x - x_0| \le \delta$ is a contraction from $S$ to itself. To see this, first note that if $\phi \in S$ and $|x - x_0| \le \delta$,

$$\|\phi_i(x) - y_{0i}| = \left| \int_{x_0}^{x} f_i(t, \phi_1(t), ..., \phi_n(t)) dt \right| \le M\delta, \ i = 1, ..., n,$$

so that $T(\phi) = ((T(\phi)_1, ..., (T(\phi))_n)$ lies in $S$. It remains to check that $T$ is a contraction map:

$$\begin{aligned} |(T(\phi))_i(x) - (T(\psi))_i(x)| &\le \int_{x_0}^{x} |f_i(t, \phi_1(t), ..., \phi_n(t)) - f_i(t, \psi_1(t), ..., \psi_n(t))| dt \\ &\le M\delta \max_i |\phi_i(x) - \psi_i(x)| \\ &\le M\delta \|\phi_i - \psi_i\|, \ i = 1, ..., n; \end{aligned}$$

so $T : S \to S$ is a contraction. It follows that the equation has a unique solution. so the system of differential equations has a unique solution.

Integral equations may be a little less familiar than differential equations although we have seen already that the two are intimately connected, so we begin with the following example.

*Example 2.3.2.* (i) Given problems in physics led to the need to "invert" the integral equation

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} f(y) dy$$

for functions $f$ and $g$ of specific kinds. This was solved – formally at least – by Fourier in 1811, who noted that this equation requires that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ixy} g(y) dy.$$

(ii) When studying generalizations of the tautochrone problem, Abel was led to the integral equation

$$g(x) = \int_a^x \frac{f(y)}{(x-y)^b} dy, \quad b \in (0,1), \quad g(a) = 0$$

for which he found the solution

$$f(y) = \frac{\sin \pi b}{\pi} \int_a^y \frac{g'(x)}{(y-x)^{1-b}} dx.$$

**Theorem 2.3.3.** *Let $k : [a,b] \times [a,b] \to \mathbf{R}$ be continuous and $\phi \in C[a,b]$. Then the Fredholm Integral Equation (FIE):*

$$f(x) = \lambda \int_a^b k(x,y) f(y) dy + \phi(x), \quad \lambda \in \mathbf{R}$$

*has a unique solution $f \in C[a,b]$ for certain $\lambda$.*

*Proof.* Define a mapping $K : C[a,b] \to C[a,b]$ by

$$K(f)(x) = \lambda \int_a^b k(x,y)f(y)dy + \phi(x), \quad \lambda \in \mathbf{R}.$$

And recall the sup-norm on $C[a,b]$:

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|.$$

Solving the FIE is equivalent to showing that $K$ has a fixed point. Now

$$\|K(f) - K(g)\|_\infty \leq |\lambda| \sup_{x,y \in [a,b]} |k(x,y)|(b-a)\|f-g\|_\infty.$$

Thus, if

$$|\lambda| \sup_{x,y \in [a,b]} |k(x,y)|(b-a) < 1,$$

then $K$ is contraction and, therefore, has a unique fixed point.

More is true.

**Theorem 2.3.4.** *Let $k : [a,b] \times [a,b] \to \mathbf{R}$ be continuous and $\phi \in C[a,b]$. Then the Volterra Integral Equation (VIE):*

$$f(x) = \lambda \int_a^x k(x,y)f(y)dy + \phi(x)$$

*has a unique solution in $C[a,b]$.*

*Proof.* To see this, it is enough to show that $K : C[a,b] \to C[a,b]$ defined by

$$K(f)(x) = \lambda \int_a^x k(x,y)f(y)dy + \phi(x),$$

has a unique fixed point. A simple calculation implies that if $f_1, f_2 \in C[a,b]$ and $M = \sup_{x,y \in [a,b]} |k(x,y)|$ then

$$|K(f_1)(x) - K(f_2)(x)| \leq |\lambda| M \|f_1 - f_2\|_\infty (x-a),$$

$$|K(K(f_1))(x) - K(K(f_2))(x)| \leq \big(|\lambda|M\big)^2 \|f_1 - f_2\|_\infty \frac{(x-a)^2}{2},$$

and

$$|K_n(f_1)(x) - K_n(f_2)(x)| \leq \big(|\lambda|M\big)^n \|f_1 - f_2\|_\infty \frac{(x-a)^n}{n!}.$$

Hence

$$\|K_n(f_1) - K_n(f_2)\|_\infty \leq \frac{\big(|\lambda|(b-a)M\big)^n}{n!} \|f_1 - f_2\|_\infty.$$

Because

$$\frac{\big(|\lambda|(b-a)M\big)^n}{n!} \to 0 \quad \text{as} \quad n \to \infty,$$

it follows that $K_n$ is a contraction mapping for large $n$ and, therefore, has a unique fixed point for any value of the parameter $\lambda$.

## Exercises

**2.1** Prove that if $C^n[0,1]$, $n \in \mathbf{N}$, is equipped with

$$\|f\| = \sup_{0\leq k\leq n} \sup_{t\in[0,1]} |f^{(k)}(t)|,$$

then it is a Banach space under this norm $\|\cdot\|$.

**2.2** Let $c_0$ denote the set of all infinite sequence $(x_j)$ of complex numbers such that $x_j \to 0$ as $j \to \infty$. Explain why $c_0$ is a Banach space under the norm $\|(x_j)\|_\infty = \sup_j |x_j|$.

**2.3** Define $T : C[0,1] \to C[0,1]$ by $T(f)(x) = \int_0^x f(t)dt$. Prove that
i) $T$ is not a contraction;
ii) $T$ has a unique fixed point;
iii) $T(T)$ is a contraction.

**2.4** Let $f : \mathbf{R} \to \mathbf{R}$ be differentiable with $|f'(x)| \leq \alpha$, where $\alpha \in [0,1)$. Prove that $f$ is a contraction mapping.

**2.5** Show that there is a unique continuous function $f : [0,1] \to \mathbf{R}$ such that

$$f(x) = \sin x + \int_0^1 f(y)\exp\big(-(x+y+1)\big)dy.$$

# 3. Linear Operators

Given two linear spaces $X$ and $Y$ over a field $\mathbf{F}$, a linear operator $T : X \to Y$ is a map $T$ from $X$ to $Y$ such that

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2), \quad \forall x_1, x_2 \in X; \quad \forall \alpha_1, \alpha_2 \in \mathbf{F}.$$

From this definition it follows that $T(0) = T(2 \cdot 0) = 2T(0)$, so that $T(0) = 0$. Also, we write $T(X) = \{Tx : x \in X\}$. If $T(x_1) = T(x_2)$ implies $x_1 = x_2$ then we say that $T$ is a one-to-one map. Clearly, if $T$ is linear then $T$ is one-to-one if and only if $T(x) = 0$ implies $x = 0$.

In this chapter, we discuss bounded operators, inverse operators, and four classical theorems in functional analysis: uniform bounded principle, open mapping theorem, closed graph theorem and Hahn-Banach theorem.

## 3.1 Bounded Operators

To begin with, we give the definition of a bounded operator.

**Definition 3.1.1.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be normed linear spaces over $\mathbf{F}$. We say that a linear operator $T : X \to Y$ is bounded provided there exists a constant $C \geq 0$ such that $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$. Define*

$$\|T\| = \|T\|_{X \to Y} = \sup_{x \in X, x \neq 0} \left\{ \frac{\|Tx\|_Y}{\|x\|_X} \right\}.$$

*Example 3.1.1.* (i) $\|T\| = \sup_{\|x\|_X = 1} \|Tx\|_Y$ follows from

$$\left\| x \|x\|_X^{-1} \right\|_X = 1 \quad \forall x \neq 0.$$

(ii) Equip $\ell_2$ with 2-norm, and let $T : \ell_2 \to \ell_2$ be given by $Tx = (0, x_1, x_2, ...)$ when $x = (x_1, x_2, ...)$. Then it is easy to check that $T$ is a bounded linear operator with $\|T\| = 1$.

(iii) Choose for both $X$ and $Y$ the space $C[0,1]$ with the sup-norm. Define $T : X \to Y$ by $(Tf)(x) = e^x f(x)$, $x \in [0,1]$. Then $T$ is bounded with $\|T\| = e$.

Recall Theorem 1.5.1 that a continuous linear transformation $T$ from $X$ to $Y$ is a linear transformation with the property:

$$\|x_n - x\|_X \to 0 \Rightarrow \|Tx_n - Tx\|_Y = 0.$$

The following result tells us that the boundedness of a linear operator is equivalent to its continuity.

**Theorem 3.1.1.** *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces and $T : X \to Y$ be a linear operator. Then the following are equivalent:*
(i) *$T$ is continuous in $X$.*
(ii) *$T$ is continuous at $0$.*
(iii) *$T$ is bounded.*
(iv) *$T$ maps bounded subsets of $X$ to bounded subsets of $Y$.*

*Proof.* (i)$\Leftrightarrow$ (ii). The direction (i)$\Rightarrow$ (ii) is trivial. Regarding (ii)$\Rightarrow$ (i), we suppose that $T$ is continuous at $0 \in X$. If $x_n \to x$ then $x_n - x \to 0$. Hence $T(x_n - x) \to T(0)$, so that $T(x_n) \to T(x)$.

(iii)$\Leftrightarrow$ (ii). If $T$ is bounded and $x_n \to 0$, then $Tx_n \to 0$ also. It follows that $T$ is continuous at 0, so by (i)$\Leftrightarrow$ (ii) $T$ is continuous in $X$. Conversely, assume that $T$ is continuous in $X$. If $T$ is not bounded, then for any $n \in \mathbf{N}$ there exists a point $x_n \in X$ with $\|Tx_n\|_Y \geq n\|x_n\|_X$. Let $y_n = \frac{x_n}{n\|x_n\|_X}$, so that $\|y_n\|_X = n^{-1} \to 0$. However, $\|Ty_n\|_Y > 1$ and $T(0) = 0$, contradicting the assumption that $T$ is continuous at 0.

(iv)$\Leftrightarrow$ (ii). Suppose (ii) is true. So $T$ is bounded with $\|T\| \leq C$ for some constant $C > 0$ due to (iii)$\Leftrightarrow$ (ii). If $S$ is a bounded subset of $X$, then there is a constant $M > 0$ such that $\|x\|_X \leq M$ $\forall x \in S$ and hence

$$\|Tx\|_Y \leq C\|x\|_X \leq CM \quad \forall x \in X.$$

That is to say, $T(S)$ is a bounded subset of $Y$. Conversely, assume that (iv) is true. Given an open ball $B_\epsilon^Y(0) = \{y \in Y : \|y\|_Y < \epsilon\}$ in $Y$, let $B_1^X(0) = \{x \in X : \|x\|_X < 1\}$ denote the open unit ball in $X$. By (iv) it follows that $T(B_1)$ is bounded in $Y$. Thus, there is a $\lambda > 0$ such that

$$T(B_1^X) \subset \lambda B_\epsilon^Y(0) = \{y \in Y : \|y\|_Y < \lambda\epsilon\}.$$

This implies $T(B_{\lambda^{-1}}^X) \subset B_\epsilon^Y$ since $T$ is linear, and so $T$ is continuous at 0.

*Example 3.1.2.* Suppose $X = \mathbf{R}^n$ with 2-norm, and $e_j = (0, ..., 1, ..., 0)$ are a basis of $\mathbf{R}^n$. Then any $x \in \mathbf{R}^n$ has the form $x = \sum_{j=1}^n x_j e_j$ and $\|x\|_2^2 = \sum_{j=1}^n |x_j|^2$. If $T : \mathbf{R}^n \to Y$ is a linear transformation, where $Y = (Y, \|\cdot\|_Y)$ is a normed linear space over $\mathbf{F}$. The Cauchy-Schwarz inequality implies

$$\|T(x)\|_Y \leq \|x\|_2 \left( \sum_{j=1}^n \|T(e_j)\|_Y^2 \right)^{\frac{1}{2}}.$$

So, $T$ is bounded with $\|T\| \leq \left( \sum_{j=1}^n \|T(e_j)\|_Y^2 \right)^{\frac{1}{2}}$.

Indeed, the last example is a special case of the following result.

**Theorem 3.1.2.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be two noremd linear spaces over $\mathbf{F}$. If $X$ is finite dimensional, then any linear transformation $T : X \to Y$ is bounded.*

*Proof.* Note that any two norms on a finite dimensional linear space $X$ over $\mathbf{F}$ are equivalent. So, we construct a new norm $\|\cdot\|$ via $\|\cdot\|_X$ and $\|\cdot\|_Y$ as follows.

$$\|x\| = \|x\|_X + \|Tx\|_Y, \quad x \in X.$$

Of course, $\|\cdot\|$ is a norm on $X$ and so it is equivalent to $\|\cdot\|_X$. This implies a constant $C > 0$ with $\|\cdot\| \le C\|\cdot\|_X$. It follows that $\|Tx\|_Y \le \|x\| \le C\|x\|_X$ for all $x \in X$. In other words, $T : X \to Y$ is bounded.

Next, we consider the space of linear operators.

**Definition 3.1.2.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be two noremd linear spaces over $\mathbf{F}$. Denote by $B(X,Y)$ the set of all bounded linear operators from $X$ to $Y$. In particular, $B(X) = B(X,X)$.*

**Theorem 3.1.3.** (i) *$B(X,Y)$ is a linear space over $\mathbf{F}$ with respect to operations:*

$$(T+S)(x) = T(x) + S(x); \quad (\alpha T)(x) = \alpha T(x), \quad x \in X;\ \alpha \in \mathbf{F}.$$

(ii) *The function $\|\cdot\| : B(X,Y) \to \mathbf{R}$, defined for every $T \in B(X,Y)$ by*

$$\|T\| = \sup_{x \in X, \|x\|_X \neq 0} \frac{\|T(x)\|_Y}{\|x\|_X},$$

*is a norm on $B(X,Y)$.*

(iii) *If $Y$ is a Banach space, then $B(X,Y)$ is a Banach space.*

*Proof.* (i). Check those conditions for a linear space with $B(X,Y)$.

(ii) We have to verify three conditions required for a norm. First, it is clear that $\|T\| \ge 0$. If $\|T\| = 0$ then $\|Tx\|_Y = 0$ for all $x \in X$, and hence $Tx = 0$ for all $x \in X$. This gives $T = 0$. Conversely, $T = 0$ implies $\|T\| = 0$.

Next,

$$\|\alpha T\| = \sup_{x \in X, \|x\|_X \neq 0} \frac{\|\alpha T(x)\|_Y}{\|x\|_X} = |\alpha|\|T\|.$$

Finally,

$$\begin{aligned}
\|T+S\| &= \sup_{x \in X, \|x\|_X \neq 0} \frac{\|T(x)+S(x)\|_Y}{\|x\|_X} \\
&\le \sup_{x \in X, \|x\|_X \neq 0} \frac{\|T(x)\|_Y}{\|x\|_X} + \sup_{x \in X, \|x\|_X \neq 0} \frac{\|S(x)\|_Y}{\|x\|_X} \\
&= \|T\| + \|S\|.
\end{aligned}$$

(iii) Let $Y$ be a Banach space. If $(T_j)$ is a Cauchy sequence in $B(X,Y)$, then it is bounded and so there is a constant $C > 0$ such that $\|T_j x\|_Y \le C\|x\|_X$ for all $x \in X$ and $j \in \mathbf{N}$. Since

$$\|T_j x - T_k x\|_Y \le \|T_j - T_k\|\|x\|_X \to 0 \quad \text{as} \quad j \ge k \to \infty,$$

the sequence $(T_j x)$ is a Cauchy sequence in $Y$. Nevertheless, $Y$ is a Banach space, so $T_j x$ converges to $y \in Y$: $y = \lim_{j\to\infty} T_j x = Tx$. Clearly, $T$ is linear, and $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$. This means $T \in B(X,Y)$.

Note that we have not yet proved that $T_j \to T$ in $\|\cdot\|$. But, since $(T_j)$ is Cauchy, for every $\epsilon > 0$ there is $N \in \mathbf{N}$ such that

$$j > k > N \Rightarrow \|T_j - T_k\| < \epsilon.$$

Consequently,

$$j > k > N \Rightarrow \|T_j x - T_k x\|_Y \le \epsilon\|x\|_X \quad \forall x \in X.$$

If $j \to \infty$ then

$$k > N \Rightarrow \|Tx - T_k x\|_Y \le \epsilon\|x\|_X.$$

That is to say, $\|T_k - T\| \le \epsilon$ as $k > N$. Thus, $\|T_k - T\| \to 0$ as $k \to \infty$.

In many situations it makes sense to multiply elements of a normed linear space together.

**Definition 3.1.3.** *Let $(X, \|\cdot\|_X)$ be a Banach space over $\mathbf{F}$. If there is a multiplication $(x,y) \mapsto xy$ from $X \times X \to X$ such that for any $x,y,z \in X$ and $\alpha \in \mathbf{F}$,*

(i) $x(yz) = (xy)z$;
(ii) $x(y+z) = xy + xz$;
(iii) $(x+y)z = xz + yz$;
(iv) $\alpha(xy) = (\alpha x)y = x(\alpha y)$;
(v) $\|xy\|_X \le \|x\|_X\|y\|_X$,

*then $X$ is called a Banach algebra.*

*Example 3.1.3.* (i) $(C[0,1], \|\cdot\|_\infty)$ is a Banach algebra with $(fg)(x) = f(x)g(x)$.

(ii) If $X = \mathbf{R}^n$ then by choosing a basis for $\mathbf{R}^n$ we may identify $B(\mathbf{R}^n)$ with the space of $n \times n$ real matrices.

(iii) If $X$ is a Banach space over $\mathbf{F}$, then $(B(X), \|\cdot\|)$ is a Banach algebra with $ST = S \circ T$ since

$$\|ST(x)\|_X = \|S(T(x))\|_X \le \|S\|\|T(x)\|_X \le \|S\|\|T\|\|x\|_X \quad \forall x \in X.$$

In the rest of this section, we focus on composition of linear transformations. We start with the following simple result.

**Theorem 3.1.4.** *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ and $(Z, \|\cdot\|_Z)$ be normed linear spaces over $\mathbf{F}$. If $T \in B(X,Y)$ and $S \in B(Y,Z)$ then $ST = S \circ T \in B(X,Z)$ with $\|ST\| \le \|S\|\|T\|$.*

*Proof.* Clearly, $ST$ is a linear transformation from $X$ to $Z$. Since $T$ and $S$ are continuous, we conclude that

$$\|(ST)x\|_Z = \|S(T(x))\|_Z \le \|S\|\|T(x)\|_Y \le \|S\|\|T\|\|x\|_X, \quad \forall x \in X.$$

This yields that $ST$ is continuous and so in $B(X,Z)$ with $\|ST\| \le \|S\|\|T\|$, as desired.

In fact, the composition defines a multiplication of two operators. Using this, we are now able to study inverses.

**Definition 3.1.4.** *Let $X = (X, \|\cdot\|_X)$ and $Y = (Y, \|\cdot\|_Y)$ be normed linear spaces over $\mathbf{F}$. We say that $T$ is invertible if there exists a linear transformation $S : Y \to X$ such that $TS = I_Y$ and $ST = I_X$, where $I_X$ and $I_Y$ are identity elements in $X$ and $Y$ respectively. In particular, $T \in B(X)$ is called invertible if there exists a linear operator $S \in B(X)$ with $ST = I = TS$, where $I \in B(X)$ is the identity operator. In this case, we write $S = T^{-1}$.*

*Example 3.1.4.* Define $T : \ell_2 \to \ell_2$ by $T(x_1, x_2, ...) = (0, x_1, x_2, ...)$. This map is continuous but not invertible, since it is clearly not onto. If $S : \ell_2 \to \ell_2$ is given by $S(x_1, x_2, ...) = (x_2, x_3, ...)$, then it is continuous but also not invertible, because it is clearly not one to one. Note that $ST = I \neq TS$.

**Theorem 3.1.5.** *Let $(X, \|\cdot\|_X)$ be a normed linear space over $\mathbf{F}$.*

(i) *If $T, S \in B(X)$ are invertible, then $ST \in B(X)$ is invertible and*

$$(ST)^{-1} = T^{-1}S^{-1}.$$

(ii) *If $(X, \|\cdot\|_X)$ is a Banach space and $T \in B(X)$ satisfies $\|T\| < 1$ then $I - T \in B(X)$ is invertible and*

$$(I-T)^{-1} = \lim_{n\to\infty} (1 + T + T^2 + \cdots + T^n) \quad \text{in} \quad \big(B(X), \|\cdot\|\big).$$

(iii) *If $(X, \|\cdot\|_X)$ is a Banach space and $\mathcal{I}$ stands for the class of all invertible operators in $B(X)$, then $\mathcal{I}$ is an open set in $B(X)$.*

*Proof.* (i) This follows from

$$(ST)T^{-1}S^{-1} = I = T^{-1}S^{-1}(ST).$$

(ii) If $x \in X$ then $\big((I + T + T^2 + \cdots + T^n)(x)\big)_n$ is a Cauchy sequence in $X$. This is because: $m > n$ implies

$$\begin{aligned}
\|\cdots - \cdots\| &= \|(I + T + T^2 + \cdots + T^m)(x) - (I + T + T^2 + \cdots + T^n)(x)\|_X \\
&= \|T^{n+1}(x) + \cdots + T^m(x)\|_X \\
&\le \|T^{n+1}(x)\|_X + \cdots + \|T^m(x)\|_X \\
&\le \Big(\|T^{n+1}\| + \cdots + \|T^m\|\Big)\|x\|_X
\end{aligned}$$

$$\begin{aligned}
&\le \Big(\sum_{i=n+1}^{\infty} \|T\|^i\Big)\|x\|_X \\
&= \frac{\|T\|^{n+1}}{1-\|T\|}\|x\|_X \\
&\to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

Note that $X$ is a Banach space. So the sequence converges to a limit $y \in X$. Let $y = Ax$. It is not hard to show that $A : X \to X$ is a linear operator on $X$. Furthermore, letting $m \to \infty$, we have

$$\|A(x) - (I + T + T^2 + \cdots + T^n)(x)\|_X \le \frac{\|T\|^{n+1}}{1-\|T\|}\|x\|_X \quad \forall x \in X,$$

so that $A - (I + T + T^2 + \cdots + T^n) \in B(X)$, and thus $A \in B(X)$. It is evident that

$$\|A - (I + T + T^2 + \cdots + T^n)\| \le \frac{\|T\|^{n+1}}{1-\|T\|} \to 0 \quad \text{as} \quad n \to \infty,$$

and so that $I + T + T^2 + \cdots + T^n \to A$ as $n \to \infty$. It remains to verify that $A = (I - T)^{-1}$. For any $x \in X$ we have

$$\begin{aligned}
\big((I-T)A\big)(x) &= \big((I-T)\lim_{n\to\infty}(I + T + T^2 + \cdots + T^n)\big)(x) \\
&= \big((I-T)\lim_{n\to\infty}(Ix + Tx + T^2x + \cdots + T^nx)\big) \\
&= \lim_{n\to\infty}\big(x - T^{n+1}(x)\big).
\end{aligned}$$

But

$$\|T^{n+1}(x)\|_X \le \|T\|^{n+1}\|x\|_X \to 0 \quad \text{as} \quad n \to \infty,$$

so that $T^{n+1}(x) \to 0$ as $n \to \infty$, and so $((I - T)A(x)) = x$. Similarly, we have $(A(I - T))(x) = x$ for any $x \in X$. Therefore $A = (I - T)^{-1}$.

(iii) If $T \in \mathcal{I}$, then $\|T^{-1}\| \ne 0$. We prove that the open ball

$$\mathcal{B} = \big\{S \in B(X) : \|T - S\| < \|T^{-1}\|^{-1}\big\}$$

is a subset of $\mathcal{I}$. To do so, it suffices to show that every element $S \in \mathcal{B}$ is invertible. Because of

$$\|(T - S)T^{-1}\| \le \|T - S\|\|T^{-1}\| < 1,$$

we obtain that

$$ST^{-1} = I - (T - S)T^{-1}$$

is invertible by (ii) above, and so that $S = (ST^{-1})T$ is invertible due to (i) above.

*Example 3.1.5.* If $X$ is a Banach space and $T \in B(X)$, then we may define an operator

$$e^T = I + T + \frac{1}{2!}T^2 + \frac{1}{3!}T^3 + \cdots,$$

which makes sense since

$$\|e^T\| \leq 1 + \|T\| + \frac{1}{2!}\|T\|^2 + \frac{1}{3!}\|T\|^3 + \cdots = \exp(\|T\|).$$

This is particularly useful in linear systems theory and control theory; if $x(t) \in \mathbf{R}^n$ then $dx/dt = Ax(t), x(0) = x_0$, where $A$ is an $n \times n$ matrix, has a solution $x(t) = e^{At}x_0$.

## 3.2 Uniform Boundedness, Open Mapping and Closed Graph

In this section, we proceed to discuss the uniform boundedness principle or the Banach-Steinhaus theorem, open mapping theorem and closed graph theorem.

The first is the so-called uniform boundedness principle.

**Theorem 3.2.1.** *Let $(X, \|\cdot\|_X)$ be a Banach space and $(Y, \|\cdot\|_Y)$ be a normed linear space. Let $\{T_\alpha\}$ be a family of bounded linear operators from $X$ to $Y$. If $\sup_\alpha \|T_\alpha x\|_Y < \infty$ for each $x \in X$, then $\sup_\alpha \|T_\alpha\| < \infty$.*

*Proof.* Suppose that there are two constants $C, \delta > 0$ and a point $x_0 \in X$ such that

$$\sup_{\alpha; \|x-x_0\|_X < \delta} \|T_\alpha x\|_Y \leq C. \tag{3.1}$$

Then it is possible to find a uniform bound on $\{\|T_\alpha\|\}$. In fact, for any $y \in X$ with $y \neq 0$, define

$$z = \frac{\delta}{2\|y\|_X}y + x_0.$$

Then $\|z - x_0\|_X < \delta$ and hence by (3.1) one gets $\|T_\alpha z\|_Y \leq C$. Furthermore, the linearity of $T_\alpha$ and the triangle inequality of norm yield

$$\frac{\delta}{2\|y\|_X}\|T_\alpha y\|_Y - \|T_\alpha x_0\|_Y \leq \left\|\frac{\delta}{2\|y\|_X}T_\alpha y + T_\alpha x_0\right\|_Y = \|T_\alpha z\|_Y \leq C,$$

which gives

$$\|T_\alpha y\|_Y \leq \frac{2(C + \|T_\alpha x_0\|_Y)}{\delta}\|y\|_X \leq \frac{4C\|y\|_X}{\delta},$$

Of course, it follows that $\|T_\alpha\| \leq \frac{4C}{\delta}$, as desired.

To end the argument we have to verify that (3.1) holds. This can be done by a contradiction argument. As a matter of fact, assume that (3.1) fails. Fix an arbitrary ball $B_0 \subset X$. By assumption there is a point $x_1 \in B_0$ such that $\|T_{\alpha_1}x_1\|_Y > 1$ for some index $\alpha_1$. Since each $T_\alpha$ is bounded and hence continuous, there exists a ball $B_{\delta_1}(x_1) \subset B_0$ such that $0 < \delta_1 < 1$ and $\|T_{\alpha_1}(x)\|_Y > 1$ for

$x \in B_{\delta_1}(x_1)$. By assumption, $\{T_\alpha x\}$ is not bounded on $B_{\delta_1}(x_1)$, so there exists a point $x_2 \in B_{\delta_1}(x_1)$ with $\|T_{\alpha_2}x_2\|_Y > 2$ for some index $\alpha_2 \neq \alpha_1$. Continue in the same way: by continuity of $T_{\alpha_2}$ there is a ball $B_{\delta_2}(x_2) \subset B_{\delta_1}(x_1)$ such that $0 < \delta_2 < 2^{-1}$ and $\|T_{\alpha_2}x\|_Y > 2$ when $x \in B_{\delta_2}(x_2)$.

Repeating this process produces points $x_3, x_4, ...$, different indices $\alpha_3, \alpha_4, ...$, and positive numbers $\delta_3, \delta_4, ...$ such that

$$B_{\delta_n}(x_n) \subset B_{\delta_{n-1}}(x_{n-1}), \quad 0 < \delta_n < \frac{1}{n} \quad \text{and} \quad \|T_{\alpha_n}x\|_Y > n \quad \forall x \in B_{\delta_n}(x_n).$$

Consequently,

$$m > n \Longrightarrow B_{\delta_m}(x_m) \subset \cdots \subset B_{\delta_n}(x_n).$$

This gives that the sequence $(x_n)$ is a Cauchy sequence and thus converges to some point $z \in X$ since $X$ is a Banach space. The continuity of $T_{\alpha_n}$ implies $\|T_{\alpha_n}z\|_Y \geq n$ which contradicts the hypothesis that $\{T_\alpha z\}$ is bounded.

**Definition 3.2.1.** *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces.*

(i) *A sequence $(T_n)$ in $B(X,Y)$ is uniformly convergent if there is a $T \in B(X,Y)$ such that* $\lim_{n\to\infty} \|T_n - T\| = 0$.

(ii) *A sequence $(T_n)$ in $B(X,Y)$ is strongly convergent if for any $x \in X$, the sequence $(T_nx)$ is convergent in $Y$. Moreover, if there is a $T \in B(X,Y)$ such that* $\lim_{n\to\infty} \|T_nx - Tx\|_Y = 0$ *for all $x \in X$, then $(T_n)$ is strongly convergent to $T$.*

*Example 3.2.1.* Clearly, the uniform convergence implies the strong convergence, but not conversely. Consider $\ell_p$, $p \in [1, \infty)$. For each $n \in \mathbf{N}$ define

$$T_nx = (x_n, x_{n+1}, ...), \quad x = (x_1, x_2, ...) \in \ell_p.$$

Then $T_n$ is in $B(\ell_p)$ with $\|T_n\| \leq 1$. Note that if $x = (x_1, x_2, ...) \in \ell_p$ then

$$\|T_nx\|_p = \left(\sum_{j=n}^{\infty} |x_j|^p\right)^{\frac{1}{p}} \to 0 \quad \text{as} \quad n \to \infty.$$

So $(T_n)$ is strongly convergent to 0. On the other hand, if $e_n = (0, ..., 0, 1, 0, ...)$ then $\|e_n\|_p = 1$ and $T_ne_n = (1, 0, 0, ...)$ and hence $\|T_n\| \geq \|T_ne_n\|_p = 1$. This shows that $(T_n)$ is not uniformly convergent to 0.

A special consequence of the uniform bounded principle that is quite useful is the following.

**Theorem 3.2.2.** *Let $(X, \|\cdot\|_X)$ be a Banach space and $(Y, \|\cdot\|_Y)$ be a normed linear space. If a sequence $(T_n)$ in $B(X,Y)$ is strongly convergent, then there exists $T \in B(X,Y)$ such that $(T_n)$ is strongly convergent to $T$.*

*Proof.* Let $(T_n)$ in $B(X,Y)$ be strongly convergent. Then for each $x \in X$ the sequence $(T_n x)$ is convergent in $Y$ and hence defines a linear operator $T$ on $X$: $Tx = \lim_{n\to\infty} T_n x$. The key is to show that $T \in B(X,Y)$. Note that $(T_n x)$ is bounded in $Y$. So, from the uniform bounded principle it turns out that there is a constant $C > 0$ such that $\sup_n \|T_n\| \le C$. Hence $\|T_n x\|_Y \le C\|x\|_X$ for all $x \in X$. This implies $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$, showing that $T$ is bounded. The definition of $T$ means that $(T_n)$ converges strongly to $T$.

The second is to establish the open mapping theorem. To do so, we need the Baire category theorem.

**Definition 3.2.2.** *Let $(X, \|\cdot\|_X)$ be a normed linear space.*

(i) *A subset $S \subset X$ is nowhere dense if $B_\epsilon(x) \cap (X \setminus \bar{S}) \neq \emptyset$ for every point $x$ in $\bar{S}$-the closure of $S$, and for every open ball $B_\epsilon(x) = \{y \in X : \|y - x\|_X < \epsilon\}$.*

(ii) *The diameter of $S \subset X$ is defined by $diam(S) = \sup_{x,y\in S} \|x - y\|_X$.*

**Theorem 3.2.3.** *Let $(X, \|\cdot\|_X)$ be a Banach space.*

(i) *If $\{F_n\}$ be a decreasing sequence of non-empty closed sets; that is, $X \supset F_n \supset F_{n+1} \quad \forall n \in \mathbf{N}$, and if $\lim_{n\to\infty} diam(F_n) = 0$, then there exists uniquely one point in $\bigcap_{n=1}^\infty F_n$.*

(ii) *$X$ cannot be written as a countable union of nowhere dense sets.*

*Proof.* (i) If $x, y \in \bigcap_{n=1}^\infty F_n$, then $\|x-y\|_X \le \mathrm{diam}(F_n) \to 0$ as $n \to \infty$ and hence $x = y$. It follows that there can be no more than one point in the intersection.

Now choose a point $x_n \in F_n$ for each $n \in \mathbf{N}$. Then $\|x_n - x_m\|_X \le \mathrm{diam}(F_n) \to 0$ as $m \ge n \to \infty$. Thus $(x_n)$ is Cauchy, so has a limit $x$ say by completeness. Each $F_n$ is closed and contains $x_m$ with $m \ge n$, so $x \in F_n$. It follows that $x \in \bigcap_{n=1}^\infty F_n$.

(ii) Suppose $X = \bigcup_{j=1}^\infty X_j$, where $X_j$ is nowhere dense; that is, $\bar{X}_j$ has empty interior. Fix a ball $B_1(x_0)$. Since $\bar{X}_1$ does not contain $B_1(x_0)$, there must be a point $x_1 \in B_1(x_0)$ with $x_1 \notin \bar{X}_1$. It follows that there is a ball $B_{r_1}(x_1)$ such that $\overline{B_{r_1}(x_1)} \subset B_1(x_0)$ and $\overline{B_{r_1}(x_1)} \bigcap \bar{X}_1 = \emptyset$. Assume without loss of generality that $r_1 < 1/2$.

Similarly, there is a point $x_2$ and a radius $r_2$ such that $\overline{B_{r_2}(x_2)} \subset B_{r_1}(x_1)$ and $\overline{B_{r_2}(x_2)} \cap \bar{X}_2 = \emptyset$, and without loss of generality $r_2 < 1/3$. Note that $\overline{B_{r_2}(x_2)} \cap \bar{X}_1 = \emptyset$ automatically since $\overline{B_{r_2}(x_2)} \subset B_{r_1}(x_1)$.

Inductively, we obtain a sequence of decreasing closed balls $\overline{B_{r_n}(x_n)}$ such that $\overline{B_{r_n}(x_n)} \cap \bar{X}_j = \emptyset$ for $1 \le j \le n$ and $r_n \to 0$ as $n \to \infty$.

Now by (i), there must be a point $x$ in $\cap_{n=1}^\infty \overline{B_{r_n}(x_n)}$, so $x \notin \bar{X}_j$ for all $j \ge 1$. This yields that $x \notin \cup_{j=1}^\infty \bar{X}_j = X$, a contradiction.

Recall that a continuous map between normed linear spaces has the property that the pre-image of any open set is open, but in general the image of an open set is not open. Bounded linear operators between Banach spaces cannot do this. This is the content of the open mapping theorem as follows.

**Theorem 3.2.4.** *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. If $T$ is a bounded linear operator from $X$ onto $Y$, then $T(U)$ is open in $Y$ whenever $U$ is open in $X$.*

*Proof.* We split the proof of the theorem into three steps. In what follows, denote by $B_r^X(x)$ and $B_r^Y(y)$ the balls of radius $r > 0$ centered at $x \in X$ and $y \in Y$, respectively.

*Step 1.* We prove that for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$T(\overline{B_{2\epsilon}^X}) \supset B_\delta^Y.$$

To see this, note that

$$X = \bigcup_{n=1}^\infty nB_\epsilon^X = \bigcup_{n=1}^\infty \{nx : x \in B_\epsilon^X\},$$

and $T \in B(X, Y)$ is surjective. So we have

$$Y = T(X) = \bigcup_{n=1}^\infty nT(B_\epsilon^X) = \bigcup_{n=1}^\infty \{ny : y \in T(B_\epsilon^X)\}.$$

Since $(Y, \|\cdot\|_Y)$ is a Banach space, we conclude that from Theorem 3.2.3 (ii) that some $nT(B_\epsilon^X)$ is nowhere dense in $Y$, and so that

$$nT(B_\epsilon^X) = \{ny : y \in T(B_\epsilon^X)\} \supset B_r^Y(z) = \{y \in Y : \|y - z\|_Y < r\}$$

for some $z \in Y$ and $r > 0$. Thus $T(\overline{B_\epsilon^X})$ must contain the ball $B_\delta^Y(y_0)$ where $y_0 = \frac{z}{n}$ and $\delta = \frac{r}{n}$. It follows that the set

$$V = \{y_1 - y_2 : \; y_1, y_2 \in B_\delta^Y(y_0)\}$$

is contained in $T(\bar{U})$, where

$$U = \{x_1 - x_2 : \; x_1, x_2 \in B_\epsilon^X\} \subset B_{2\epsilon}^X.$$

Thus, $T(\overline{B_{2\epsilon}^X}) \supset V$. Any point $y \in B_\delta^Y$ can be written as $y = (y + y_0) - y_0$, so $B_\delta^Y \subset V$, as desired.

*Step 2.* We further prove that for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$T(B_{2\epsilon}^X) \supset B_\delta^Y.$$

To do so, choose $(\epsilon_n)$ with $\epsilon_n > 0$ and $\sum_{n=1}^\infty \epsilon_n < \epsilon$. By Step 1 there is a sequence $(\delta_n)$ such that $T(\overline{B_{2\epsilon_n}^X}) \supset B_{\delta_n}^Y$. Without loss of generality, we may assume that $\lim_{n\to\infty} \delta_n = 0$.

Let $y \in B_{\delta_1}^Y$. Then there is a point $x_1 \in B_{2\epsilon_1}^X$ with $\|y - Tx_1\|_Y < \delta_2$. Since $y - Tx_1 \in B_{\delta_2}^Y$, we conclude that there is a point $x_2 \in B_{2\epsilon_2}^X$ such that $\|y - Tx_1 - Tx_2\|_Y < \delta_3$. Continuing, we obtain a sequence $(x_n)$ such that $x_n \in B_{2\epsilon_n}^X$ and

$$\Big\|y - T\big(\sum_{k=1}^n x_k\big)\Big\|_Y < \delta_{n+1}.$$

Since $\|x_n\|_X < 2\epsilon_n$, we conclude that $\sum_{n=1}^{\infty} x_n$ is absolutely convergent and hence convergent to $x = \sum_{n=1}^{\infty} x_n$. This implies

$$\|x\|_X \leq \sum_{n=1}^{\infty} \|x_n\|_X < 2 \sum_{n=1}^{\infty} \epsilon_n < 2\epsilon.$$

The map $T$ is continuous, so $y = Tx$ since $\delta_n \to 0$. In other words, for any $y \in B_\delta^Y$ ($\delta = \delta_1$) we have found a point $x \in B_{2\epsilon}^X$ such that $Tx = y$, implying the desired inclusion.

*Step 3.* We prove that if $G$ is open in $X$ then for any point $x \in G$, there exists a $\delta > 0$ such that

$$B_\delta^Y\big(Tx\big) \subset T(G).$$

In fact, if $x \in G$, then there exists an $\epsilon > 0$ such that $B_{2\epsilon}^X(x) \subset G$. By Step 2, we have $T(B_{2\epsilon}^X) \supset B_\delta^Y$ for some $\delta > 0$. Hence

$$T(G) \supset T(x + B_{2\epsilon}^X) = T(x) + T(B_{2\epsilon}^X) \supset Tx + B_\delta^Y = B_\delta^Y(Tx).$$

Of course, this step completes the proof of the theorem.

As an application of the open mapping theorem we establish a general property of inverse maps.

**Definition 3.2.3.** *Let $T : X \to Y$ be an injective linear operator. Define the inverse of $T$, $T^{-1}$ by requiring that $T^{-1}y = x$ if and only if $Tx = y$. Then the domain of $T^{-1}$ is a linear subspace of $Y$ and $T^{-1}$ is a linear operator. Moreover, $T^{-1}Tx = x\, \forall x \in X$ and $TT^{-1}y = y$ for all $y$ in the domain of $T^{-1}$.*

**Lemma 3.2.1.** *Let $X$ and $Y$ be Banach spaces, and let $T$ be an injective bounded linear map from $X$ to $Y$. Then $T^{-1}$ is a bounded linear map.*

*Proof.* It suffices to prove the continuity of $T^{-1}$. Since $T = (T^{-1})^{-1}$ maps Banach space $X$ to Banach space $Y$, we conclude from Theorem 3.2.4 that $T$ maps open sets in $X$ to open sets in $Y$. This amounts to saying that $T^{-1}$ is continuous.

**Corollary 3.2.1.** *Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(1)}$ be two norms defined on a Banach space $X$. If there is a constant $C_1 > 0$ such that $\|x\|_{(1)} \leq C_1\|x\|_{(2)}$ $\forall x \in X$, then there exists another constant $C_2 > 0$ such that $\|x\|_{(2)} \leq C_2\|x\|_{(1)}$ $\forall x \in X$. Consequently, both norms are equivalent.*

*Proof.* Consider the identity operator: $I : (X, \|\cdot\|_{(2)}) \to (X, \|\cdot\|_{(1)})$; $Ix = x\, \forall x \in X$. Clearly, $I$ is bounded. By Lemma 3.2.1 $I^{-1}$ is also bounded, giving the norm inequality in the other direction.

**Definition 3.2.4.** *Given two normed linear spaces $X$ and $Y$, let $T : X \to Y$ be a linear operator. Then we define the graph of $T$ to be*

$$G(T) = \{(x, y) \in X \times Y : y = Tx\}.$$

*Moreover we say that $G(T)$ is closed if $G(T)$ is a closed subset of $X \times Y$.*

The forthcoming result is called the closed graph theorem.

**Theorem 3.2.5.** *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Then a linear operator $T : X \to Y$ is bounded if and only if $G(T)$ is closed.*

*Proof.* We initially observe that $X \times Y$ is a Banach space under (among others) the norm $\|(x, y)\| = \|x\|_X + \|y\|_Y$ for $x \in X$ and $y \in Y$. Addition and scalar multiplication are defined in the expected manner:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad \text{and} \quad \alpha(x, y) = (\alpha x, \alpha y).$$

The completeness of $(X \times Y, \|\cdot\|)$ follows readily from the hypothesis.

On the one hand, suppose that $T : X \to Y$ is bounded. To prove that $G(T)$ is closed subset of $X \times Y$ it is enough to show that $G(T)$ is (sequentially) closed. Accordingly, assume that $x_n \to x$ in $X$ and $Tx_n \to y$ in $Y$. The boundedness of $T$ implies $Tx_n \to Tx$ in $Y$. Note that $Y$ is Banach space. So $Tx = y$. This means $(x, y) \in G(T)$. Thus $G(T)$ is closed.

On the other hand, suppose that $G(T)$ is closed. To verify that $T$ is bounded, we consider the projection map $P$ from $G(T)$ onto $X$ via: $P(x, Tx) = x$. Clearly, $P$ is linear, bijective and bounded. From Lemma 3.2.1 it turns out that $P^{-1}$ is bounded linear map from $X$ to $G(T)$, so there is a constant $C > 0$ such that

$$\|x\|_X + \|Tx\|_Y = \|(x, Tx)\| = \|P^{-1}x\| \le C\|x\|_X \quad \forall x \in X.$$

Consequently, $T$ is bounded.

## 3.3 Hahn-Banach Theorem

**Definition 3.3.1.** *Let $X$ be a linear space over $\mathbf{F}$. A linear map from $X$ to $\mathbf{F}$ is called a linear functional on $X$. If $X$ is a normed linear space, then $B(X, \mathbf{F}) = X^*$ is called the dual space of $X$.*

The question is whether or not $X^*$ consists only of zero functional for a given normed linear space $X$. This is answered in great generality using the Hahn-Banach theorem. First, let us give the following Hahn-Banach lemma.

**Lemma 3.3.1.** *Let $X$ be a real linear space, and $p : X \to \mathbf{R}$ a continuous function with*

$$p(x + y) \le p(x) + p(y) \quad \text{and} \quad p(\lambda x) = \lambda p(x) \quad \forall \lambda \ge 0, x, y \in X.$$

*If $Y$ is a subspace of $X$ and $f$ is a real-valued linear functional on $Y$ with $f(x) \le p(x)$ $\forall x \in Y$, then there is a real-valued linear functional $F$ on $X$ such that $F(x) = f(x)$ $\forall x \in Y$ and $F(x) \le p(x)$ $\forall x \in X$.*

*Proof.* Assume that $\mathcal{K}$ is the set of all pairs $(Y_\alpha, g_\alpha)$ in which $Y_\alpha$ is a linear subspace of $X$ containing $Y$, and $g_\alpha$ is a real linear functional on $Y_\alpha$ with

$$g_\alpha(x) = f(x) \quad \forall x \in Y, \quad g_\alpha(x) \le p(x) \quad \forall x \in Y_\alpha.$$

Make $\mathcal{K}$ into a partially ordered set by defining the relation $(Y_\alpha, g_\alpha) \preceq (Y_\beta, g_\beta)$ if $Y_\alpha \subset Y_\beta$ and $g_\alpha = g_\beta$ on $Y_\alpha$. Clearly, any totally ordered subset $\{(Y_\lambda, g_\lambda)\}$ (for which at least one of $(Y_\alpha, g_\alpha) \preceq (Y_\beta, g_\beta)$ and $(Y_\beta, g_\beta) \preceq (Y_\alpha, g_\alpha)$ holds) has an upper bound $\bigcup_\lambda Y_\lambda$ on which the functional is given by $g_\lambda$ on each $Y_\lambda$. By Zorn's lemma – if $S$ is a partially ordered set in which every totally ordered subset has an upper bound then $S$ has a maximal element, we find that there is a maximal element $(Y_0, g_0)$ in $\mathcal{K}$. The proof will be completed if $Y_0 = X$ and hence $F = g_0$.

If $Y_0 \neq X$, then there is $y_1 \in X \setminus Y_0$. Let $Y_1$ be the linear space spanned by $Y_0$ and $y_1$; that is,

$$Y_1 = \{x = y + \lambda y_1 : \ y \in Y_0, \ \lambda \in \mathbf{R}\}.$$

Note that if $x, y \in Y_0$ then

$$g_0(y) - g_0(x) = g_0(y - x) \le p(y - x) \le p(y + y_1) + p(-y_1 - x)$$

and hence

$$-p(-y_1 - x) - g_0(x) \le p(y + y_1) - g_0(y).$$

It follows that

$$A = \sup_{x \in Y_0} \big\{ -p(-y_1 - x) - g_0(x) \big\} \le \inf_{y \in Y_0} \big\{ p(y + y_1) - g_0(y) \big\} = B.$$

Now for any number $c \in [A, B]$ define $g_1(y + \lambda y_1) = g_0(y) + \lambda c$. Then $g_1$ is clearly linear, and $g_1(y) = g_0(y) \le p(y)$ when $y \in Y_0$. Moreover, if $\lambda > 0$ and $y \in Y_0$, then

$$\begin{aligned} g_1(y + \lambda y_1) &= \lambda \Big( g_0(\frac{y}{\lambda}) + c \Big) \\ &\le \lambda \Big( g_0(\frac{y}{\lambda}) + p(\frac{y}{\lambda} + y_1) - g_0(\frac{y}{\lambda}) \Big) \\ &= \lambda p(\frac{y}{\lambda} + y_1) \\ &= p(y + \lambda y_1), \end{aligned}$$

whereas if $\lambda < 0$, then

$$\begin{aligned} g_1(y + \lambda y_1) &= |\lambda| \Big( g_0(\frac{y}{|\lambda|}) - c \Big) \\ &\le |\lambda| \Big( g_0(\frac{y}{|\lambda|}) - g_0(\frac{y}{|\lambda|}) + p(-y_1 + \frac{y}{|\lambda|}) \Big) \\ &= |\lambda| p\big( - y_1 + \frac{y_1}{|\lambda|} \big) \\ &= p(y + \lambda y_1). \end{aligned}$$

Consequently,

$$g_1(y+\lambda y_1) = g_0(y) + \lambda c \le p(y+\lambda y_1) \quad \forall \lambda \in \mathbf{R},\ y \in Y_0.$$

This to say, $(Y_1, g_1) \in \mathcal{K}$ and $(Y_0, g_0) \preceq (Y_1, g_1)$ with $Y_0 \neq Y_1$. This contradicts the maximality of $(Y_0, g_0)$.

The following is the Hahn-Banach theorem over $\mathbf{R}$.

**Theorem 3.3.1.** *Let $(X, \|\cdot\|_X)$ be a real normed space, and $Y$ a linear subspace of $X$. Then to any $f \in B(Y, \mathbf{R})$ there corresponds an $F \in B(X, \mathbf{R})$ such that*

$$\|F\| = \|f\| \quad \textit{and} \quad F(y) = f(y) \quad \forall y \in Y.$$

*Proof.* Given $f \in B(Y, \mathbf{R})$, let $p(x) = \|f\|\|x\|_X$ $\forall x \in X$. Then

$$f(x) \le \|f\|\|x\|_X = p(x) \quad \forall x \in Y.$$

And hence from Lemma 3.3.1 it turns out that there is an extension $F \in B(X, \mathbf{R})$ with $F = f$ on $Y$ and $F \le p$ on $X$. It is clear that

$$\|f\| = \sup_{y\in Y,\ \|y\|_X=1} |f(y)| = \sup_{y\in Y,\ \|y\|_X=1} |F(y)| \le \|F\|.$$

In order to verify the reverse inequality, we write $F(x) = \theta|F(x)|$ for $\theta = \pm 1$. Then

$$|F(x)| = \theta F(x) = F(\theta x) \le p(\theta x) = \|f\|\|\theta x\|_X = \|f\|\|x\| \quad \forall x \in X;$$

that is, $\|F\| \le \|f\|$. Therefore $\|F\| = \|f\|$. The proof is complete.

As one of the most important results in functional analysis, the Hahn-Banach theorem has many useful consequences of which some are given below.

**Corollary 3.3.1.** *Let $(X, \|\cdot\|_X)$ be a real normed linear space. Then*

(i) *If $Y$ is a linear subspace of $X$ and $x_0 \in X$ satisfies $\inf_{y\in Y} \|y - x_0\|_X = d > 0$, then there is an $F \in B(X, \mathbf{R})$ such that*

$$F(x_0) = 1, \quad \|F\| = d^{-1}, \quad F(y) = 0 \quad \forall y \in Y.$$

(ii) *If $Y$ is a linear subspace of $X$ and is not dense in $X$, then there is a nonzero $F \in B(X, \mathbf{R})$ such that $F(y) = 0$ $\forall y \in Y$.*

(iii) *If $x \neq 0$ in $X$ then there is an $F \in B(X, \mathbf{R})$ such that $\|F\| = 1$ and $F(x) = \|x\|_X$.*

(iv) *If $y, z \in X$ and $y \neq z$, then there is an $F \in B(X, \mathbf{R})$ such that $F(y) \neq F(z)$.*

(v) $\|x\|_X = \sup_{F\neq 0} \frac{|F(x)|}{\|F\|} = \sup_{\|F\|=1} |F(x)|$.

(vi) *If $\mathcal{N}_F = \{x \in X : F(x) = 0\}$, then there exists a one-dimensional subspace $Y$ of $X$ such that $X = \mathcal{N}_F + Y$ and $\mathcal{N}_F \cap Y = \{0\}$.*

*Proof.* (i) Let $Y_1$ be the linear space spanned by $Y$ and $x_0$. Since $x_0 \notin Y$, every point $x \in Y_1$ may be written uniquely as $x = y + \lambda x_0$, with $y \in Y$, $\lambda \in \mathbf{R}$. Define a linear functional $f \in B(Y_1, \mathbf{R})$ by $f(y + \lambda x_0) = \lambda$. Then $f(y) = 0$ and $f(x_0) = 1$. If $\lambda \neq 0$ and $x = y + \lambda x_0$, then

$$\|x\|_X = \|y + \lambda x_0\|_X = |\lambda| \left\| \lambda^{-1} y + x_0 \right\|_X \geq |\lambda| d = |f(x)| d,$$

and hence $\|f\| \leq d^{-1}$. Pick a sequence $(y_n)$ in $Y$ with $\|x_0 - y_n\|_X \to d$ as $n \to \infty$. Then

$$1 = f(x_0 - y_n) \leq \|f\| \|x_0 - y_n\|_X \to d\|f\|,$$

so $\|f\| \geq d^{-1}$. Therefore $\|f\| = d^{-1}$. Accordingly, a direct application of the Hahn-Banach theorem produces an $F \in B(X, \mathbf{R})$ such that $F(x) = f(x)$ as $x \in Y_1$ and $\|F\| = \|f\|$, as desired.

(ii) Since $Y$ is not dense in $X$, we conclude that there is an $x_0 \in X$ such that $\inf_{y \in Y} \|y - x_0\|_X = d > 0$. An application of (i) produces the conclusion in (ii).

(iii) Just apply (i) with $Y = \{0\}$ to get $f \in B(X, \mathbf{R})$ such that $\|f\| = \|x\|_X^{-1}$, $f(x) = 1$. We may then take $F = \|x\|_X f$.

(iv) Apply (iii) to $x = y - z$.

(v) Clearly, we have $\sup_{\|F\|=1} |F(x)| \leq \|x\|_X$. By (iii), for $x \neq 0$ there is an $f$ such that $f = \|x\|_X$ and $\|f\| = 1$, so $\sup_{\|F\|=1} |F(x)| = \|x\|_X$.

(vi) If $F \neq 0$ then there is a point $x_0 \neq 0$ such that $F(x_0) = 1$. Note that any element $x \in X$ can then be written as $x = x - \lambda x_0 + \lambda x_0$ with $\lambda = F(x)$. So, if $Y = \{\lambda x_0 : \lambda \in \mathbf{R}\}$ then the desired decomposition follows right away. It is clear that $Y$ is the one-dimensional space spanned by $x_0$. If $x \in \mathcal{N}_F \cap Y$, then $x = \lambda x_0$ and $0 = F(x) = \lambda F(x_0) = \lambda$ and hence $x = 0$. This completes the proof.

Of course, Lemma 3.3.1, Theorem 3.3.1 and Corollary 3.3.1 are valid for $\mathbf{C}$.

## Exercises

**3.1** Prove that if $C[0,1]$ is equipped with the sup-norm and $T : C[0,1] \to \mathbf{R}$ is given by $T(f) = f(0)$, then $T$ is bounded with $\|T\| = 1$.

**3.2** (i) Suppose the infinite matrix $(a_{i,j})$ satisfies $\sup_{i \in \mathbf{N}} \sum_{j=1}^{\infty} |a_{i,j}| < \infty$. Define an operator

$$T : x = (x_1, x_2, ...) \mapsto y = Tx = \left( \sum_{j=1}^{\infty} a_{1,j} x_j, \sum_{j=1}^{\infty} a_{2,j} x_j, ... \right).$$

Prove that $T : \ell_\infty \to \ell_\infty$ is bounded and $\|T\| = \sup_{i \in \mathbf{N}} \sum_{j=1}^{\infty} |a_{i,j}|$.

(ii) Let $T : \ell_2 \to \ell_2$ be defined by $T(x) = (0, x_1, x_2, ...)$ for $x = (x_1, x_2, ...) \in \ell_2$. Prove that $\|Tx\|_2 = \|x\|_2$.

(iii) Let $R^1[a,b]$ be the space of all Riemann integrable functions $f$ on $[a,b]$ with $\|f\|_1 = \int_a^b |f(x)| dx < \infty$. Define $Tf(x) = \int_a^x f(t) dt$. Prove $T$ is a bounded linear operator from $R^1[a,b]$ to itself with $\|T\| = b - a$.

**3.3** Let $C(0,1)$ be the space of all real-valued continuous functions on $(0,1)$. Equip $C(0,1)$ with 2-norm: $\|f\|_2 = \left(\int_0^1 |f(t)|^2 dt\right)^{\frac{1}{2}}$. Define $T : C(0,1) \to C(0,1)$ by $T(f)(t) = tf(t)$ for $t \in (0,1)$. Prove that $T$ is bounded but not invertible.

**3.4** Show by an example that the uniform bounded principle does not hold once the completeness is dropped.

**3.5** (i) Prove $\lim_{n\to\infty} \int_0^{2\pi} |\sin(n+\frac{1}{2})x||\sin\frac{x}{2}|^{-1}dx = \infty$.

(ii) Given a Riemann-integrable function $f : (0, 2\pi) \to \mathbf{R}$, let its Fourier series be

$$s(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx}, \quad \text{where} \quad a_m = \frac{1}{2\pi}\int_0^{2\pi} f(y)e^{-imy}dy.$$

Extend the definition of $f$ to make it $2\pi$periodic. Define the $n$-th partial sum of the Fourier series to be $s_n(x) = \sum_{m=-n}^{n} a_m e^{imx}$. Prove

$$s_n(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x+y)\frac{\sin(n+\frac{1}{2})y}{\sin\frac{y}{2}}dy \quad \forall x \in (0, 2\pi).$$

(iii) Let $X$ be the Banach space of continuous functions $f : [0, 2\pi] \to \mathbf{R}$ with $f(0) = f(2\pi)$, with the sup-norm. Prove that the linear operator $T_n : X \to \mathbf{R}$ defined by

$$T_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}dx$$

is bounded, and

$$\|T_n\| = \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right| dx.$$

(iv) Prove that there exists a continuous function $f : [0, 2\pi] \to \mathbf{R}$ with $f(0) = f(2\pi)$ such that its Fourier series diverges at $x = 0$.

**3.6** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be given by $T(x,y) = (x, 0)$. Prove that $T$ is linear, bounded, but not onto, and cannot map open sets to open sets in $\mathbf{R}^2$.

**3.7** Let $X = C^1[0,1]$ and $Y = C[0,1]$, both equipped with the sup-norm. Prove:

(i) $X$ is not complete.

(ii) The map $(d/dx) : X \to Y$ is closed but not bounded.

**3.8** Let $(X, \|\cdot\|)$ be normed linear space over $\mathbf{C}$. Prove:

(i) If $f$ is a complex linear functional and $u = \Re f$ then $u$ is a real linear functional and $f(x) = u(x) - iu(ix)$ for all $x \in X$. Conversely, if $u$ is a real linear functional and $f(x) = u(x) - iu(ix)$, then $f$ is a complex linear functional. In this case, $\|u\| = \|f\|$.

(ii) If $Y$ a linear subspace, $p$ is a seminorm on $X$ (i.e., a nonnegative function on $X$ with properties: $p(x+y) \le p(x)+p(y)$ and $p(\alpha x) = |\alpha|p(x)$ for all $x, y \in X$ and $\alpha \in \mathbf{C}$) and $f$ is a complex linear functional on $Y$ such that $|f(x)| \le p(x)$ for $x \in Y$, then there is a linear functional $F$ on $X$ such that $|F(x)| \le p(x)$ $\forall x \in X$ and $F(x) = f(x)$ $\forall x \in Y$.

# 4. Lebesgue Measures, Integrals and Spaces

In this chapter, we set forth the basic concepts of Lebesgue measure and its integration, but also deal with $L_p$ spaces – an interesting and important class of Banach spaces of functions whose norm are defined in terms of Lebesgue integrals.

## 4.1 Measurable Sets and Functions

We begin with the notion of a $\sigma$-algebra.

**Definition 4.1.1.** *A non-empty collection* $\mathbf{A}$ *of subsets of a set* $X$ *is called a* $\sigma$*-algebra provided:*

(i) $\emptyset$, $X$ *belong to* $\mathbf{A}$*;*

(ii) $S \in \mathbf{A}$ *implies* $S^c = X \setminus S \in \mathbf{A}$*;*

(iii) *If* $(S_j)$ *is a sequence of sets in* $\mathbf{A}$*, then* $\bigcup_{j=1}^{\infty} S_j$ *belongs to* $\mathbf{A}$*.*

Note that these assumptions imply that if $(S_j)$ is a sequence of sets in $\mathbf{A}$, then $\bigcap_{j=1}^{\infty} S_j \in \mathbf{A}$, and if $S_1, S_2 \in \mathbf{A}$ then $S_1 \setminus S_2 = S_1 \cap S_2^c \in \mathbf{A}$.

It is easy to see that any family $\mathcal{F}$ of subsets of $X$ can be extended to a sigma-algebra just taking the sigma-algebra consisting of all subsets of $X$. Among all these extensions there is a special one. Consider all the sigma-algebras that contain $\mathcal{F}$ and take their intersection, denoted $\Sigma$, i.e., a subset $S \subset X$ is in $\Sigma$ if and only if $S$ is in every sigma-algebra containing $\mathcal{F}$. Clearly, $\Sigma$ is indeed a sigma-algebra. Actually, it is the smallest sigma-algebra containing $\mathcal{F}$–the sigma-algebra generated by $\mathcal{F}$. An important example is the sigma-algebra $\mathcal{B}$ of Borel sets of $\mathbf{R}^n$ which is generated by the open subsets of $\mathbf{R}^n$. Alternatively, it is generated by the open balls in $\mathbf{R}^n$. But, it has been proved that $\mathcal{B}$ does not contain all subsets of $\mathbf{R}^n$.

**Definition 4.1.2.** *Given a* $\sigma$*-algebra* $\mathbf{A}$*, a measure* $\mu$ *on* $\mathbf{A}$ *is a function from* $\mathbf{A}$ *to* $[0, \infty]$ *such that*

(i) $\mu(\emptyset) = 0$*;*

(ii) $\mu(S) \geq 0$ *for any* $S \in \mathbf{A}$*;*

(iii) $\mu$ *is countably additive in the sense that if* $(S_j)$ *is any sequence of disjoint sets in* $\mathbf{A}$*, then*

$$\mu\Big(\bigcup_{j=1}^{\infty} S_j\Big) = \sum_{j=1}^{\infty} \mu(S_j).$$

Because $\mu$ is allowed to take $\infty$, $\sum_{j=1}^{\infty} \mu(S_j)$ may be a divergent series. If a measure does not take $\infty$, then it is called a finite measure. In addition, if $\mathbf{A} = \mathcal{B}$ then $\mu$ is called a Borel measure.

**Lemma 4.1.1.** *Let $\mu$ be a measure defined on a $\sigma$-algebra $\mathbf{A}$.*

(i) *If $S_1, S_2 \in \mathbf{A}$ and $S_1 \subset S_2$, then $\mu(S_1) \le \mu(S_2)$ and hence $\mu(S_2 \setminus S_1) = \mu(S_2) - \mu(S_1)$ whenever $\mu(S_1) < \infty$;*

(ii) *If $(S_j)$ is an increasing sequence in $\mathbf{A}$, then*

$$\mu\Big(\bigcup_{j=1}^{\infty} S_j\Big) = \lim_{j\to\infty} \mu(S_j);$$

(iii) *If $(S_j)$ is an decreasing sequence in $\mathbf{A}$ and if $\mu(S_1) < \infty$, then*

$$\mu\Big(\bigcap_{j=1}^{\infty} S_j\Big) = \lim_{j\to\infty} \mu(S_j).$$

*Proof.* (i) Since $S_2 = (S_2 \setminus S_1) \cup S_1$ and $S_2 \setminus S_1 = S_2 \cap S_1^c$, we conclude that if $\mu(S_1) = \infty$ then $\mu(S_2) = \infty$ and hence $\mu(S_1) \le \mu(S_2)$. If otherwise $\mu(S_1) < \infty$ then

$$\mu(S_2) = \mu(S_1) + \mu(S_2 \setminus S_1) \ge \mu(S_1)$$

and so $\mu(S_2 \setminus S_1) = \mu(S_2) - \mu(S_1)$.

(ii) Since $(S_j)$ is an increasing sequence in $\mathbf{A}$, we have

$$\bigcup_{j=1}^{\infty} S_j = S_1 \cup (S_2 \setminus S_1) \cup (S_3 \setminus S_2) \cup \cdots$$

and accordingly, if letting $S_0 = \emptyset$ then

$$\mu\Big(\bigcup_{j=1}^{\infty} S_j\Big) = \sum_{j=0}^{\infty} \mu(S_{j+1} \setminus S_j) = \lim_{n\to\infty} \sum_{j=0}^{n} \mu(S_{j+1} \setminus S_j) = \lim_{n\to\infty} \mu(S_{n+1}).$$

(iii) Note that

$$\left(\bigcap_{j=1}^{\infty} S_j\right)^c = \bigcup_{j=1}^{\infty} S_j^c$$

and that $(S_j^c)$ is increasing when $(S_j)$ is decreasing. This, together with (ii) and (i), yields the desired formula.

**Definition 4.1.3.** (i) *An ordered pair $(X, \mathbf{A})$ consisting of a set $X$ and a $\sigma$-algebra $\mathbf{A}$ of subsets of $X$ is called a measurable space. Any set in $\mathbf{A}$ is called a measurable set; more exactly, $\mathbf{A}$-measurable set.* (ii) *A measure space is a triple $(X, \mathbf{A}, \mu)$ consisting of a set $X$, a $\sigma$-algebra $\mathbf{A}$ of subsets of $X$, and a measure $\mu$ defined on $\mathbf{A}$.*

(iii) *Given a measure space $(X, \mathbf{A}, \mu)$, two functions are said to be equal $\mu$-almost everywhere, denoted $f = g$, $\mu$-a.e. provided $\mu(\{x \in X : f(x) \ne g(x)\}) = 0$. Similarly, $f = \lim_{n\to\infty} f_n$, $\mu$-a.e., if there is a set $S$ such that $\mu(S) = 0$ and $f_n(x) \to f(x)$ for each $x \in X \setminus S$.*

*Example 4.1.1.* Given a non-empty set $X$, let $2^X$ be the set of all subsets of $X$. Then $2^X$ is a $\sigma$-algebra and hence $(X, 2^X)$ is a measurable space. Moreover, let $P$ be a fixed element in $X$. Define a function $\mu$ on $2^X$ as follows:

$$\mu(S) = \begin{cases} 0, & \text{if} \quad P \notin S \\ 1, & \text{if} \quad P \in S \end{cases}$$

Then $\mu$ is a finite measure which is called the unit measure concentrated at $P$. By definition, $(X, 2^X, \mu)$ is a measure space.

**Definition 4.1.4.** *Given two points $(a_1, ..., a_n), (b_1, ..., b_n) \in \mathbf{R}^n$ with $a_j \leq b_j$ for $j = 1, ..., n$. Then*

$$I = (a_1, b_1) \times \cdots (a_n, b_n) = \{(x_1, ..., x_n) : a_j < x_j < b_j,\ j = 1, ..., n\}$$

*is called an open interval in $\mathbf{R}^n$. Any or all of the $<$ signs may be replaced by $\leq$, with corresponding changes made in the interval notation. Such subsets of $\mathbf{R}^n$ are called the intervals of $\mathbf{R}^n$. In particular, $\{(x_1, ..., x_n) : a_j < x_j \leq b_j,\ j = 1, ..., n\}$ or $\{(x_1, ..., x_n) : a_j \leq x_j < b_j,\ j = 1, ..., n\}$ is called a half-open intervals, and $\{(x_1, ..., x_n) : a_j \leq x_j \leq b_j,\ j = 1, ..., n\}$ is called a closed interval. For an interval $I$, let $m(I) = (b_1 - a_1) \cdots (b_n - a_n)$. It is clear that if $n = 1, 2, 3$ then $m(I)$ is the length, area, volume of $I$. If $S$ is an arbitrary subset of $\mathbf{R}^n$, then the Lebesgue outer measure of $S$ is defined by*

$$m^*(S) = \inf \sum_{j=1}^{\infty} m(I_j),$$

*where the infimum is taken over all countable collections $\{I_j\}$ of open intervals such that $S \subset \bigcup_{j=1}^{n} I_j$.*

To see whether or not $m^*$ is a measure on $2^{\mathbf{R}^n}$, we first establish the following result.

**Theorem 4.1.1.** *The set function $m^*$ on $2^{\mathbf{R}^n}$ satisfies:*

(i) $m^*(\emptyset) = 0$*;*

(ii) $0 \leq m^*(S) \leq m^*(T)\ \forall S \subset T \subset \mathbf{R}^n$*;*

(iii) *For any sequence $(S_j)$ of subsets of $\mathbf{R}^n$,*

$$m^*\Big(\bigcup_{j=1}^{\infty} S_j\Big) \leq \sum_{j=1}^{\infty} m^*(S_j).$$

*Proof.* (i) It is obvious since $\emptyset$ is a subset of any open interval $I_{j^{-1}}$ with $m(I_{j^{-1}}) = j^{-n} \to 0$.

(ii) $0 \leq m^*(S)$ follows right away from the definition. If $S \subset T$, then any open covering $(I_j)$ of $T$ must cover $S$, and so

$$m^*(S) \leq \sum_{j=1}^{\infty} m(I_j),$$

giving the desired inequality.

(iii) Given $\epsilon > 0$, for each $j \in \mathbf{N}$ there is a sequence of open intervals $(I_{i,j})$ such that

$$S_j \subset \bigcup_{i=1}^{\infty} I_{i,j} \quad \text{and} \quad \sum_{i=1}^{\infty} m(I_{i,j}) \le m^*(S_j) + 2^{-j}\epsilon.$$

Then

$$\bigcup_{j=1}^{\infty} S_j \subset \bigcup_{i,j=1}^{\infty} I_{i,j}$$

and

$$\begin{aligned}\sum_{i,j=1}^{\infty} m(I_{i,j}) &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} m(I_{i,j}) \\ &\le \sum_{j=1}^{\infty} m^*(S_j) + \sum_{j=1}^{\infty} 2^{-j}\epsilon \\ &= \sum_{j=1}^{\infty} m^*(S_j) + \epsilon.\end{aligned}$$

Thus by (ii) it follows that

$$m^*\Big(\bigcup_{j=1}^{\infty} S_j\Big) \le \sum_{i,j=1}^{\infty} m(I_{i,j}) \le \sum_{j=1}^{\infty} m^*(S_j) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we conclude that (iii) is true.

*Example 4.1.2.* (i) If $E = \mathbf{Q}^n \cap [0,1]^n$ then $m^*E = 0$. It suffices to verify the case $n = 1$. Let $E = \{r_1, r_2, ....\}$ and $I_i = (r_i - 2^{-(i+1)}\epsilon, r_i + 2^{-(i+1)}\epsilon)$ for $\epsilon > 0$. Then

$$m^*(E) \le \sum_{i=1}^{\infty} m(I_i) = \epsilon \to 0.$$

(ii) If $S$ is an interval $I \subset \mathbf{R}^n$, then $m^*(S) = m(I)$. Clearly, $m^*(S) \le m(I)$. To see the converse inequality, let $\epsilon > 0$ be arbitrary, there is a covering $(I_j)$ such that $I \subset \cup_{j=1}^{\infty} I_j$ with $\sum_{j=1}^{\infty} m(I_j) < m^*(I) + \epsilon$. A simple geometric argument yields

$$m(I) = m\big(\cup_{j=1}^{\infty} (I \cap I_j)\big) \le \sum_{j=1}^{\infty} m\big(I \cap I_j\big) \le \sum_{j=1}^{\infty} m(I_j) \le m^*(I) + \epsilon,$$

yielding $m(I) \ge m^*(I)$.

(iii) $m^*$ does not satisfy (iii) of the definition of a measure – there is a sequence of disjoint subsets $E_j$ of $\mathbf{R}^n$ such that

$$m^*\Big(\bigcup_{j=1}^{\infty} E_j\Big) \ne \sum_{j=1}^{\infty} m^*(E_j).$$

For simplicity, let us consider $n = 1$ only. We say that $x \sim y$ if $x - y \in \mathbf{Q}$. For $x \in [0,1]$ let $E(x) = \{x + r \in [0,1] : r \in \mathbf{Q}\}$. It is clear that $x \in E(x)$. If $E(x) \neq E(y)$, then we say that both $E(x)$ and $E(y)$ are different. Note that it is possible to have $E(x) = E(y)$ even if $x \neq y$. However, if $E(x) \neq E(y)$, then $E(x) \cap E(y) = \emptyset$ – In fact, if $E(x) \cap E(y) \neq \emptyset$ then there is $z \in E(x) \cap E(y)$ and hence $z = x + r_x = y + r_y$ where $r_x, r_y \in \mathbf{Q}$. This yields $y = x + r_x - r_y$. Whenever $u \in E(y)$, we have $u = y + r_u = x + r_x - r_y + r_u$ where $r_u \in \mathbf{Q}$, and so $u \in E(x)$. Similarly, $u \in E(x) \Rightarrow u \in E(y)$. Thus $E(x) = E(y)$ contradicting the given condition $E(x) \neq E(y)$. This tells us that $[0,1]$ is decomposed into the union of all disjoint sets $E(x)$, and so there is a set $S \subset [0,1]$ which contains exactly one point from each equivalence class determined by $\sim$. It is clear that

$$[0,1] \subset \bigcup_{r \in \mathbf{Q} \cap [-1,1]} (S + r) \subset [-1,2]$$

as well as

$$(S + r) \cap (S + s) = \emptyset \quad \text{if} \quad r, s \in \mathbf{Q} \quad \text{and} \quad r \neq s.$$

Note that $m^*(S + r) = m^*(S)$ by definition. So if $m^*$ is countably subadditive then it follows that

$$1 = m^*([0,1]) \leq m^*\Big(\bigcup_{r \in \mathbf{Q} \cap [-1,1]} (S + r)\Big) = \sum_{r \in \mathbf{Q} \cap [-1,1]} m^*(S) \leq m^*([-1,2]) = 3.$$

This is a contradiction. Thus

$$m^*\Big(\bigcup_{r \in \mathbf{Q} \cap [-1,1]} (S + r)\Big) \neq \sum_{r \in \mathbf{Q} \cap [-1,1]} m^*(S + r).$$

**Definition 4.1.5.** *A subset $S$ of $\mathbf{R}^n$ is said to be $m^*$-measurable provided for each subset $T$ of $\mathbf{R}^n$ one has $m^*(T) = m^*(T \cap S) + m^*(T \cap S^c)$.*

**Theorem 4.1.2.** *Let $S \subset \mathbf{R}^n$. Then the following statements are equivalent:*
(i) *$S$ is $m^*$-measurable;*
(ii) *For any $A \subset S$ and $B \subset S^c$ one has $m^*(A \cup B) = m^*(A) + m^*(B)$;*
(iii) *$S^c$ is $m^*$-measurable.*

*Proof.* (i)$\Leftrightarrow$(ii). Take $T = A \cup B$. Then $T \cap S = A$ and $T \cap S^c = B$. So, if (i) holds, then

$$m^*(A \cup B) = m^*(T) = m^*(T \cap S) + m^*(T \cap S^c) = m^*(A) + m^*(B),$$

reaching (ii). Conversely, if (ii) holds, then $A = T \cap S$ and $B = T \cap S^c$ give $A \cup B = T$ and

$$m^*(T) = m^*(A \cup B) = m^*(A) + m^*(B) = m^*(T \cap S) + m^*(T \cap S^c).$$

(i)$\Leftrightarrow$(iii). This follows from

$$m^*(T \cap S) + m^*(T \cap S^c) = m^*(T \cap (S^c)^c) + m^*(T \cap S^c).$$

**Theorem 4.1.3.** *Let $\mathbf{M}_n$ be the class of all $m^*$-measurable subsets of $\mathbf{R}^n$. Then*

(i) *If $S_1, S_2 \in \mathbf{M}$ then $S_1 \cup S_2, S_1 \cap S_2, S_1 \setminus S_2 \in \mathbf{M}$. Moreover, if $S_1 \cap S_2 = \emptyset$ then*

$$m^*\big(T \cap (S_1 \cup S_2)\big) = m^*(T \cap S_1) + m^*(T \cap S_2) \quad \forall T \in 2^{\mathbf{R}^n}.$$

(ii) *$\mathbf{M}_n$ is a $\sigma$-algebra.*

(iii) *$m^*$ is countably additive and hence a measure on $\mathbf{M}_n$.*

*Proof.* (i) Let $S_1, S_2 \in \mathbf{M}_n$. Then for any $T \in 2^{\mathbf{R}^n}$ we use Theorem 4.1.2 to get

$$\begin{aligned} m^*(T) &= m^*(T \cap S_1) + m^*(T \cap S_1^c) \\ &= m^*(T \cap S_1) + m^*\big((T \cap S_1^c) \cap S_2\big) + m^*\big((T \cap S_1^c) \cap S_2^c\big) \\ &= m^*(T \cap S_1) + m^*\big((T \cap S_1^c) \cap S_2\big) + m^*\big(T \cap (S_1 \cup S_2)^c\big) \\ &= m^*(T \cap \big(S_1 \cup (S_1^c \cap S_2)\big) + m^*\big(T \cap (S_1 \cup S_2)^c\big) \\ &= m^*\big(T \cap (S_1 \cup S_2)\big) + m^*\big(T \cap (S_1 \cup S_2)^c\big), \end{aligned}$$

so that $S_1 \cup S_2 \in \mathbf{M}_n$.

Since

$$S_1 \cap S_2 = \big((S_1 \cap S_2)^c\big)^c = \big(S_1^c \cup S_2^c\big)^c,$$

we conclude from Theorem 4.1.2 (iii) and the forgoing (i) that $S_1 \cap S_2 \in \mathbf{M}_n$. This result implies $S_1 \setminus S_2 = S_1 \cap S_2^c \in \mathbf{M}_n$.

The last result of (i) follows immediately from Theorem 4.1.2 (ii).

(ii) To see that $\mathbf{M}_n$ is a $\sigma$-algebra, it suffices to check that if $S_j \in \mathbf{M}_n$, $j \in \mathbf{N}$ are disjoint then $\cup_{j=1}^{\infty} S_j \in \mathbf{M}_n$. This is because of

$$\bigcup_{j=1}^{\infty} S_j = S_1 \cup \big(S_2 \setminus S_1\big) \cup \big(S_3 \setminus (S_1 \cup S_2)\big) \cup \big(S_4 \setminus (S_1 \cup S_2 \cup S_3)\big) \cup \cdots.$$

Now for any $T \in 2^{\mathbf{R}^n}$ and $k \in \mathbf{N}$, one gets from (i) that for $E_k = \bigcup_{j=1}^{k} S_j$ and $E_\infty = \bigcup_{j=1}^{\infty} S_j$,

$$\begin{aligned} m^*(T) &= m^*(T \cap E_k) + m^*(T \cap (E_k)^c) \\ &\geq m^*(T \cap E_k) + m^*(T \cap (E_\infty)^c) \\ &= \sum_{j=1}^{k} m^*(T \cap S_j) + m^*(T \cap (E_\infty)^c). \end{aligned}$$

Letting $k \to \infty$ and using the property (iii) of $m^*$, one obtains

$$\begin{aligned} m^*(T) &\geq \sum_{j=1}^{\infty} m^*(T \cap S_j) + m^*(T \cap (E_\infty)^c) \\ &\geq m^*(T \cap E_\infty) + m^*(T \cap (E_\infty)^c). \end{aligned}$$

On the other hand, note that

$$T = (T \cap E_\infty) \cup (T \cap (E_\infty)^c).$$

So

$$m^*(T) \le m^*(T \cap E_\infty) + m^*(T \cap (E_\infty)^c).$$

Therefore $E_\infty$ is $m^*$-measurable.

(iii) As a product of the previous argument and the property of $m^*$ one finds that $m^*$ is countably additive: $m^*(E_\infty) = \sum_{j=1}^\infty m^*(S_j)$. Of course, here $(S_j)$ are assumed to be disjoint $m^*$-measurable subsets of $\mathbf{R}^n$. This implies $m^*$ is a measure on $\mathbf{M}_n$.

*Example 4.1.3.* (i) If $m^*(E) = 0$ then $E$ is $m^*$-measurable; Any subset of $m^*$-zero set is $m^*$-measurable; countable union of $m^*$-zero sets is $m^*$-measurable.

(ii) Any interval in $\mathbf{R}^n$ is $m^*$-measurable.

(iii) Any open or closed subset of $\mathbf{R}^n$ is $m^*$-measurable.

(iv) The set $S$ constructed in Example 4.1.2 is a non-$m^*$-measurable subset of $\mathbf{R}^n$.

We shall now take up the theory of extended real-valued measurable functions with domains in $\mathbf{R}^n$.

**Definition 4.1.6.** *An extended real-valued function $f$ defined on $E \in \mathbf{M}_n$ is called Lebesgue measurable provided $\{x \in E : f(x) > a\}$ belongs to $\mathbf{M}_n$ for any $a \in \mathbf{R}$.*

Note that measurability does not require a measure at all. Moreover, if $f$ is an extended complex-valued function, then we say that $f$ is Lebesgue measurable provided its real and imaginary parts are Lebesgue measurable in the previous sense.

**Lemma 4.1.2.** *Let $E \in \mathbf{M}_n$ and $f : E \to [-\infty, \infty]$. Then the following statements are equivalent:*

(i) *The set $\{x \in E : f(x) > a\}$ belongs to $\mathbf{M}_n$ for each $a \in \mathbf{R}$;*

(ii) *The set $\{x \in E : f(x) \ge a\}$ belongs to $\mathbf{M}_n$ for each $a \in \mathbf{R}$;*

(iii) *The set $\{x \in E : f(x) < a\}$ belongs to $\mathbf{M}_n$ for each $a \in \mathbf{R}$;*

(iv) *The set $\{x \in E : f(x) \le a\}$ belongs to $\mathbf{M}_n$ for each $a \in \mathbf{R}$.*

*Proof.* Since $\mathbf{M}_n$ is a $\sigma$-algebra, we conclude that (i)⇔(iv) and (ii)⇔(iii) right away. Note that

$$\{x \in E : f(x) \ge a\} = \bigcap_{j=1}^\infty \{x \in E : f(x) > a - j^{-1}\}.$$

So (i) implies (ii). Also

$$\{x \in E : f(x) > a\} = \bigcup_{j=1}^\infty \{x \in E : f(x) \ge a + j^{-1}\}$$

Then (ii) yields (i).

*Example 4.1.4.* Let $E \in \mathbf{M}_n$ and $f : E \to [-\infty, \infty]$ be Lebesgue measurable. Then $\{x \in E : f(x) = a\} \in \mathbf{M}_n$ for any $a \in \mathbf{R} \cup \{-\infty, \infty\}$, but not conversely. The first part follows from the foregoing lemma and the following formulas:

$$\{x \in E : f(x) = a\} = \{x \in E : f(x) \le a\} \bigcap \{x \in E : f(x) \ge a\} \quad \text{if} \quad a \in \mathbf{R},$$

$$\{x \in E : f(x) = \infty\} = \bigcap_{j=1}^{\infty} \{x \in E : f(x) > j\}$$

and

$$\{x \in E : f(x) = -\infty\} = \bigcap_{j=1}^{\infty} \{x \in E : f(x) < -j\}.$$

Regarding the second part, consider a non measurable set $S \subset (0,1)$ and define

$$f(x) = \begin{cases} x, & \text{if} \quad x \in S, \\ -x, & \text{if} \quad x \notin S. \end{cases}$$

Clearly, this function is one-to-one and so $\{x \in (0,1) : f(x) = a\} \in \mathbf{M}_1$ for any $a \in [-\infty, \infty]$, but since $\{x \in (0,1) : f(x) > 0\} = S \notin \mathbf{M}_1$, $f$ is not Lebesgue measurable.

**Theorem 4.1.4.** (i) *If $f$ and $g$ are real-valued Lebesgue measurable functions on $E \in \mathbf{M}_n$ and $c$ is a real number, then so are $f+g$, $fg$, $cf$, $|f|$, $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$.*

(ii) *If $(f_j)$ is a sequence of real-valued Lebesgue measurable functions on $E \in \mathbf{M}_n$, then the following four functions:*

$$(\inf f_j)(x) = \inf_{j \in \mathbf{N}} f_j(x), \quad (\sup f_j)(x) = \sup_{j \in \mathbf{N}} f_j(x),$$

$$(\liminf f_j)(x) = \sup_{j \in \mathbf{N}} \inf_{k \ge j} f_k(x), \quad (\limsup f_j)(x) = \inf_{j \in \mathbf{N}} \sup_{k \ge j} f_k(x),$$

*are Lebesgue measurable on $E \in \mathbf{M}_n$.*

*Proof.* Since $(f+g)(x) < a$ is equivalent to $f(x) < a - g(x)$, the density of rational numbers in $\mathbf{R}$ implies, the equivalence amounts to the existence of a rational number $r$ such that $f(x) < r < a - g(x)$. This yields that $\{x \in E : (f+g)(x) < a\}$ equals

$$\bigcup_{r \in \mathbf{Q}} \left( \Big\{ x \in E : f(x) < r \Big\} \bigcap \Big\{ x \in E : g(x) < a - r \Big\} \right).$$

Then $f + g$ is Lebesgue measurable on $E$.

If $c = 0$ then $\{x \in E : cf(x) > a\}$ is either $\emptyset$ or $E$, and hence Lebesgue measurable on $E$. If $c > 0$ or $c < 0$ then

$$\{x \in E : cf(x) > a\} = \{x \in E : f(x) > a/c\} \in \mathbf{M}_n$$

or

$$\{x \in E : cf(x) > a\} = \{x \in E : f(x) < a/c\} \in \mathbf{M}_n,$$

and hence $cf$ is Lebesgue measurable on $E$.

In order to see that $fg$ and $|f|$ are Lebesgue measurable on $E$, we just observe three equalities:

$$fg = \frac{(f+g)^2 - (f-g)^2}{4},$$

$$\{x \in E : f^2(x) > |a|\} = \{x \in E : f(x) > \sqrt{|a|}\} \cup \{x \in E : f(x) < -\sqrt{|a|}\},$$

and

$$\{x \in E : |f|(x) > a\} = \{x \in E : f(x) > a\} \cup \{x \in E : f(x) < -a\}.$$

To check that $f_+$ and $f_-$ are Lebesgue measurable on $E$, we note that

$$f_+ = \frac{|f| + f}{2} \quad \text{and} \quad f_- = \frac{|f| - f}{2}.$$

(ii) It suffices to verify that $\inf f_j$ is Lebesgue measurable on $E$. In fact, for $a \in \mathbf{R}$, we have

$$\{x \in E : \inf f_j(x) > a\} = \bigcap_{j=1}^{\infty} \{x \in E : f_j(x) > a\} \in \mathbf{M}_n,$$

as desired.

**Definition 4.1.7.** *A real-valued function with only a finite number of elements in its range is called a simple function. In particular, let*

$$1_E(x) = \begin{cases} 1, & \text{if} \quad x \in E, \\ 0, & \text{if} \quad x \notin E. \end{cases}$$

*be the characteristic function $1_E$ of a set $E \subset \mathbf{R}^n$.*

Every simple function can be written as a finite linear combination of characteristic functions. More precisely, if the range of the simple function $s$ is $\{c_1, ..., c_k\}$, then $s(x) = \sum_{j=1}^{k} c_j 1_{E_j}(x)$, where $E_j = \{x \in \mathbf{R}^n : s(x) = c_j\}$. The function $s$ is Lebesgue measurable if and only if $E_1, ..., E_k \in \mathbf{M}_n$.

**Theorem 4.1.5.** *Let $E \in \mathbf{M}_n$ and $f : E \to [-\infty, \infty]$. Then $f$ is Lebesgue measurable on $E$ if and only if there exists a sequence $(s_j)$ of simple functions on $E$ such that $\lim_{j\to\infty} s_j(x) = f(x)$ and $|s_1(x)| \le |s_2(x)| \le ...$ for any $x \in E$*

*Proof.* It is enough to verify the necessity.

Case 1: $f \ge 0$. For any $j \in \mathbf{N}$ let

$$s_j(x) = \begin{cases} k2^{-j}, & \text{if} \quad k2^{-j} \le f(x) < (k+1)2^{-j}, \quad k = 0, 1, ..., j2^j - 1, \\ j, & \qquad\qquad\qquad\qquad \text{if} \quad j \le f(x). \end{cases}$$

Then $s_j$ is a simple function on $E$ and is of monotone property: $s_j(x) \le s_{j+1}(x)$ for any $x \in E$.

If $f(x) = \infty$ then $s_j(x) = j$ and hence $\lim_{j\to\infty} s_j(x) = f(x)$.

If $f(x) < \infty$ then there is an $N \in \mathbf{N}$ such that $f(x) < N$ and so $j \ge N$ implies that there is a $k = 0, 1, ..., j2^j - 1$ such that $f(x) \in [k2^{-j}, (k+1)2^{-j})$ and then $s_j(x) = k2^{-j}$. Of course, we have

$$j \ge N \Rightarrow |f(x) - s_j(x)| < 2^{-j}.$$

Consequently, $\lim_{j\to\infty} s_j(x) = f(x)$.

Case 2: $f \not\ge 0$. In this case, we know that $f_+, f_- \ge 0$, $f = f_+ - f_-$ and $|f| = f_+ + f_-$. By Case 1, we have simple functions $(s_{j,+})$ and $(s_{j,-})$ corresponding to $f_+$ and $f_-$ respectively. It is not hard to see that $s_j = s_{j,+} - s_{j,-}$ are the desired simple functions.

## 4.2 Integrals and Their Convergence

We now have the machinery to develop Lebesgue integrals and three fundamental convergence results of integration theory: monotone convergence theorem, Fatou's lemma and Lebesgue dominated convergence theorem.

**Definition 4.2.1.** *Let $E \in \mathbf{M}_n$ and $m = m^*$ on $\mathbf{M}_n$. Then*

(i) *The Lebesgue integral of a Lebesgue measurable simple function $s(x) = \sum_{j=1}^{k} c_j 1_{E_j}(x)$ on $E$ is defined by*

$$\int_E s dm = \sum_{j=1}^{k} c_j m(E \cap E_j).$$

(ii) *The Lebesgue integral of a Lebesgue measurable function $f : E \to [0, \infty]$ is defined by*

$$\int_E f dm = \sup\left\{\int_E s dm : s \text{ is simple} \quad \text{and} \quad 0 \le s \le f\right\}.$$

(iii) *The Lebesgue integral of a Lebesgue measurable function $f : E \to [-\infty, \infty]$ is defined by*

$$\int_E f dm = \int_E f_+ dm - \int_E f_- dm.$$

*If at least one of $\int_E f_+ dm$ and $\int_E f_- dm$ is finite, then we say that $f$ has Lebesgue integral on $E$, and moreover if $\int_E f dm$ is finite then $f$ is Lebesgue integrable and all such functions are denoted by $L(E)$.*

**Theorem 4.2.1.** *Let $E \in \mathbf{M}_n$. Then*
(i) $L(E)$ *is a linear space over* $\mathbf{R}$.
(ii) $f \in L(E) \Longrightarrow |f| \in L(E)$ *with* $\left|\int_E f dm\right| \le \int_E |f| dm$.
(iii) $f \in L(E)$ *and* $f(x) \ge 0 \, \forall x \in E \Longrightarrow \int_E f dm \ge 0$.
(iv) $f \in L(E) \Longrightarrow \int_S f dm = 0 \ \ \forall S \subset E$ *with* $m(S) = 0$.
(v) $S, T \in \mathbf{M}$, $T \subset S \subset E$ *and* $m(S \setminus T) = 0 \Longrightarrow \int_S f dm = \int_T f dm$.

*Proof.* (i) follows from the Lebesgue monotone convergence theorem which will be discussed later on. The proof of (ii) deponds on (i). However, (iii), (iv) and (v) just follow from Definition 4.2.1.

The following property shows that the Lebesgue integrals are countably additive.

**Theorem 4.2.2.** *Let $E \in \mathbf{M}_n$, $0 \le f \in L(E)$ and $E_1, E_2, ... \subset E$. If $(E_j) \subset \mathbf{M}_n$ are mutually disjoint, then*

$$\int_{\cup_{j=1}^\infty E_j} f dm = \sum_{j=1}^\infty \Big(\int_{E_j} f dm\Big).$$

*Proof.* If $f = 1_E$, then by the countable additivity of $m$,

$$\int_{\cup_{j=1}^\infty E_j} f dm = m\Big(\bigcup_{j=1}^\infty E_j \cap E\Big) = \sum_{j=1}^\infty \Big(\int_{E_j} f dm\Big).$$

Of course, if $f$ is a simple function, then the result is still true.

Next, if $f \ge 0$ is arbitrary Lebesgue integrable on $E$, then by definition, for any $\epsilon > 0$ we may choose a simple function $s \le f$ so that

$$\begin{aligned}\int_{\cup_{j=1}^\infty E_j} f dm &\le \epsilon + \int_{\cup_{j=1}^\infty E_j} s dm \\ &= \epsilon + \sum_{j=1}^\infty \int_{E_j} s dm \\ &\le \epsilon + \sum_{j=1}^\infty \int_{E_j} f dm.\end{aligned}$$

This yields

$$\int_{\cup_{j=1}^\infty E_j} f dm \le \sum_{j=1}^\infty \Big(\int_{E_j} f dm\Big).$$

On the other hand, for any $k \in \mathbf{N}$ let $s_j$, $j = 1, ..., k$ be simple functions satisfying $0 \le s_j \le f$ as well as

$$\int_{E_j} s_j dm \ge -\frac{\epsilon}{k} + \int_{E_j} f dm.$$

Let $s = \max_{j=1,...,k} s_j$. Then $s$ is a simple function obeying $0 \le s \le f$. Clearly,

$$\int_{E_j} sdm \geq -\frac{\epsilon}{k} + \int_{E_j} fdm.$$

This, together with the definition of $s$ and the first part, infers

$$\begin{aligned}\int_{\cup_{j=1}^{\infty} E_j} fdm &\geq \int_{\cup_{j=1}^{k} E_j} sdm \\ &= \int_{E_1} sdm + \cdots \int_{E_k} sdm \\ &\geq -\epsilon + \int_{E_1} fdm + \cdots + \int_{E_k} fdm.\end{aligned}$$

Note that $\epsilon > 0$ and $k \in \mathbf{N}$ are arbitrary. So

$$\int_{\cup_{j=1}^{\infty}} fdm \geq \sum_{j=1}^{\infty} \int_{E_j} fdm.$$

The desired equality follows right away.

The Lebesgue monotone convergence theorem reads as

**Theorem 4.2.3.** *Let $E \in \mathbf{M}_n$ and $(f_j)$ be a sequence of Lebesgue measurable functions from $E$ to $[0, \infty]$ such that $f_1 \leq f_2 \leq \cdots$ on $E$. If $f = \lim_{j\to\infty} f_j$ on $E$, then*

$$\lim_{j\to\infty} \int_E f_j dm = \int_E fdm.$$

*Proof.* From monotonicity of $(f_j)$ it follows that $\left(\int_E f_j dm\right)$ is non-decreasing and hence $A = \lim_{j\to\infty} \int_E f_j dm$ exists: note that $A$ is allowed to be $\infty$. Since $f_j \leq f$ on $E$, we conclude that $A \leq \int_E fdm$. The proof will be concluded by proving $\int_E fdm \leq A$. To do this we take a number $\eta \in (0, 1)$ and a simple function $s$ obeying $0 \leq s \leq f$ on $E$. Let $E_j = \{x \in E : f_j(x) \geq \eta s(x)\}$. Then $E_1 \subset E_2 \subset E_3 \subset \cdots$ and $E = \cup_{j=1}^{\infty} E_j$. This implies

$$A \geq \lim_{j\to\infty} \int_{E_j} f_j dm \geq \eta \lim_{j\to\infty} \int_{E_j} sdm.$$

By Theorem 4.2.2 we get that

$$\begin{aligned}\int_E sdm &= \int_{E_1} sdm + \sum_{j=2}^{\infty} \int_{E_j \setminus E_{j-1}} sdm \\ &= \int_{E_1} sdm + \lim_{k\to\infty} \sum_{j=2}^{k} \int_{E_j \setminus E_{j-1}} sdm \\ &= \lim_{k\to\infty} \int_{E_k} sdm.\end{aligned}$$

and so that $A \geq \eta \int_E sdm$. Letting $\eta \to 1$, we obtain $A \geq \int_E sdm$. Taking the supremum over all such simple functions gives $A \geq \int_E fdm$. We are done.

*Proof of Theorem 4.2.1 (i).* It is enough to prove

$$\int_E (f+g)dm = \int_E f dm + \int_E g dm \quad \forall f, g \in L(E), \quad E \in \mathbf{M}.$$

Clearly, this is valid for simple functions. Also, we may assume $f, g \geq 0$ on $E$. In this case, we use Theorem 4.1.5 to obtain two non-decreasing sequences of simple functions $(s_j)$ and $(t_j)$ such that

$$\lim_{j\to\infty} s_j = f \quad \text{and} \quad \lim_{j\to\infty} t_j = g \quad \text{on} \quad E.$$

An application of Theorem 4.2.3 implies

$$\begin{aligned}\int_E (f+g)dm &= \lim_{j\to\infty} \int_E (s_j + t_j)dm \\ &= \lim_{j\to\infty} \int_E s_j dm + \lim_{j\to\infty} \int_E t_j dm \\ &= \int_E f dm + \int_E g dm.\end{aligned}$$

The following is the Fatou's lemma.

**Lemma 4.2.1.** *Given $E \in \mathbf{M}_n$, let $(f_j)$ be a sequence of Lebesgue measurable functions from $E$ to $[0, \infty]$. If $f = \liminf_{j\to\infty} f_j$ on $E$, then*

$$\int_E f dm \leq \lim \inf_{j\to\infty} \int_E f_j dm.$$

*Proof.* For each $k \in \mathbf{N}$ let $g_k = \inf_{j\geq k} f_j$. Then by Theorem 4.1.4, we see that $g_k$ is Lebesgue measurable, and $(g_k)$ is non-decreasing with $f = \lim_{k\to\infty} g_k = \sup_{k\in\mathbf{N}} g_k$ on $E$. By Theorem 4.2.3 and the fact that $g_k \leq f_j$ for each $j \geq k$, we obtain

$$\int_E f dm = \lim_{k\to\infty} \int_E g_k dm \leq \lim \inf_{j\to\infty} \int_E f_j dm,$$

as desired.

The most useful general result about Lebesgue integration is the following Lebesgue's dominated convergence theorem.

**Theorem 4.2.4.** *Given $E \in \mathbf{M}_n$, let $(f_j)$ be a sequence of Lebesgue measurable functions from $E$ to $[-\infty, \infty]$ and $\lim_{j\to\infty} f_j = f$ a.e. on $E$. If there is a function $g \in L(E)$ such that $|f_j| \leq g$ a.e. on $E$, then*

$$\lim_{j\to\infty} \int_E f_j dm = \int_E f dm.$$

*Proof.* It is obvious that $g \in L(E)$ and $|f_j| \le g$ a.e. on $E$ imply $f_j \in L(E)$. From Fatou's lemma it follows that

$$\int_E |f|dm = \int_E \lim_{j\to\infty} |f_j|dm \le \lim\inf_{j\to\infty} \int_E |f_j|dm \le \int_E gdm.$$

Because of $f_j + g \ge 0$ a.e. on $E$, by Fatou's lemma again it follows that

$$\begin{aligned}\int_E fdm + \int_E gdm &= \int_E \lim\inf_{j\to\infty} (f_j+g)dm \\ &\le \lim\inf_{j\to\infty} \int_E (f_j+g)dm \\ &= \lim\inf_{j\to\infty} \int_E f_j dm + \int_E gdm.\end{aligned}$$

This gives

$$\int_E fdm \le \lim\inf_{j\to\infty} \int_E f_j dm.$$

Also $g - f_j \ge 0$ a.e. on $E$, a similar argument yields

$$\lim\inf_{j\to\infty} \int_E f_j dm \le \lim\sup_{j\to\infty} \int_E f_j dm \le \int_E fdm.$$

The last two lines of inequalities are combined to derive the desired result.

As a direct consequence of the Lebesgue's dominated convergence theorem, we have the following the bounded convergence theorem.

**Corollary 4.2.1.** *Given $E \in \mathbf{M}_n$ with $m(E) < \infty$, let $(f_j)$ be a sequence of Lebesgue measurable functions from $E$ to $[-\infty, \infty]$ and $\lim_{j\to\infty} f_j = f$ a.e. on $E$. If there is a constant $M > 0$ such that $|f_j| \le M$ a.e. on $E$, then*

$$\lim_{j\to\infty} \int_E f_j dm = \int_E fdm.$$

*Proof.* Since $m(E) < \infty$, we conclude that $\int_E Mdm < \infty$. This, together with Theorem 4.2.4, deduces the desired limit result.

## 4.3 $L_p$-Spaces and Their Completeness

After this discussion of Lebesgue measures and integrals we proceed to a survey of the linear vector spaces formed by equivalence classes of Lebesgue integrable functions.

**Definition 4.3.1.** *Let $E \in \mathbf{M}_n$. If $p \in [1, \infty)$ then $L_p(E)$ is defined to be the class of all Lebesgue measurable functions $f : E \to [-\infty, \infty]$ satisfying*

$$\|f\|_p = \left(\int_E |f|^p dm\right)^{\frac{1}{p}} < \infty.$$

Note that

$$E = \bigcup_{j=1}^{\infty} \{x \in E : |f(x)| \geq j^{-1}\} \cup \{x \in E : f(x) = 0\}.$$

So, $\|f\|_p = 0$ is equivalent to $f = 0$ a.e. on $E$. This indicates that $L_p(E)$ really consists of equivalence classes of functions rather than of functions, via the equivalence relation $\sim$: $f \sim g \Leftrightarrow f = g$ a.e on $E$. However, we conform to standard malpractice by referring to these equivalence classes as functions.

In order to see more properties of $L_p(E)$, we need the following Hölder's inequality and Minkowski's inequality.

**Theorem 4.3.1.** *Let $E \in \mathbf{M}_n$.*

(i) *Hölder's inequality: If $f \in L^p(E)$, $g \in L^{\frac{p}{p-1}}(E)$, $p \in (1, \infty)$, then $fg \in L_1(E)$ with*

$$\|fg\|_1 \leq \|f\|_p \|g\|_{\frac{p}{p-1}}.$$

(ii) *Minkowski's inequality: If $f, g \in L_p(E)$, $p \in [1, \infty)$, then $f + g \in L_p(E)$ with*

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

*Proof.* (i) Let $q = p/(p-1)$ and $\phi(t) = t^{\frac{1}{p}}$ for $t \geq 0$. Since $p^{-1} \in (0, 1)$, $\phi''(r) < 0$ for all $r > 0$ and $\phi$ is concave. Hence $\phi(t) \leq \phi(1) + \phi'(1)(t-1)$, or

$$t^{\frac{1}{p}} \leq 1 + \frac{t-1}{p} = \frac{t}{p} + \frac{1}{q}.$$

Setting $t = u^p v^{-q}$, where $u \geq 0$ and $v > 0$, we find since $1 - q = -q/p$ that

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

Obviously, this inequality also holds when $v = 0$.

If $\|f\|_p = 0$ or $\|g\|_q = 0$ then the inequality is evident. So we may assume that $\|f\|_p > 0$ and $\|g\|_q > 0$. The first paragraph shows that

$$\frac{|fg|}{\|f\|_p \|g\|_q} \leq \frac{|f|^p}{p\|f\|_p^p} + \frac{|g|^q}{q\|g\|_q^q}.$$

Taking Lebesgue integration on both sides of the estimate gives the desired inequality.

(ii) It is enough to verify the case $p > 1$. Without loss of generality, we may assume that $\|f + g\|_p \neq 0$. By Hölder inequality with $|f|$ (or $|g|$) and $|f + g|^{p-1}$, we achieve

$$\begin{aligned}
\|f + g\|_p^p &= \int_E |f + g||f + g|^{p-1} dm \\
&\leq \int_E |f||f + g|^{p-1} dm + \int_E |g||f + g|^{p-1} dm \\
&\leq \|f\|_p \big\| |f + g|^{p-1} \big\|_q + \|g\|_p \big\| |f + g|^{p-1} \big\|_q \\
&= \|f + g\|_p^{p-1} (\|f\|_p + \|g\|_p).
\end{aligned}$$

**Theorem 4.3.2.** *Let $E \in \mathbf{M}_n$ and $p \in [1, \infty)$. Then $(L_p(E), \|\cdot\|_p)$ is a Banach space.*

*Proof.* Theorem 4.3.1 implies that $(L_p(E), \|\cdot\|_p)$ is linear.

Next, let's verify that $\|\cdot\|_p$ is a norm on $L_p(E)$. It is known that $\|f\|_p \ge 0$ for which the equality holds if and only if $f = 0$ a.e. on $E$. Also it is trivial to get $\|\alpha f\|_p = |\alpha|\|f\|_p$ for $\alpha \in \mathbf{R}$. The triangle inequality for $\|\cdot\|_p$ follows from the Minkowski inequality.

The nontrivial part is to check the completeness. By Theorem 2.1.1 one suffices to verify that if $\sum_{j=1}^{\infty} \|f_j\|_p < \infty$ then $\sum_{j=1}^{\infty} f_j$ converges in $\|\cdot\|_p$. Now for each $k \in \mathbf{N}$ let $g_k = \sum_{j=1}^{k} |f_j|$. Then the Minkowski's inequality gives

$$\|g_k\|_p \le \sum_{j=1}^{k} \|f_j\|_p \le \sum_{j=1}^{\infty} \|f_j\|_p < \infty.$$

Since $(g_k)$ is non-decreasing, there is an extended real-valued function $g$ such that $\lim_{k\to\infty} g_k = g$ pointwise on $E$. Clearly, $g$ is Lebesgue measurable, and this, together with Fatou's lemma, implies

$$\left(\int_E g^p dm\right)^{\frac{1}{p}} \le \lim\inf_{k\to\infty} \left(\int_E g_k^p dm\right)^{\frac{1}{p}} \le \sum_{j=1}^{\infty} \|f_j\|_p < \infty.$$

In particular, this estimate shows that $g$ is finite a.e. on $E$. For each $x$ such that $g(x)$ is finite, the series $\sum_{j=1}^{\infty} f_j(x)$ is absolutely convergent. Let

$$s(x) = \begin{cases} 0, & \text{if} \quad g(x) \text{ is infinite} \\ \sum_{j=1}^{\infty} f_j(x), & \text{if} \quad g(x) \text{ is finite.} \end{cases}$$

This function equals to the limit of the partial sums $s_k(x) = \sum_{j=1}^{k} f_j(x)$ a.e. on $E$, and hence is itself Lebesgue measurable. Since $|s_k| \le g$ a.e. on $E$, we conclude that $|s| \le g$ a.e. on $E$. Of course, $s \in L_p(E)$ and $|s_k - s|^p \le 2^p g^p$. We can now use Lebesgue's dominated convergence theorem to obtain

$$\lim_{k\to\infty} \|s_k - s\|_p^p = \lim_{k\to\infty} \int_E (s_k - s)^p dm = 0.$$

**Definition 4.3.2.** *A simple function $s = \sum_{j=1}^{k} c_j 1_{E_j}$ is called a step function if each of the sets $E_j$ has finite Lebesgue measure.*

In what follows, we show the density of all step functions in the Lebesgue spaces.

**Theorem 4.3.3.** *Let $E \in \mathbf{M}_n$ and $p \in [1, \infty)$. Then the step functions are dense in $(L_p(E), \|\cdot\|_p)$.*

*Proof.* Let $f \in L_p(E)$. Since $f = f_+ - f_-$, it is enough to consider $f \geq 0$ a.e. on $E$. In this case, there is a sequence of simple functions: $(s_j)$ such that

$$0 \leq s_1 \leq s_2 \leq \cdots \leq f, \quad \lim s_j = f \quad \text{a.e. on} \quad E.$$

Each of these simple functions is actually a step function. Moreover,

$$(f - s_1)^p \geq (f - s_2)^p \geq \cdots \geq 0, \quad \lim (f - s_j)^p = 0 \quad \text{a.e. on} \quad E.$$

Lebesgue's dominated convergence theorem now tells us that

$$\lim_{j \to \infty} \|f - s_j\|_p = \lim_{j \to \infty} \Big( \int_E |f - s_j|^p \Big)^{\frac{1}{p}} = 0.$$

We are done.

To complete the picture of $L_p$ spaces, we consider a space corresponding to the limiting case $p = \infty$. For a Lebesgue measurable function $f$ on $E \in \mathbf{M}_n$, let

$$\|f\|_\infty = \inf\{a : m(\{x \in E : |f(x)| > a \geq 0\}) = 0\},$$

with the convention $\inf \emptyset = \infty$. Note that the infimum is actually attained since

$$\{x \in E : |f(x)| > a\} = \bigcup_{j=1}^{\infty} \{x \in E : |f(x)| > a + j^{-1} \geq 0\}$$

and if the sets on the right hand side are empty, so is the one on the left hand side. $\|f\|_\infty$ is sometimes called the essential supremum of $f$ and written

$$\|f\|_\infty = \mathrm{esssup}_{x \in E} |f(x)|.$$

However, this is not the same as $\sup_{x \in E} |f(x)|$: for example, if

$$f(x) = \begin{cases} 1, & \text{if} \quad x \in \mathbf{Q} \cap [0,1], \\ 0, & \text{if} \quad x \in \mathbf{Q}^c \cap [0,1]. \end{cases}$$

then $\sup_{x \in [0,1]} |f(x)| = 1$ and $\mathrm{esssup}_{x \in [0,1]} |f(x)| = 0$. Of course, $f = 0$ a.e. on $[0,1]$, for $m(\mathbf{Q} \cap [0,1]) = 0$.

**Definition 4.3.3.** *Let $E \in \mathbf{M}_n$. Then $L_\infty(E)$ is defined to be the class of all Lebesgue measurable functions $f : E \to [-\infty, \infty]$ with $\|f\|_\infty < \infty$, with the usual convention that two functions are equal a.e. on $E$ define the same element of $L_\infty(E)$.*

The results and their proofs given above for $1 \leq p < \infty$ can readily extend to the value $p = \infty$.

**Theorem 4.3.4.** *Let $E \in \mathbf{M}_n$.*
(i) *If $f \in L_1(E)$ and $g \in L_\infty(E)$ then $fg \in L_1(E)$ with $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$;*
(ii) *$(L_\infty(E), \|\cdot\|_\infty)$ is a Banach space;*
(iii) *The bounded step functions are dense in $(L_\infty(E), \|\cdot\|_\infty)$.*

*Proof.* It is left to the reader for an exercise.

## Exercises

**4.1** Prove:

(i) Every half-open interval $I$ in $\mathbf{R}^n$ is Lebesgue measurable and $m^*(I) = m(I)$.

(ii) Every open set in $\mathbf{R}^n$ is the union of a countable collection of disjoint half-open intervals.

(iii) Every open set in in $\mathbf{R}^n$ is Lebesgue measurable and so is a closed set.

**4.2** Let $E \subset \mathbf{R}^n$ be Lebesgue measurable. Prove:

(i) For any $\epsilon > 0$, there is a sequence of open subsets $(G_j)$ of $\mathbf{R}^n$ such that $E \subset G = \cap_{j=1}^{\infty} G_j$ and $m(G \setminus E) < \epsilon$.

(ii) For any $\epsilon > 0$, there is a sequence of closed subsets $(F_j)$ of $\mathbf{R}^n$ such that $E \supset F = \cup_{j=1}^{\infty} F_j$ and $m(E \setminus F) < \epsilon$.

**4.3** We say that an extended real-valued function $f$ on $E \subset \mathbf{R}^n$ is continuous at $x_0 \in E$ if $y_0 = f(x_0)$ is finite, and for any open ball $V(y_0) \subset \mathbf{R}$ centered at $y_0$ there is an open ball $U(x_0) \subset \mathbf{R}^n$ centered at $x_0$ such that $f(U(x_0) \cap E) \subset V(y_0)$. If $f$ is continuous at any point in $E$, we say that $f$ is continuous on $E$. Now, let $f$ be an extended real-valued continuous on a Lebesgue measurable set $E \subset \mathbf{R}^n$. Prove that $f$ is Lebesgue measurable on $E$; that is, for any $a \in \mathbf{R}$, the set $E[f > a] = \{x \in E : \ f(x) > a\}$ is Lebesgue measurable.

**4.4** Let $f$ be an extended real-valued function on $\mathbf{R}^n$. Prove that $1/f$ is Lebesgue measurable if $f$ is Lebesgue measurable.

**4.5** Show that monotone convergence in the monotone convergence theorem cannot be replaced by pointwise convergence in the monotone convergence theorem and that the dominating function $g$ is needed in Lebesgue's dominated convergence theorem.

**4.6** Prove by example that the strict inequality in the Fatou lemma can occur.

**4.7** (i) Show that if $f$ is Riemann integrable on $[0, 1]$, then $f$ is Lebesgue integrable on $[0, 1]$ and $\int_{[0,1]} f dm = \int_0^1 f(x)dx$.

(ii) Construct a function such that it is Lebesgue integrable on $[0, 1]$ but not Riemann integrable.

# 5. Hilbert Spaces

As a straightforward generalization of finite-dimensional Euclidean spaces and the most important Banach spaces, the Hilbert spaces and the most refined analysis on these spaces will be discussed in this chapter.

## 5.1 Definition and Basic Properties

**Definition 5.1.1.** *Let $X$ be a complex linear space.*

(i) *A inner product on $X$ is a map $(x, y) \mapsto \langle x, y\rangle$ from $X \times X$ to $\mathbf{C}$ such that*

(a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle \ \forall x, y, z \in X$ *and* $\forall \alpha, \beta \in \mathbf{C}$*;*

(b) $\langle y, x\rangle = \overline{\langle x, y\rangle} \ \forall x, y \in X$*;*

(c) $\langle x, x\rangle \geq 0 \ \forall x \in X$ *and* $\langle x, x\rangle = 0$ *if and only if* $x = 0$*.*

(ii) *If $X$ is equipped with an inner product, then $X$ is said to be an inner product space or a pre-Hilbert space.*

(iii) *A Hilbert space is a complete, complex, inner product space.*

*Example 5.1.1.* (i) $\mathbf{C}^n$ is a Hilbert space under the inner product $\langle x, y\rangle = \sum_{j=1}^{n} x_j \bar{y_j}$ for $x = (x_1, ..., x_n), \ y = (y_1, ..., y_n) \in \mathbf{C}^n$.

(ii) Given $E \in \mathbf{M}_n$, let $L_2(E)$ be the set of all Lebesgue measurable functions $f : \mathbf{R}^n \to \mathbf{C}$ such that $\|f\|_2 = \big(\int_E |f|^2 dm\big)^{\frac{1}{2}} < \infty$. It is easy to see that the formula $\langle f, g\rangle = \int_E f\bar{g} dm$ defines an inner product on $L_2(E)$. In fact, $L_2(E)$ is a Hilbert space.

This example motivates the following result.

**Theorem 5.1.1.** *Let $X$ be a complex linear space equipped with the inner product $\langle\cdot,\cdot\rangle$. If $\|x\| = \sqrt{\langle x, x\rangle} \ \forall x \in X$, then we have*

(i) *Schwarz's Inequality: $|\langle x, y\rangle| \leq \|x\|\|y\| \ \forall x, y \in X$, with equality if and only if $x$ and $y$ are linearly dependent.*

(ii) *Norm: $\|\cdot\|$ defines a norm on $X$;*

(iii) *Parallelogram Law: $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \ \forall x, y \in X$.*

*Proof.* (i) If $\langle x, y\rangle = 0$, then there is nothing to argue. If $\langle x, y\rangle \neq 0$, then $x \neq 0$ and $y \neq 0$ holds. Let $\alpha = \langle x, y\rangle$ and $z = \alpha y$. Then for $t \in \mathbf{R}$ we have

$$0 \leq \langle x - tz, x - tz\rangle = \|x\|^2 - 2t|\langle x, y\rangle|^2 + t^2|\langle x, y\rangle|^2\|y\|^2.$$

The last expression is a quadratic function of $t$ whose absolute minimum occurs at $t = \|y\|^{-2}$. Substituting this value for $t$, we get

$$0 \le \|x - tz\|^2 = \|x\|^2 - \|y\|^{-2}|\langle x, y\rangle|^2$$

with equality if and only if $x - tz = x - \alpha t y = 0$, from which the desired result is immediate.

(ii) It is obvious that $\|x\| = 0$ if and only if $x = 0$ and that $\|\lambda x\| = |\lambda|\|x\|$. As for the triangle inequality, (i) is applied to imply that

$$\|x + y\|^2 = \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2 \le (\|x\| + \|y\|)^2.$$

(iii) This follows directly from expanding the inner products defining $\|x+y\|^2$ and $\|x - y\|^2$.

*Remark 5.1.1.* If a norm $\|\cdot\|$ on $X$ satisfies the parallelogram law above, then

$$\langle x, y\rangle = 4^{-1}\sum_{n=1}^{4} i^n\|x + i^n y\|^2 \quad (i^2 = -1),$$

defines an inner product on $X$. As a matter of fact, this $\langle\cdot,\cdot\rangle$ yields

$$\langle x, x\rangle = \|x\|^2 + \frac{i|1+i|^2}{4}\|x\|^2 - \frac{i|1-i|^2}{4}\|x\|^2 = \|x\|^2 \quad \forall x \in X.$$

To verify that $\langle\cdot,\cdot\rangle$ is actually an inner product, one suffices to prove that Definition 5.1.1 (i) (a) holds for $\alpha = \beta = 1$ and $\langle\lambda\cdot,\cdot\rangle = \lambda\langle\cdot,\cdot\rangle$ for any $\lambda \in \mathbf{C}$. To the former, we use the parallelogram law to achieve

$$\|u + v + w\|^2 + \|u + v - w\|^2 = 2\|u + v\|^2 + 2\|w\|^2$$

and

$$\|u - v + w\|^2 + \|u - v - w\|^2 = 2\|u - v\|^2 + 2\|w\|^2.$$

Hence

$$(\|u+v+w\|^2 - \|u-v+w\|^2) + (\|u+v-w\|^2 - \|u-v-w\|^2) = 2\|u+v\|^2 - 2\|u-v\|^2.$$

This infers

$$\Re\langle u + w, v\rangle + \Re\langle u - w, v\rangle = 2\Re\langle u, v\rangle.$$

The relation with $\Re$ replaced by $\Im$ is proved similarly. So

$$\langle u + w, v\rangle + \langle u - w, v\rangle = 2\langle u, v\rangle.$$

When $u = w$, one has $\langle 2u, v\rangle = 2\langle u, v\rangle$. Taking $u + w = x, u - w = y, v = z$, one gets

$$\langle x, z\rangle + \langle y, z\rangle = 2\langle\frac{x+y}{2}, z\rangle = \langle x + y, z\rangle.$$

To reach the latter, one notes that for any $m \in \mathbf{N}$,

$$\langle mx, y\rangle = \langle (m-1)x + x, y\rangle = \langle (m-1)x, y\rangle + \langle x, y\rangle = \cdots = m\langle x, y\rangle.$$

Thus for any $k \in \mathbf{N}$,

$$k\langle \frac{x}{k}, y\rangle = \langle x, y\rangle \quad \text{and} \quad \langle \frac{x}{k}, y\rangle = k^{-1}\langle x, y\rangle.$$

Consequently, for any $r = m/k$,

$$r\langle x, y\rangle = m\langle \frac{x}{k}, y\rangle = \langle \frac{m}{k}, y\rangle = \langle rx, y\rangle.$$

Since $\langle x, y\rangle$ is continuous functional in $x$, one concludes that $\lambda\langle x, y\rangle = \langle \lambda x, y\rangle$ for any $\lambda > 0$. If $\lambda < 0$ then

$$\lambda\langle x, y\rangle - \langle \lambda x, y\rangle = \lambda\langle x, y\rangle - |\lambda|\langle -x, y\rangle = \lambda\langle 0, y\rangle = 0.$$

Also, it is not hard to see that $i\langle x, y\rangle = \langle ix, y\rangle$. Finally, for any $\lambda = \mu + i\nu \in \mathbf{C}$,

$$\lambda\langle x, y\rangle = \mu\langle x, y\rangle + i\langle \nu x, y\rangle = \langle (\mu + i\nu)x, y\rangle,$$

as desired.

## 5.2 Orthogonality, Orthogonal Complement and Conjugate Spaces

First of all, let us consider orthogonality.

**Definition 5.2.1.** *Given a Hilbert space $X$. Let $x, y \in X$.*

(i) *The angle between $x$ and $y$ is defined by*

$$\theta_{x,y} = \begin{cases} 0, & \text{if} \quad x \text{ or } y = 0, \\ \arccos \frac{\Re\langle x,y\rangle}{\|x\|\|y\|}, & \text{otherwise.} \end{cases}$$

(ii) *$x$ and $y$ are called orthogonal provided $\langle x, y\rangle = 0$.*

(iii) *For any subset $S$ of $X$, $S^\perp = \{x \in X : \ \langle x, y\rangle = 0 \quad \forall y \in S\}$ is called the orthogonal complement of $S$.*

The following result has natural geometric and finite-dimensional antecedents.

**Theorem 5.2.1.** *A closed convex subset $M$ of a Hilbert space $X$ contains a unique element of smallest norm; that is, there exists exactly one $x \in M$ such that $\|x\| = \inf_{y\in M}\|y\|$.*

*Proof.* Let $\delta = \inf_{y\in M}\|y\|$ and let $(x_j)$ be any sequence in $M$ such that $\lim_{j\to\infty}\|x_j\| = \delta$. It is clear that such sequences exist. Using the parallelogram law we may write

$$\|x_k + x_j\|^2 + \|x_k - x_j\|^2 = 2(\|x_k\|^2 + \|x_j\|^2).$$

Since $M$ is convex, we conclude that $2^{-1}(x_j + x_k) \in M$ and hence

$$\|x_k + x_j\|^2 = 4\|2^{-1}(x_k + x_j)\|^2 \geq 4\delta^2.$$

The above statements imply $\|x_j - x_k\| \to 0$ as $j, k \to \infty$. Namely, $(x_j)$ is a Cauchy sequence and so there is a point $x \in M$ such that $x_j \to x$ in $X$. The continuity of the norm leads to $\|x\| = \delta$. Finally, $x$ is unique since given $y \in M$ with $\|y\| = \delta$, the forgoing argument can be applied to the sequence $x, y, x, y, x, y, ....$ to show that it is Cauchy, which can only be the case if $x = y$.

Next, we recall the definition of the direct sum.

**Definition 5.2.2.** *Given a linear space $Z$ and subspaces $X$ and $Y$, $Z$ is said to be the direct sum of $X$ and $Y$, denoted $Z = X \oplus Y$, provided every $z \in Z$ can be expressed uniquely in the form $z = x + y$, $x \in X$ and $y \in Y$ and $X \cap Y = \{0\}$.*

We are about to prove a theorem about decomposing Hilbert space into a direct sum of mutually orthogonal closed subspaces. Before doing so, we need the following result.

**Lemma 5.2.1.** *Let $M$ be a proper closed subspace of a Hilbert space $X$. Then there exists a (clearly not unique) nonzero $z \in X$ such that $\langle z, y\rangle = 0$ for all $y \in M$.*

*Proof.* Given any $x \in X$, the set $x + M$ is a closed convex set. Thus by Theorem 5.2.1 there exists a unique $z \in X$ such that $z \in x + M$ and $\|z\| = \inf_{y \in M} \|x + y\|$. We shall show that $\langle z, y\rangle = 0$ for all $y \in M$ and if we choose $x \notin M$ then $z \neq 0$. In fact, given any $y \in M$ and any $\alpha \in \mathbf{C}$, we have $z + \alpha y \in x + M$. By our choice of $z$ we then have $\|z + \alpha y\|^2 \geq \|z\|^2$ and, expanding,

$$|\alpha|^2\|y\|^2 + 2\Re(\bar{\alpha}\langle z, y\rangle) \geq 0 \quad \forall \alpha \in \mathbf{C}.$$

Choose $\theta \in [0, 2\pi)$ with

$$e^{-i\theta}\langle z, y\rangle = |\langle z, y\rangle|$$

and let $\alpha = te^{i\theta}$, $t \in \mathbf{R}$. The inequality obtained by substituting for $\alpha$ is then

$$t^2\|y\|^2 + 2t|\langle z, y\rangle| \geq 0 \quad \forall t \in \mathbf{R}.$$

By choosing $t$ negative and letting $t$ approach 0 we get $\langle z, y\rangle = 0$ since otherwise the negative term would eventually dominate the positive term and contradict the inequality. Since $y \in M$ was arbitrary, we thus have $\langle z, y\rangle = 0$ for all $y \in M$.

**Theorem 5.2.2.** *If $M$ is a closed subspace of a Hilbert space $X$, then $X = M \oplus M^{\perp}$.*

*Proof.* Given $x \in X$, apply the procedure in the proof of Lemma 5.2.1 to obtain $z \in X$ such that $z \in x + M$ and $\langle z, y\rangle = 0$ for all $y \in M$. Then $z \in M^{\perp}$ and $z = x - y$ for some $y \in M$. Hence $x = y + z$, $y \in M$, and $z \in M^{\perp}$. Noting that $M \cap M^{\perp} = \{0\}$ since

$$x \in M \cap M^{\perp} \Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0,$$

we thus have $X = M \oplus M^{\perp}$. Observe that $M^{\perp}$ is closed. Indeed, if $x_j \in M^{\perp}$ and $x_j \to x$ in $X$, then for any $y \in M$ we have by Schwarz's inequality,

$$|\langle y, x \rangle| \leq |\langle y, x_j - x \rangle| + |\langle y, x_j \rangle| \leq \|y\| \|x_j - x\| \to 0$$

as $j \to \infty$. So $x \in M^{\perp}$.

Finally, we consider the conjugate space of a Hilbert space. Below is the classical Riesz representation theorem.

**Theorem 5.2.3.** *Given a Hilbert space $X$ and $y \in X$, define $L_y : X \to \mathbf{C}$ by $L_y(x) = \langle x, y \rangle$. Then $L_y \in X^*$ and $\|L_y\| = \|y\|$. Conversely, for every $f \in X^*$ there exists a unique $y \in X$ such that $f = L_y$.*

*Proof.* Clearly, $L_y \in X^*$ with $\|L_y\| \leq \|y\|$. Note that $\|y\|^2 = L_y(y) \leq \|L_y\| \|y\|$. So $\|L_y\| = \|y\|$. Conversely, let $f \in X^*$ be given. If $f = 0$, then $f = L_y$ where $y = 0$. If $f \neq 0$, then we may assume without loss of generality that $\|f\| = 1$ since $f/\|f\| = L_y \Rightarrow f = L_{\|f\|y}$. For such an $f$, let

$$M = \{x \in X : f(x) = 0\}.$$

Then it is closed subspace of $X$. If $M = X$ then $f = 0$ and hence $f(x) = \langle x, 0 \rangle = L_0(x)$. If $M \neq X$ then by Lemma 5.2.1 we can obtain a unique nonzero $y \in M^{\perp}$ such that $f(y) \neq 0$. Now for any $x \in X$, we have

$$x - \frac{f(x)}{f(y)} y \in M \quad \text{and} \quad \langle x - \frac{f(x)}{f(y)} y, y \rangle = 0.$$

Accordingly,

$$\langle x, y \rangle = f(x) \langle \frac{y}{f(y)}, y \rangle.$$

By taking $z = f(\bar{y}) \|y\|^{-2} y$, we further get $f(x) = \langle x, z \rangle$, and then $f = L_z$. To see the uniqueness, assume that there is another point $w \in X$ such that $f = L_w$. Then $\langle x, w - z \rangle = 0$ for all $x \in X$, and hence

$$\|w - z\|^2 = \langle w - z, w - z \rangle = 0 \quad \text{and} \quad w = z.$$

We are done.

**Corollary 5.2.1.** *Let $X$ be a Hilbert space. Then the map $\sigma : X \to X^*$ given by $(\sigma x)(y) = \langle y, x \rangle$ is an isometric embedding from $X$ onto $X^*$.Moreover, $X^*$ is also a Hilbert space.*

*Proof.* The first part has just been proved in Theorem 5.2.3. Furthermore it is easily seen that $\sigma(x_1 + x_2) = \sigma(x_1) + \sigma(x_2)$ so the isometry is additive. Since

$$\sigma(\alpha x)(y) = \langle y, \alpha x \rangle = \bar{\alpha} \langle y, x \rangle = \bar{\alpha} \sigma(x)(y),$$

we conclude that $\sigma(\alpha x) = \bar{\alpha} \sigma(x)$. Note that the inherited inner product on $X^*$ must be defined by $\langle \sigma(x), \sigma(y) \rangle = \langle y, x \rangle$ so that its action on the second variable will be conjugate linear. Of course, this yields that $X^*$ is a Hilbert space.

## 5.3 Orthonormal Bases

**Definition 5.3.1.** *Let $X$ be a Hilbert space. A set $S \subset X$ is called orthogonal if any two different elements in $S$ are orthogonal. An orthonormal set is an orthogonal set consisting entirely of elements of norm* 1.

A well-known and constructive result about orthonormal sets is the Gram-Schmidt orthonormalization as follows.

**Theorem 5.3.1.** *Given any countable linearly independent set $\{x_j\}$ of a Hilbert space $X$, an orthonormal set $\{e_j\}$ can be constructed so that*

$$\mathit{span}(\{e_j\}_{j=1}^n) = \mathit{span}(\{x_j\}_{j=1}^n \quad \forall n \in \mathbf{N}.$$

*Proof.* Define $e_1 = x_1/\|x_1\|$, and, proceeding inductively, if $\{e_j\}_{j=1}^n$ are successfully defined, let $e_{n+1} = y_{n+1}/\|y_{n+1}\|$ where

$$y_{n+1} = x_{n+1} - \sum_{j=1}^n \langle x_{n+1}, e_j\rangle e_j.$$

Then $\|y_{n+1}\| \neq 0$ since otherwise we would have

$$x_{n+1} \in \operatorname{span}(\{e_j\}_{j=1}^n) = \operatorname{span}(\{x_j\}_{j=1}^n,$$

contradicting the linear independence of $\{x_j\}$. It is clear that

$$\operatorname{span}(\{e_j\}_{j=1}^{n+1}) = \operatorname{span}(\{x_j\}_{j=1}^{n+1})$$

since this is true for $\{e_j\}_{j=1}^n$ and $\{x_j\}_{j=1}^n$. Finally, for $j \leq n$,

$$\langle e_{n+1}, e_j\rangle = \frac{\langle x_{n+1}, e_j\rangle - \langle x_{n+1}, e_j\rangle\langle e_j, e_j\rangle}{\|y_{n+1}\|} = 0,$$

so the set $\{e_j\}$ obtained by this inductive construction is an orthonormal set with the desired property.

*Example 5.3.1.* Given $a < b, a, b \in [-\infty, \infty]$ and a function $\omega : (a,b) \to (0,\infty)$ with the property that the Riemann integral $\int_a^b t^n \omega(t)dt$ is finite for all $n \in \mathbf{N}$, define the Hilbert space $L_p^\omega(a,b)$ to be the linear space of Lebesgue measurable functions $f$ on $(a,b)$ with $\|f\|_\omega = \langle f, f\rangle_\omega^{\frac{1}{2}}$ where

$$\langle f, g\rangle_\omega = \int_{(a,b)} f\bar{g}\omega dm.$$

It may be verified that the linearly independent set $\{1, t, t^2, ...\}$ has a linear span dense in $L_2^\omega(a,b)$. The Gram-Schmidt orthonormalization process may be applied to this set to produce various families of classical orthonormal functions:

(i) If $\omega = 1$ and $a = -1, b = 1$, then the process generates the Legendre polynomials;

(ii) If $\omega(t) = (1 - t^2)^{-\frac{1}{2}}$ and $a = -1, b = 1$, then the process generates the Tchebychev polynomials;

(iii) If $\omega(t) = t^{q-1}(1-t)^{p-q}$ (with $q > 0$ and $p - q > -1$) and $a = 0, b = 1$, then the process generates the Jacobi polynomials;

(iv) If $\omega(t) = e^{-t^2}$ and $a = -\infty, b = \infty$, then the process generates the Hermite polynomials;

(v) If $\omega(t) = e^{-t}$ and $a = 0, b = \infty$, then the process generates the Laguerre polynomials;

**Lemma 5.3.1.** *Let $\{e_j\}_{j=1}^n$ be a finite orthonormal set of a Hilbert space $X$. Then*

(i) $\sum_{j=1}^n |\langle x, e_j\rangle|^2 \le \|x\|^2 \quad \forall x \in X$;

(ii) $\langle x - \sum_{j=1}^n \langle x, e_j\rangle e_j, e_k\rangle = 0 \quad \forall x \in X$ *and* $k = 1, 2, ..., n$.

*Proof.* (i) This follows from

$$0 \le \|x - \sum_{j=1}^n \langle x, e_j\rangle e_j\|^2 = \|x\|^2 - \sum_{j=1}^n |\langle x, e_j\rangle|^2,$$

where the last equality is obtained by expanding in the usual fashion and using orthonormality of $\{e_j\}_{j=1}^n$.

(ii) This follows from

$$\langle x - \sum_{j=1}^n \langle x, e_j\rangle e_j, e_k\rangle = \langle x, e_k\rangle - \langle x, e_k\rangle\langle e_k, e_k\rangle = 0.$$

**Corollary 5.3.1.** *Let $I$ be any index set and $\{e_j\}_{j\in I}$ be an orthonormal set of a Hilbert space $X$. Then $S = \{e_j :\ \langle x, e_j\rangle \ne 0\}$ is countable for any $x \in X$ .*

*Proof.* For each $n \in \mathbf{N}$, define

$$S_n = \{e_j :\ |\langle x, e_j\rangle|^2 > \|x\|^2/n\}.$$

By Lemma 5.3.1 each $S_n$ contains at most $n - 1$ members. Since $S = \bigcup_{n=1}^\infty S_n$, we conclude that $S$ is countable.

Given an arbitrary orthonormal set $\{e_j\}_{j\in I}$, we would like to extend the results of the last lemma. Using this lemma and its corollary, we denote by $\sum_j |\langle x, e_j\rangle|^2$ and $\sum_j \langle x, e_j\rangle e_j$ the series $\sum_{k=1}^\infty |\langle x, e_{j_k}\rangle|^2$ and $\sum_{j_k} \langle x, e_{j_k}\rangle e_{j_k}$ respectively, where we restrict ourselves to the countable number of $e_{j_k}$ for which $\langle x, e_{j_k}\rangle \ne 0$. The following result assures that both series are well defined.

**Theorem 5.3.2.** *Let $\{x_k\}_{k=1}^\infty$ be an orthonormal sequence in a Hilbert space $X$, and let $\{c_j\}$ be any sequence of scalars. Then $\sum_{k=1}^\infty c_k x_k$ is convergent in $X$ if and only if $\sum_{k=1}^\infty |c_k|^2 < \infty$, and if so*

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\|^2 = \sum_{k=1}^{\infty} |c_k|^2.$$

*Moreover,* $\sum_{k=1}^{\infty} c_k x_k$ *is independent of the order in which its terms are arranged.*

*Proof.* From the orthonormality of $\{x_k\}_{k=1}^{\infty}$ it follows that for $m > n$,

$$\left\| \sum_{k=n}^{m} c_k x_k \right\|^2 = \sum_{k=n}^{m} |c_k|^2.$$

This, together with the completeness of $X$, implies the 'iff' part of the theorem. If $n = 1$ and $m \to \infty$, then the desired equality follows.

To complete the proof, we assume that $\sum_{k=1}^{\infty} |c_k|^2 < \infty$ and let $y = \sum_{l=1}^{\infty} c_{k_l} x_{k_l}$ be a rearrangement of $x = \sum_{k=1}^{\infty} c_k x_k$. Then

$$\|x - y\|^2 = \langle x, x\rangle + \langle y, y\rangle - \langle x, y\rangle - \langle y, x\rangle,$$

and $\langle x, x\rangle = \langle y, y\rangle = \sum_{k=1}^{\infty} |c_k|^2$. If

$$s_m = \sum_{k=1}^{m} c_k x_k \quad \text{and} \quad t_m = \sum_{l=1}^{m} c_{k_l} x_{k_l},$$

then

$$\langle x, y\rangle = \lim_{m\to\infty} \langle s_m, t_m\rangle = \sum_{k=1}^{\infty} |c_k|^2.$$

Note that $\langle y, x\rangle = \overline{\langle x, y\rangle} = \langle x, y\rangle$. So it follows that $\|x - y\| = 0$ and hence $x = y$. We are done.

**Theorem 5.3.3.** *Let $I$ be any index set and $\{e_j\}_{j\in I}$ be an orthonormal set of a Hilbert space $X$.*

(i) *(Bessel's inequality)* $\sum_{j\in I} |\langle x, e_j\rangle|^2 \le \|x\|^2 \quad \forall x \in X$

(ii) $\langle x - \sum_{j\in I} \langle x, e_j\rangle e_j, e_k\rangle = 0 \quad \forall x \in X$ *and* $k \in I$.

*Proof.* (i) This follows from

$$\sum_{j\in I} |\langle x, e_j\rangle|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |\langle x, e_{j_k}\rangle|^2 \le \|x\|^2.$$

(ii) By using continuity of the inner product in its left-hand variable (which follows from Schwarz's inequality) we obtain

$$\begin{aligned}
\langle x - \sum_{j\in I} \langle x, e_j\rangle e_j, e_k\rangle &= \langle x, e_k\rangle - \langle \sum_{j\in I} \langle x, e_j\rangle e_j, e_k\rangle \\
&= \langle x, e_k\rangle - \langle \lim_{n\to\infty} \sum_{m=1}^{n} \langle x, e_{j_m}\rangle e_j, e_k\rangle \\
&= \langle x, e_k\rangle - \lim_{n\to\infty} \langle \sum_{m=1}^{n} \langle x, e_{j_m}\rangle e_j, e_k\rangle \\
&= \langle x, e_k\rangle - \langle x, e_k\rangle = 0.
\end{aligned}$$

**Definition 5.3.2.** *An orthonormal set $\{e_j\}$ of a Hilbert space $X$ is called an orthonormal basis for $X$ provided*

$$\langle x, e_j \rangle = 0 \quad \forall e_j \Rightarrow x = 0.$$

Less formally, this definition says that an orthonormal set is an orthonormal basis if it is impossible to adjoin an additional nonzero element to the set while still preserving its orthonormality.

**Theorem 5.3.4.** *Let $\{e_j\}$ be an orthonormal set in a Hilbert space $X$. Then the following are equivalent:*
(i) *$\{e_j\}$ is an orthonormal basis;*
(ii) *The closed linear span of $\{e_j\}$ is $X$;*
(iii) *(Parseval's identity) $\|x\|^2 = \sum_j |\langle x, e_j \rangle|^2 \quad \forall x \in X$.*

*Proof.* First, we prove that (i) implies (ii) and (iii). If (i) is true, then Definition 5.3.2 and Theorem 5.3.3 (ii) yield

$$x = \sum_j \langle x, e_j \rangle e_j \quad \forall x \in X.$$

Clearly, (iii) follows from Theorem 5.3.2.

Next, we prove that each of (ii) and (iii) implies (i).

(ii)⇒(i). Suppose that $y \in X$ satisfies $\langle y, e_j \rangle = 0 \ \forall e_j$. To prove $y = 0$, consider $S = \{x \in X : \ \langle y, x \rangle = 0\}$. It is easy to see that $S$ is linear subspace of $X$. Since $e_j \in S$, it follows that $S$ must contain the linear span of $\{e_j\}$. On the other hand, $S$ is closed in view of the continuity of the inner product, and so $S$ must contain the closure of the linear span of $\{e_j\}$. Hence $S = X$ by (ii). In particular, we have $y \in S$ and so $\langle y, y \rangle = 0$, whence $y = 0$ as required.

(iii)⇒(i). Suppose on the contrary that $\{e_j\}$ does not form an orthonormal basis of $X$. Then there exists a nonzero $x \in X$ such that $\langle x, e_j \rangle = 0 \quad \forall e_j$. Then

$$0 \neq \|x\|^2 = \sum_j |\langle x, e_j \rangle|^2 = 0.$$

This is a contradiction.

*Example 5.3.2.* $\{e^{int}\}_{n \in \mathbf{Z}}$ is an orthonormal basis for $L_2[0, 2\pi]$. This leads to the classical Fourier analysis.

**Theorem 5.3.5.** *Every Hilbert space has an orthonormal basis. Any orthonormal basis in a separable Hilbert space is countable.*

*Proof.* Let $X$ be a Hilbert space, and consider the collection $E$ of orthonormal subsets of $X$. From Theorem 5.3.1 it is seen that $E$ is nonempty and can be partially ordered under inclusion. If $F$ is any totally ordered subcollection of $E$, the set $U = \bigcup_{S \in F} S$ is member of $E$ and an upper bound for $F$. By Zorn's lemma there is a maximal orthonormal set $M$. Since $M$ is maximal, we conclude from Definition 5.3.2 that $M$ is an orthonormal basis.

If $X$ is separable, and if $\{e_j\}$ is an uncountable orthonormal basis, then for $j \neq k$, we have

$$\|e_j - e_k\|^2 = \|e_j\|^2 + \|e_k\|^2 = 2,$$

and so the open balls $B_{1/2}(e_j)$ (with center $e_j$ and radius $1/2$) are mutually disjoint. If $\{x_j\}_{j=1}^{\infty}$ is a dense countable sequence in $X$, because $\{e_j\}$ is uncountable, there is a ball $B_{1/2}(e_{j_0})$ that does not contain any of the points $x_j$. Hence $e_{j_0}$ is not in the closure of $\{x_j\}$. This contradicts the density of $\{x_j\}$ in $X$. We are done.

**Corollary 5.3.2.** *Any two infinite dimensional separable Hilbert spaces are isometrically isomorphic.*

*Proof.* Suppose $X$ and $Y$ are two such spaces. Theorem 5.3.5 tells us that there are sequences $\{x_j\}$ and $\{y_j\}$ that form orthonormal bases for $X$ and $Y$ respectively. If $x \in X$ and $y \in Y$, then

$$x = \sum_{j=1}^{\infty} \langle x, x_j \rangle x_j \quad \text{and} \quad y = \sum_{j=1}^{\infty} \langle y, y_j \rangle y_j.$$

Define a map $T : X \to Y$ by $Tx = y$ if $\langle x, x_j \rangle = \langle y, y_j \rangle$. It is clear that $T$ is linear and one-to-one, and it maps $X$ onto $Y$ since $(\langle x, x_j \rangle)_{j=1}^{\infty}$ and $(\langle y, y_j \rangle)_{j=1}^{\infty}$ run through all of $\ell_2$. Also

$$\|Tx\|^2 = \sum_{j=1}^{\infty} |\langle y, y_j \rangle|^2 = \sum_{j=1}^{\infty} |\langle x, x_j \rangle|^2 = \|x\|^2,$$

so $T$ is isometrically isomorphic. The proof is complete.

## 5.4 Adjoint Operators

**Definition 5.4.1.** *Let $X$ be a Hilbert space. Then the adjoint operator $T^*$ of a $T \in B(X)$ is defined by*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad \forall x, y \in X.$$

Of course, this adjoint operator is unique since for each fixed $y \in X$, $\langle Tx, y \rangle$ is a continuous linear functional on $X$. Thus by the Riesz theorem there exists a unique $z \in X$ such that $\langle Tx, y \rangle = L_z(x) = \langle x, z \rangle$, and we define $T^* y = z$. The linearity of $T^*$ follows from the following calculation:

$$\begin{aligned} \langle x, T^*(\alpha y_1 + \beta y_2) \rangle &= \langle Tx, \alpha y_1 + \beta y_2 \rangle \\ &= \bar{\alpha} \langle Tx, y_1 \rangle + \bar{\beta} \langle Tx, y_2 \rangle \\ &= \langle x, \alpha T^* y_1 + \beta T^* y_2 \rangle \end{aligned}$$

**Theorem 5.4.1.** *Let $X$ be a Hilbert space. Then the operator $*$ maps $B(X)$ to itself, and has the following properties for all $T, S \in B(X)$ and $\alpha, \beta \in \mathbf{C}$:*
(i) $(\alpha T + \beta S)^* = \bar{\alpha} T^* + \bar{\beta} S^*$;
(ii) $(TS)^* = S^* T^*$;
(iii) $T^{**} = T$;
(iv) $\|T^*\| = \|T\|$;
(v) $\|T^* T\| = \|T\|^2$.

*Proof.* First of all, we verify that $*$ maps $B(X)$ to $B(X)$. The linearity of $T^*$ follows from the following calculation:

$$\begin{aligned}\langle x, T^*(\alpha y_1 + \beta y_2)\rangle &= \langle Tx, \alpha y_1 + \beta y_2\rangle \\ &= \bar{\alpha}\langle Tx, y_1\rangle + \bar{\beta}\langle Tx, y_2\rangle \\ &= \langle x, \alpha T^* y_1 + \beta T^* y_2\rangle\end{aligned}$$

By the definition of the operator norm, we have

$$\begin{aligned}\|T^*\| &= \sup_{\|y\|=1} \|T^* y\| \\ &\leq \sup_{\|x\|=1, \|y\|=1} |\langle x, T^* y\rangle| \\ &= \sup_{\|x\|=1, \|y\|=1} |\langle Tx, y\rangle| \\ &\leq \sup_{\|x\|=1} \|Tx\| \\ &= \|T\|.\end{aligned}$$

This implies $*$ is bounded on $B(X)$.

Next, we check those five properties.

(i) For any $x, y \in X$, we have

$$\langle x, (\alpha T + \beta S)^* y\rangle = \alpha\langle Tx, y\rangle + \beta\langle Sx, y\rangle = \langle x, (\bar{\alpha} T^* + \bar{\beta} S^*) y\rangle.$$

(ii) This follows from

$$\langle x, (TS)^* y\rangle = \langle TSx, y\rangle = \langle Sx, T^* y\rangle = \langle x, S^* T^* y\rangle.$$

(iii) This follows from

$$\langle x, T^{**} y\rangle = \langle T^* x, y\rangle = \overline{\langle y, T^* x\rangle} = \overline{\langle Ty, x\rangle} = \langle x, Ty\rangle.$$

(iv) It is known that $\|T^*\| \leq \|T\|$ and $T^{**} = T$. So $\|T\| \leq \|T^*\|$. This gives $\|T\| = \|T^*\|$.

(v) By (iv), we get

$$\|T^* T\| \leq \|T^*\| \|T\| = \|T\|^2.$$

For the reverse inequality, we note that

$$\|Tx\|^2 = \langle T^* Tx, x\rangle \leq \|T^* Tx\| \|x\| \leq \|T^* T\| \|x\|^2.$$

So, $\|T\|^2 \leq \|T^* T\|$, which completes the proof.

*Example 5.4.1.* Let $\ell_2$ be the Hilbert space of square summable complex-valued sequences. If $T$ is the operator on $\ell_2$ defined by $T(x_1, x_2, ...) = (0, x_1, x_2, ...)$, then $T^*(x_1, x_2, ...) = (x_2, x_3, ...)$. Clearly, $\|T\| = \|T^*\| = 1$.

## 5.5 Self-adjoint, Normal, Unitary, and Projective Operators

**Definition 5.5.1.** *Let $X$ be a Hilbert space. An operator $T \in B(X)$ is said to be self-adjoint provided $T = T^*$.*

**Theorem 5.5.1.** *Let $X$ be a Hilbert space. Then $T \in B(X)$ is self-adjoint if and only if $\langle Tx, x\rangle$ is real for all $x \in X$.*

*Proof.* If $T \in B(X)$ is self-adjoint, then $\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$ for any $x \in X$. Hence $\langle Tx, x\rangle$ is real for all $x \in X$.

Conversely, If $f(x) = \langle Tx, x\rangle$, then

$$f(x + y) = f(x) + f(y) + \langle Ty, x\rangle + \langle Tx, y\rangle$$

and

$$f(x + iy) = f(x) + f(y) + i\langle Ty, x\rangle - i\langle Tx, y\rangle.$$

Since $f(x)$ is real-valued, we conclude from the last two identities that there are $r, s \in \mathbf{R}$ such that

$$\langle Ty, x\rangle + \langle Tx, y\rangle = r \quad \text{and} \quad \langle Ty, x\rangle - \langle Tx, y\rangle = is.$$

This yields

$$\langle Ty, x\rangle = \frac{r + is}{2} \quad \text{and} \quad \langle Tx, y\rangle = \frac{r - is}{2},$$

and thus

$$\langle Ty, x\rangle = \overline{\langle Tx, y\rangle} = \overline{\langle x, T^*y\rangle} = \langle T^*y, x\rangle.$$

Of course, $T = T^*$. We are done.

*Example 5.5.1.* Given a Hilbert space $X$, let $T \in B(X)$ be such that $\langle Tx, x\rangle = 0 \quad \forall x \in X$. Then for $x, y \in X$ and $\alpha \in \mathbf{C}$ we have

$$0 = \langle T(\alpha x + y, \alpha x + y\rangle = \alpha\langle Tx, y\rangle + \bar{\alpha}\langle Ty, x\rangle.$$

Taking $\alpha = i, 1$ respectively, we get

$$\langle Tx, y\rangle - \langle Ty, x\rangle = 0 \quad \text{and} \quad \langle Tx, y\rangle + \langle Ty, x\rangle = 0,$$

which implies $\langle Tx, y\rangle = 0$ and so $\langle Tx, Tx\rangle = 0$. Thus $T = 0$.

**Definition 5.5.2.** *Let $X$ be a Hilbert space. An operator $T \in B(X)$ is said to be:*

(i) *Normal if $TT^* = T^*T$.*
(ii) *Positive if $\langle Tx, x\rangle \geq 0 \quad \forall x \in X$;*

The following result shows that the normal operators correspond to complex numbers.

**Theorem 5.5.2.** *Given a Hilbert space $X$, let $T \in B(X)$. Then the following are equivalent:*
(i) *$T$ is normal;*
(ii) *$T = T_1 + iT_2$ where $T_1$ and $T_2$ are self-adjoint and $T_1T_2 = T_2T_1$;*
(iii) $\|Tx\| = \|T^*x\| \quad \forall x \in X$.

*Proof.* (i)⇒(ii). Put

$$T_1 = \frac{T+T^*}{2} \quad \text{and} \quad T_2 = \frac{T-T^*}{2i}.$$

It is easy to see that $T_1$ and $T_2$ are self-adjoint and commute.

(ii)⇒(iii). Using the given decomposition of $T$ together with $T_1T_2 = T_2T_1$, we have

$$\begin{aligned}
\|Tx\|^2 &= \langle x, T^*Tx\rangle \\
&= \langle x, (T_1 - iT_2)(T_1 + iT_2)x\rangle \\
&= \langle x, (T_1T_1 + T_2T_2)x\rangle \\
&= \langle x, (T_1 + iT_2)(T_1 - iT_2)x\rangle \\
&= \langle x, TT^*x\rangle \\
&= \|T^*x\|^2.
\end{aligned}$$

(iii)⇒(i). For any $x \in X$, we have

$$\langle (TT^* - T^*T)x, x\rangle = \langle TT^*x, x\rangle - \langle T^*Tx, x\rangle = \|T^*x\|^2 - \|Tx\|^2 = 0.$$

So, $TT^* - T^*T = 0$ and $T$ is normal.

Below is the basic structure of a positive operator.

**Theorem 5.5.3.** *Given a Hilbert space $X$, let $I \in B(X)$ be the identity operator. If $T \in B(X)$ is a positive operator, then*
(i) *$T + I$ is invertible;*
(ii) *No negative real number can belong to the spectrum of $T$:*

$$\sigma(T) = \{\lambda \in C : \; T - \lambda I \quad \textit{is not invertible}\}.$$

*Proof.* (i) Since $T \in B(X)$ is a positive operator, we conclude that

$$(T + I)x = 0 \Rightarrow \langle (T + I)x, x\rangle = 0 \Rightarrow 0 \le \langle Tx, x\rangle = -\|x\|^2 \Rightarrow x = 0$$

and injectivity of $T + I$ follows.

Surjectivity will follows if we can prove that $M = \text{range}(T + I)$ is both dense and closed. The argument used for injectivity applies to $x \in M^\perp$ to infer that $\langle (T + I)x, x\rangle = 0 \; \forall x \in M^\perp$ and so $M^\perp = 0$. On the other hand, for any $x \in X$, we have

$$\|(T+I)x\|^2 = \|Tx\|^2 + 2\langle Tx, x\rangle + \|x\|^2$$

and hence

$$\|x\| \le \|(T+I)x\|.$$

Consequently, if $\big((T+I)x_j\big)$ is a Cauchy sequence in $M$, then $(x_j)$ is a Cauchy sequence in $X$ and hence it is convergent. This shows that $M$ is complete and thereby closed in $X$. From Theorem 5.2.2 it turns out that $X = M \oplus M^\perp = M$; that is, $T+I$ is surjective.

(ii) Assume $r \in \sigma(T)$, $r < 0$. Then $-r^{-1}T + I = -r^{-1}(T - rI)$ is not invertible. Since $-r^{-1}T$ is a positive operator, we conclude from (i) that $T - rI$ is invertible, a contradiction.

**Definition 5.5.3.** *Let $X$ be a Hilbert spaces. Then $T \in B(X)$ is called unitary provided $T$ is an isometric ($\|Tx\| = \|x\|$) isomorphism of $X$ onto itself.*

*Example 5.5.2.* The operator $T$ defined by $Tx = (x_2, x_1, ...)$ for any $x = (x_1, x_2, ...) \in \ell_2$ is a unitary operator on $\ell_2$.

We have a characterization of the unitary operators as follows.

**Theorem 5.5.4.** *Let $X$ be a Hilbert spaces. For $T \in B(X)$ the following are equivalent:*

(i) *$T$ is unitary;*
(ii) *$TT^* = T^*T = I$;*
(iii) *$\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in X$, and $T$ is surjective.*

*Proof.* (i)$\Rightarrow$(iii). If $T$ is unitary, then $\|Tx\| = \|x\|$ and hence

$$\begin{aligned}\langle Tx, Ty\rangle &= 4^{-1}\sum_{n=1}^{4} i^n \|Tx + i^n Ty\|^2 \\ &= 4^{-1}\sum_{n=1}^{4} i^n \|T(x + i^n y)\|^2 \\ &= 4^{-1}\sum_{n=1}^{4} i^n \|x + i^n y\|^2 \\ &= \langle x, y\rangle.\end{aligned}$$

Surjectivity of $T$ follows the definition of a unitary operator.

(iii)$\Rightarrow$(ii). Clearly, (iii) implies that $T$ is injective. Moreover

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \langle x, x\rangle.$$

This, along with Example 5.5.1, yields

$$\langle (T^*T - I)x, x\rangle = 0 \Longrightarrow T^*T = I \Rightarrow TT^* = I.$$

(ii)$\Rightarrow$(i). If (ii) is true, then $T$ is surjective and

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, x\rangle = \|x\|^2.$$

This implies that $T$ is injective, and thereby $T$ is unitary.

*Example 5.5.3.* Let $T = 2iI$ on a given Hilbert space $X$, where $I$ is the identity operator on $X$. Then $T^* = -2iI$ and $TT^* = T^*T = 4I$. Thus, $T$ is normal operator but not unitary nor self-adjoint.

The final class of operators that we consider is the important class of projection operators.

**Definition 5.5.4.** *Let $X$ be a Hilbert space. Then an operator $T \in B(X)$ is called a projection provided $T^2 = T$.*

*Example 5.5.4.* Let $M$ be a closed subspace of a Hilbert space $X$. Then $X = M \oplus M^\perp$; that is, any $x \in X$ there are unique $y \in M$ and $z \in M^\perp$ such that $x = y + z$. If $P_M(x) = y$ then $P_M$ is a projection. In fact,

$$P_M^2(x) = P_M(y) = y = P_M(x) \Rightarrow P_M^2 = P_M.$$

The fact that $P_M \in B(X)$ follows from

$$\|x\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2 \geq \|y\|^2 = \|P_M(x)\|^2 \Rightarrow \|P_M\| \leq 1.$$

Traditionally, $P_M$ is called the orthogonal projection of $X$ onto $M$. Moreover, if $M \neq \{0\}$ then $P_M \neq 0$ and hence for any $x \in M \setminus \{0\}$ we have $x = x + 0$ and $P_M x = x$, giving $\|P_M\| = 1$.

The following theorem singles out the above orthogonal projections as a very important subclass.

**Theorem 5.5.5.** *Let $X$ be a Hilbert space. If $T \in B(X)$ is a projection, then the following are equivalent:*
(i) *$T$ is positive;*
(ii) *$T$ is self-adjoint;*
(iii) *$T$ is normal;*
(iv) *$T$ is the orthogonal projection on its range $T(X)$.*

*Proof.* Since (i)⇒(ii)⇒(iii) are straightforward, we only verify (iii)⇒(iv)⇒(i).

(iii)⇒(iv). Assume that (iii) holds. To reach (iv), let $M = \mathrm{range}(T)$.

We first prove that $M$ is closed. Given $(y_j)_{j=1}^\infty$ in $M$ with $y_j \to y$, we have $y_j = Tx_j$, $x_j \in X$. Since $T$ is a projection, we can conclude from $y_j \to y$ that

$$y_j = Tx_j = T^2 x_j = Ty_j \to Ty \Rightarrow y = Ty \in M$$

and hence $M$ is closed.

Next, given $x \in X$ write $x = y + z$, $y \in M$ and $z \in M^\perp$. We must verify that $Tx = y$. Because $Tx = Ty + Tz$, it suffices to prove that $Ty = y$ and $Tz = 0$. Since $y \in M$, there is $w \in X$ such that $y = Tw$ and so

$$Ty = T^2 w = Tw = y.$$

By definition of $M$ we have $Tz \in M$, and if we can also prove that $Tz \in M^\perp$, then we can conclude that $Tz = 0$. Accordingly, given any $u \in M$, then there is $v \in X$ such that $u = Tv$ and

$$\langle Tz, u\rangle = \langle Tz, Tv\rangle = \langle z, T^*Tv\rangle = \langle z, TT^*v\rangle = 0$$

due to $z \in M^\perp$ and $TT^*v \in M$. That is to say, $Tz \in M^\perp$.

(iv)⇒(i). For any $x \in X$, let

$$x = y + z, \quad y \in M = T(X) \quad \text{and} \quad z \in M^\perp.$$

Then

$$\langle Tx, x\rangle = \langle y, y+z\rangle = \langle y, y\rangle + \langle y, z\rangle = \langle y, y\rangle \geq 0.$$

Hence $T$ is positive and the proof is complete.

Last of all, it is worth pointing out that there is a natural one-to-one correspondence between projection operators $T$ and direct sum decompositions $X = M \oplus N$, $M$ and $N$ closed. As a matter of fact, if $T$ is a projection, then $X = M \oplus N$ where $M = T(X)$ and $N = (I - T)(X)$. Conversely, if $X = M \oplus N$, then we define $T$ as was done for the orthogonal projection but using $N$ instead of $M^\perp$. The orthogonal projections are precisely those that arise when $N = M^\perp$, and these are generally the projections of interest in Hilbert space theory.

## 5.6 Compact Operators

In this section, we consider compact operators on Banach spaces which are stronger than boundedness, in particular, prove the spectral theorem for compact self-adjoint operators on Hilbert spaces.

First of all, we need the definition of spectrum of a linear operator on Banach space.

**Definition 5.6.1.** *Let $X$ be a Banach space and $T \in B(X)$. Then* (i)

$$\sigma(T) = \{\lambda \in \mathbf{C} : \lambda I - T \quad \text{is not invertible}\}$$

*is called the spectrum of $T$, where $I \in B(X)$ is the identity operator.*

(ii) *If $Tx = \lambda x$, then $x$ and $\lambda$ are called an eigenvector and eigenvalue of $T$ respectively.*

*Example 5.6.1.* (i) It is clear that $\sigma(I) = \{1\}$ since $\lambda I - I$ is invertible if and only if $\lambda \neq 1$.

(ii) If $X$ is a finite dimensional normed vector space, then we can identify every continuous linear operator $T$ on $X$ with a square matrix $A$ and hence $\lambda I - T$ is invertible if and only if $\lambda I - A$ is invertible, where the second $I$ stands for the identity matrix. This yields that $\sigma(T)$ consists exactly of all the eigenvalues of $T$, and these are precisely the eigenvalues of $A$.

(iii) If $Tx = (0, x_1, x_2, ...)$ for $x \in \ell_2$ then $T$ has no eigenvalues.

**Theorem 5.6.1.** *Let $X$ be a Banach space and $T \in B(X)$. Then*

(i) *Every eigenvalue of $T$ belongs to $\sigma(T)$.*

(ii) *If $|\lambda| > \|T\|$ then $\lambda \notin \sigma(T)$; that is, $\sigma(T) \subseteq \{\lambda \in \mathbf{C} : |\lambda| \leq \|T\|\}$.*

(iii) *$\sigma(T)$ is closed in $\mathbf{C}$.*

*Proof.* (i) $(\lambda I - T)x = 0$ implies $\lambda I - T$ has non-trivial kernel and cannot thus be invertible.

(ii) $|\lambda| > \|T\|$ yields $\|\lambda^{-1}Y\| < 1$. This, together with Theorem 3.1.5 (ii), infers that $I - \lambda^{-1}T$ is invertible, and so $\lambda I - T$ is invertible, whence $\lambda \notin \sigma(T)$.

(iii) Suppose $f(\lambda) = \lambda I - T$. This is a map from $\mathbf{C}$ to $B(X)$. It is clear that

$$\|f(\lambda_1) - f(\lambda_2)\| = |\lambda_1 - \lambda_2| \quad \forall \lambda_1, \lambda_2 \in \mathbf{C},$$

and so that $f$ is continuous. By Theorem 3.1.5 (iii), we have that $S$, the set of all non-invertible linear operators in $B(X)$, is closed. Therefore, $\sigma(T) = f^{-1}(S)$ is closed due to the continuity of $f$.

To see the existence that points of $\sigma(T)$ lie on the unit circle centered at origin, we need to consider the compact operators.

**Definition 5.6.2.** *Let $(X, \|\cdot\|_X)$ be a Banach space. Then a linear operator $T$ on $X$ is called compact provide that for every bounded sequence $(x_j)_{j=1}^i nfty$ in $X$, the sequence $\big(T(x_j)\big)_{j=1}^\infty$ has a convergent subsequence.*

*Example 5.6.2.* (i) A compact linear operator $T$ on $X$ must be continuous. For otherwise, there exists a bounded sequence $(x_j)_{j=1}^i nfty$ in $X$ such that $\lim_{j\to\infty} \|Tx_j\| = \infty$, so $\big(T(x_j)\big)_{j=1}^\infty$ cannot have any convergent subsequence.

(ii) If $T$ has finite rank; that is, $\dim\big(T(X)\big) < \infty$, then $T$ is compact.

**Lemma 5.6.1.** *Let $X$ be a Banach space. If $T_j \in B(X)$ is compact and $\lim_{j\to\infty} \|T_j - T\| = 0$ then $T$ is a compact operator on $X$.*

*Proof.* Suppose $(x_j)_{j=1}^\infty$ is a bounded sequence in $X$. Since $T_1$ is compact, there is a subsequence $(x_{1,m})_{m=1}^\infty$ out of $(x_j)_{j=1}^\infty$ such that $(T_1 x_{1,m})_{m=1}^\infty$ is convergent. Also since $T_2$ is compact, there is a subsequence $(x_{2,m})_{m=1}^\infty$ out of $(x_{1,m})_{m=1}^\infty$ such that $(T_2 x_{2,m})_{m=1}^\infty$ is convergent. Continuing this process, we can obtain subsequence $(x_{n+1,m})_{m=1}^\infty$ out of $(x_{n,m})_{m=1}^\infty$ such that $(T_{n+1} x_{n+1,m})_{m=1}^\infty$ is convergent. It follows that $(T_n x_{m,m})_{m=1}^\infty$ is convergent for any $n \in \mathbf{N}$.

Note that $(x_{m,m})_{m=1}^\infty$ is bounded in $X$. So there is a constant $c > 0$ such that $\sup_{m\in\mathbf{N}} \|x_{m,m}\|_X \le c$. Given $\epsilon > 0$, $\lim_{j\to\infty} \|T_j - T\| = 0$ implies that there is an $N \in \mathbf{N}$ such that $\|T_N - T\| < \frac{\epsilon}{3c}$. Since $(T_N x_{m,m})_{m=1}^\infty$ is convergent in $X$, there is an $N_1 \in \mathbf{N}$ such that

$$k, l > N_1 \Longrightarrow \|T_N x_{k,k} - T_N x_{l,l}\| < \frac{\epsilon}{3}.$$

With this, we achieve

$$\begin{aligned}
&\|Tx_{k,k} - Tx_{l,l}\|_X \\
&\le \|Tx_{k,k} - T_N x_{k,k}\|_X + \|T_N x_{k,k} - T_N x_{l,l}\|_X + \|T_N x_{l,l} - Tx_{l,l}\|_X \\
&\le \|T - T_N\|\|x_{k,k}\|_X + \frac{\epsilon}{3} + \|T - T_N\|\|x_{l,l}\|_X \\
&< \frac{\epsilon c}{3c} + \frac{\epsilon}{3} + \frac{\epsilon}{3c} = \epsilon,
\end{aligned}$$

and so $(Tx_{m,m})_{m=1}^{\infty}$ is convergent in $X$ owing to the fact that $X$ is a Banach space under the norm $\|\cdot\|_X$.

*Example 5.6.3.* (i) Suppose $X$ is a Hilbert space with orthonormal basis $\{e_j\}_{j=1}^{\infty}$ and $T \in B(X)$ is defined by $Te_j = \lambda_j e_j$. Then $T$ is compact on $X$ if and only if $\lim_{j\to\infty}\lambda_j = 0$. To see this, if $\lim_{j\to\infty}\lambda_j = 0$, then for each $n \in \mathbf{N}$ let $T_n e_j = Te_j$ if $j \le n$ and $T_n e_j = 0$ if $j > n$. Then $T_n$ has finite rank and hence is compact, and it is straightforward to see that

$$\|T - T_n\| = \sup_{j>n} |\lambda_j|.$$

Thus $T$ is compact due to Lemma 5.6.1. Conversely, if there is some $\epsilon > 0$ with $\{j : |\lambda_j| > \epsilon\}$ being infinite, then

$$\|Te_j - Te_k\|^2 = |\lambda_j|^2 + ||\lambda_k|^2 \ge 2\epsilon^2.$$

Therefore $\{Te_j\}_{j=1}^{\infty}$ has no convergent subsequence, giving that $T$ is not compact.

(ii) Suppose $X$ is a Hilbert space with orthonormal basis $\{e_j\}_{j=1}^{\infty}$. If $T \in B(X)$, then $T$ is called a Hilbert-Schmidt operator provided $\sum_{j,k=1}^{\infty} |\langle Te_j, e_k\rangle|^2 < \infty$. Then $T$ is compact. To see this, just define $T_n e_j = e_j$ if $j \le n$ and $T_n e_j = 0$ if $j > n$ and prove $\lim_{n\to\infty}\|T - T_n\| = 0$.

Before reaching the spectral theorem for compact operators, we also need one more auxiliary result on self-adjoint operators.

**Lemma 5.6.2.** *Let $X$ be a Hilbert space and $T \in B(X)$.*
(i) *If $T$ is self-adjoint and $TM \subset M$, then $TM^{\perp} \subset M^{\perp}$.*
(ii) $\|T\| = \sup\{|\langle Tx, y\rangle| : \|x\|, \|y\| \le 1\}$.
(iii) *If $T$ is self-adjoint, then $\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| \le 1\}$*

*Proof.* (i) If $y \in M^{\perp}$ then for any $x \in M$ one has $Tx \in M$ and hence $\langle x, Ty\rangle = \langle Tx, y\rangle = 0$.

(ii) It is clear that the Schwarz inequality implies

$$\|T\| \ge \sup\{|\langle Tx, y\rangle| : \|x\|, \|y\| \le 1\}.$$

For the reverse inequality, we may just consider the case $T \ne 0$. Regarding any $y$ with $Ty \ne 0$, we have

$$\left|\left\langle Ty, \frac{Ty}{\|Ty\|}\right\rangle\right| = \|Ty\|.$$

Taking the supremum over all $y$ with $\|y\| = 1$ yields $\|T\|$ and then

$$\|T\| \le \sup\{|\langle Tx, y\rangle| : \|x\|, \|y\| \le 1\}.$$

(iii) Let $\kappa = \sup\{|\langle Tx, x\rangle| : \|x\| \le 1\}$. It is easy to see that $\kappa \le \|T\|$. By (ii), we suffice to verify $|\langle Tx, y\rangle| \le \kappa\|x\|\|y\|$ for all $x, y \in X$. As this estimate is unchanged if we multiply $y$ by a complex number of modulus 1, we may assume $\langle Tx, y\rangle \in \mathbf{R}$. Since $T = T^*$, we can conclude that

$$\langle T(x+y),(x+y)\rangle = \langle Tx,x\rangle + 2\langle Tx,y\rangle + \langle Ty,y\rangle$$

and

$$\langle T(x-y),(x-y)\rangle = \langle Tx,x\rangle - 2\langle Tx,y\rangle + \langle Ty,y\rangle$$

Subtracting the last equation from the one preceding it, we obtain

$$4\langle Tx,y\rangle = \langle T(x+y),(x+y)\rangle - \langle T(x-y),(x-y)\rangle.$$

This, together with the definition of $\kappa$ and the parallelogram law, derives

$$|\langle Tx,y\rangle| \le \frac{\kappa}{4}(\|x+y\|^2 + \|x-y\|^2) \le \frac{\kappa}{2}(\|x\|^2 + \|y\|^2).$$

For any $\lambda > 0$, we apply the inequality to $\sqrt{\lambda}x, y/\sqrt{\lambda}$, and thus have

$$|\langle Tx,y\rangle| \le \frac{\kappa}{2}(\lambda\|x\|^2 + \lambda^{-1}\|y\|^2) \quad \forall x,y \in X.$$

In particular, $\lambda = \|y\|\|x\|^{-1}$ (of course, $x \neq 0$ is assumed, otherwise, there is nothing to argue) produces $|\langle Tx,y\rangle| \le \kappa$, as desired.

Now, it is time to state the spectral theorem for compact operators.

**Theorem 5.6.2.** *Let $X$ be a Hilbert space and $T \in B(X)$ be compact self-adjoint. Then*

(i) *At least one of $\pm\|T\|$ is an eigenvalue of $T$, and so belongs to $\sigma(T)$.*

(ii) *$X$ has an orthonormal basis consisting of eigenvectors of $T$.*

*Proof.* (i) The assertion is obviously true whenever $T = 0$. So, we just consider the case $T \neq 0$. Noticing Lemma 5.6.2 (iii), we get a sequence $(x_j)_{j=1}^\infty$ in $X$ such that $\|x_j\| = 1$ and $\lim_{j\to\infty} |\langle Tx_j, x_j\rangle| = \|T\|$. Since $T$ is self-adjoint, we conclude that $\langle Tx_j, x_j\rangle$ is real. Replacing $(x_j)_{j=1}^\infty$ by a subsequence if necessary, we may therefore assume that $\lim_{j\to\infty}\langle Tx_j, x_j\rangle = \lambda$, where $\lambda = \pm\|T\|$. Then a simple calculation gives

$$\|Tx_j - \lambda x_j\|^2 \le \|T\|^2\|x_j\|^2 - 2\lambda\langle Tx_j,x_j\rangle + \lambda^2\|x_j\|^2 = 2\lambda^2 - 2\lambda\langle Tx_j,x_j\rangle \to 0$$

as $j \to \infty$. It follows that $Txj - \lambda x_j \to 0$ in $X$. However, since $T$ is compact too, we conclude that there is a subsequence $(x_{j_k})_{k=1}^\infty$ out of $(x_j)_{j=1}^\infty$ such that $(Tx_{j_k})_{k=1}^\infty$ is convergent to $y$ in $X$. This implies $\lambda x_{j_k} \to y$ as $k \to \infty$. The continuity of $T$ further infers $\lambda Tx_{j_k} \to Ty$ as $k \to \infty$. Thus $T(y) = \lambda y$ with

$$\|y\| = \lim_{k\to\infty} \|\lambda x_{j_k}\| = |\lambda| = \|T\| \neq 0.$$

Namely, $\lambda$ and $y$ are eigenvalue and eigenvector of $T$.

(ii) By Zorn's lemma we can choose an orthonormal set of eigenvectors of $T$ which is maximal among all orthonormal sets of eigenvectors. Let $M$ be the closure of the span of these vectors. It suffices to prove $M = X$. It is clear that $TM \subseteq M$, and so $TM^\perp \subseteq M^\perp$ by Lemma 5.6.2 (i). Then the restriction of $T$ on $M^\perp$ belongs to $B(M^\perp)$ and is self-adjoint. Since $X = M \oplus M^\perp$, the proof will be complete if one shows $M^\perp = \{0\}$. Suppose now $M^\perp \neq \{0\}$. Then by the preceding argument for (i), there is an eigenvector of $T$ in $M^\perp$. This clearly contradicts the maximality property of the orthonormal set generating $M$.

## Exercises

**5.1** Let $C[-1,1]$ be the space of all continuous complex-valued functions on $[-1,1]$ with the 2-norm and the inner product $\langle f, g\rangle = \int_{-1}^{1} f(x)\overline{g(x)}dx$. Prove that $C[-1,1]$ is an inner product space but not a Hilbert space.

**5.2** Suppose $\ell_p$, $1 \le p < \infty$, is the space of all complex-valued sequences with $p$-norm. Prove that $\ell_2$ is a Hilbert space under the inner product $\langle x, y\rangle = \sum_{j=1}^{\infty} x_j \bar{y_j}$ for $x = (x_j)_{j=1}^{\infty}$ and $y = (y_j)_{j=1}^{\infty}$, but $\ell_p$ is not a Hilbert space according to $p$-norm $\|x\|_p = \big(\sum_{j=1}^{\infty} |x_j|^p\big)^{\frac{1}{p}}$ if $p \neq 2$.

**5.3** Let $X$ be a Hilbert space and $M$ be a closed subspace of $X$. Prove $(M^\perp)^\perp = M$.

**5.4** Let $L_2[-1,1]$ be the class of all Lebesgue measurable complex-valued functions that are square integrable on $[-1,1]$ with inner product $\langle f, g\rangle = \int_{[-1,1]} f\bar{g}dm$.

(i) Let $M = \{f \in L_2[-1,1] : \ f(x) = 0 \quad \forall x \in [-1,0]\}$. Find $M^\perp$;

(ii) Let $M_{odd} = \{f \in L_2[-1,1] : \ f(-x) = -f(x) \quad \forall x \in [-1,1]\}$ and $M_{even} = \{f \in L_2[-1,1] : \ f(-x) = f(x) \quad \forall x \in [-1,1]\}$. Prove $L_2[-1,1] = M_{odd} \oplus M_{even}$.

**5.5** Let $L_2[-\pi,\pi]$ be the Hilbert space of all Lebesgue measurable complex-valued functions that are square integrable on $[-\pi,\pi]$ with inner product $\langle f, g\rangle = \int_{[-\pi,\pi]} f\bar{g}dm$. Prove

(i) $\big((2\pi)^{-\frac{1}{2}} e^{inx}\big)_{n\in\mathbf{Z}}$ is an orthonormal sequence.

(ii)

$$\big((2\pi)^{-\frac{1}{2}}, \pi^{-\frac{1}{2}}\cos t, \pi^{-\frac{1}{2}}\sin t, \pi^{-\frac{1}{2}}\cos 2t, \pi^{-\frac{1}{2}}\sin 2t, ...\big)$$

is an orthonormal basis of $L_2[-\pi,\pi]$.

**5.6** Given $\phi \in C[-\pi,\pi]$, let $T : L_2[-\pi,\pi] \to L_2[-\pi,\pi]$ be given by $T(f) = \phi f$.

(i) Calculate $T^*$ using the inner product defined above.

(ii) Prove that if $\phi$ is real-valued then $T$ is self-adjoint.

(iii) Find a condition on $\phi$ such that $T$ is respectively unitary, positive, or a projection.

**5.7** Suppose $X = C[0,1]$ and $T \in B(X)$ is given by $Tx(t) = tx(t)$. Prove $\sigma(T) = [0,1]$.

**5.8** Suppose that $X$ is a Hilbert space, and that $T \in B(X)$. Prove $\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}$.

**5.9** Suppose that $X$ is a Banach space, and that $T \in B(X)$ is a compact operator. Prove that $TS$ and $ST$ are compact for any $S \in B(X)$.

# References

1. John B. Conway, A Course in Functional Analysis, Graduate Texts in Mathematics, Springer-Verlag New York Inc., 1985.
2. WWL Chen, Linear Functional Analysis, http://www.maths.mq.edu.au/ wchen /lnlfafold/lnlfa.html
3. Ruth F. Curtain and A. J. Pritchard, Functional Analysis in Modern Applied Mathematics, Mathematics in Science and Engineering 132, Academic Press Inc. (London) Ltd., 1977.
4. Carl L. DeVito, Functional Analysis, Mathematics 81, Academic Press Inc., 1978.
5. Gerald B. Folland, Real Analysis, Modern Techniques and Their Applications, Pure & Applied Mathematics, A Wiley-Interscience Series of Texts, Monographs, and Tracts, John Wiley & Sons, Inc., 1984.
6. Avner Friedman, Foundations of Modern Analysis, Hold, Rinehart and Winston, Inc., 1970.
7. Einar Hille, Methods in Classical and Functional Analysis, Addison-Wesley Publishing Company, Inc., 1972.
8. Douglas S. Kurtz and Charles W. Swartz, Theories of Integrations, The Integrals of Riemann, Lebesgue, Henstock-Kurzweil, and Mcshane, Series in Real Analysis 9, World Scientific Publishing Co. Pte. Ltd., 2004.
9. Edward W. Packel, Functional Analysis, A Short Course, Intext, Inc., 1974.
10. Karen Saxe, Beginning Functional Analysis, Undergraduate Texts in Mathematics, Springer-Verlag New York, Inc., 2002.
11. Martin Schechter, Principles of Functional Analysis, Academic Press, Inc., 1971.
12. Eric M. Vestrup, The Theory of Measures and Integration, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 2003.
13. T. B. Ward, Functional Analysis, http://www.mth.uea.ac.uk/ h720/sheets /functionalanalysis/
14. Robert J. Zimmer, Essential Results of Functional Analysis, The University of Chicago Press, 1990.

# Solutions to Exercises

1. *Normed Linear Spaces*

Ex.1.1. It suffices to verify the inequality for continuous functions. Let $\phi(t) = t^{\frac{1}{p}}$ for $t \geq 0$. Since $p^{-1} \in (0,1)$, $\phi''(s) < 0$ for all $s > 0$ and $\phi$ is concave. Hence $\phi(t) \leq \phi(1) + \phi'(1)(t-1)$, or

$$t^{\frac{1}{p}} \leq 1 + \frac{t-1}{p} = \frac{t}{p} + \frac{1}{q}.$$

Setting $t = u^p v^{-q}$, where $u \geq 0$ and $v > 0$, we find since $1 - q = -q/p$ that

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}. \tag{5.1}$$

Obviously, this inequality also holds when $v = 0$.
If $\|f\|_p = 0$ then $f = 0$ on $[0,1]$ (since $f$ is continuous) and both sides of Hölder's inequality are 0. Similarly both sides are 0 if $\|g\|_q = 0$. Suppose $\|f\|_p > 0$, $\|g\|_q > 0$, and let $f_1 = \|f\|_p^{-1} f$, $g_1 = \|g\|_q^{-1} q$. Then $\|f_1\|_p = 1 = \|g_1\|_q$, and setting $u = |f_1|$ and $v = |g_1|$ in (1) one has

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} = \int_0^1 |f_1||g_1| \leq \frac{1}{p} + \frac{1}{q} = 1.$$

This proves Hölder's inequality.
For vectors:

$$\Big(\sum_{j=1}^n |x_j + y_j|^p\Big)^{\frac{1}{p}} \leq \Big(\sum_{j=1}^n |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_{j=1}^n |y_j|^p\Big)^{\frac{1}{p}}.$$

For sequences $(x_j)$ and $(y_j)$ with $\sum_j |x_j|^p < \infty$ and $\sum_j |y_j|^p < \infty$:

$$\Big(\sum_j |x_j + y_j|^p\Big)^{\frac{1}{p}} \leq \Big(\sum_j |x_j|^p\Big)^{\frac{1}{p}} + \Big(\sum_j |y_j|^p\Big)^{\frac{1}{p}}.$$

For continuous functions $f, g$ with $\int_0^1 |f|^p < \infty$, $\int_0^1 |g|^p < \infty$:

$$\Big(\int_0^1 |f+g|^p\Big)^{\frac{1}{p}} \leq \Big(\int_0^1 |f|^p\Big)^{\frac{1}{p}} + \Big(\int_0^1 |g|^p\Big)^{\frac{1}{p}}.$$

Ex.1.2. It suffices to verify the inequality for continuous functions. Note that

$$\int_0^1 |f+g|^p \le \int_0^1 |f||f+g|^{p-1} + \int_0^1 |g||f+g|^{p-1}. \tag{5.2}$$

However, by Hölder's inequality,

$$\int_0^1 |f||f+g|^{p-1} \le \Big(\int_0^1 |f|^p\Big)^{\frac{1}{p}} \Big(\int_0^1 |f+g|^{(p-1)q}\Big)^{\frac{1}{q}}.$$

Since $(p-1)q = p$, one concludes from estimating similarly the last term in (2) that

$$\int_0^1 |f+g|^p \le (\|f\|_p + \|g\|_p)\Big(\int_0^1 |f+g|^p\Big)^{\frac{1}{q}}.$$

If $\|f+g\|_p = 0$, then both sides are 0. Otherwise one divides both sides by $\Big(\int_0^1 |f+g|^p\Big)^{\frac{1}{q}}$. This yields Minkowski's inequality.

Ex.1.3. (i) $n-1$, since the following $n-1$ vectors $(1,-1,0,...0)$, $(0,1,-1,0,...,0)$,..., $(0,0,0,...,1,-1)$ are linearly independent; (ii) $\infty$, since $1,t,t^2,...,t^n$ are linearly independent for any $n\in\mathbf{N}$; (iii) $\infty$, since $1,t,t^2,...,t^n$ are linearly independent for any $n\in\mathbf{N}$; (iii) $\infty$.

Ex.1.4. Let $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n\}$ be a basis of $X$. Given $\mathbf{x}\in X$, $x$ has a representation as $\mathbf{x} = \sum_{j=1}^n a_j\mathbf{x}_j, \quad a_j\in\mathbf{R}$. Since the set of coefficients $a_1,...,a_n$ are the only ones that will give us $\mathbf{x}$, by the linear independence of $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_n\}$, we can use them to define a map $T: X\to\mathbf{R}^n$ by

$$T\mathbf{x} = \begin{pmatrix} a_1\\ \cdot\\ \cdot\\ \cdot\\ a_n \end{pmatrix}.$$

Clearly, $T$ maps $X$ onto $\mathbf{R}^n$ and it is linear:

$$T(\mathbf{x}+\mathbf{y}) = T\mathbf{x}+T\mathbf{y} \text{ and } T(\lambda\mathbf{x}) = \lambda T\mathbf{x}.$$

To verify that $T$ is $1-1$, we suffice to prove that $T\mathbf{x}=0 \Rightarrow \mathbf{x}=0$, but this follows from the definition of $T$. Suppose $\|\cdot\|_X$ is any norm on $X$. Now define another norm $\|\cdot\|_X$ as follows: $|||x|||_X = \|Tx\|_{\mathbf{R}^n}$-any given norm on $\mathbf{R}^n$. By Theorem 1.3.1 we see that $\|\cdot\|_X$ and $|||\mathbf{x}|||_X$ are equivalent and so that $\|\cdot\|_X$ is equivalent to $\|T\mathbf{x}\|_{\mathbf{R}^n}$. Therefore, $X$ and $\mathbf{R}^n$ are topologically isomorphic.

Ex.1.5. Obviously, $\|f\|_p \le \|f\|_\infty$ when $p\in[1,\infty)$ and $f\in C[0,1]$, but there is no such a constant $\kappa>0$ that $\|f\|_\infty \le \kappa\|f\|_p$, $\forall f\in C[0,1]$. For $k-2\in\mathbf{N}$ let

$$f(t) = \begin{cases} k^{\frac{2}{p}}t^{\frac{1}{p}}, & 0\le t\le \frac{1}{k},\\ k^{\frac{2}{p}}(\frac{2}{k}-t)^{\frac{1}{p}}, & \frac{1}{k}\le t\le\frac{2}{k},\\ 0, & \frac{2}{k}\le t\le 1 \end{cases}$$

It is easy to see that $\|f\|_\infty = k^{\frac{1}{p}} \to \infty$ (as $k\to\infty$) and $\|f\|_p = 1$. This verifies the nonexistence of the above constant $\kappa>0$.

Ex.1.6. Note that $|\|x\| - \|a\|| \le \|x-a\|$. So $\|\cdot\|: V\to\mathbf{R}$ is continuous. Also, regarding the continuity of the vector addition and the scalar multiplication, we naturally assume that the norm defined on $X \times Y$ is given by $\|\cdot\|_X + \|\cdot\|_Y$. Therefore, the desired continuity follows from:

$$\|(x+y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\|$$

and

$$\|\lambda x - \lambda_0 x_0\| \leq |\lambda|\|x - x_0\| + |\lambda - \lambda_0|\|x_0\|.$$

Ex.1.7. It is clear that $\ell_0$ is a linear subspace of $\ell_\infty$. Note that $a = (1, \frac{1}{2}, \frac{1}{3}, ...) \in \ell_\infty$. For every $n \in \mathbf{N}$, let $x_n = (1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{n}, 0, 0, ...) \in \ell_0$. Then

$$\|x_n - a\|_\infty = \left\|\left(0, ..., 0, \frac{1}{n+1}, \frac{1}{n+2}, ....\right)\right\| = \frac{1}{n+1} \to 0$$

as $n \to \infty$. It follows that $(x_n)$ converges in $\ell_\infty$, but the limit $a$ does not belong to $\ell_0$. Hence $\ell_0$ is not closed in $\ell_\infty$.

Ex.1.8. (i) follows from the definition right away.
(ii) By a neighborhood of a point, one means an open ball centered at this point. So, $f : X \to Y$ between two normed linear spaces is continuous at $p \in X$ iff for every neighborhood $V$ of $f(p)$ there is a neighborhood $U$ of $p$ such that $f(U) \subset V$. Let $f : X \to Y$ be continuous and $U \subset Y$ be open. Let $p$ be any point of $f^{-1}(U)$ and $V$ be a neighborhood of $f(p)$ such that $V \subset U$. Since $f$ is continuous, there is a neighborhood $B$ of $p$ such that $f(B) \subset V$. Then $B \subset f^{-1}(U)$ which shows that $f^{-1}(U)$ is open. Conversely, let $f^{-1}(U)$ be open for each open set $U$. Let $p$ be any point of $X$, and $V$ be any neighborhood of $f(p)$. Since $V$ is open, $f^{-1}(V)$ is open and contains $p$. If $U$ is a neighborhood of $p$ such that $U \subset f^{-1}(V)$. Then $f(U) \subset V$ which shows that $f$ is continuous at $p$. Since this is true for every $p \in X$, $f$ is continuous on $X$.
Let $f(x) = \frac{1}{1+x^2}$. Then $f\big((-1,1)\big) = (\frac{1}{2}, 1]$ which is not open even though $(-1,1)$ is open.
(iii) Let $U$ be an open ball in $X$, namely, $U = \{x \in X : \|x - x_0\|_X < r\}$ for some $x_0 \in X$ and $r > 0$. If $x_1, x_2 \in U$ and $x = tx_1 + (1-t)x_2$ for $t \in [0,1]$, then $\|x_k - x_0\|_X < r$, $k = 1, 2$ and hence

$$\|x - x_0\|_X = \|t(x_1 - x_0) + (1-t)(x_2 - x_0)\|_X \leq t\|x_1 - x_0\|_X + (1-t)\|x_2 - x_0\|_X < r.$$

That is to say, $x \in U$ and $U$ is convex. Similarly, any closed ball in $X$ is convex.

Ex.1.9. The 'if' part has been proved in the text. So it suffices to verify the 'only-if' part. Now assume that $Y$ is not closed and let $x$ be a point that is not in $Y$ but is in the closure of $Y$. If $(y_j)$ is a sequence of points of $Y$ that converges to $x$ then

$$\|x + Y\| = \inf_{z \in x+Y} \|z\|_X \leq \inf_{j \in \mathbf{N}} \|x - y_j\|_X$$

Since the right side infimum is zero, $\|x + Y\| = 0$ and consequently $x \in Y$– a contradiction. Therefore $\|\cdot\|$ could not be a norm on $X/Y$. We are done.

2. *Banach Spaces*

Ex.2.1. If $(f_j)$ is Cauchy in $C^n[0,1]$, then for any $\epsilon > 0$ there exists $N \in \mathbf{N}$ such that

$$m, n \geq N \Rightarrow |f_m^{(k)}(x) - f_n^{(k)}(x)| \leq \|f_m - f_n\|_\infty < \epsilon \quad \text{for all} \quad x \in [0,1].$$

So for each integer $k \in [0, n]$, one has that $(f_j^{(k)}(x))$ is Cauchy sequence in $\mathbf{R}$. Therefore it converges to some real number $g_k(x)$ for every $x \in [0,1]$; this defines a new function $g_k$ such that $f_j^{(k)} \to g_k$ pointwise on $[0,1]$. Moreover, $(f_j^{(k)})$ converges to $g_k$ uniformly on $[0,1]$, and that $g_k \in C[0,1]$ – this follows readily from the last estimate when letting $m \to \infty$. In particular, $k = 0$ implies that $(f_j)$ converges uniformly on $[0,1]$ to $g_0 = f$, and $f$ is differentiable with $f' = g_1$. Furthermore, the same reasoning yields $f^{(k)} = g_k$ on $[0,1]$. So, $f \in C^n[0,1]$.

Ex.2.2. Suppose $(x^{(j)} = (x_1^{(j)}, x_2^{(j)}, ...)$ is Cauchy in $c_0$. Then for any $\epsilon > 0$ there exists $N \in \mathbf{N}$ such that

$$m, n \geq N \Rightarrow |x_j^{(m)} - x_j^{(n)}| \leq \|x^{(m)} - x^{(n)}\|_\infty < \epsilon \quad \text{for any} \quad j \in \mathbf{N}.$$

So, for each $j$, $(x_j^{(m)})$ is Cauchy in $\mathbf{C}$. By the completeness of $\mathbf{C}$, there is $c_j \in \mathbf{C}$ such that $(x_j^{(m)})$ converges to $c_j$ as $m \to \infty$. It turns out from the above estimates that $c = (c_j)$ belongs to the space $c_0$. Thus, $c_0$ is complete and a Banach space.

Ex.2.3. i) Taking $f = 1$ and $g = 0$ one has $\|T(f) - T(g)\|_\infty = 1 = \|f - g\|_\infty$;
ii) If $T(f) = f$, then $\int_0^x f = f(x)$ and hence $f = 0$ which gives the uniqueness;
iii) If $1_E$ denotes the characteristic function of the set $E$, then

$$T(T(f(x))) = \int_0^x \Big( \int_0^y f(t)dt \Big) dy = \int_0^x (x - t) f(t) dt$$

and hence

$$\|T(T(f)) - T(T(g))\|_\infty \leq \frac{1}{2} \|f - g\|_\infty.$$

Ex.2.4. It follows from Theorem 2.2.1 with $|f(b) - f(a)| = |f'(c)(b - a)| \leq \alpha |b - a|$ for some $c \in [a, b]$.

Ex.2.5. Define the operator $T$ on $C[0,1]$ by

$$T(f)(x) = \sin x + \int_0^1 f(y) \exp\big(-(x + y + 1)\big) dy.$$

Let $f, g \in C[0,1]$. We have

$$\begin{aligned} \|T(f) - T(g)\|_\infty &\leq \sup_{x \in [0,1]} \int_0^1 |f(y) - g(y)| \exp\big(-(x + y + 1)\big) dy \\ &\leq \sup_{x \in [0,1]} |f(x) - g(x)| \exp(-x) \int_0^1 \exp(-(y + 1)) dy \\ &\leq \|f - g\|_\infty (e^{-1} - e^{-2}), \end{aligned}$$

where $e^{-1} - e^{-2} \in (0,1)$. Hence $T$ is a contraction mapping. Therefore, by Theorem 2.2.1, there is a unique $f \in C[0,1]$ with $T(f) = f$.

3. *Linear Operators*

Ex.3.1. It is clear that $T$ is linear. Furthermore, if $f \in C[0,1]$ then

$$|T(f)| = |f(0)| \leq \|f\|_\infty,$$

giving the boundedness of $T$ with $\|T\| \leq 1$. If $f(x) = 1$ then

$$\|f\|_\infty = 1 \quad \text{and} \quad Tf = 1.$$

This yields $\|T\| = 1$.

Ex.3.2. (i) Clearly,

$$\|Tx\|_\infty \le \sup_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}||x_j| \le \|x\|_\infty \sup_{1\le i\le n} \sum_{j=1}^{\infty} |a_{i,j}|$$

and so

$$\|T\| \le \sup_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}.$$

To get the equality, just take $x_j = \mathrm{sgn} a_{i_0,j}$ where

$$\sum_{j=1}^{n} |a_{i_0,j}| = \sup_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}|.$$

Then $\|x\|_\infty = 1$ and $y_i = \sum_{j=1} |a_{i_0,j}|$. So

$$\|T\| \ge \|Tx\|_\infty = \sup_{1\le i\le n} \Big| \sum_{j=1}^{n} a_{i,j} \mathrm{sgn} a_{i_0,j} \Big| \ge \Big| \sum_{j=1}^{n} a_{i_0,j} \mathrm{sgn} a_{i_0,j} \Big| = \sup_{1\le i\le n} \sum_{j=1}^{n} |a_{i,j}|,$$

giving the desired equality.
(ii) It follows from the definition of $\|\cdot\|_2$.
(iii) It is clear that for $f \in R^1[a,b]$,

$$\begin{aligned}\|Tf\|_1 &= \int_a^b \Big| \int_a^x f(t)dt \Big| dx \\ &\le \int_a^b \int_a^x |f(t)|dtdx \\ &\le \int_a^b \int_a^b |f(t)|dtdx \\ &= (b-a)\|f\|_1,\end{aligned}$$

and so that $T$ is bounded with $\|T\| \le b-a$. To see the equality, for any $n \in \mathbf{N}$ with $a + n^{-1} < b$ let

$$f_n(x) = \begin{cases} n, \text{ if } & t \in [a, a+n^{-1}] \\ 0, \text{ if } & t \in (a+n^{-1}, b]. \end{cases}$$

It is easy to check $\|f_n\|_1 = 1$ and

$$\|Tf_n\|_1 = \int_a^{a+n^{-1}} n(x-a)dx + \int_{a+n^{-1}}^b dx = (b-a) - (2n)^{-1}.$$

So $\|T\| \ge \sup_{n\in\mathbf{N}} \|Tf_n\|_1 = b-a$. Therefore $\|T\| = b-a$.

Ex.3.3. The boundedness is obvious. Since $tf(t) = tg(t)$ (for $t \in (0,1)$) implies $f(t) = g(t)$, we conclude that $T$ is 1-1. But, $T$ is not onto. In fact, if $Tf = 1$ on $(0,1)$ then the only possible candidate for $f$ is $f(t) = t^{-1}$. It is clear that $f \in C(0,1)$ with $\|f\|_2 = \infty$. From this it turns out that $T$ is not invertible.

Ex.3.4. Take $X = Y = E_\infty$ and equip it with 2-norm: $\|x\|_2 = \left(\sum_{j=1}^\infty |x_j|^2\right)^{\frac{1}{2}})$. Here $x \in E_\infty$ if and only if $x \in \ell_\infty$ and it has only finitely many nonzero entries. Then $X$ and $Y$ are not complete. Now define $T_n x = (0, ..., 0, nx_n, 0, ...)$ for $x = (x_1, x_2, ...)$. Then $T_n$ is bounded with $\|T_n\| = n \to \infty$. On the other hand, if $x \in E_\infty$ then there is an $N \in \mathbf{N}$ such that $n > N$, $x_n = 0$ and hence $\|T_n x\|_2 = n|x_n| = 0$ and if $n \le N$ then $\|T_n x\|_2 = n|x_n| \le n\|x\|_2 \le N\|x\|_2$. Hence for each $x \in X$ we have $\sup_n \|T_n x\|_2 < \infty$. Clearly, the uniform bounded principle fails.

Ex.3.5. (i) Since $|\sin x| \le |x|$, we conclude that the integral is not less than

$$\int_0^{2\pi} \frac{2}{x}|\sin(n+\frac{1}{2})x|dx.$$

Note that

$$k\pi + \frac{\pi}{6} \le (n+\frac{1}{2})x \le k\pi + \frac{\pi}{3} \Rightarrow |\sin(n+\frac{1}{2})x| \ge \frac{1}{2}, \quad k \in \mathbf{N}.$$

So

$$\int_0^{2\pi} \frac{2}{x}|\sin(n+\frac{1}{2})x|dx \ge \sum_{k=0}^{2n}\left(\frac{\pi(k+\frac{1}{3})}{n+\frac{1}{2}}\right)^{-1} \to \infty \quad \text{as} \quad n \to \infty.$$

(ii) It follows from the change of variable that

$$s_n(x) = \frac{1}{2\pi}\int_{-x}^{2\pi - x} f(x+z)\Big(\sum_{m=-n}^{n} e^{-imz}\Big)dz.$$

This yields that $\overline{s_n(x)} = s_n(x)$ and hence $s_n(x)$ is real-valued. Note that $e^{ix} = \cos x + i\sin x$ and $f$ is $2\pi$-periodic. So

$$\begin{aligned} s_n(x) &= \frac{1}{2\pi}\int_{-x}^{2\pi - x} f(x+z)\Big(\sum_{m=-n}^{n} \cos mz\Big)dz \\ &= \frac{1}{2\pi}\int_{-x}^{2\pi - x} f(x+z)\Big(1 + 2\sum_{m=1}^{n} \cos mz\Big)dz \\ &= \frac{1}{2\pi}\int_{-x}^{2\pi - x} f(x+z)\frac{\sin(n+2^{-1})z}{\sin 2^{-1}z}dz \\ &= \frac{1}{2\pi}\int_{0}^{2\pi} f(x+y)\frac{\sin(n+2^{-1})y}{\sin 2^{-1}y}dy. \end{aligned}$$

(iii) For any $f \in X$, we have

$$|T_n(f)| \le \frac{1}{2\pi}\|f\|_\infty \int_0^{2\pi}\left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right|dx,$$

and then

$$\|T_n\| \le \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right|dx.$$

To prove the equality actually holds, we may assume

$$q_n(x) = \frac{\sin(n+2^{-1})x}{\sin 2^{-1}x}$$

and $g_n(x) = \mathrm{sgn} q_n(x)$, that is,

$$g_n(x) = \begin{cases} 1, & g_n(x) > 0, \\ 0, & g_n(x) = 0, \\ -1, & g_n(x) < 0. \end{cases}$$

Then $|g_n(x)| = g_n(x)q_n(x)$. Though $g_n$ is not continuous, for any $\epsilon > 0$, there is a continuous function $f_n$ such that

$$\frac{1}{2\pi}\left|\int_0^{2\pi} \big(f_n(x) - g_n(x)\big)q_n(x)dx\right| < \epsilon.$$

This can be easily realized since $q_n$ is continuous on $[0, 2\pi]$. In fact, it is enough to use piecewise-defined segments to connect the discontinuous points of $g_n$ so that $f_n$ is sufficiently close to $g_n$. Then $\|f_n\|_\infty = \max_{x\in[0,2\pi]} |f_n(x)| = 1$, but

$$\begin{aligned}
|T_n f_n| &= \frac{1}{2\pi}\int_0^{2\pi} f_n(x)q_n(x)dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} \big(f_n(x) - g_n(x)\big)q_n(x)dx + \frac{1}{2\pi}\int_0^{2\pi} g_n(x)q_n(x)dx \\
&\geq \frac{1}{2\pi}\left|\int_0^{2\pi} g_n(x)q_n(x)dx\right| - \frac{1}{2\pi}\left|\int_0^{2\pi} \big(f_n(x) - g_n(x)\big)q_n(x)dx\right| \\
&\geq \int_0^{2\pi} |q_n(x)|dx - \epsilon.
\end{aligned}$$

Obviously, this implies

$$\|T_n\| \geq \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right| dx,$$

and so

$$\|T_n\| = \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\right| dx.$$

(iv) It is clear that $T_n(f) = s_n(0)$ for all $f \in X$. Moreover, for fixed $f \in X$, if the Fourier series of $f$ converges at 0, then $\{T_n f\}$ is bounded as $n$ varies since each element is just a partial sum of a convergent series. Thus if the Fourier series of $f$ converges at 0 for all $f \in X$, then for each $f \in X$ the set $\{T_n f\}$ is bounded. By the uniform bounded principle, this implies that $\{\|T_n\|\}$ is bounded, which contradicts (i) and (iii).

Ex.3.6. Here, suppose that the norm on $\mathbf{R}^2$ is 2-norm.
Linear: $T\big(\alpha(x_1, y_1) + \beta(x_2, y_2)\big) = (\alpha x_1 + \beta x_2, 0) = \alpha T(x_1, y_1) + \beta T(x_2, y_2)$.
Bounded: $\|T(x, y)\|_2 = |x| \leq \|(x, y)\|_2$.
Not onto: $(x, 1)$ has no preimage under $T$.
An example: $S = (0, 1) \times (0, 1)$ is an open set of $\mathbf{R}^2$, but $TS = \{(x, 0) : x \in (0, 1)\}$ is not open set of $\mathbf{R}^2$.

Ex.3.7. (i) Consider $f_n(x) = \sqrt{(x - 2^{-1})^2 + n^{-2}}$. Then $(f_n)$ is convergent to $f(x) = |x - 2^{-1}|$ uniformly on $[0, 1]$; that is,

$$\big|f_n(x) - |x - 2^{-1}|\big| = \frac{n^{-2}}{\sqrt{(x - 2^{-1})^2 + n^{-2}} + \sqrt{(x - 2^{-1})^2}} \leq n^{-1} \to 0.$$

So,

$$\lim_{n\to\infty} \|f_n - f\|_\infty = 0.$$

However, $f$ is not in $C^1[0,1]$. Namely, $X$ is not complete under the sup-norm.
(ii) To prove that $d/dx$ is closed, let $(x_n)$ be a sequence in $C^1[0,1]$ such that $x_n \to x, Tx_n \to y$. Since the convergence in $C[0,1]$ means the uniform convergence, we conclude that $Tx_n$ converges to $y$ uniformly on $[0,1]$ and $y \in C[0,1]$. Of course, $x \in C[0,1]$ and $y(t) = x'(t)$; that is, $T$ is closed. But, it is not bounded since if $x_n(t) = \sin n\pi t$ then

$$\|Tx_n\| = \max_{t\in[0,1]} |\cos n\pi t| n\pi = n\pi \to \infty.$$

Ex.3.8. (i) Let $u = \mathrm{Re} f$. Then $u$ is clearly real linear and $\mathrm{Im} f(x) = -\mathrm{Re}(if(x)) = -u(ix)$, so $f(x) = u(x) - iu(ix)$. If $u$ is real linear and $f(x) = u(x) - iu(ix)$, then $f$ is clearly linear over $\mathbf{R}$, and $f(ix) = u(ix) - iu(-x) = u(ix) + iu(x) = if(x)$, so $f$ is also linear over $\mathbf{C}$. Finally, since $|u(x)| = |Ref(x)| \le |f(x)|$ we have $\|u\| \le \|f\|$. On the other hand, if $f(x) \neq 0$, let $\alpha = \mathrm{sgn} f(x)$, where

$$\mathrm{sgn} f(x) = \begin{cases} \exp\Big(-i \arg f(x)\Big), & \text{if} \quad f(x) \neq 0, \\ \quad 0, & \text{if} \quad f(x) = 0. \end{cases}$$

Then

$$|f(x)| = \alpha f(x) = f(\alpha x) = u(\alpha x)$$

because $f(\alpha x)$ is real, so

$$|f(x)| \le \|u\|\|\alpha x\| = \|u\|\|x\|,$$

whence $\|f\| \le \|u\|$.
(ii) Let $u = \mathrm{Re} f$. Then by the real Hahn-Banach theorem there is a real extension $U$ of $u$ to $X$ such that $|U(x)| \le p(x)$ for all $x \in X$. Set $F(x) = U(x) - iU(ix)$ as in (i). Then $F$ is a complex linear extension of $f$, and as in the proof of (i), if $\alpha = \overline{\mathrm{sgn} f(x)}$, then we have

$$|F(x)| = \alpha F(x) = F(\alpha x) \le p(\alpha x) = p(x).$$

4. *Lebesgue Measures, Integrals and Spaces*

Ex.4.1. (i) Let $H = \{x = (x_1, ..., x_n) \in \mathbf{R}^n : \ x_j > c\}$ or $\{x = (x_1, ..., x_n) \in \mathbf{R}^n : x_j < c\}$. Then we have to prove that for any $T \subset \mathbf{R}^n$ one has

$$m^*(T) = m^*(T \cap H) + m^*(T \setminus H) = m^*(T \cap H) + m^*(T \cap H^c),$$

it suffices to verify that

$$m^*(T) \ge m^*(T \cap H) + m^*(T \cap H^c)$$

since $m^*$ is subadditive. If $m^*(T) = \infty$, then there is nothing to prove. We assume $m^*(T) < \infty$. In case $(I_j)$ is a sequence of open intervals such that $T \subset \bigcup_{j=1}^\infty I_j$, then

$$T \cap H \subset \bigcup_{j=1}^\infty I_j \cap H \quad \text{and} \quad T \cap H^c \subset \bigcup_{j=1}^\infty I_j \cap H^c.$$

Hence, by monotonicity and countable sub-additivity of $m^*$,

$$m^*(T \cap H) + m^*(T \cap H^c) \leq \sum_{j=1}^{\infty} m^*(I_j \cap H) + m^*(I_j \cap H^c).$$

If we show that for every open interval $I \subset \mathbf{R}^n$ one has

$$m^*(I) = m^*(I \cap H) + m^*(I \cap H^c), \tag{5.3}$$

then

$$m^*(T \cap H) + m^*(T \cap H^c) \leq \sum_{j=1}^{\infty} m^*(I_j)$$

and hence

$$m^*(T \cap H) + m^*(T \cap H^c) \leq \inf_{T \subset \bigcup I_j} \sum_{j=1}^{\infty} m^*(I_j) = m^*(T),$$

giving the desired inequality.
Suppose now $I = \{x = (x_1, ..., x_n) : a_j < x_j < b_j,\ j = 1, ..., n\}$. Without loss of generality, we may assume that $H$ is given by $\{x = (x_1, ..., x_n) :\ x_1 < c\}$. Since (1) is obviously true if $c < a_1$ or $c \geq b_1$, let us suppose $a_1 < c < b_1$. Consequently, $I \cap H$ is an open interval given by

$$\{x = (x_1, ..., x_n) :\ a_1 < x_1 < c \quad \text{and} \quad a_j < x_j < b_j, j = 2, ..., n\}.$$

$I \cap H^c$ is the set given by

$$\{x = (x_1, ..., x_n) :\ c \leq x_1 < b_1 \quad \text{and} \quad a_j < x_j < b_j, j = 2, ..., n\}.$$

It can be checked easily that

$$m^*(I \cap H) = (c - a_1)(b_2 - a_2) \cdots (b_n - a_n),$$
$$m^*(I \cap H^c) = (b_1 - c)(b_2 - a_2) \cdots (b_n - a_n),$$

and

$$m^*(I \cap H) + m^*(I \cap H^c) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n) = m^*(I).$$

(ii) Let $H_j = \{x = (x_1, ..., x_n) \in \mathbf{R}^n :\ c_j < x_j\}$. Then $H_j$ is Lebesgue measurable due to (i), and hence $H_j^c$ is Lebesgue measurable too. Without loss of generality, we may assume that the half-open interval is $I = \{x = (x_1, ..., x_n) \in \mathbf{R}^n :\ b_j < x_j \leq c_j,\ j = 1, ..., n\}$. Then

$$I = \bigcap_{j=1}^{n} (J_j \cap H_j^c) \quad \text{where} \quad J_j = \{x = (x_1, ..., x_n) \in \mathbf{R}^n :\ b_j < x_j\}.$$

This implies that $I$ is Lebesgue measurable since so are $J_j$ and $H_j^c$.
(iii) Let $O \subset \mathbf{R}^n$ be an open set. For each $k \in \mathbf{N}$, the hyperplanes

$$x_j = l2^{-k},\ l \in \mathbf{Z};\ j = 1, ..., n, \tag{5.4}$$

partition $\mathbf{R}^n$ into a countable collection of disjoint half-open intervals. Let $I_1^1, I_1^2, I_1^3, ...$ be a collection of such intervals generated by (2) for $k = 1$ that are contained in $O$. We use recursion to define suitable $I_k^i$'s. For $k > 1$, let $I_k^1, I_k^2, I_k^3, ...$

be the collection of half-open intervals generated by (2) which are contained in $O$ but not contained in any interval $I_q^p$ with $1 \le q < k$. If $x \in O$, then $x$ is an interior point; so there is a partition of $\mathbf{R}^n$ given by (2) such that the interval containing $x$ is contained in $O$. Therefore

$$O \subset \bigcup_{k=1}^{\infty} \bigcup_{i} I_k^i.$$

Since $I_k^i \subset O$ for each $i, k$, we have

$$O \supset \bigcup_{k=1}^{\infty} \bigcup_{i} I_k^i.$$

This is clearly a countable collection of half-open intervals, and we have constructed them so they are disjoint. (iv) It follows from (iii) and (ii) right away.

Ex.4.2. (i) First, assume that $E$ is bounded. Then $m^*(E) = m(E)$ is finite. Given $\epsilon > 0$, then by definition of $m^*$, there is a countable collection of open intervals $(I_j)$ such that

$$E \subset \bigcup I_j \quad \text{and} \quad \sum_{j=1}^{\infty} m(I_j) < m(E) + \epsilon.$$

Let $V = \bigcup I_j$. Then $V$ is open set in $\mathbf{R}^n$. Moreover, $m(V) \le \sum_{j=1}^{\infty} m(I_j) < m(E) + \epsilon$. Since $E \subset V$ and $m(E) < \infty$, we conclude that $m(V \setminus E) = m(V) - m(E) < \epsilon$. Secondly, suppose that $E$ is an unbounded Lebesgue measurable set in $\mathbf{R}^n$. For each $k \in \mathbf{N}$, let

$$J_k = \{x = (x_1, ..., x_n) \in \mathbf{R}^n : \ |x_j| < k, \ j = 1, ..., n\}.$$

Of course, $J_k$ is bounded but also Lebesgue measurable. Note that $\mathbf{R}^n = \bigcup_{k=1}^{\infty} J_k$. So $E = \bigcup_{k=1}^{\infty} E \cap J_k$. The previous results now apply to the bounded Lebesgue measurable set $E \cap J_k$: For $\epsilon > 0$ and $k \in \mathbf{N}$, there is an open set $V_k$ such that $J_k \cap E \subset V_k$ and $m\big(V_k \setminus (E \cap J_k)\big) < 2^{-k}\epsilon$. If $V = \bigcup_{k=1}^{\infty} V_k$, then $V$ is open and $E \subset V$ and

$$m(V \setminus E) = m\Big(\bigcup_{k=1}^{\infty} V_k - \bigcup_{k=1}^{\infty} E \cap J_k\Big) \le m\Big(\bigcup_{k=1}^{\infty} \big(V_k \setminus (E \cap J_k)\big)\Big) \le \sum_{k=1}^{\infty} \epsilon 2^{-k} = \epsilon.$$

(ii) To establish this assertion, let $F = \mathbf{R}^n \setminus E$, then $F$ is Lebesgue measurable. Hence for each $\epsilon > 0$ there is an open set $V$ in $\mathbf{R}^n$ such that $F \subset V$ and $m(V \setminus F) < \epsilon$. Set $U = \mathbf{R}^n \setminus V$. Then $U$ is closed and $U \subset E$. Since $F \cap U = \emptyset$ and $E \setminus U = (\mathbf{R}^n \setminus U) \setminus F = V \setminus F$, we conclude that $m(E \setminus U) < \epsilon$.

Ex.4.3. Let $a \in \mathbf{R}$ and $x \in E[f > a]$. Since $f$ is continuous on $E$, we conclude that there is an open ball $U(x) \subset \mathbf{R}^n$ centered at $x$ such that $U(x) \cap E \subset E[f > a]$. Put $O = \bigcup_{x \in E[f>a]} U(x)$. Then

$$E[f > a] \subset O \cap E[f > a] \quad \text{and hence} E[f > a] = O \cap E.$$

Since $O$ is open, we get that $O \cap E$ is Lebesgue measurable.

Ex.4.4. Note that

$$\{x \in \mathbf{R}^n : \ 1/f > a\} = \begin{cases} \{x \in \mathbf{R}^n : \ f > 0\} \cap \{x \in \mathbf{R}^n : \ f < 1/a\}, & \text{if} \quad a > 0, \\ \{x \in \mathbf{R}^n : \ f > 0\} \setminus \{x \in \mathbf{R}^n : \ f = \infty\}, & \text{if} \quad a = 0, \\ \{x \in \mathbf{R}^n : \ f > 0\} \cup \{x \in \mathbf{R}^n : \ f < 1/a\}, & \text{if} \quad a < 0. \end{cases}$$

Then the desired result follows right away.

Ex.4.5. For $k \in \mathbf{N}$ let

$$f_k(x) = \begin{cases} 0, & \text{if} \quad 0 \le x < \frac{1}{2k}, \\ k, & \text{if} \quad \frac{1}{2k} \le x \le \frac{1}{k}, \\ 0, & \text{if} \quad \frac{1}{k} < x \le 1. \end{cases}$$

Since $\liminf_{k\to\infty} f_k = 0$, we conclude that $\int_{[0,1]} \liminf_{k\to\infty} f_k dm = 0$. But,

$$\int_{[0,1]} f_k dm = \int_{[0,\frac{1}{2k})} 0dm + \int_{[\frac{1}{2k},\frac{1}{k}]} kdm + \int_{(\frac{1}{k},1]} 0dm = 2^{-1}$$

implies

$$\liminf_{k\to\infty} \int_{[0,1]} f_k dm = 2^{-1} > 0 = \int_{[0,1]} \liminf_{k\to\infty} f_k dm = 0..$$

Ex.4.6. (i) For each $k \in \mathbf{N}$, partition $[0,1]$ into $k$ subintervals with the equal width $k^{-1}$. Let $P_n$ denote this partition and $0 = x_0 < x_1 < \cdots < x_k = 1$ denote the points of $P_k$. Define

$$g_k(x) = \sum_{j=1}^{k} m_j 1_{[x_{j-1},x_j)}(x) \quad \text{and} \quad h_k(x) = \sum_{j=1}^{k} M_j 1_{[x_{j-1},x_j)}(x).$$

Here and henceforth,

$$m_j = \inf\{f(x) : \ x \in [x_{j-1}, x_j]\} \quad \text{and} \quad M_j = \sup\{f(x) : \ x \in [x_{j-1}, x_j]\}.$$

It is clear that $(g_k)$ and $(h_k)$ are nondecreasing and nonincreasing sequences respectively. Put

$$g(x) = \lim_{k\to\infty} g_k(x) \quad \text{and} \quad h(x) = \lim_{k\to\infty} h_k(x).$$

Then $g$ and $h$ are Lebesgue integrable that satisfy

$$g(x) \le f(x) \le h(x) \quad \text{a.e. on}[0,1].$$

Note that

$$\int_{[0,1]} g_k dm = \sum_{j=1}^{k} m_j(x_j - x_{j-1}) \quad \text{and} \quad \int_{[0,1]} h_k dm = \sum_{j=1}^{k} M_j(x_j - x_{j-1})$$

but also

$$h_k - g_k \ge 0 \quad \text{a.e. on } [0,1] \quad \text{and} \quad \lim_{k\to\infty}(h_k - g_k) = h - g.$$

So, by Lebesgue's monotone convergence theorem it follows that

$$\begin{aligned} 0 &\le \int_{[0,1]} (h-g)dm \\ &= \lim_{k\to\infty} \int_{[0,1]} (h_k - g_k)dm \\ &= \lim_{k\to\infty} \int_{[0,1]} h_k dm - \lim_{k\to\infty} \int_{[0,1]} g_k dm \\ &= \lim_{k\to\infty} \sum_{j=1}^{n} M_j(x_j - x_{j-1}) - \lim_{k\to\infty} \sum_{j=1}^{n} m_j(x_j - x_{j-1}) \\ &= 0. \end{aligned}$$

Here, we have used the fact that $f$ is Riemann integrable on $[0,1]$. The above estimates tell us that $h-g=0$ a.e. on $[0,1]$ and so that $f=g=h$ a.e. on $[0,1]$. Consequently, $f$ is Lebesgue measurable, and

$$\begin{aligned}\int_{[0,1]} f dm &= \lim_{k\to\infty}\int_{[0,1]} g_k dm \\ &= \lim_{k\to\infty}\sum_{j=1}^{n} M_j(x_j - x_{j-1}) \\ &= \int_0^1 f(x)dx.\end{aligned}$$

(ii) Consider $1_{\mathbf{Q}}$ where $\mathbf{Q}$ is the set of all rational numbers in $[0,1]$. It is well-known that this function is not Riemann integrable on $[0,1]$. Since $m(\mathbf{Q})=0$, we conclude that this function is Lebesgue integrable with $\int_{[0,1]} 1_{\mathbf{Q}} dm = 0$.

5. *Hilbert Spaces*

Ex.5.1. It is easy to check that the inner product $\langle f,g\rangle = \int_{-1}^{1} f(x)\overline{g(x)}dx$ satisfies (a), (b) and (c) of Definition 5.1.1, and so $C[-1,1]$ is an inner product space. Since the norm equipped with $C[-1,1]$ is the 2-norm: $\|f\|_2 = \int_{-1}^{1}|f(x)|^2dx$, we can conclude that this space is not complete. In fact, for each $j\in\mathbf{N}$ let

$$f_j(x) = \begin{cases} -1, & \text{if} \quad x\in[-1,0), \\ jx, & \text{if} \quad x\in[-j^{-1}, j^{-1}], \\ 1, & \text{if} \quad x\in[j^{-1},1]. \end{cases}$$

then $f_j\in C[-1,1]$. Observe that for each $m,n\in\mathbf{N}$ satisfying $m>n$, one has

$$\begin{aligned}\|f_m - f_n\|_2^2 &= 2\left(\int_0^{m^{-1}}(mx-nx)^2dx + \int_{m^{-1}}^{n^{-1}}(1-nx)^2dx\right) \\ &= 2\left(\frac{(m-n)^2}{3m^3} + \frac{1}{3n} - \frac{1}{m} + \frac{n}{m^2} - \frac{n^2}{3m^3}\right) \\ &< \frac{6}{m} + \frac{1}{n} < \frac{7}{n} \to 0 \quad \text{as} \quad n\to\infty.\end{aligned}$$

This means that $(f_j)$ is a Cauchy sequence in $C[-1,1]$. On the other hand, $(f_j)$ converges pointwise to

$$f(x) = \begin{cases} -1, & \text{if} \quad x\in[-1,0), \\ 0, & \text{if} \quad x=0, \\ 1, & \text{if} \quad x\in(0,1]. \end{cases}$$

which does not belong to $C[-1,1]$. Therefore, $C[-1,1]$ is not a Hilbert space.

Ex.5.2. It is clear that $\langle x,y\rangle = \sum_{j=1}^{\infty} x_j\bar{y}_j$ is an inner product defined on $\ell_2$. From Example 2.1.1 (iii) we have seen that $\ell_2$ is complete under the norm $\|x\|_2 = \sqrt{\langle x,x\rangle}$. So, $\ell_2$ is a Hilbert space. To obtain that $\ell_p$ is not a Hilbert space, one takes $x = (1,1,0,0,...)$ and $y=(1,-1,0,0,...)$. Then $x,y\in\ell_p$ and $\|x\|_p = \|y\|_p = 2^{\frac{1}{p}}$ but $\|x+y\|_p = \|x-y\|_p = 2$. Hence the parallelogram law fails at $p\neq 2$; that is to say, if $p \neq 2$ then $\|\cdot\|_p$ does not induce an inner product, and hence $\ell_p$ is not a Hilbert space.

Ex.5.3. Obviously, $M \subset (M^\perp)^\perp$. To get the reverse inclusion, let $x \in (M^\perp)^\perp$. Then the projection theorem (Theorem 5.2.2) implies that there exist $y \in M \subset (M^\perp)^\perp$ and $z \in M^\perp$ such that $x = y+z$. Because $x \in (M^\perp)^\perp$ and $(M^\perp)^\perp$ is a linear space, one has $z = x - y \in (M^\perp)^\perp$ and so $z \in M^\perp \cap (M^\perp)^\perp = \{0\}$, i.e., $z = 0$ and $x = y$. This gives $(M^\perp)^\perp \subset M$.

Ex.5.4. (i) Given $g \in M^\perp$. Then for any $f \in M$ one has

$$0 = \int_{[-1,1]} f\bar{g}dm = \int_{[0,1]} f\bar{g}dm.$$

If $f$ is taken to respectively be 0 a.e. on $[-1,0]$ and $g$ a.e. on $(0,1]$, then $\int_0^1 |g(x)|^2dx = 0$ and hence $g = 0$ a.e. on $(0,1]$. It follows that $M^\perp = \{g \in L_2[-1,1]: \ g(x) = 0 \quad \text{a.e on} \quad (0,1]\}$.
(ii) Any function $f \in L_2[-1,1]$ can be written as $f_{even} + f_{odd}$ where

$$f_{even}(x) = \frac{f(x)+f(-x)}{2} \quad \text{and} \quad f_{odd}(x) = \frac{f(x)-f(-x)}{2}.$$

Clearly, $f_{even} \in M_{even}$ and $f_{odd} \in M_{odd}$. Moreover, if $f \in M_{even} \cap M_{odd}$, then $f(x) = -f(x)$ and hence $f(x) = 0$. In other words, $M_{even} \cap M_{odd} = \{0\}$. This implies the desired direct sum decomposition.

Ex.5.5. (i) This follows from

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}dx = \begin{cases} 1 & \text{if} \quad n = m \\ 0 & \text{if} \quad n \neq m. \end{cases}$$

(ii) First of all, a calculation gives

$$\int_{-\pi}^{\pi} 1dx = 2\pi, \quad \int_{-\pi}^{\pi} \cos^2 nxdx = \int_{-\pi}^{\pi} \sin^2 nxdx = \pi \quad \forall n \in \mathbf{N};$$

$$\int_{-\pi}^{\pi} \cos nxdx = \int_{-\pi}^{\pi} \sin nxdx = 0 \quad \forall n \in \mathbf{N};$$

$$\int_{-\pi}^{\pi} \cos mx \cos nxdx = \int_{-\pi}^{\pi} \sin mx \sin nxdx = 0 \quad \forall m \neq n, m, n \in \mathbf{N};$$

and

$$\int_{-\pi}^{\pi} \sin mx \cos nxdx = 0 \quad \forall m, n \in \mathbf{N}.$$

Now, let $\{e_j\}_{j=1}^\infty$ be this orthonormal set. According to the definition of orthonormal basis, we must prove that

$$\langle f, e_j\rangle = 0 \quad \Rightarrow f = 0 \quad \text{a.e. on} \quad [-\pi,\pi].$$

To do so, let us first consider the case that $f$ is continuous and real-valued. If $f \neq 0$ a.e. on $[-\pi,\pi]$, then there is an $x_0 \in [-\pi,\pi]$ at which $|f|$ achieves a maximum, and we may assume $f(x_0) > 0$. Thus, there is a $\delta > 0$ such that

$$f(x) > \frac{f(x_0)}{2} \quad \forall x \in (x_0 - \delta, x_0 + \delta).$$

If
$$g(x) = 1 + \cos(x_0 - x) - \cos\delta,$$
then
$$1 < g(x) \quad \forall x \in (x_0 - \delta, x_0 + \delta) \quad \text{and} |g(x)| \le 1 \quad \forall x \in [-\pi, \pi] \setminus (x_0 - \delta, x_0 + \delta).$$
Note that $\langle f, e_j \rangle = 0$. So for any $n \in \mathbf{N}$,
$$\begin{aligned} 0 = \langle f, g^n \rangle &= \int_{-\pi}^{\pi} f(x) g^n(x) dx \\ &= \int_{-\pi}^{x_0 - \delta} f(x) g^n(x) dx + \int_{x_0 - \delta}^{x_0 + \delta} f(x) g^n(x) dx + \int_{x_0 + \delta}^{\pi} f(x) g^n(x) dx. \end{aligned}$$

Using the properties of $g$ above, we see that
$$\left| \int_{-\pi}^{x_0 - \delta} f(x) g^n(x) dx \right|, \quad \left| \int_{x_0 + \delta}^{\pi} f(x) g^n(x) dx \right| \le 2\pi f(x_0)$$
and
$$\int_{x_0 - \delta}^{x_0 + \delta} f(x) g^n(x) dx \ge \int_a^b f(x) g^n(x) dx \quad \forall [a, b] \subset (x_0 - \delta, x_0 + \delta).$$
Since $g$ is continuous on $[a, b]$, we can conclude that $g$ achieves a minimum value, $\kappa > 1$, there. This implies
$$4\pi f(x_0) \ge \int_a^b f(x) g^n(x) dx \ge \frac{f(x_0)}{2} \kappa^n (b - a) \to \infty \quad \text{as} \quad n \to \infty.$$
This is a contradiction. Thus, $f = 0$ on $[-\pi, \pi]$.
If $f$ is continuous but not real-valued, then our hypothesis gives
$$\int_{-\pi}^{\pi} f(x) e^{-ikx} dx = 0 \quad \text{and} \quad \int_{-\pi}^{\pi} \overline{f(x)} e^{-ikx} dx = 0 \quad \forall k = 0, \pm 1, \pm 2, ....$$
Hence
$$\int_{-\pi}^{\pi} \Re f(x) e_j(x) dx = 0 \quad \text{and} \quad \int_{-\pi}^{\pi} \Im f(x) e_j(x) dx = 0.$$
By the first part, we get that $\Re f(x) = 0 = \Im f(x)$ and so $f = 0$ on $[-\pi, \pi]$.
Finally, we no longer assume that $f$ is continuous. But, $f$ generates a continuous function
$$F(x) = \int_{-\pi}^{x} f(t) dt.$$
Integration by parts yields
$$\int_{-\pi}^{\pi} F(x) \sin(kx) dx = \frac{1}{k} \int_{-\pi}^{\pi} f(x) \cos(kx) dx = 0.$$
Similarly,
$$\int_{-\pi}^{\pi} F(x) \cos(kx) dx = 0.$$

This infers that $F$ and $F-C$ for every constant $C$, is orthogonal to each of the non-constant members of $\{e_j\}$. For $(2\pi)^{-1/2}$, let

$$C_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)dx.$$

Then $\langle F - C_0, e_j\rangle = 0$ for all $j \in \mathbf{N}$. Since $F$ is continuous, we conclude from the first two parts that $F - C_0 = 0$. Of course, $f = F' = 0$ a.e. on $[-\pi,\pi]$.

Ex.5.6. (i) From $\langle Tf, g\rangle = \langle f, T^*g\rangle$ we have

$$\int_{[-\pi,\pi]} \phi f\bar{g}dm = \int_{[-\pi,\pi]} f\overline{T^*g}dm$$

and so $T^*g = \bar{\phi}g$ by uniqueness of $T^*$.
(ii) It is clear that if $\phi$ is real-valued, then $\bar{\phi} = \phi$ and hence $T^* = T$.
(iii) If $|\phi| = 1$ a.e. on $[-\pi,\pi]$ then $T$ is unitary; If $\phi \geq 0$ a.e. on $[-\pi,\pi]$ then $\langle Tf, f\rangle \geq 0$ and hence $T$ is positive; If $\phi = \pm 1$ a.e. on $[-\pi,\pi]$ then $T^2 f = \phi^2 f = f$ and hence $T$ is projection.

Ex.5.7. Consider $(\lambda I - T)x = y$. Thus $x(t) = y(t)/(\lambda - t)$ provided $t \neq \lambda$ for $t \in [0,1]$. So there is a unique solution $x \in C[0,1]$ except when $t = \lambda$ for $t \in [0,1]$ and so $\sigma(T) = [0,1]$.

Ex.5.8. It follows from $(\lambda I - T)^* = \bar{\lambda}I - T^*$.

Ex.5.9. Consider a bounded sequence $(x_j)$ in $X$. Since $T$ is compact, the sequence $(Tx_j)$ has a convergent subsequence, say, $(Tx_{j_k})$ which is convergent to $y$ in $X$. Then

$$\|STx_{j_k} - Sy\| \leq \|S\|\|Tx_{j_k} - y\| \to 0 \quad \text{as} \quad k \to \infty.$$

This shows that $ST$ is compact. Note that $(Sx_j)$ is bounded. So $(TSx_j)$ has a convergent subsequence, showing that $TS$ is compact.