FUNCTIONAL ANALYSIS NOTES
(2011)

Mr. Andrew Pinchuck

Department of Mathematics (Pure & Applied)

Rhodes University

Contents

Introduction

These course notes are adapted from the original course notes written by Prof. Sizwe Mabizela when he last gave this course in 2006 to whom I am indebted. I thus make no claims of originality but have made several changes throughout. In particular, I have attempted to motivate these results in terms of applications in science and in other important branches of mathematics.

Functional analysis is the branch of mathematics, specifically of analysis, concerned with the study of vector spaces and operators acting on them. It is essentially where linear algebra meets analysis. That is, an important part of functional analysis is the study of vector spaces endowed with topological structure. Functional analysis arose in the study of tansformations of functions, such as the Fourier transform, and in the study of differential and integral equations. The founding and early development of functional analysis is largely due to a group of Polish mathematicians around Stefan Banach in the first half of the 20th century but continues to be an area of intensive research to this day. Functional analysis has its main applications in differential equations, probability theory, quantum mechanics and measure theory amongst other areas and can best be viewed as a powerful collection of tools that have far reaching consequences.

As a prerequisite for this course, the reader must be familiar with linear algebra up to the level of a standard second year university course and be familiar with real analysis. The aim of this course is to introduce the student to the key ideas of functional analysis. It should be remembered however that we only scratch the surface of this vast area in this course. We examine normed linear spaces, Hilbert spaces, bounded linear operators, dual spaces and the most famous and important results in functional analysis such as the Hahn-Banach theorem, Baires category theorem, the uniform boundedness principle, the open mapping theorem and the closed graph theorem. We attempt to give justifications and motivations for the ideas developed as we go along.

Throughout the notes, you will notice that there are exercises and it is up to the student to work through these. In certain cases, there are statements made without justification and once again it is up to the student to rigourously verify these results. For further reading on these topics the reader is referred to the following texts:

- G. BACHMAN, L. NARICI, Functional Analysis, Academic Press, N.Y. 1966.
- E. KREYSZIG, Introductory Functional Analysis, John Wiley & sons, New York-Chichester-Brisbane-Toronto, 1978.
- G. F. SIMMONS, Introduction to topology and modern analysis, McGraw-Hill Book Company, Singapore, 1963.
- A. E. TAYLOR, Introduction to Functional Analysis, John Wiley & Sons, N. Y. 1958.

I have also found Wikipedia to be quite useful as a general reference.

Chapter 1

Linear Spaces

1.1 Introducton

In this first chapter we review the important notions associated with vector spaces. We also state and prove some well known inequalities that will have important consequences in the following chapter.

Unless otherwise stated, we shall denote by $\mathbb{R}$ the field of real numbers and by $\mathbb{C}$ the field of complex numbers. Let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$.

1.1.1 Definition

A **linear space** *over a field* $\mathbb{F}$ *is a nonempty set* X *with two operations*

$$\begin{aligned} + &: X \times X \to X \quad \text{(called addition), and} \\ \cdot &: \mathbb{F} \times X \to X \quad \text{(called multiplication)} \end{aligned}$$

satisfying the following properties:

[1] $x + y \in X$ *whenever* $x, y \in X$;

[2] $x + y = y + x$ *for all* $x, y \in X$;

[3] There exists a unique element in X, *denoted by 0, such that* $x + 0 = 0 + x = x$ *for all* $x \in X$;

[4] Associated with each $x \in X$ *is a unique element in* X, *denoted by* $-x$, *such that* $x + (-x) = -x + x = 0$;

[5] $(x + y) + z = x + (y + z)$ *for all* $x, y, z \in X$;

[6] $\alpha \cdot x \in X$ *for all* $x \in X$ *and for all* $\alpha \in \mathbb{F}$;

[7] $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ *for all* $x, y \in X$ *and all* $\alpha \in \mathbb{F}$;

[8] $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$ *for all* $x \in X$ *and all* $\alpha, \beta \in \mathbb{F}$;

[9] $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ *for all* $x \in X$ *and all* $\alpha, \beta \in \mathbb{F}$;

[10] $1 \cdot x = x$ *for all* $x \in X$.

We emphasize that a linear space is a quadruple $(X, \mathbb{F}, +, \cdot)$ where X is the underlying set, $\mathbb{F}$ a field, $+$ addition, and $\cdot$ multiplication. When no confusion can arise we shall identify the linear space $(X, \mathbb{F}, +, \cdot)$ with the underlying set X. To show that X is a linear space, it suffices to show that it is closed under addition and scalar multiplication operations. Once this has been shown, it is easy to show that all the other axioms hold.

1.1.2 Definition

A **real** *(resp.* **complex***)* **linear space** *is a linear space over the real (resp. complex) field.*

A linear space is also called a **vector space** and its elements are called **vectors**.

1.1.3 Examples

[1] For a fixed positive integer n, let $X = \mathbb{F}^n = \{x = (x_1,\ x_2,\ \ldots,\ x_n) : x_i \in \mathbb{F}, i = 1,\ 2, \ldots, n\}$ – the set of all n-tuples of real or complex numbers. Define the operations of addition and scalar multiplication pointwise as follows: For all $x = (x_1,\ x_2,\ \ldots,\ x_n)$, $y = (y_1,\ y_2,\ \ldots,\ y_n)$ in $\mathbb{F}^n$ and $\alpha \in \mathbb{F}$,

$$\begin{aligned} x + y &= (x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n) \\ \alpha \cdot x &= (\alpha x_1,\ \alpha x_2,\ \ldots,\ \alpha x_n). \end{aligned}$$

Then $\mathbb{F}^n$ is a linear space over $\mathbb{F}$.

[2] Let $X = \mathcal{C}[a, b] = \{\ x : [a, b] \to \mathbb{F} \mid x$ is continuous $\}$. Define the operations of addition and scalar multiplication pointwise: For all $x, y \in X$ and all $\alpha \in \mathbb{R}$, define

$$\left.\begin{aligned} (x + y)(t) &= x(t) + y(t) \text{ and} \\ (\alpha \cdot x)(t) &= \alpha x(t) \end{aligned}\right\} \text{ for all } t \in [a, b].$$

Then $\mathcal{C}[a, b]$ is a real vector space.

Sequence Spaces: Informally, a sequence in X is a list of numbers indexed by $\mathbb{N}$. Equivalently, a sequence in X is a function $x : \mathbb{N} \to X$ given by $n \mapsto x(n) = x_n$. We shall denote a sequence $x_1,\ x_2,\ \ldots$ by

$$x = (x_1,\ x_2,\ \ldots) = (x_n)_1^\infty.$$

[3] The sequence space **s**. Let **s** denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers. Define the operations of addition and scalar multiplication pointwise: For all $x = (x_1,\ x_2,\ \ldots)$, $y = (y_1,\ y_2,\ \ldots) \in \mathbf{s}$ and all $\alpha \in \mathbb{F}$, define

$$\begin{aligned} x + y &= (x_1 + y_1,\ x_2 + y_2,\ \ldots) \\ \alpha \cdot x &= (\alpha x_1,\ \alpha x_2,\ \ldots). \end{aligned}$$

Then **s** is a linear space over $\mathbb{F}$.

[4] **The sequence space ℓ_∞.** Let $\ell_\infty = \ell_\infty(\mathbb{N})$ denote the set of all bounded sequences of real or complex numbers. That is, all sequences $x = (x_n)_1^\infty$ such that

$$\sup_{i \in \mathbb{N}} |x_i| < \infty.$$

Define the operations of addition and scalar multiplication pointwise as in example (3). Then ℓ_∞ is a linear space over $\mathbb{F}$.

[5] **The sequence space $\ell_p = \ell_p(\mathbb{N}),\quad 1 \le p < \infty$.** Let ℓ_p denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers satisfying the condition

$$\sum_{i=1}^{\infty} |x_i|^p < \infty.$$

Define the operations of addition and scalar multiplication pointwise: For all $x = (x_n)$, $y = (y_n)$ in ℓ_p and all $\alpha \in \mathbb{F}$, define

$$\begin{aligned} x + y &= (x_1 + y_1,\ x_2 + y_2,\ \ldots) \\ \alpha \cdot x &= (\alpha x_1,\ \alpha x_2,\ \ldots). \end{aligned}$$

Then ℓ_p is a linear space over $\mathbb{F}$.

Proof. Let $x = (x_1,\ x_2,\ \ldots),\ y = (y_1,\ y_2,\ \ldots) \in \ell_p$. We must show that $x + y \in \ell_p$. Since, for each $i \in \mathbb{N}$,

$$|x_i + y_i|^p \leq [2\max\{|x_i|, |y_i|\}]^p \leq 2^p \max\{|x_i|^p, |y_i|^p\} \leq 2^p \left(|x_i|^p + |y_i|^p\right),$$

it follows that

$$\sum_{i=1}^{\infty} |x_i + y_i|^p \leq 2^p \left(\sum_{i=1}^{\infty} |x_i|^p + \sum_{i=1}^{\infty} |y_i|^p \right) < \infty.$$

Thus, $x + y \in \ell_p$. Also, if $x = (x_n) \in \ell_p$ and $\alpha \in \mathbb{F}$, then

$$\sum_{i=1}^{\infty} |\alpha x_i|^p = |\alpha|^p \sum_{i=1}^{\infty} |x_i|^p < \infty.$$

That is, $\alpha \cdot x \in \ell_p$.

[6] **The sequence space $\mathbf{c} = \mathbf{c}(\mathbb{N})$.** Let $\mathbf{c}$ denote the set of all convergent sequences $x = (x_n)_1^\infty$ of real or complex numbers. That is, $\mathbf{c}$ is the set of all sequences $x = (x_n)_1^\infty$ such that $\lim_{n\to\infty} x_n$ exists. Define the operations of addition and scalar multiplication pointwise as in example (3). Then $\mathbf{c}$ is a linear space over $\mathbb{F}$.

[7] **The sequence space $\mathbf{c}_0 = \mathbf{c}_0(\mathbb{N})$.** Let $\mathbf{c}_0$ denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers which converge to zero. That is, $\mathbf{c}_0$ is the space of all sequences $x = (x_n)_1^\infty$ such that $\lim_{n\to\infty} x_n = 0$. Define the operations of addition and scalar multiplication pointwise as in example (3). Then $\mathbf{c}_0$ is a linear space over $\mathbb{F}$.

[8] **The sequence space $\ell_0 = \ell_0(\mathbb{N})$.** Let ℓ_0 denote the set of all sequences $x = (x_n)_1^\infty$ of real or complex numbers such that $x_i = 0$ for all but finitely many indices i. Define the operations of addition and scalar multiplication pointwise as in example (3). Then ℓ_0 is a linear space over $\mathbb{F}$.

1.2 Subsets of a linear space

Let X be a linear space over $\mathbb{F}$, $x \in X$ and A and B subsets of X and $\lambda \in \mathbb{F}$. We shall denote by

$$\begin{aligned} x + A &:= \{x + a : a \in A\}, \\ A + B &:= \{a + b : a \in A,\ b \in B\}, \\ \lambda A &:= \{\lambda a : a \in A\}. \end{aligned}$$

1.3 Subspaces and Convex Sets

1.3.1 Definition
A subset M of a linear space X is called a **linear subspace** *of X if*

(a) $x + y \in M$ for all $x, y \in M$, and

(b) $\lambda x \in M$ for all $x \in M$ and for all $\lambda \in \mathbb{F}$.

Clearly, a subset M of a linear space X is a linear subspace if and only if $M + M \subset M$ and $\lambda M \subset M$ for all $\lambda \in \mathbb{F}$.

1.3.2 Examples

[1] Every linear space X has at least two distinguished subspaces: $M = \{0\}$ and $M = X$. These are called the **improper subspaces** of X. All other subspaces of X are called the **proper subspaces**.

[2] Let $X = \mathbb{R}^2$. Then the nontrivial linear subspaces of X are straight lines through the origin.

[3] $M = \{x = (0,\ x_2,\ x_3, \ldots, x_n) : x_i \in \mathbb{R}, i = 2, 3, \ldots, n\}$ is a subspace of $\mathbb{R}^n$.

[4] $M = \{x : [-1, 1] \to \mathbb{R}, x$ continuous and $x(0) = 0\}$ is a subspace of $\mathcal{C}[-1, 1]$.

[5] $M = \{x : [-1, 1] \to \mathbb{R}, x$ continuous and $x(0) = 1\}$ is not a subspace of $\mathcal{C}[-1, 1]$.

[6] Show that $\mathbf{c}_0$ is a subspace of $\mathbf{c}$.

1.3.3 Definition
Let K be a subset of a linear space X. The **linear hull** *of K, denoted by lin(K) or span(K), is the intersection of all linear subspaces of X that contain K.*

The linear hull of K is also called the **linear subspace of X spanned (or generated) by K**.

It is easy to check that the intersection of a collection of linear subspaces of X is a linear subspace of X. It therefore follows that the linear hull of a subset K of a linear space X is again a linear subspace of X. In fact, the linear hull of a subset K of a linear space X is the smallest linear subspace of X which contains K.

1.3.4 Proposition
Let K be a subset of a linear space X. Then the linear hull of K is the set of all finite linear combinations of elements of K. That is,

$$\mathit{lin}(K) = \left\{ \sum_{j=1}^{n} \lambda_j x_j \mid x_1,\ x_2,\ \ldots,\ x_n \in K,\ \lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n \in \mathbb{F},\ n \in \mathbb{N} \right\}.$$

Proof. Exercise. ■

1.3.5 Definition

[1] A subset $\{x_1,\ x_2,\ \ldots,\ x_n\}$ *of a linear space* X *is said to be* **linearly independent** *if the equation*

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$

only has the trivial solution $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. *Otherwise, the set* $\{x_1,\ x_2,\ \ldots,\ x_n\}$ *is* **linearly dependent**.

[2] A subset K *of a linear space* X *is said to be* **linearly independent** *if every finite subset* $\{x_1,\ x_2,\ \ldots,\ x_n\}$ *of* K *is linearly independent.*

1.3.6 Definition

If $\{x_1,\ x_2,\ \ldots,\ x_n\}$ *is a linearly independent subset of* X *and* $X = \text{lin}\{x_1,\ x_2,\ \ldots,\ x_n\}$, *then* X *is said to have* **dimension** n. *In this case we say that* $\{x_1,\ x_2,\ \ldots,\ x_n\}$ *is a* **basis** *for the linear space* X. *If a linear space* X *does not have a finite basis, we say that it is infinite-dimensional.*

1.3.7 Examples

[1] The space $\mathbb{R}^n$ has dimension n. Its standard basis is $\{e_1,\ e_2,\ \ldots,\ e_n\}$, where, for each $j = 1,\ 2,\ \ldots,\ n$, e_j is an n-tuple of real numbers with 1 in the j-th position and zeroes elsewhere; i.e.,

$$e_j = (0,\ 0,\ \ldots, 1,\ 0,\ \ldots,\ 0), \text{ where } 1 \text{ occurs in the } j\text{-th position.}$$

[2] The space $\mathbb{P}_n$ of polynomials of degree at most n has dimension $n+1$. Its standard basis is $\{1,\ t,\ t^2,\ \ldots,\ t^n\}$.

[3] The function space $\mathcal{C}[a,b]$ is infinite-dimensional.

[4] The spaces ℓ_p, with $1 \le p \le \infty$, are infinite-dimensional.

1.3.8 Definition

Let K *be a subset of a linear space* X. *We say that*

(a) K *is* **convex** *if* $\lambda x + (1-\lambda)y \in K$ *whenever* $x, y \in K$ *and* $\lambda \in [0,1]$;

(b) K *is* **balanced** *if* $\lambda x \in K$ *whenever* $x \in K$ *and* $|\lambda| \le 1$;

(c) K *is* **absolutely convex** *if* K *is convex and balanced.*

1.3.9 Remark

[1] It is easy to verify that K is absolutely convex if and only if $\lambda x + \mu y \in K$ whenever $x, y \in K$ and $|\lambda| + |\mu| \le 1$.

[2] Every linear subspace is absolutely convex.

1.3.10 Definition

Let S be a subset of the linear space X. The **convex hull** of S, denoted co(S), is the intersection of all convex sets in X which contain S.

Since the intersection of convex sets is convex, it follows that co(S) is the smallest convex set which contains S. The following result is an alternate characterization of co(S).

1.3.11 Proposition
Let S be a nonempty subset of a linear space X. Then $co(S)$ is the set of all convex combinations of elements of S. That is,

$$co(S) = \left\{ \sum_{j=1}^{n} \lambda_j x_j \mid x_1,\ x_2,\ \ldots,\ x_n \in S,\ \lambda_j \geq 0\ \forall\ j = 1,\ 2,\ \ldots,\ n,\ \sum_{j=1}^{n} \lambda_j = 1,\ n \in \mathbb{N} \right\}.$$

Proof. Let C denote the set of all convex combinations of elements of S. That is,

$$C = \left\{ \sum_{j=1}^{n} \lambda_j x_j \mid x_1,\ x_2,\ \ldots,\ x_n \in S,\ \lambda_j \geq 0\ \forall\ j = 1,\ 2,\ \ldots,\ n,\ \sum_{j=1}^{n} \lambda_j = 1,\ n \in \mathbb{N} \right\}.$$

Let $x, y \in C$ and $0 \leq \lambda \leq 1$. Then $x = \sum_{1}^{n} \lambda_i x_i$, $y = \sum_{1}^{m} \mu_i y_i$, where $\lambda_i, \mu_i \geq 0$, $\sum_{1}^{n} \lambda_i = 1$, $\sum_{1}^{m} \mu_i = 1$, and $x_i, y_i \in S$. Thus

$$\lambda x + (1-\lambda)y = \sum_{1}^{n} \lambda\lambda_i x_i + \sum_{1}^{m} (1-\lambda)\mu_i y_i$$

is a linear combination of elements of S, with nonnegative coefficients, such that

$$\sum_{1}^{n} \lambda\lambda_i + \sum_{1}^{m} (1-\lambda)\mu_i = \lambda \sum_{1}^{n} \lambda_i + (1-\lambda) \sum_{1}^{m} \mu_i = \lambda + (1-\lambda) = 1.$$

That is, $\lambda x + (1-\lambda)y \in C$ and C is convex. Clearly $S \subset C$. Hence $\mathrm{co}(S) \subset C$.

We now prove the inclusion $C \subset \mathrm{co}(S)$. Note that, by definition, $S \subset \mathrm{co}(S)$. Let $x_1,\ x_2 \in S$, $\lambda_1 \geq 0,\ \lambda_2 \geq 0$ and $\lambda_1 + \lambda_2 = 1$. Then, by convexity of $\mathrm{co}(S)$, $\lambda_1 x_1 + \lambda_2 x_2 \in \mathrm{co}(S)$. Assume that $\sum_{1}^{n-1} \lambda_i x_i \in \mathrm{co}(S)$ whenever $x_1,\ x_2,\ \ldots,\ x_{n-1} \in S$, $\lambda_j \geq 0$, $j = 1,\ 2,\ \ldots,\ n-1$ and $\sum_{j=1}^{n-1} \lambda_j = 1$. Let $x_1,\ x_2,\ \ldots,\ x_n \in S$ and $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$ be such that $\lambda_j \geq 0$, $j = 1,\ 2,\ \ldots,\ n$ and $\sum_{j=1}^{n} \lambda_j = 1$. If $\sum_{j=1}^{n-1} \lambda_j = 0$, then $\lambda_n = 1$. Hence $\sum_{1}^{n} \lambda_j x_j = \lambda_n x_n \in \mathrm{co}(S)$. Assume that $\beta = \sum_{j=1}^{n-1} \lambda_j > 0$. Then $\frac{\lambda_j}{\beta} \geq 0$ for all $j = 1,\ 2,\ \ldots,\ n-1$ and $\sum_{j=1}^{n-1} \frac{\lambda_j}{\beta} = 1$. By the induction assumption, $\sum_{j=1}^{n-1} \frac{\lambda_j}{\beta} x_j \in \mathrm{co}(S)$. Hence

$$\sum_{j=1}^{n} \lambda_j x_j = \beta \left(\sum_{j=1}^{n-1} \frac{\lambda_j}{\beta} x_j \right) + \lambda_n x_n \in \mathrm{co}(S).$$

Thus $C \subset \mathrm{co}(S)$. ■

1.4 Quotient Space

Let M be a linear subspace of a linear space X over $\mathbb{F}$. For all $x, y \in X$, define

$$x \equiv y (\mathrm{mod}\ M) \iff x - y \in M.$$

It is easy to verify that $\equiv$ defines an equivalence relation on X.

For $x \in X$, denote by

$$[x] := \{y \in X : x \equiv y(\text{mod } M)\} = \{y \in X : x - y \in M\} = x + M,$$

the *coset* of x with respect to M. The **quotient space** X/M consists of all the equivalence classes $[x]$, $x \in X$. The quotient space is also called a **factor space**.

1.4.1 Proposition

Let M be a linear subspace of a linear space X over $\mathbb{F}$. For $x, y \in X$ and $\lambda \in \mathbb{F}$, define the operations

$$[x] + [y] = [x + y] \text{ and } \lambda \cdot [x] = [\lambda \cdot x].$$

Then X/M is a linear space with respect to these operations.

Proof. Exercise. ■

Note that the linear operations on X/M are equivalently given by: For all $x, y \in X$ and $\lambda \in \mathbb{F}$,

$$(x + M) + (y + M) = x + y + M \text{ and } \lambda(x + M) = \lambda x + M.$$

1.4.2 Definition

Let M be a linear subspace of a linear space X over $\mathbb{F}$. The **codimension** *of M in X is defined as the dimension of the quotient space X/M. It is denoted by codim$(M) = \dim(X/M)$.*

Clearly, if $X = M$, then $X/M = \{0\}$ and so codim$(X) = 0$.

1.5 Direct Sums and Projections

1.5.1 Definition

Let M and N be linear subspaces of a linear space X over $\mathbb{F}$. We say that X is a **direct sum** *of M and N if*

$$X = M + N \text{ and } M \cap N = \{0\}.$$

If X is a direct sum of M and N, we write $X = M \oplus N$. In this case, we say that M (resp. N) is an **algebraic complement** *of N (resp. M).*

1.5.2 Proposition

Let M and N be linear subspaces of a linear space X over $\mathbb{F}$. If $X = M \oplus N$, then each $x \in X$ has a unique representation of the form $x = m + n$ for some $m \in M$ and $n \in N$.

Proof. Exercise. ■

Let M and N be linear subspaces of a linear space X over $\mathbb{F}$ such that $X = M \oplus N$. Then codim(M) =dim(N). Also, since $X = M \oplus N$, dim(X) =dim(M)+dim(N). Hence

$$\dim(X) = \dim(M) + \text{codim}(M).$$

It follows that if dim$(X) < \infty$, then codim(M) =dim$(X)-$dim(M).

The operator $P : X \to X$ is called an **algebraic projection** if P is linear (i.e., $P(\alpha x + y) = \alpha P x + P y$ for all $x, y \in X$ and $\alpha \in \mathbb{F}$) and $P^2 = P$, i.e., P is idempotent.

1.5.3 Proposition
Let M and N be linear subspaces of a linear space X over $\mathbb{F}$ such that $X = M \oplus N$. Define $P : X \to X$ by $P(x) = m$, where $x = m + n$, with $m \in M$ and $n \in N$. Then P is an algebraic projection of X onto M along N. Moreover $M = P(X)$ and $N = (I - P)(X) = \ker(P)$.

Conversely, if $P : X \to X$ is an algebraic projection, then $X = M \oplus N$, where $M = P(X)$ and $N = (I - P)(X) = \ker(P)$.

Proof. **Linearity of P**: Let $x = m_1 + n_1$ and $y = m_2 + n_2$, where $m_1, m_2 \in M$ and $n_1, n_2 \in N$. For $\alpha \in \mathbb{F}$,

$$P(\alpha x + y) = P((\alpha m_1 + m_2) + (\alpha n_1 + n_2)) = \alpha m_1 + m_2 = \alpha P x + P y.$$

Idempotency of P: Since $m = m + 0$, with $m \in M$ and $0 \in N$, we have that $Pm = m$ and hence $P^2 x = Pm = m = Px$. That is, $P^2 = P$.
Finally, $n = x - m = (I - P)x$. Hence $N = (I - P)(X)$. Also, $Px = 0$ if and only if $x \in N$, i.e., $\ker(P) = N$.

Conversely, let $x \in X$ and set $m = Px$ and $n = (I - P)x$. Then $x = m + n$, where $m \in M$ and $n \in N$. We show that this representation is unique. Indeed, if $x = m_1 + n_1$ where $m_1 \in M$ and $n_1 \in N$, then $m_1 = Pu$ and $n_1 = (I - P)v$ for some $u, v \in X$. Since $P^2 = P$, it follows that $Pm_1 = m_1$ and $Pn_1 = 0$. Hence $m = Px = Pm_1 + Pn_1 = Pm_1 = m_1$. Similarly $n = n_1$. □

1.6 The Hölder and Minkowski Inequalities

We now turn our attention to three important inequalities. The first two are required mainly to prove the third which is required for our discussion about normed linear spaces in the subsequent chapter.

1.6.1 Definition
Let p and q be positive real numbers. If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, or if $p = 1$ and $q = \infty$, or if $p = \infty$ and $q = 1$, then we say that p and q are **conjugate exponents**.

1.6.2 Lemma
(Young's Inequality). *Let p and q be conjugate exponents, with $1 < p, q < \infty$ and $\alpha, \beta \geq 0$. Then*

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

Proof. If $p = 2 = q$, then the inequality follows from the fact that $(\alpha - \beta)^2 \geq 0$. Notice also, that if $\alpha = 0$ or $\beta = 0$, then the inequality follows trivially.

If $p \neq 2$, then consider the function $f : [0, \infty) \to \mathbb{R}$ given by

$$f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta, \text{ for fixed } \beta > 0.$$

Then, $f'(\alpha) = \alpha^{p-1} - \beta = 0$ when $\alpha^{p-1} = \beta$. That is, when $\alpha = \beta^{\frac{1}{p-1}} = \beta^{\frac{q}{p}} > 0$. We now apply the second derivative test to the critical point $\alpha = \beta^{\frac{q}{p}}$.

$$f''(\alpha) = (p - 1)\alpha^{p-2} > 0, \quad \text{for all } \alpha \in (0, \infty).$$

Thus, we have a global minimum at $\alpha = \beta^{\frac{q}{p}}$. It is easily verified that

$$0 = f(\beta^{\frac{q}{p}}) \leq f(\alpha) = \frac{\alpha^p}{p} + \frac{\beta^q}{q} - \alpha\beta \Leftrightarrow \alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q},$$

for each $\alpha \in [0, \infty)$.

■

1.6.3 Theorem
(Hölder's Inequality for sequences). *Let $(x_n) \in \ell_p$ and $(y_n) \in \ell_q$, where $p > 1$ and $1/p + 1/q = 1$. Then*

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}.$$

Proof. If $\sum_{k=1}^{\infty} |x_k|^p = 0$ or $\sum_{k=1}^{\infty} |y_k|^q = 0$, then the inequality holds. Assume that $\sum_{k=1}^{\infty} |x_k|^p \neq 0$ and $\sum_{k=1}^{\infty} |y_k|^q \neq 0$. Then for $k = 1, 2, \ldots$, we have, by Lemma 1.6.2, that

$$\frac{|x_k|}{\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}}} \cdot \frac{|y_k|}{\left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} \frac{|x_k|^p}{\sum_{k=1}^{\infty} |x_k|^p} + \frac{1}{q} \frac{|y_k|^q}{\sum_{k=1}^{\infty} |y_k|^q}.$$

Hence,

$$\frac{\sum_{k=1}^{\infty} |x_k y_k|}{\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

That is,

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}}.$$

■

1.6.4 Theorem
(Minkowski's Inequality for sequences). *Let $p > 1$ and (x_n) and (y_n) sequences in ℓ_p. Then*

$$\left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}.$$

Proof. Let $q = \dfrac{p}{p-1}$. If $\sum_{k=1}^{\infty} |x_k + y_k|^p = 0$, then the inequality holds. We therefore assume that $\sum_{k=1}^{\infty} |x_k + y_k|^p \neq 0$. Then

$$\begin{aligned}
\sum_{k=1}^{\infty} |x_k + y_k|^p &= \sum_{k=1}^{\infty} |x_k + y_k|^{p-1}|x_k + y_k| \\
&\leq \sum_{k=1}^{\infty} |x_k + y_k|^{p-1}|x_k| + \sum_{k=1}^{\infty} |x_k + y_k|^{p-1}|y_k| \\
&\leq \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} \left[\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}\right] \\
&= \left(\sum_{k=1}^{\infty} |x_k + y_k|^{p}\right)^{\frac{1}{q}} \left[\left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}\right].
\end{aligned}$$

Dividing both sides by $\left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{\frac{1}{q}}$, we have

$$\left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{\infty} |x_k + y_k|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |y_k|^p\right)^{\frac{1}{p}}.$$

■

1.6.5 Exercise

[1] Show that the set of all $n \times m$ real matrices is a real linear space.

[2] Show that a subset M of a linear space X is a linear subspace if and only if $\alpha x + \beta y \in M$ for all $x, y \in M$ and all $\alpha, \beta \in \mathbb{F}$.

[3] Prove Proposition 1.3.4

[4] Prove Proposition 1.4.1.

[5] Prove Proposition 1.5.2.

[6] Show that $\mathbf{c}_0$ is a linear subspace of the linear space ℓ_∞.

[7] Which of the following subsets are linear subspaces of the linear space $\mathcal{C}[-1,1]$?

(a) $M_1 = \{x \in \mathcal{C}[-1,1] \,:\, x(-1) = x(1)\}$.

(b) $M_2 = \{x \in \mathcal{C}[-1,1] \,:\, \int_{-1}^{1} x(t)dt = 1\}$.

(c) $M_3 = \{x \in \mathcal{C}[-1,1] \,:\, |x(t_2) - x(t_1)| \leq |t_2 - t_1|$ for all $t_1, t_2 \in [-1,1]\}$.

[8] Show that if $\{M_\lambda\}$ is a family of linear subspaces of a linear space X, then $M = \cap_\lambda M_\lambda$ is a linear subspace of X.

If M and N are linear subspaces of a linear space X, under what condition(s) is $M \cup N$ a linear subspace of X?

Chapter 2

Normed Linear Spaces

2.1 Preliminaries

For us to have a meaningful notion of convergence it is necessary for the Linear space to have a notion distance and therefore a topology defined on it. This leads us to the definition of a norm which induces a metric topology in a natural way.

2.1.1 Definition

A **norm** *on a linear space X is a real-valued function $\|\cdot\| : X \to \mathbb{R}$ which satisfies the following properties: For all $x, y \in X$ and $\lambda \in \mathbb{F}$,*

N1. $\|x\| \geq 0$;

N2. $\|x\| = 0 \iff x = 0$;

N3. $\|\lambda x\| = |\lambda|\|x\|$;

N4. $\|x + y\| \leq \|x\| + \|y\|$ *(Triangle Inequality).*

A **normed linear space** *is a pair $(X, \|\cdot\|)$, where X is a linear space and $\|\cdot\|$ a norm on X. The number $\|x\|$ is called the* **norm** *or* **length** *of x.*

Unless there is some danger of confusion, we shall identify the normed linear space $(X, \|\cdot\|)$ with the underlying linear space X.

2.1.2 Examples

(Examples of normed linear spaces.)

[1] Let $X = \mathbb{F}$. For each $x \in X$, define $\|x\| = |x|$. Then $(X, \|\cdot\|)$ is a normed linear space. We give the proof for $X = \mathbb{C}$. Properties N1 -N3 are easy to verify. We only verify N4. Let $x, y \in \mathbb{C}$. Then

$$\begin{aligned}
\|x + y\|^2 = |x + y|^2 &= (x + y)\overline{(x + y)} = (x + y)(\overline{x} + \overline{y}) = x\overline{x} + y\overline{x} + x\overline{y} + y\overline{y} \\
&= |x|^2 + \overline{x\overline{y}} + x\overline{y} + |y|^2 = |x|^2 + 2\Re e(x\overline{y}) + |y|^2 \\
&\leq |x|^2 + 2|x\overline{y}| + |y|^2 = |x|^2 + 2|x||\overline{y}| + |y|^2 \\
&= |x|^2 + 2|x||y| + |y|^2 \\
&= (|x| + |y|)^2 = (\|x\| + \|y\|)^2 .
\end{aligned}$$

Taking the positive square root both sides yields N4. ■

[2] Let n be a natural number and $X = \mathbb{F}^n$. For each $x = (x_1,\ x_2, \ldots, x_n) \in X$, define

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad \text{for} \quad 1 \leq p < \infty, \text{ and}$$
$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|.$$

Then $(X, \|\cdot\|_p)$ and $(X, \|\cdot\|_\infty)$ are normed linear spaces. We give a detailed proof that $(X, \|\cdot\|_p)$ is a normed linear space for $1 \leq p < \infty$.

N1. For each $1 \leq i \leq n$,

$$|x_i| \geq 0 \quad \Rightarrow \quad \sum_{i=1}^{n} |x_i|^p \geq 0 \quad \Rightarrow \quad \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \geq 0 \quad \Rightarrow \quad \|x\|_p \geq 0.$$

N2. For any $x \in X$,

$$\begin{aligned} \|x\|_p = 0 \quad &\iff \quad \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = 0 \\ &\iff \quad |x_i|^p = 0 \quad \text{for all } i = 1, 2, 3, \ldots, n \\ &\iff \quad x_i = 0 \quad \text{for all } i = 1, 2, 3, \ldots, n \iff x = 0. \end{aligned}$$

N3. For any $x \in X$ and any $\lambda \in \mathbb{F}$,

$$\begin{aligned} \|\lambda x\|_p &= \left(\sum_{i=1}^{n} |\lambda x_i|^p\right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \\ &= |\lambda| \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = |\lambda| \|x\|_p. \end{aligned}$$

N4. For any $x, y \in X$,

$$\begin{aligned} \|x + y\|_p &= \left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \\ &\leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \quad \text{(by Minkowski's Inequality)} \\ &= \|x\|_p + \|y\|_p. \end{aligned}$$

[3] Let $X = \mathcal{B}[a, b]$ be the set of all bounded real-valued functions on $[a, b]$. For each $x \in X$, define

$$\|x\|_\infty = \sup_{a\leq t\leq b} |x(t)|.$$

Then $(X, \|\cdot\|_\infty)$ is a normed linear space. We prove the triangle inequality: For any $t \in [a, b]$ and any $x, y \in X$,

$$|x(t) + y(t)| \leq |x(t)| + |y(t)| \leq \sup_{a\leq t\leq b} |x(t)| + \sup_{a\leq t\leq b} |y(t)| = \|x\|_\infty + \|y\|_\infty.$$

Since this is true for all $t \in [a,b]$, we have that

$$\|x+y\|_\infty = \sup_{a\le t\le b} |x(t)+y(t)| \le \|x\|_\infty + \|y\|_\infty.$$

[4] Let $X = \mathcal{C}[a,b]$. For each $x \in X$, define

$$\begin{aligned} \|x\|_\infty &= \sup_{a\le t\le b} |x(t)| \\ \|x\|_2 &= \left(\int_a^b |x(t)|^2 dt\right)^{\frac{1}{2}}. \end{aligned}$$

Then $(X, \|\cdot\|_\infty)$ and $(X, \|\cdot\|_2)$ are normed linear spaces.

[5] Let $X = \ell_p,\ 1 \le p < \infty$. For each $x = (x_i)_1^\infty \in X$, define

$$\|x\|_p = \left(\sum_{i\in\mathbb{N}} |x_i|^p\right)^{\frac{1}{p}}.$$

Then $(X, \|\cdot\|_p)$ is a normed linear space.

[6] Let $X = \ell_\infty,\ c$ or c_0. For each $x = (x_i)_1^\infty \in X$, define

$$\|x\| = \|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|.$$

Then X is a normed linear space.

[7] Let $X = \mathcal{L}(\mathbb{C}^n)$ be the linear space of all $n \times n$ complex matrices. For $A \in \mathcal{L}(\mathbb{C}^n)$, let $\tau(A) = \sum_{i=1}^{n} (A)_{ii}$ be the trace of A. For $A \in \mathcal{L}(\mathbb{C}^n)$, define

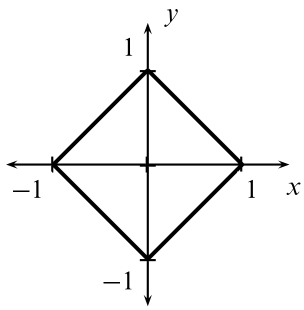
$$\|A\|_2 = \sqrt{\tau(A^*A)} = \sqrt{\sum_{i=1}^{n}\sum_{k=1}^{n} \overline{(A)_{ki}}(A)_{ki}} = \sqrt{\sum_{i=1}^{n}\sum_{k=1}^{n} |(A)_{ki}|^2},$$

where A^* is the conjugate transpose of the matrix A.

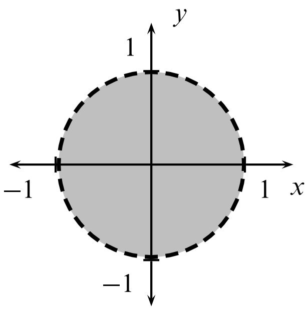
Notation

Let a be an element of a normed linear space$(X, \|\cdot\|)$ and $r > 0$.

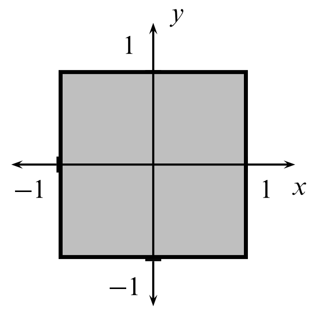
$$\begin{aligned} B(a,r) &= \{x \in X \mid \|x-a\| < r\} \text{ (Open ball with centre } a \text{ and radius } r); \\ B[a,r] &= \{x \in X \mid \|x-a\| \le r\} \text{ (Closed ball with centre } a \text{ and radius } r); \\ S(a,r) &= \{x \in X \mid \|x-a\| = r\} \text{ (Sphere with centre } a \text{ and radius } r). \end{aligned}$$



$\|(x, y)\|_1 = 1$



$\|(x, y)\|_2 < 1$



$\|(x, y)\|_\infty \leq 1$

Equivalent Norms

2.1.3 Definition
Let $\|\cdot\|$ and $\|\cdot\|_0$ be two different norms defined on the same linear space X. We say that $\|\cdot\|$ is **equivalent** *to $\|\cdot\|_0$ if there are positive numbers α and β such that*

$$\alpha\|x\| \leq \|x\|_0 \leq \beta\|x\|, \quad \textit{for all} \quad x \in X.$$

2.1.4 Example
Let $X = \mathbb{F}^n$. For each $x = (x_1, x_2, \ldots, x_n) \in X$, let

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \text{ and } \quad \|x\|_\infty = \max_{1\leq i\leq n} |x_i|.$$

We have seen that $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are norms on X. We show that these norms are equivalent.

Equivalence of $\|\cdot\|_1$ and $\|\cdot\|_\infty$: Let $x = (x_1, x_2, \ldots, x_n) \in X$. For each $k = 1, 2, \ldots, n$,

$$|x_k| \leq \sum_{i=1}^{n} |x_i| \Rightarrow \max_{1\leq k\leq n} |x_k| \leq \sum_{i=1}^{n} |x_i| \iff \|x\|_\infty \leq \|x\|_1.$$

Also, for $k = 1, 2, \ldots, n$,

$$|x_k| \leq \max_{1\leq k\leq n} |x_k| = \|x\|_\infty \Rightarrow \sum_{i=1}^{n} |x_i| \leq \sum_{i=1}^{n} \|x\|_\infty = n\|x\|_\infty \iff \|x\|_1 \leq n\|x\|_\infty.$$

Hence, $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$.

We now show that $\|\cdot\|_2$ is equivalent to $\|\cdot\|_\infty$. Let $x = (x_1, x_2, \ldots, x_n) \in X$. For each $k = 1, 2, \ldots, n$,

$$|x_k| \leq \|x\|_\infty \Rightarrow |x_k|^2 \leq (\|x\|_\infty)^2 \quad \Rightarrow \quad \sum_{i=1}^{n} |x_i|^2 \leq \sum_{i=1}^{n} (\|x\|_\infty)^2 = n(\|x\|_\infty)^2$$
$$\iff \quad \|x\|_2 \leq \sqrt{n}\|x\|_\infty.$$

Also, for each $k = 1, 2, \ldots, n$,

$$|x_k| \le \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} = \|x\|_2 \Rightarrow \max_{1\le k\le n} |x_k| \le \|x\|_2 \iff \|x\|_\infty \le \|x\|_2.$$

Consequently, $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$, which proves equivalence of the norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$.

It is, of course, obvious now that all the three norms are equivalent to each other. We shall see later that *all* norms on a finite-dimensional normed linear space are equivalent.

2.1.5 Exercise

Let $\mathcal{N}(X)$ denote the set of norms on a linear space X. For $\|\cdot\|$ and $\|\cdot\|_0$ in $\mathcal{N}(X)$, define a relation $\simeq$ by

$$\|\cdot\| \simeq \|\cdot\|_0 \text{ if and only if } \|\cdot\| \text{ is equivalent to } \|\cdot\|_0.$$

Show that $\simeq$ is an equivalence relation on $\mathcal{N}(X)$, i.e., $\simeq$ is reflexive, symmetric, and transitive.

Open and Closed Sets

2.1.6 Definition

A subset S of a normed linear space $(X, \|\cdot\|)$ is **open** *if for each $s \in S$ there is an $\epsilon > 0$ such that $B(s,\epsilon) \subset S$.*

A subset F of a normed linear space $(X, \|\cdot\|)$ is **closed** *if its complement $X \setminus F$ is open.*

2.1.7 Definition

Let S be a subset of a normed linear space $(X, \|\cdot\|)$. We define the **closure** of S, denoted by $\overline{S}$, to be the intersection of all closed sets containing S.

It is easy to show that S is closed if and only if $S = \overline{S}$.

Recall that a **metric** on a set X is a real-valued function $d : X \times X \to \mathbb{R}$ which satisfies the following properties: For all $x, y, z \in X$,

M1. $d(x, y) \ge 0$;

M2. $d(x, y) = 0 \iff x = y$;

M3. $d(x, y) = d(y, x)$;

M4. $d(x, z) \le d(x, y) + d(y, z)$.

2.1.1 Theorem

(a) If $(X, \|\cdot\|)$ is a normed linear space, then

$$d(x, y) = \|x - y\|$$

defines a metric on X. Such a metric d is said to be **induced** *or* **generated** *by the norm $\|\cdot\|$. Thus, every normed linear space is a metric space, and unless otherwise specified, we shall henceforth regard any normed linear space as a metric space with respect to the metric induced by its norm.*

(b) If d is a metric on a linear space X satisfying the properties: For all $x, y, z \in X$ and for all $\lambda \in \mathbb{F}$,

(i) $d(x, y) = d(x + z, y + z)$ *(Translation Invariance)*

(ii) $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ *(Absolute Homogeneity),*

then

$$\|x\| = d(x, 0)$$

defines a norm on X.

Proof. (a) We show that $d(x,y) = \|x-y\|$ defines a metric on X. To that end, let $x, y, z \in X$.

M1. $d(x,y) = \|x-y\| \geq 0$ by N1.

M2.

$$\begin{aligned} d(x,y) = 0 \quad &\iff \quad \|x-y\| = 0 \quad \iff \quad x - y = 0 \text{ by N2} \\ &\iff \quad x = y. \end{aligned}$$

M3.

$$\begin{aligned} d(x,y) = \|x-y\| = \|(-1)(y-x)\| &= |-1|\|y-x\| \qquad \text{by N3} \\ &= \|y-x\| = d(y,x). \end{aligned}$$

M4.

$$\begin{aligned} d(x,z) = \|x-z\| = \|(x-y)+(y-z)\| &\leq \|x-y\| + \|y-z\| \quad \text{by N4} \\ &= d(x,y) + d(y,z). \end{aligned}$$

(b) Exercise. ∎

It is clear from Theorem 2.1.1, that a metric d on a linear space X is induced by a norm on X if and only if d is translation-invariant and positive homogeneous.

2.2 Quotient Norm and Quotient Map

We now want to introduce a norm on a quotient space. Let M be a closed linear subspace of a normed linear space X over $\mathbb{F}$. For $x \in X$, define

$$\|[x]\| := \inf_{y \in [x]} \|y\|.$$

If $y \in [x]$, then $y - x \in M$ and hence $y = x + m$ for some $m \in M$. Hence

$$\|[x]\| = \inf_{y \in [x]} \|y\| = \inf_{m \in M} \|x+m\| = \inf_{m \in M} \|x - m\| = d(x, M).$$

2.2.1 Proposition
Let M be a closed linear subspace of a normed linear space X over $\mathbb{F}$. The quotient space X/M is a normed linear space with respect to the norm

$$\|[x]\| := \inf_{y \in [x]} \|y\|, \quad \textit{where } [x] \in X/M.$$

Proof.

N1. It is clear that for any $x \in X$, $\|[x]\| = d(x, M) \geq 0$.

N2. For any $x \in X$,

$$\|[x]\| = 0 \iff d(x,M) = 0 \iff x \in \overline{M} = M \iff x + M = M = [0].$$

N3. For any $x, y \in X$ and $\lambda \in \mathbb{F} \setminus \{0\}$,

$$\begin{aligned} \|\lambda[x]\| &= \|[\lambda x]\| = d(\lambda x, M) = \inf_{y \in M} \|\lambda x - y\| = \inf_{y \in M} \left\| \lambda \left(x - \frac{y}{\lambda} \right) \right\| \\ &= |\lambda| \inf_{z \in M} \|x - z\| = |\lambda| d(x, M) = |\lambda| \|[x]\|. \end{aligned}$$

N4. Let $x, y \in X$. Then

$$\begin{aligned}
\|[x]+[y]\| = \|[x+y]\| &= d(x+y, M) = \inf_{z\in M} \|x+y-z\| \\
&= \inf_{z_1,z_2\in M} \|x+y-(z_1+z_2)\| \\
&= \inf_{z_1,z_2\in M} \|(x-z_1)+(y-z_2)\| \\
&\leq \inf_{z_1,z_2\in M} \|x-z_1\| + \|y-z_2\| \\
&= \inf_{z_1\in M} \|x-z_1\| + \inf_{z_2\in M} \|y-z_2\| \\
&= d(x,M)+d(y,M) = \|[x]\| + \|[y]\|.
\end{aligned}$$

The norm on X/M as defined in Proposition 2.2.1 is called the **quotient norm** on X/M.

Let M be a closed subspace of the normed linear space X. The mapping Q_M from $X \to X/M$ defined by

$$Q_M(x) = x + M, \quad x \in X,$$

is called the **quotient map** (or **natural embedding**) of X onto X/M.

2.3 Completeness of Normed Linear Spaces

Now that we have established that every normed linear space is a metric space, we can deploy on a normed linear space all the machinery that exists for metric spaces.

2.3.1 Definition
Let $(x_n)_{n=1}^{\infty}$ be a sequence in a normed linear space $(X, \|\cdot\|)$.

(a) $(x_n)_{n=1}^{\infty}$ is said to **converge** *to x if given $\epsilon > 0$ there exists a natural number $N = N(\epsilon)$ such that*

$$\|x_n - x\| < \epsilon \text{ for all } n \geq N.$$

Equivalently, $(x_n)_{n=1}^{\infty}$ converges to x if

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

If this is the case, we shall write

$$x_n \to x \text{ or } \lim_{n\to\infty} x_n = x.$$

Convergence in the norm is called **norm convergence** *or* **strong convergence**.

(b) $(x_n)_{n=1}^{\infty}$ is called a **Cauchy sequence** *if given $\epsilon > 0$ there exists a natural number $N = N(\epsilon)$ such that*

$$\|x_n - x_m\| < \epsilon \text{ for all } n, m \geq N.$$

Equivalently, (x_n) is Cauchy if

$$\lim_{n,m\to\infty} \|x_n - x_m\| = 0.$$

In the following lemma we collect some elementary but fundamental facts about normed linear spaces. In particular, it implies that the operations of addition and scalar multiplication, as well as the norm and distance functions, are continuous.

2.3.2 Lemma

Let C be a closed set in a normed linear space $(X, \|\cdot\|)$ over $\mathbb{F}$, and let (x_n) be a sequence contained in C such that $\lim_{n\to\infty} x_n = x \in X$. Then $x \in C$.

Proof. Exercise.

2.3.3 Lemma

Let X be a normed linear space and A a nonempty subset of X.

[1] $|d(x, A) - d(y, A)| \leq \|x - y\|$ for all $x, y \in X$;

[2] $|\,\|x\| - \|y\|\,| \leq \|x - y\|$ for all $x, y \in X$;

[3] If $x_n \to x$, then $\|x_n\| \to \|x\|$;

[4] If $x_n \to x$ and $y_n \to y$, then $x_n + y_n \to x + y$;

[5] If $x_n \to x$ and $\alpha_n \to \alpha$, then $\alpha_n x_n \to \alpha x$;

[6] The closure of a linear subspace in X is again a linear subspace;

[7] Every Cauchy sequence is bounded;

[8] Every convergent sequence is a Cauchy sequence.

Proof. (1). For any $a \in A$,

$$d(x, A) \leq \|x - a\| \leq \|x - y\| + \|y - a\|,$$

so $d(x, A) \leq \|x - y\| + d(y, A)$ or $d(x, A) - d(y, A) \leq \|x - y\|$. Interchanging the roles of x and y gives the desired result.

(2) follows from (1) by taking $A = \{0\}$.

(3) is an obvious consequence of (2).

(4), (5) and (8) follow from the triangle inequality and, in the case of (5), the absolute homogeneity.

(6) follows from (4) and (5).

(7). Let (x_n) be a Cauchy sequence in X. Choose n_1 so that $\|x_n - x_{n_1}\| \leq 1$ for all $n \geq n_1$. By (2), $\|x_n\| \leq 1 + \|x_{n_1}\|$ for all $n \geq n_1$. Thus

$$\|x_n\| \leq \max\{\, \|x_1\|, \|x_2\|, \|x_3\|, \ldots, \|x_{n_1 - 1}\|, 1 + \|x_{n_1}\|\}$$

for all n.

(8) Let (x_n) be a sequence in X which converges to $x \in X$ and let $\epsilon > 0$. Then there is a natural number N such that $\|x_n - x\| < \frac{\epsilon}{2}$ for all $n \geq N$. For all $n, m \geq N$,

$$\|x_n - x_m\| \leq \|x_n - x\| + \|x - x_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus, (x_n) is a Cauchy sequence in X. ∎

2.3.4 Proposition

Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$. A Cauchy sequence in X which has a convergent subsequence is convergent.

Proof. Let (x_n) be a Cauchy sequence in X and (x_{n_k}) its subsequence which converges to $x \in X$. Then, for any $\epsilon > 0$, there are positive integers N_1 and N_2 such that

$$\|x_n - x_m\| < \frac{\epsilon}{2} \text{ for all } n, m \geq N_1$$

and

$$\|x_{n_k} - x\| < \frac{\epsilon}{2} \text{ for all } k \geq N_2.$$

Let $N = \max\{N_1,\ N_2\}$. If $k \geq N$, then since $n_k \geq k$,

$$\|x_k - x\| \leq \|x_k - x_{n_k}\| + \|x_{n_k} - x\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence $x_n \to x$ as $n \to \infty$. ■

2.3.5 Definition

A metric space (X, d) is said to be **complete** if every Cauchy sequence in X converges in X.

2.3.6 Definition

A normed linear space that is complete with respect to the metric induced by the norm is called a **Banach space**.

2.3.1 Theorem

Let $(X, \|\cdot\|)$ be a Banach space and let M be a linear subspace of X. Then M is complete if and only if the M is closed in X.

Proof. Assume that M is complete. We show that M is closed. To that end, let $x \in \overline{M}$. Then there is a sequence (x_n) in M such that $\|x_n - x\| \to 0$ as $n \to \infty$. Since (x_n) converges, it is Cauchy. Completeness of M guarantees the existence of an element $y \in M$ such that $\|x_n - y\| \to 0$ as $n \to \infty$. By uniqueness of limits, $x = y$. Hence $x \in M$ and, consequently, M is closed.

Assume that M is closed. We show that M is complete. Let (x_n) be a Cauchy sequence in M. Then (x_n) is a Cauchy sequence in X. Since X is complete, there is an element $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. But then $x \in M$ since M is closed. Hence M is complete. ■

2.3.7 Examples

[1] Let $1 \leq p < \infty$. Then for each positive integer n, $(\mathbb{F}^n, \|\cdot\|_p)$ is a Banach space.

[2] For each positive integer n, $(\mathbb{F}^n, \|\cdot\|_\infty)$ is a Banach space.

[3] Let $1 \leq p < \infty$. The sequence space ℓ_p is a Banach space. Because of the importance of this space, we give a detailed proof of its completeness.

The classical sequence space ℓ_p is complete.

Proof. Let $(x_n)_1^\infty$ be a Cauchy sequence in ℓ_p. We shall denote each member of this sequence by

$$x_n = (x_n(1), x_n(2), \ldots).$$

Then, given $\epsilon > 0$, there exists an $N(\epsilon) = N \in \mathbb{N}$ such that

$$\|x_n - x_m\|_p = \left(\sum_{i=1}^{\infty} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} < \epsilon \quad \text{for all} \quad n, m \geq N.$$

For each fixed index i, we have

$$|x_n(i) - x_m(i)| < \epsilon \quad \text{for all} \quad n, m \geq N.$$

That is, for each fixed index i, $(x_n(i))_1^\infty$ is a Cauchy sequence in $\mathbb{F}$. Since $\mathbb{F}$ is complete, there exists $x(i) \in \mathbb{F}$ such that

$$x_n(i) \to x(i) \text{ as } n \to \infty.$$

Define $x = (x(1), x(2), \ldots)$. We show that $x \in \ell_p$, and $x_n \to x$. To that end, for each $k \in \mathbb{N}$,

$$\left(\sum_{i=1}^{k} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} \leq \|x_n - x_m\|_p = \left(\sum_{i=1}^{\infty} |x_n(i) - x_m(i)|^p\right)^{\frac{1}{p}} < \epsilon.$$

That is,

$$\sum_{i=1}^{k} |x_n(i) - x_m(i)|^p < \epsilon^p, \text{ for all } k = 1, 2, 3, \ldots.$$

Keep k and $n \geq N$ fixed and let $m \to \infty$. Since we are dealing with a finite sum,

$$\sum_{i=1}^{k} |x_n(i) - x(i)|^p \leq \epsilon^p.$$

Now letting $k \to \infty$, then for all $n \geq N$,

$$\sum_{i=1}^{\infty} |x_n(i) - x(i)|^p \leq \epsilon^p, \tag{2.3.7.1}$$

which means that $x_n - x \in \ell_p$. Since $x_n \in \ell_p$, we have that $x = (x - x_n) + x_n \in \ell_p$. It also follows from (2.3.7.1) that $x_n \to x$ as $n \to \infty$. ■

[4] The space ℓ_0 of all sequences $(x_i)_1^\infty$ with only a finite number of nonzero terms is an incomplete normed linear space. It suffices to show that ℓ_0 is not closed in ℓ_2 (and hence not complete). To that end, consider the sequence $(x_i)_1^\infty$ with terms

$$\begin{aligned}
x_1 &= (1, 0, 0, 0, \ldots) \\
x_2 &= (1, \frac{1}{2}, 0, 0, 0, \ldots) \\
x_3 &= (1, \frac{1}{2}, \frac{1}{2^2}, 0, 0, 0, \ldots) \\
&\vdots \\
x_n &= (1, \frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^{n-1}}, 0, 0, 0, \ldots) \\
&\vdots
\end{aligned}$$

This sequence $(x_i)_1^\infty$ converges to

$$x = (1, \frac{1}{2}, \frac{1}{2^2}, \ldots, \frac{1}{2^{n-1}}, \frac{1}{2^n}, \frac{1}{2^{n+1}}, \ldots).$$

Indeed, since $x - x_n = (0, 0, 0, \ldots, 0, \frac{1}{2^n}, \frac{1}{2^{n+1}}, \ldots)$, it follows that

$$\|x_n - x\|^2 = \sum_{k=n}^{\infty} \frac{1}{2^{2k}} \to 0 \text{ as } n \to \infty.$$

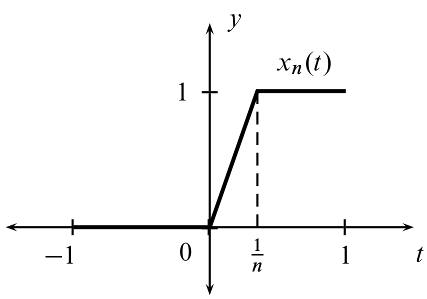
That is, $x_n \to x$ as $n \to \infty$, but $x \notin \ell_0$. ■

[5] The space $\mathcal{C}_2[-1,1]$ of continuous real-valued functions on $[-1,1]$ with the norm

$$\|x\|_2 = \left(\int_{-1}^{1} x^2(t)\,dt\right)^{1/2}$$

is an incomplete normed linear space.
To see this, it suffices to show that there is a Cauchy sequence in $\mathcal{C}_2[-1,1]$ which converges to an element which does not belong to $\mathcal{C}_2[-1,1]$. Consider the sequence $(x_n)_1^\infty \in \mathcal{C}_2[-1,1]$ defined by

$$x_n(t) = \begin{cases} 0 & \text{if } \; -1 \le t \le 0 \\ nt & \text{if } \; 0 \le t \le \frac{1}{n} \\ 1 & \text{if } \; \frac{1}{n} \le t \le 1. \end{cases}$$



We show that $(x_n)_1^\infty$ is a Cauchy sequence in $\mathcal{C}_2[-1,1]$. To that end, for positive integers m and n such that $m > n$,

$$\begin{aligned} \|x_n - x_m\|_2^2 &= \int_{-1}^{1} [x_n(t) - x_m(t)]^2 \; dt \\ &= \int_0^{1/m} [nt - mt]^2 \; dt + \int_{1/m}^{1/n} [1 - nt]^2 \; dt \\ &= \int_0^{1/m} [m^2t^2 - 2mnt^2 + n^2t^2] \; dt + \int_{1/m}^{1/n} [1 - 2nt + n^2t^2] \; dt \\ &= (m^2 - 2mn + n^2) \left. \frac{t^3}{3} \right|_0^{1/m} + \left. \left(t - nt^2 + n^2\frac{t^3}{3} \right) \right|_{1/m}^{1/n} \\ &= \frac{m^2 - 2mn + n^2}{3m^2n} = \frac{(m-n)^2}{3m^2n} \to 0 \quad \text{as} \quad n, m \to \infty. \end{aligned}$$

Define

$$x(t) = \begin{cases} 0 & \text{if } \; -1 \le t \le 0 \\ 1 & \text{if } \; 0 < t \le 1. \end{cases}$$

Then $x \notin \mathcal{C}_2[-1,1]$, and

$$\|x_n - x\|_2^2 = \int_{-1}^{1} [x_n(t) - x(t)]^2 \, dt = \int_{0}^{\frac{1}{n}} [nt-1]^2 \, dt = \frac{1}{3n} \to 0 \quad \text{as} \quad n \to \infty.$$

That is, $x_n \to x$ as $n \to \infty$. ■

2.4 Series in Normed Linear Spaces

Let (x_n) be a sequence in a normed linear space $(X, \|\cdot\|)$. To this sequence we associate another sequence (s_n) of *partial sums*, where $s_n = \sum_{k=1}^{n} x_k$.

2.4.1 Definition
Let (x_n) be a sequence in a normed linear space $(X, \|\cdot\|)$. If the sequence (s_n) of partial sums converges to s, then we say that the series $\sum_{k=1}^{\infty} x_k$ **converges** *and that its* **sum** *is s. In this case we write $\sum_{k=1}^{\infty} x_k = s$.*
The series $\sum_{k=1}^{\infty} x_k$ is said to be **absolutely convergent** *if $\sum_{k=1}^{\infty} \|x_k\| < \infty$.*

We now give a series characterization of completeness in normed linear spaces.

2.4.1 Theorem
A normed linear space $(X, \|\cdot\|)$ is a Banach space if and only if every absolutely convergent series in X is convergent.

Proof. Let X be a Banach space and suppose that $\sum_{j=1}^{\infty} \|x_j\| < \infty$. We show that the series $\sum_{j=1}^{\infty} x_j$ converges. To that end, let $\epsilon > 0$ and for each $n \in \mathbb{N}$, let $s_n = \sum_{j=1}^{n} x_j$. Let K be a positive integer such that $\sum_{j=K+1}^{\infty} \|x_j\| < \epsilon$. Then, for all $m > n > K$, we have

$$\|s_m - s_n\| = \left\| \sum_{1}^{m} x_j - \sum_{1}^{n} x_j \right\| = \left\| \sum_{n+1}^{m} x_j \right\| \leq \sum_{n+1}^{m} \|x_j\| \leq \sum_{n+1}^{\infty} \|x_j\| \leq \sum_{K+1}^{\infty} \|x_j\| < \epsilon.$$

Hence the sequence (s_n) of partial sums forms a Cauchy sequence in X. Since X is complete, the sequence (s_n) converges to some element $s \in X$. That is, the series $\sum_{j=1}^{\infty} x_j$ converges.

Conversely, assume that $(X, \|\cdot\|)$ is a normed linear space in which every absolutely convergent series converges. We show that X is complete. Let (x_n) be a Cauchy sequence in X. Then there is an $n_1 \in \mathbb{N}$ such that $\|x_{n_1} - x_m\| < \frac{1}{2}$ whenever $m > n_1$. Similarly, there is an $n_2 \in \mathbb{N}$ with $n_2 > n_1$ such that $\|x_{n_2} - x_m\| < \frac{1}{2^2}$ whenever $m > n_2$. Continuing in this way, we get natural numbers $n_1 < n_2 < \cdots$ such that $\|x_{n_k} - x_m\| < \frac{1}{2^k}$ whenever $m > n_k$. In particular, we have that for each $k \in \mathbb{N}$, $\|x_{n_{k+1}} - x_{n_k}\| < 2^{-k}$. For each $k \in \mathbb{N}$, let $y_k = x_{n_{k+1}} - x_{n_k}$. Then

$$\sum_{k=1}^{n} \|y_k\| = \sum_{k=1}^{n} \|x_{n_{k+1}} - x_{n_k}\| < \sum_{k=1}^{n} \frac{1}{2^k}.$$

Hence, $\sum_{k=1}^{\infty} \|y_k\| < \infty$. That is, the series $\sum_{k=1}^{\infty} y_k$ is absolutely convergent, and hence, by our assumption, the series $\sum_{k=1}^{\infty} y_k$ is convergent in X. That is, there is an $s \in X$ such that $s_j = \sum_{k=1}^{j} y_k \to s$ as $j \to \infty$. It follows that

$$s_j = \sum_{k=1}^{j} y_k = \sum_{k=1}^{j} [x_{n_{k+1}} - x_{n_k}] = x_{n_{j+1}} - x_{n_1} \overset{j \to \infty}{\longrightarrow} s.$$

Hence $x_{n_{j+1}} \overset{j \to \infty}{\longrightarrow} s + x_{n_1}$. Thus, the subsequence (x_{n_k}) of (x_n) converges in X. But if a Cauchy sequence has a convergent subsequence, then the sequence itself also converges (to the same limit as the subsequence). It thus follows that the sequence (x_n) also converges in X. Hence X is complete. ■

We now apply Theorem 2.4.1 to show that if M is a closed linear subspace of a Banach space X, then the quotient space X/M, with the quotient norm, is also a Banach space.

2.4.2 Theorem
Let M be a closed linear subspace of a Banach space X. Then the quotient space X/M is a Banach space when equipped with the quotient norm.

Proof. Let $([x_n])$ be a sequence in X/M such that $\sum_{j=1}^{\infty} \|[x_j]\| < \infty$. For each $j \in \mathbb{N}$, choose an element $y_j \in M$ such that

$$\|x_j - y_j\| \leq \|[x_j]\| + 2^{-j}.$$

It now follows that $\sum_{j=1}^{\infty} \|x_j - y_j\| < \infty$, i.e., the series $\sum_{j=1}^{\infty} (x_j - y_j)$ is absolutely convergent in X. Since X is complete, the series $\sum_{j=1}^{\infty} (x_j - y_j)$ converges to some element $z \in X$. We show that the series $\sum_{j=1}^{\infty} [x_j]$ converges to $[z]$. Indeed, for each $n \in \mathbb{N}$,

$$\begin{aligned}
\left\| \sum_{j=1}^{n} [x_j] - [z] \right\| &= \left\| \left[\sum_{j=1}^{n} x_j \right] - [z] \right\| = \left\| \left[\sum_{j=1}^{n} x_j - z \right] \right\| \\
&= \inf_{m \in M} \left\| \sum_{j=1}^{n} x_j - z - m \right\| \\
&\leq \left\| \sum_{j=1}^{n} x_j - z - \sum_{j=1}^{n} y_j \right\| \\
&= \left\| \sum_{j=1}^{n} (x_j - y_j) - z \right\| \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Hence, every absolutely convergent series in X/M is convergent, and so X/M is complete. ■

2.5 Bounded, Totally Bounded, and Compact Subsets of a Normed Linear Space

2.5.1 Definition
A subset A of a normed linear space $(X, \|\cdot\|)$ is **bounded** *if $A \subset B[x, r]$ for some $x \in X$ and $r > 0$.*

It is clear that A is bounded if and only if there is a $C > 0$ such that $\|a\| \leq C$ for all $a \in A$.

2.5.2 Definition
*Let A be a subset of a normed linear space $(X, \|\cdot\|)$ and $\epsilon > 0$. A subset $A_\epsilon \subset X$ is called an ϵ-***net** *for A if for each $x \in A$ there is an element $y \in A_\epsilon$ such that $\|x - y\| < \epsilon$. Simply put, $A_\epsilon \subset X$ is an ϵ-net for A if each element of A is within an ϵ distance to some element of A_ϵ.*
A subset A of a normed linear space $(X, \|\cdot\|)$ is **totally bounded** *(or* **precompact***) if for any $\epsilon > 0$ there is a finite ϵ-net $F_\epsilon \subset X$ for A. That is, there is a finite set $F_\epsilon \subset X$ such that*

$$A \subset \bigcup_{x \in F_\epsilon} B(x, \epsilon).$$

The following proposition shows that total boundedness is a stronger property than boundedness.

2.5.3 Proposition
Every totally bounded subset of a normed linear space $(X, \|\cdot\|)$ is bounded.

Proof. This follows from the fact that a finite union of bounded sets is also bounded. ■

The following example shows that boundedness does not, in general, imply total boundedness.

2.5.4 Example
Let $X = \ell_2$ and consider $B = B(X) = \{x \in X \mid \|x\| \leq 1\}$, the closed unit ball in X. Clearly, B is bounded. We show that B is not totally bounded. Consider the elements of B of the form: for $j \in \mathbb{N}$, $e_j = (0, 0, \ldots, 0, 1, 0, \ldots)$, where 1 occurs in the j-th position. Note that $\|e_i - e_j\|_2 = \sqrt{2}$ for all $i \neq j$. Assume that an ϵ-net $B_\epsilon \subset X$ existed for $0 < \epsilon < \frac{\sqrt{2}}{2}$. Then for each $j \in \mathbb{N}$, there is an element $y_j \in B_\epsilon$ such that $\|e_j - y_j\| < \epsilon$. This says that for each $j \in \mathbb{N}$, there is an element $y_j \in B_\epsilon$ such that $y_j \in B(e_j, \epsilon)$. But the balls $B(e_j, \epsilon)$ are disjoint. Indeed, if $i \neq j$, and $z \in B(e_i, \epsilon) \cap B(e_j, \epsilon)$, then by the triangle inequality

$$\sqrt{2} = \|e_i - e_j\|_2 \leq \|e_i - z\| + \|z - e_j\| < 2\epsilon < \sqrt{2},$$

which is absurd. Since the balls $B(e_j, \epsilon)$ are (at least) countably infinite, there can be no finite ϵ-net for B.

In our definition of total boundedness of a subset $A \subset X$, we required that the finite ϵ-net be a subset of X. The following proposition suggests that the finite ϵ-net may actually be assumed to be a subset of A itself.

2.5.5 Proposition

A subset A of a normed linear space $(X, \|\cdot\|)$ is totally bounded if and only if for any $\epsilon > 0$ there is a finite set $F_\epsilon \subset A$ such that

$$A \subset \bigcup_{x \in F_\epsilon} B(x, \epsilon).$$

Proof. Exercise. ■

We now give a characterization of total boundedness.

2.5.1 Theorem

A subset K of a normed linear space $(X, \|\cdot\|)$ is totally bounded if and only if every sequence in K has a Cauchy subsequence.

Proof. Assume that K is totally bounded and let (x_n) be an infinite sequence in K. There is a finite set of points $\{y_{11},\ y_{12},\ \ldots, y_{1r}\}$ in K such that

$$K \subset \bigcup_{j=1}^{r} B(y_{1j}, \frac{1}{2}).$$

At least one of the balls $B(y_{1j}, \frac{1}{2}),\ j = 1,\ 2, \ldots,\ r$, contains an infinite subsequence (x_{n1}) of (x_n). Again, there is a finite set $\{y_{21},\ y_{22},\ \ldots, y_{2s}\}$ in K such that

$$K \subset \bigcup_{j=1}^{s} B(y_{2j}, \frac{1}{2^2}).$$

At least one of the balls $B(y_{2j}, \frac{1}{2^2}),\ j = 1,\ 2, \ldots,\ s$, contains an infinite subsequence (x_{n2}) of (x_{n1}). Continuing in this way, at the m-th step, we obtain a subsequence (x_{nm}) of $(x_{n(m-1)})$ which is contained in a ball of the form $B\left(y_{mj}, \frac{1}{2^m}\right)$.

<u>**Claim:**</u> The diagonal subsequence (x_{nn}) of (x_n) is Cauchy. Indeed, if $m > n$, then both x_{nn} and x_{mm} are in the ball of radius 2^{-n}. Hence, by the triangle inequality,

$$\|x_{nn} - x_{mm}\| < 2^{1-n} \ \to\ 0 \text{ as } n \ \to\ \infty.$$

Conversely, assume that every sequence in K has a Cauchy subsequence and that K is *not* totally bounded. Then, for some $\epsilon > 0$, no finite ϵ-net exists for K. Hence, if $x_1 \in K$, then there is an $x_2 \in K$ such that $\|x_1 - x_2\| \geq \epsilon$. (Otherwise, $\|x_1 - y\| < \epsilon$ for all $y \in K$ and consequently $\{x_1\}$ is a finite ϵ-net for K, a contradiction.) Similarly, there is an $x_3 \in K$ such that

$$\|x_1 - x_3\| \geq \epsilon \text{ and } \|x_2 - x_3\| \geq \epsilon.$$

Continuing in this way, we obtain a sequence (x_n) in K such that $\|x_n - x_m\| \geq \epsilon$ for all $m \neq n$. Therefore (x_n) cannot have a Cauchy subsequence, a contradiction. ■

2.5.6 Definition

A normed linear space $(X, \|\cdot\|)$ is **sequentially compact** *if every sequence in X has a convergent subsequence.*

2.5.7 Remark

It can be shown that on a metric space, compactness and sequential compactness are equivalent. Thus, it follows, that on a normed linear space, we can use these terms interchangeably.

2.5.2 Theorem
A subset of a normed linear space is sequentially compact if and only if it is totally bounded and complete.

Proof. Let K be a sequentially compact subset of a normed linear space $(X, \|\cdot\|)$. We show that K is totally bounded. To that end, let (x_n) be a sequence in K. By sequential compactness of K, (x_n) has a subsequence (x_{n_k}) which converges in K. Since every convergent sequence is Cauchy, the subsequence (x_{n_k}) of (x_n) is Cauchy. Therefore, by Theorem 2.5.1, K is totally bounded.

Next, we show that K is complete. Let (x_n) be a Cauchy sequence in K. By sequential compactness of K, (x_n) has a subsequence (x_{n_k}) which converges in K. But if a subsequence of a Cauchy sequence converges, so does the full sequence. Hence (x_n) converges in K and so K is complete.

Conversely, assume that K is a totally bounded and complete subset of a normed linear space $(X, \|\cdot\|)$. We show that K is sequentially compact. Let (x_n) be a sequence in K. By Theorem 2.5.1, (x_n) has a Cauchy subsequence (x_{n_k}). Since K is complete, (x_{n_k}) converges in K. Hence K is sequentially compact. ■

2.5.8 Corollary
A subset of a Banach space is sequentially compact if and only if it is totally bounded and closed.

Proof. Exercise. ■

2.5.9 Corollary
A sequentially compact subset of a normed linear space is closed and bounded.

Proof. Exercise. ■

We shall see that in finite-dimensional spaces the converse of Corollary 2.5.9 also holds.

2.5.10 Corollary
A closed subset F of a sequentially compact normed linear space $(X, \|\cdot\|)$ is sequentially compact.

Proof. Exercise. ■

2.6 Finite Dimensional Normed Linear Spaces

The theory for finite-dimensional normed linear spaces turns out to be much simpler than that of their infinite-dimensional counterparts. In this section we highlight some of the special aspects of finite-dimensional normed linear spaces.

The following Lemma is crucial in the analysis of finite-dimensional normed linear spaces.

2.6.1 Lemma
Let $(X, \|\cdot\|)$ be a finite-dimensional normed linear space with basis $\{x_1, x_2, \ldots, x_n\}$. Then there is a constant $m > 0$ such that for every choice of scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, we have

$$m \sum_{j=1}^{n} |\alpha_j| \leq \left\| \sum_{j=1}^{n} \alpha_j \, x_j \right\|.$$

Proof. If $\sum_{j=1}^{n} |\alpha_j| = 0$, then $\alpha_j = 0$ for all $j = 1, 2, \ldots, n$ and the inequality holds for any $m > 0$. Assume that $\sum_{j=1}^{n} |\alpha_j| \neq 0$. We shall prove the result for a set of scalars $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ that satisfy the condition $\sum_{j=1}^{n} |\alpha_j| = 1$. Let

$$A = \{(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{F}^n \mid \sum_{j=1}^{n} |\alpha_j| = 1\}.$$

Since A is a closed and bounded subset of $\mathbb{F}^n$, it is compact. Define $f : A \to \mathbb{R}$ by

$$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = \left\| \sum_{j=1}^{n} \alpha_j \, x_j \right\|.$$

Since for any $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(\beta_1, \beta_2, \ldots, \beta_n)$ in A

$$\begin{aligned}
|f(\alpha_1, \alpha_2, \ldots, \alpha_n) - f(\beta_1, \beta_2, \ldots, \beta_n)| &= \left| \left\| \sum_{j=1}^{n} \alpha_j \, x_j \right\| - \left\| \sum_{j=1}^{n} \beta_j \, x_j \right\| \right| \\
&\leq \left\| \sum_{j=1}^{n} \alpha_j \, x_j - \sum_{j=1}^{n} \beta_j \, x_j \right\| \\
&= \left\| \sum_{j=1}^{n} (\alpha_j - \beta_j) x_j \right\| \leq \sum_{j=1}^{n} |\alpha_j - \beta_j| \|x_j\| \\
&\leq \max_{1 \leq j \leq n} \|x_j\| \sum_{j=1}^{n} |\alpha_j - \beta_j|,
\end{aligned}$$

f is continuous on A. Since f is a continuous function on a compact set A, it attains its minimum on A, i.e., there is an element $(\mu_1, \mu_2, \ldots, \mu_n) \in A$ such that

$$f(\mu_1, \mu_2, \ldots, \mu_n) = \inf\{f(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid (\alpha_1, \alpha_2, \ldots, \alpha_n) \in A\}.$$

Let $m = f(\mu_1, \mu_2, \ldots, \mu_n)$. Since $f \geq 0$, it follows that $m \geq 0$. If $m = 0$, then

$$\left\| \sum_{j=1}^{n} \mu_j \, x_j \right\| = 0 \;\; \Rightarrow \sum_{j=1}^{n} \mu_j \, x_j = 0.$$

Since the set $\{x_1, x_2, \ldots, x_n\}$ is linearly independent, $\mu_j = 0$ for all $j = 1, 2, \ldots, n$. This is a contradiction since $(\mu_1, \mu_2, \ldots, \mu_n) \in A$. Hence $m > 0$ and consequently for all $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in A$,

$$0 < m \leq f(\alpha_1, \alpha_2, \ldots, \alpha_n) \iff m \sum_{j=1}^{n} |\alpha_j| \leq \left\| \sum_{j=1}^{n} \alpha_j \, x_j \right\|.$$

Now, let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be *any* collection of scalars and set $\beta = \sum_{j=1}^{n} |\alpha_j|$. If $\beta = 0$, then the inequality holds vacuously. If $\beta > 0$, then $\left(\frac{\alpha_1}{\beta}, \frac{\alpha_2}{\beta}, \dots, \frac{\alpha_n}{\beta}\right) \in A$ and consequently

$$\left\|\sum_{j=1}^{n} \alpha_j x_j\right\| = \left\|\sum_{j=1}^{n} \frac{\alpha_j}{\beta} x_j\right\| \beta = f\left(\frac{\alpha_1}{\beta}, \frac{\alpha_2}{\beta}, \dots, \frac{\alpha_n}{\beta}\right) \beta \geq m\beta = m \sum_{j=1}^{n} |\alpha_j|.$$

That is, $m \sum_{j=1}^{n} |\alpha_j| \leq \left\|\sum_{j=1}^{n} \alpha_j x_j\right\|$. ■

2.6.1 Theorem

Let X be a finite-dimensional normed linear space over $\mathbb{F}$. Then all norms on X are equivalent.

Proof. Let $\{x_1, x_2, \dots, x_n\}$ be a basis for X and $\|\cdot\|_0$ and $\|\cdot\|$ be any two norms on X. For any $x \in X$ there is a set of scalars $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ such that $x = \sum_{j=1}^{n} \alpha_j x_j$. By Lemma 2.6.1, there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_j| \leq \left\|\sum_{j=1}^{n} \alpha_j x_j\right\| = \|x\|.$$

By the triangle inequality

$$\|x\|_0 \leq \sum_{j=1}^{n} |\alpha_j| \|x_j\|_0 \leq M \sum_{j=1}^{n} |\alpha_j|,$$

where $M = \max_{1 \leq j \leq n} \|x_j\|_0$. Hence

$$\|x\|_0 \leq M\left(\frac{1}{m}\|x\|\right) \Rightarrow \frac{m}{M}\|x\|_0 \leq \|x\| \iff \alpha\|x\|_0 \leq \|x\| \text{ where } \alpha = \frac{m}{M}.$$

Interchanging the roles of the norms $\|\cdot\|_0$ and $\|\cdot\|$, we similarly get a constant β such that $\|x\| \leq \beta\|x\|_0$. Hence, $\alpha\|x\|_0 \leq \|x\| \leq \beta\|x\|_0$ for some constants α and β. ■

2.6.2 Theorem

Every finite-dimensional normed linear space $(X, \|\cdot\|)$ is complete.

Proof. Let $\{x_1, x_2, \dots, x_n\}$ be a basis for X and let (z_k) be a Cauchy sequence in X. Then, given any $\epsilon > 0$, there is a natural number N such that

$$\|z_k - z_\ell\| < \epsilon \text{ for all } k, \ell > N.$$

Also, for each $k \in \mathbb{N}$, $z_k = \sum_{j=1}^{n} \alpha_{kj} x_j$. By Lemma 2.6.1, there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_{kj} - \alpha_{\ell j}| \leq \|z_k - z_\ell\|.$$

Hence, for all $k, \ell > N$ and all $j = 1, 2, \dots, n$,

$$|\alpha_{kj} - \alpha_{\ell j}| \leq \frac{1}{m}\|z_k - z_\ell\| < \frac{\epsilon}{m}.$$

That is, for each $j = 1, 2, \ldots, n$, $(\alpha_{kj})_k$ is a Cauchy sequence of numbers. Since $\mathbb{F}$ is complete, $\alpha_{kj} \to \alpha_j$ as $k \to \infty$ for each $j = 1, 2, \ldots, n$. Define $z = \sum_{j=1}^{n} \alpha_j x_j$. Then $z \in X$ and

$$\|z_k - z\| = \left\| \sum_{j=1}^{n} \alpha_{kj} x_j - \sum_{j=1}^{n} \alpha_j x_j \right\| = \left\| \sum_{j=1}^{n} (\alpha_{kj} - \alpha_j) x_j \right\| \le \sum_{j=1}^{n} |\alpha_{kj} - \alpha_j| \|x_j\| \to 0$$

as $k \to \infty$. That is, the sequence (z_k) converges to $z \in X$. hence X is complete. ■

2.6.2 Corollary
Every finite-dimensional normed linear space X is closed.

Proof. Exercise. ■

2.6.3 Theorem
In a finite-dimensional normed linear space $(X, \|\cdot\|)$, a subset $K \subset X$ is sequentially compact if and only if it is closed and bounded.

Proof. We have seen (Corollary 2.5.9), that a compact subset of a normed linear space is closed and bounded.

Conversely, assume that a subset $K \subset X$ is closed and bounded. We show that K is compact. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for X and let (z_k) be any sequence in K. Then for each $k \in \mathbb{N}$, $z_k = \sum_{j=1}^{n} \alpha_{kj} x_j$. Since K is bounded, there is a positive constant M such that $\|z_k\| \le M$ for all $k \in \mathbb{N}$. By Lemma 2.6.1, there is an $m > 0$ such that

$$m \sum_{j=1}^{n} |\alpha_{kj}| \le \left\| \sum_{j=1}^{n} \alpha_{kj} x_j \right\| = \|z_k\| \le M.$$

It now follows that $|\alpha_{kj}| \le \frac{M}{m}$ for each $j = 1, 2, \ldots, n$, and for all $k \in \mathbb{N}$. That is, for each fixed $j = 1, 2, \ldots, n$, the sequence $(\alpha_{kj})_k$ of numbers is bounded. Hence the sequence $(\alpha_{kj})_k$ has a subsequence $(\alpha_{k_r j})$ which converges to α_j for $j = 1, 2, \ldots, n$. Setting $z = \sum_{j=1}^{n} \alpha_j x_j$, we have that

$$\|z_{k_r} - z\| = \left\| \sum_{j=1}^{n} \alpha_{k_r j} x_j - \sum_{j=1}^{n} \alpha_j x_j \right\| \le \sum_{j=1}^{n} |\alpha_{k_r j} - \alpha_j| \|x_j\| \to 0 \text{ as } r \to \infty.$$

That is, $z_{k_r} \to z$ as $r \to \infty$. Since K is closed, $z \in K$. Hence K is compact. ■

2.6.3 Lemma
(Riesz's Lemma). *Let M be a closed proper linear subspace of a normed linear space $(X, \|\cdot\|)$. Then for each $0 < \epsilon < 1$, there is an element $z \in X$ such that $\|z\| = 1$ and*

$$\|y - z\| > 1 - \epsilon \quad \textit{for all} \quad y \in M.$$

Proof. Choose $x \in X \setminus M$ and define

$$d = d(x, M) = \inf_{m \in M} \|x - m\|.$$

Since M is closed, $d > 0$. By definition of infimum, there is a $m \in M$ such that

$$d \leq \|x - m\| < d + \epsilon d = d(1 + \epsilon).$$

Take $z = -\left(\dfrac{m - x}{\|m - x\|}\right)$. Then $\|z\| = 1$ and for any $y \in M$,

$$\begin{aligned} \|y - z\| &= \left\| y + \left(\frac{m - x}{\|m - x\|}\right)\right\| = \frac{\|y(\|m - x\|) + m - x\|}{\|m - x\|} \\ &\geq \frac{d}{\|m - x\|} > \frac{d}{d(1 + \epsilon)} = \frac{1}{1 + \epsilon} = 1 - \frac{\epsilon}{1 + \epsilon} > 1 - \epsilon. \end{aligned}$$

We now give a topological characterization of the algebraic concept of finite dimensionality.

2.6.4 Theorem

A normed linear space $(X, \|\cdot\|)$ is finite-dimensional if and only its closed unit ball $B(X) = \{x \in X \mid \|x\| \leq 1\}$ is compact.

Proof. Assume that $(X, \|\cdot\|)$ is finite-dimensional normed linear space. Since the ball $B(X)$ is closed and bounded, it is compact.

Assume that the closed unit ball $B(X) = \{x \in X \mid \|x\| \leq 1\}$ is compact. Then $B(X)$ is totally bounded. Hence there is a finite $\frac{1}{2}$-net $\{x_1, x_2, \ldots, x_n\}$ in $B(X)$. Let M =lin$\{x_1, x_2, \ldots, x_n\}$. Then M is a finite-dimensional linear subspace of X and hence closed.

Claim: $M = X$. If M is a proper subspace of X, then, by Riesz's Lemma there is an element $x_0 \in B(X)$ such that $d(x_0, M) > \frac{1}{2}$. In particular, $\|x_0 - x_k\| > \frac{1}{2}$ for all $k = 1, 2, \ldots, n$. However this contradicts the fact that $\{x_1, x_2, \ldots, x_n\}$ is a $\frac{1}{2}$-net in $B(X)$. Hence $M = X$ and, consequently, X is finite-dimensional. ■

We now give another argument to show that boundedness does not imply total boundedness. Let $X = \ell_2$ and $B(X) = \{x \in X \mid \|x\|_2 \leq 1\}$. It is obvious that $B(X)$ is bounded. We show that $B(X)$ is not totally bounded. Since X is complete and $B(X)$ is a closed subset of X, $B(X)$ is complete. If $B(X)$ were totally bounded, then $B(X)$ would, according to Theorem 2.26, be compact. By Theorem 2.6.4, X would be finite-dimensional. But this is false since X is infinite-dimensional.

2.7 Separable Spaces and Schauder Bases

2.7.1 Definition

(a) *A subset S of a normed linear space $(X, \|\cdot\|)$ is said to be* **dense** *in X if $\overline{S} = X$; i.e., for each $x \in X$ and $\epsilon > 0$, there is a $y \in S$ such that $\|x - y\| < \epsilon$.*

(b) *A normed linear space $(X, \|\cdot\|)$ is said to be* **separable** *if it contains a countable dense subset.*

2.7.2 Examples

[1] The real line $\mathbb{R}$ is separable since the set $\mathbb{Q}$ of rational numbers is a countable dense subset of $\mathbb{R}$.

[2] The complex plane $\mathbb{C}$ is separable since the set of all complex numbers with rational real and imaginary parts is a countable dense subset of $\mathbb{C}$.

[3] The sequence space ℓ_p, where $1 \leq p < \infty$, is separable. Take M to be the set of all sequences with rational entries such that all but a finite number of the entries are zero. (If

the entries are complex, take for M the set of finitely nonzero sequences with rational real and imaginary parts.) It is clear that M is countable. We show that M is dense in ℓ_p. Let $\epsilon > 0$ and $x = (x_n) \in \ell_p$. Then there is an N such that

$$\sum_{k=N+1}^{\infty} |x_k|^p < \frac{\epsilon}{2}.$$

Now, for each $1 \leq k \leq N$, there is a rational number q_k such that $|x_k - q_k|^p < \frac{\epsilon}{2N}$. Set $q = (q_1, q_2, \ldots, q_N, 0, 0, \ldots)$. Then $q \in M$ and

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

Hence M is dense in ℓ_p.

[4] The sequence space ℓ_∞, with the supremum norm, is not separable. To see this, consider the set M of elements $x = (x_n)$, in which x_n is either 0 or 1. This set is uncountable since we may consider each element of M as a binary representation of a number in the interval $[0, 1]$. Hence there are uncountably many sequences of zeroes and ones. For any two distinct elements $x, y \in M$, $\|x - y\|_\infty = 1$. Let each of the elements of M be a centre of a ball of radius $\frac{1}{4}$. Then we get uncountably many nonintersecting balls. If A is any dense subset of ℓ_∞, then each of these balls contains a point of A. Hence A cannot be countable and, consequently, ℓ_∞ is not separable.

2.7.1 Theorem

A normed linear space $(X, \|\cdot\|)$ is separable if and only if it contains a countable set B such that $\overline{\text{lin}}(B) = X$.

Proof. Assume that X is separable and let A be a countable dense subset of X. Since the linear hull of A, $\text{lin}(A)$, contains A and A is dense in X, we have that $\text{lin}(A)$ is dense in X, that is, $\overline{\text{lin}}(A) = X$.

Conversely, assume that X contains a countable set B such that $\overline{\text{lin}}(B) = X$. Let $B = \{x_n \mid n \in \mathbb{N}\}$. Assume first that $= \mathbb{R}$, and put

$$C = \left\{ \sum_{j=1}^{n} \lambda_j x_j \mid \lambda_j \in \mathbb{Q}, j = 1, 2, \ldots, n, \ n \in \mathbb{N} \right\}.$$

We first show that C is a countable subset of X. The set $\mathbb{Q} \times B$ is countable and consequently, the family $\mathcal{F}$ of all finite subsets of $\mathbb{Q} \times B$ is also countable. The mapping

$$\{(\lambda_1, x_1), (\lambda_2, x_2), \ldots, (\lambda_n, x_n)\} \mapsto \sum_{j=1}^{n} \lambda_j x_j$$

maps $\mathcal{F}$ onto C. Hence C is countable.

Next, we show that C is dense in X. Let $x \in X$ and $\epsilon > 0$. Since $\overline{\text{lin}}(B) = X$, we can find an $n \in \mathbb{N}$, points $x_1, x_2, \ldots, x_n \in B$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$ such that

$$\left\| x - \sum_{j=1}^{n} \lambda_j x_j \right\| < \frac{\epsilon}{2}.$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for each $\lambda_i \in \mathbb{R}$, we can find a $\mu_i \in \mathbb{Q}$ such that

$$|\lambda_i - \mu_i| < \frac{\epsilon}{2n(1 + \|x_i\|)} \text{ for all } i = 1, 2, \dots, n.$$

Hence,

$$\begin{aligned}\left\|x - \sum_{j=1}^{n} \mu_j x_j\right\| &\leq \left\|x - \sum_{j=1}^{n} \lambda_j x_j\right\| + \left\|\sum_{j=1}^{n} \lambda_j x_j - \sum_{j=1}^{n} \mu_j x_j\right\| \\ &< \frac{\epsilon}{2} + \sum_{j=1}^{n} |\lambda_j - \mu_j| \|x_j\| \\ &< \frac{\epsilon}{2} + \sum_{j=1}^{n} \frac{\epsilon\|x_j\|}{2n(1 + \|x_j\|)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

This shows that C is dense in X.
If $\mathbb{F} = \mathbb{C}$, the set C is that of finite linear combinations with coefficients being those complex numbers with rational real and imaginary parts.

We now give another argument based on Theorem 2.7.1 to show that the sequence space ℓ_p, where $1 \leq p < \infty$, is separable. Let $e_n = (\delta_{nm})_{m \in \mathbb{N}}$, where

$$\delta_{nm} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $e_n \in \ell_p$. Let $\epsilon > 0$ and $x = (x_n) \in \ell_p$. Then there is a natural number N such that

$$\sum_{k=n+1}^{\infty} |x_k|^p < \epsilon^p \text{ for all } n \geq N.$$

Now, if $n \geq N$, then

$$\left\|x - \sum_{j=1}^{n} x_j e_j\right\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p\right)^{1/p} < \epsilon.$$

Hence $\overline{\text{lin}}(\{e_n \mid n \in \mathbb{N}\}) = \ell_p$. Of course, the set $\{e_n \mid n \in \mathbb{N}\}$ is countable.

2.7.3 Definition

A sequence (b_n) in a Banach space $(X, \|\cdot\|)$ is called a **Schauder basis** *if for any $x \in X$, there is a unique sequence (α_n) of scalars such that*

$$\lim_{n\to\infty} \left\|x - \sum_{j=1}^{n} \alpha_j b_j\right\| = 0.$$

In this case we write $x = \sum_{j=1}^{\infty} \alpha_j b_j$.

2.7.4 Remark

It is clear from Definition 2.7.3 that (b_n) is a Schauder basis if and only if $X = \overline{\text{lin}}\{b_n \mid n \in \mathbb{N}\}$ and every $x \in X$ has a unique expansion $x = \sum_{j=1}^{\infty} \alpha_j b_j$.

Uniqueness of this expansion clearly implies that the set $\{b_n \mid n \in \mathbb{N}\}$ is linearly independent.

2.7.5 Examples

[1] For $1 \leq p < \infty$, the sequence (e_n), where $e_n = (\delta_{nm})_{m\in\mathbb{N}}$, is a Schauder basis for ℓ_p.

[2] (e_n) is a Schauder basis for c_0.

[3] $(e_n) \cup \{e\}$, where $e = (1,\ 1,\ 1, \ldots)$ (the constant 1 sequence), is a Schauder basis for c.

[4] ℓ_∞ has no Schauder basis.

2.7.6 Proposition

If a Banach space $(X, \|\cdot\|)$ has a Schauder basis, then it is separable.

Proof. Let (b_n) be a Schauder basis for X. Then $\{b_n \mid n \in \mathbb{N}\}$ is countable and $\overline{\text{lin}}(\{b_n \mid n \in \mathbb{N}\}) = X$. ■

Schauder bases have been constructed for most of the well-known Banach spaces. Schauder conjectured that every separable Banach space has a Schauder basis. This conjecture, known as the Basis Problem, remained unresolved for a long time until Per Enflo in 1973 answered it in the negative. He constructed a separable reflexive Banach space with no basis.

2.7.7 Exercise

[1] Let X be a normed linear space over $\mathbb{F}$. Show that X is finite-dimensional if and only if every bounded sequence in X has a convergent subsequence.

[2] Complete the proof of Theorem 2.1.1.

[3] Prove Lemma 2.3.2.

[4] Prove the claims made in [1] and [2] of Example 2.3.7.

[5] Prove Theorem 2.5.5.

[6] Prove Corollary 2.5.8.

[7] Prove Corollary 2.5.9.

[8] Prove Corollary 2.5.10.

[9] Prove Corollary 2.6.2.

[10] Is $(\mathcal{C}[a,b], \|\cdot\|_1)$ complete? What about $(\mathcal{C}[a,b], \|\cdot\|_\infty)$? Fully justify both answers.

Chapter 3

Hilbert Spaces

3.1 Introduction

In this chapter we introduce an inner product which is an abstract version of the dot product in elementary vector algebra. Recall that if $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ are any two vectors in $\mathbb{R}^3$, then the dot product of x and y is $x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3$. Also, the length of the vector x is $\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{x \cdot x}$.

It turns out that Hilbert spaces are a natural generalization of finite-dimensional Euclidean spaces. Hilbert spaces arise naturally and frequently in mathematics, physics, and engineering, typically as infinite-dimensional function spaces.

3.1.1 Definition

Let X be a linear space over a field $\mathbb{F}$. An **inner product** *on X is a scalar-valued function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ such that for all $x, y, z \in X$ and for all $\alpha, \beta \in \mathbb{F}$, we have*

IP1. $\langle x, x \rangle \geq 0$;

IP2. $\langle x, x \rangle = 0 \iff x = 0$;

IP3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ *(The bar denotes complex conjugation.);*

IP4. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$;

IP5. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.

An **inner product space** $(X, \langle \cdot, \cdot \rangle)$ *is a linear space X together with an inner $\langle \cdot, \cdot \rangle$ product defined on it. An inner product space is also called* **pre-Hilbert space**.

3.1.2 Examples

Examples of inner product spaces.

[1] Fix a positive integer n. Let $X = \mathbb{F}^n$. For $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in X, define

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Since this is a finite sum, $\langle \cdot, \cdot \rangle$ is well-defined. It is easy to show that $(X, \langle \cdot, \cdot \rangle)$ is an inner product space. The space $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) with this inner product is called the **Euclidean n-space** (resp. **unitary n-space**) and will be denoted by $\ell_2(n)$.

[2] Let $X = \ell_0$, the linear space of finitely non-zero sequences of real or complex numbers. For $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ in X, define

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

Since this is essentially a finite sum, $\langle\cdot,\cdot\rangle$ is well-defined. It is easy to show that $(X, \langle\cdot,\cdot\rangle)$ is an inner product space.

[3] Let $X = \ell_2$, the space of all sequences $x = (x_1, x_2, \ldots)$ of real or complex numbers with $\sum_{1}^{\infty} |x_i|^2 < \infty$. For $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ in X, define

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

In order to show that $\langle\cdot,\cdot\rangle$ is well-defined we first observe that if a and b are real numbers, then

$$0 \le (a-b)^2, \quad \text{whence} \quad ab \le \frac{1}{2}(a^2 + b^2).$$

Using this fact, we have that

$$|x_i \overline{y_i}| = |x_i||\overline{y_i}| \le \frac{1}{2}\left(|x_i|^2 + |y_i|^2\right) \quad \Rightarrow \quad \sum_{i=1}^{\infty} |x_i \overline{y_i}| \le \frac{1}{2}\left(\sum_{i=1}^{\infty} |x_i|^2 + \sum_{i=1}^{\infty} |y_i|^2\right) < \infty.$$

Hence, $\langle\cdot,\cdot\rangle$ is well-defined (i.e., the series converges).

[4] Let $X = \mathcal{C}[a, b]$, the space of all continuous complex-valued functions on $[a, b]$. For $x, y \in X$, define

$$\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}\, dt.$$

We shall denote by $\mathcal{C}_2[a, b]$ the linear space $\mathcal{C}[a, b]$ equipped with this inner product.

[5] Let $X = \mathcal{L}(\mathbb{C}^n)$ be the linear space of all $n \times n$ complex matrices. For $A \in \mathcal{L}(\mathbb{C}^n)$, let $\tau(A) = \sum_{i=1}^{n} (A)_{ii}$ be the trace of A. For $A, B \in \mathcal{L}(\mathbb{C}^n)$, define

$$\langle A, B\rangle = \tau(B^* A), \quad \text{where} \quad B^* \text{ denotes conjugate transpose of matrix } B.$$

Show that $(\mathcal{L}(\mathbb{C}^n), \langle\cdot,\cdot\rangle)$ is an inner product space.

It should be mentioned that we could consider real Hilbert spaces but there are powerful methods that can be applied by using the more general complex Hilbert spaces.

3.1.1 Theorem

(Cauchy-Bunyakowsky-Schwarz Inequality). *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Then for all $x, y \in X$,*

$$|\langle x, y\rangle| \le \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

Moreover, given any $x, y \in X$, the equality

$$|\langle x, y\rangle| = \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}$$

holds if and only if x and y are linearly dependent.

Proof. If $x = 0$ or $y = 0$, then the result holds vacuously. Assume that $x \neq 0$ and $y \neq 0$. For any $\alpha \in \mathbb{F}$, we have

$$\begin{aligned} 0 \leq \langle x - \alpha y, x - \alpha y \rangle &= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\alpha} \langle x, y \rangle + \alpha \overline{\alpha} \langle y, y \rangle \\ &= \langle x, x \rangle - \overline{\alpha} \langle x, y \rangle - \alpha [\langle y, x \rangle - \overline{\alpha} \langle y, y \rangle]. \end{aligned}$$

Now choosing $\alpha = \dfrac{\langle x, y \rangle}{\langle y, y \rangle}$, we have

$$0 \leq \langle x, x \rangle - \frac{\langle y, x \rangle \langle x, y \rangle}{\langle y, y \rangle} = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle},$$

whence

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$

Assume that $|\langle x, y \rangle| = \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}$. We show that x and y are linearly dependent. If $x = 0$ or $y = 0$, then x and y are obviously linearly dependent. We therefore assume that $x \neq 0$ and $y \neq 0$. Then $\langle y, y \rangle \neq 0$. With $\alpha = \dfrac{\langle x, y \rangle}{\langle y, y \rangle}$, we have that

$$\langle x - \alpha y, x - \alpha y \rangle = \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} = 0.$$

That is,

$$\langle x - \alpha y, x - \alpha y \rangle = 0, \quad \Rightarrow \quad x = \alpha y.$$

That is, x and y are linearly dependent.

Conversely, assume that x and y are linearly dependent. Without loss of generality, $x = \lambda y$ for some $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} |\langle x, y \rangle| &= |\langle \lambda y, y \rangle| = |\lambda| |\langle y, y \rangle| = |\lambda| \langle y, y \rangle \\ &= |\lambda| \sqrt{\langle y, y \rangle} \sqrt{\langle y, y \rangle} = \sqrt{|\lambda|^2 \langle y, y \rangle} \sqrt{\langle y, y \rangle} = \sqrt{\langle \lambda y \lambda y \rangle} \sqrt{\langle y, y \rangle} \\ &= \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}. \end{aligned}$$

■

3.1.2 Theorem
Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over a field $\mathbb{F}$. For each $x \in X$, define

$$\|x\| := \sqrt{\langle x, x \rangle}. \tag{3.1.2.1}$$

Then $\| \cdot \|$ defines a norm on X. That is, $(X, \| \cdot \|)$ is a normed linear space over $\mathbb{F}$.

Proof. Let $x, y \in X$ and $\lambda \in \mathbb{F}$. Then

N1. $\|x\| = \sqrt{\langle x, x \rangle} \geq 0$;

N2. $\|x\| = 0 \iff \sqrt{\langle x, x \rangle} = 0 \iff \langle x, x \rangle = 0 \iff x = 0$, by IP2.

N3. $\|\lambda x\| = \sqrt{\langle \lambda x \lambda x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} = |\lambda| \|x\|$.

N4.

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle &= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle y,x\rangle} + \langle y,y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\ &= \|x\|^2 + 2\Re e\langle x,y\rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} + \|y\|^2 \text{ (by Theorem 3.1.1)} \\ &= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

Taking the positive square root both sides yields

$$\|x+y\| \leq \|x\| + \|y\|. \qquad \blacksquare$$

In view of (3.1.2.1), the Cauchy-Bunyakowsky-Schwarz Inequality now becomes

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

Any inner product space can thus be made into a normed linear space in a natural way: by defining the norm as in (3.1.2.1). The norm advertised in (3.1.2.1) is called the **inner product norm** or a **norm induced or generated** by the inner product.

A natural question arises: Is every normed linear space an inner product space? If the answer is NO, how then does one recognise among all normed linear spaces those that are inner product spaces in disguise, i.e., those whose norms are induced by an inner product?

These questions will be examined later.

3.1.3 Theorem

(Polarization Identity). *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Then for all $x, y \in X$,*

$$\langle x,y\rangle = \frac{\|x+y\|^2}{4} - \frac{\|x-y\|^2}{4} \quad \textit{if} \quad \mathbb{F} = \mathbb{R}, \quad \textit{and}$$

$$\langle x,y\rangle = \frac{\|x+y\|^2}{4} - \frac{\|x-y\|^2}{4} + i\left(\frac{\|x+yi\|^2}{4} - \frac{\|x-yi\|^2}{4}\right) \quad \textit{if} \quad \mathbb{F} = \mathbb{C}.$$

Proof. Assume that $\mathbb{F} = \mathbb{R}$. Then

$$\begin{aligned}\|x+y\|^2 - \|x-y\|^2 &= \langle x+y, x+y\rangle - \langle x-y, x-y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle - \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle - \langle y,y\rangle \\ &= 4\langle x,y\rangle, \quad \text{since} \quad \langle x,y\rangle = \langle y,x\rangle.\end{aligned}$$

The case when $\mathbb{F} = \mathbb{C}$ is proved analogously and is left as an exercise. $\blacksquare$

3.1.4 Theorem

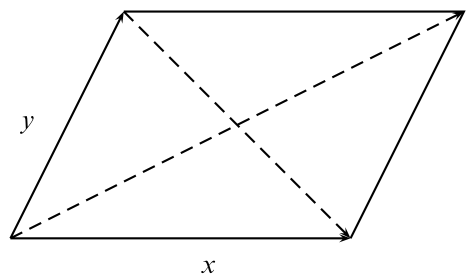
(Parallelogram Identity). *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Then for all $x, y \in X$,*

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2. \qquad (3.1.4.1)$$

Proof.

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

The geometric interpretation of the Parallelogram Identity is evident: the sum of the squares of the lengths of the diagonals of a parallelogram is equal to the sum of the squares of the lengths of the four sides.



The following theorem asserts that the Parallelogram Identity (Theorem 3.1.4) distinguishes inner product spaces among all normed linear spaces. It also answers the question posed after Theorem 3.1.2. That is, a normed linear space is an inner product space if and only if its norm satisfies the Parallelogram Identity.

3.1.5 Theorem
A normed linear space X over a field $\mathbb{F}$ is an inner product space if and only if the Parallelogram Identity

$$\|x-y\|^2+\|x+y\|^2=2\|x\|^2+2\|y\|^2 \qquad (PI)$$

holds for all $x, y \in X$.

Proof. "$\Rightarrow$". We have already shown (Theorem 3.1.4) that if X is an inner product space, then the parallelogram identity (PI) holds in X.

"$\Leftarrow$". Let X be a normed linear space in which the parallelogram identity (PI) holds. We shall only consider the case $\mathbb{F}=\mathbb{R}$. The polarization identity (Theorem 3.1.3) gives us a hint as to how we should define an inner product: For all $x, y \in X$, define

$$\langle x, y\rangle = \left\|\frac{x+y}{2}\right\|^2 - \left\|\frac{x-y}{2}\right\|^2.$$

We claim that $\langle\cdot,\cdot\rangle$ is an inner product on X.

IP1. $\langle x, x\rangle = \left\|\frac{x+x}{2}\right\|^2 - \left\|\frac{x-x}{2}\right\|^2 = \|x\|^2 \geq 0.$

IP2. $\langle x, x\rangle = 0 \iff \|x\|^2 = 0 \iff x = 0.$

IP3. $\langle x, y\rangle = \left\|\frac{x+y}{2}\right\|^2 - \left\|\frac{x-y}{2}\right\|^2 = \left\|\frac{y+x}{2}\right\|^2 - \left\|\frac{y-x}{2}\right\|^2 = \langle y, x\rangle = \overline{\langle y, x\rangle}$ since $\mathbb{F}=\mathbb{R}$.

IP5. Replace x by $u+v$ and y by $w+v$ in the parallelogram identity:

$$\|u+w+2v\|^2+\|u-w\|^2=2\|u+v\|^2+2\|w+v\|^2. \qquad (3.1.5.1)$$

Replace x by $u-v$ and y by $w-v$ in the parallelogram identity:

$$\|u+w-2v\|^2+\|u-w\|^2=2\|u-v\|^2+2\|w-v\|^2. \qquad (3.1.5.2)$$

Subtract (3.1.5.2) from (3.1.5.1):

$$\|u+w+2v\|^2-\|u+w-2v\|^2=2\left[\|u+v\|^2-\|u-v\|^2+\|v+w\|^2-\|v-w\|^2\right].$$

Use the definition of $\langle\cdot,\cdot\rangle$,

$$4\langle u+w, 2v\rangle = 8[\langle u,v\rangle + \langle w,v\rangle] \quad\Rightarrow\quad \frac{1}{2}\langle u+w,2v\rangle = \langle u,v\rangle + \langle w,v\rangle. \tag{3.1.5.3}$$

Take $w = 0$:

$$\frac{1}{2}\langle u,2v\rangle = \langle u,v\rangle. \tag{3.1.5.4}$$

Now replace u by $x+y$ and v by z in (3.1.5.4) and use (3.1.5.3) to get

$$\langle x+y,z\rangle = \frac{1}{2}\langle x+y,2z\rangle = \langle x,z\rangle + \langle y,z\rangle.$$

IP4. We show that $\langle\lambda x, y\rangle = \lambda\langle x,y\rangle$ for all $\lambda\in\mathbb{R}$ and all $x,y\in X$. If $\lambda = n$ is a nonzero integer, then using IP5,

$$\langle nx,y\rangle = n\langle x,y\rangle \quad\Rightarrow\quad n\left\langle \frac{x}{n},y\right\rangle = \left\langle\frac{nx}{n},y\right\rangle = \langle x,y\rangle.$$

That is,

$$\left\langle\frac{x}{n},y\right\rangle = \frac{1}{n}\langle x,y\rangle.$$

If λ is a rational number, $\lambda = \frac{p}{q}$, say. Then

$$\left\langle\frac{p}{q}x,y\right\rangle = p\left\langle\frac{x}{q},y\right\rangle = \frac{p}{q}\langle x,y\rangle.$$

If $\lambda\in\mathbb{R}$, then there is a sequence (r_k) of rational numbers such that $r_k\to\lambda$ as $k\to\infty$. Using continuity of the norm, we have that

$$\begin{aligned}
\langle\lambda x,y\rangle = \langle\lim_{k\to\infty} r_k x,y\rangle &= \frac{1}{4}\left\|\lim_{k\to\infty} r_k x + y\right\|^2 - \frac{1}{4}\left\|\lim_{k\to\infty} r_k x - y\right\|^2\\
&= \frac{1}{4}\lim_{k\to\infty}\|r_k x + y\|^2 - \frac{1}{4}\lim_{k\to\infty}\|r_k x - y\|^2\\
&= \lim_{k\to\infty}\left(\left\|\frac{r_k x + y}{2}\right\|^2 - \left\|\frac{r_k x - y}{2}\right\|^2\right)\\
&= \lim_{k\to\infty}\langle r_k x,y\rangle\\
&= \lim_{k\to\infty} r_k\langle x,y\rangle = \lambda\langle x,y\rangle.
\end{aligned}$$

Thus, $\langle\lambda x,y\rangle = \lambda\langle x,y\rangle$ for all $\lambda\in\mathbb{R}$ and all $x,y\in X$. ■

3.1.3 Corollary

Let $(X,\|\cdot\|)$ be a normed linear space over a field $\mathbb{F}$. If every two-dimensional linear subspace of X is an inner product space over $\mathbb{F}$, then X is an inner product space.

3.1.4 Examples

[1] Let $X = \ell_p$, for $p\neq 2$. Then X is **not** an inner product space. We show that the norm on ℓ_p, $p\neq 2$ does not satisfy the parallelogram identity. Take $x = (1,1,0,0,\ldots)$ and $y = (1,-1,0,0,\ldots)$ in ℓ_p. Then

$$\|x\| = 2^{\frac{1}{p}} = \|y\| \qquad \text{and} \qquad \|x+y\| = 2 = \|x-y\|.$$

Thus,

$$\|x+y\|^2 + \|x-y\|^2 = 8 \neq 2\|x\|^2 + 2\|y\|^2 = 4\cdot 2^{\frac{2}{p}}.$$

[2] The normed linear space $X = \mathcal{C}[a,b]$, with the supremum norm $\|\cdot\|_\infty$ is **not** an inner product space. We show that the norm

$$\|x\|_\infty = \max_{a \le t \le b} |x(t)|$$

does not satisfy the parallelogram identity. To that end, take

$$x(t) = 1 \qquad \text{and} \qquad y(t) = \frac{t-a}{b-a}.$$

Since

$$x(t) + y(t) = 1 + \frac{t-a}{b-a} \quad \text{and} \quad x(t) - y(t) = 1 - \frac{t-a}{b-a},$$

we have that

$$\|x\| = 1 = \|y\|, \qquad \text{and} \quad \|x+y\| = 2, \quad \|x-y\| = 1.$$

Thus,

$$\|x+y\|^2 + \|x-y\|^2 = 5 \ne 2\|x\|^2 + 2\|y\|^2 = 4.$$

3.2 Completeness of Inner Product Spaces

The mathematical concept of a Hilbert space, named after David Hilbert, generalizes the notion of Euclidean space. Hilbert spaces, as the following definition states, are inner product spaces which in addition are required to be complete, a property that stipulates the existence of enough limits in the space to allow the techniques of calculus to be used.

The earliest Hilbert spaces were studied from this more abstract point of view in the first decade of the 20th century by David Hilbert, Erhard Schmidt, and Frigyes Riesz. They are indispensable tools in the theories of partial differential equations, quantum mechanics, Fourier analysis which includes applications to signal processing, and ergodic theory which forms the mathematical underpinning of the study of thermodynamics.

3.2.1 Definition
Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. If X is complete with respect to the norm induced by the inner product $\langle\cdot,\cdot\rangle$, then we say that X is a **Hilbert space**.

3.2.2 Examples

[1] The classical space ℓ_2 is a Hilbert space.

[2] ℓ_0 is an incomplete inner product space.

[3] The space $\mathcal{C}[-1,1]$ is an incomplete inner product space.

3.3 Orthogonality

3.3.1 Definition
Two elements x and y in an inner product space $(X, \langle\cdot,\cdot\rangle)$ are said to be **orthogonal**, *denoted by $x \perp y$, if*

$$\langle x, y\rangle = 0.$$

The set $M \subset X$ is called **orthogonal** *if it consists of non-zero pairwise orthogonal elements.*
If M is a subset of X such that $\langle x, m\rangle = 0$ for all $m \in M$, then we say that x is orthogonal to M and write $x \perp M$. We shall denote by

$$M^\perp = \{x \in X : \langle x, m\rangle = 0 \ \forall\ m \in M\}$$

the set of all elements in X that are orthogonal to M. The set $M^\perp$ is called the **orthogonal complement** *of M.*

3.3.2 Proposition
Let M and N be subsets of an inner product space $(X, \langle\cdot,\cdot\rangle)$. Then

[1] $\{0\}^\perp = X$ and $X^\perp = \{0\}$;

[2] $M^\perp$ is a closed linear subspace of X;

[3] $M \subset (M^\perp)^\perp = M^{\perp\perp}$;

[4] If M is a linear subspace, then $M \cap M^\perp = \{0\}$;

[5] If $M \subset N$, then $N^\perp \subset M^\perp$;

[6] $M^\perp = (\text{lin}M)^\perp = (\overline{\text{lin}}M)^\perp$.

Proof.

[1] Exercise.

[2] Let $x, y \in M^\perp$, and $\alpha, \beta \in \mathbb{F}$. Then for each $z \in M$,

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle = 0.$$

Hence, $\alpha x + \beta y \in M^\perp$. That is, $M^\perp$ is a subspace of X. To show that $M^\perp$ is closed, let $x \in \overline{M^\perp}$. Then there exists a sequence (x_n) in $M^\perp$ such that $x_n \to x$ as $n \to \infty$. Thus, for all $y \in M$,

$$\langle x, y\rangle = \lim_n \langle x_n, y\rangle = 0,$$

whence $x \in M^\perp$.

[3] Exercise.

[4] Exercise.

[5] Let $x \in N^\perp$. Then $\langle x, y\rangle = 0$ for all $y \in N$. In particular, $\langle x, y\rangle = 0$ for all $y \in M$ since $M \subset N$. Thus, $x \in M^\perp$.

[6] Since $M \subset \text{lin}M \subset \overline{\text{lin}}M$, we have, by [5], that $(\overline{\text{lin}}M)^\perp \subset (\text{lin}M)^\perp \subset M^\perp$. It remains to show that $M^\perp \subset (\overline{\text{lin}}M)^\perp$. To that end, let $x \in M^\perp$. Then $\langle x, y\rangle = 0$ for all $y \in M$, and consequently $\langle x, y\rangle = 0$ for all $y \in \text{lin}M$. If $z \in \overline{\text{lin}}M$, then there exists a sequence (z_n) in $\text{lin}M$ such that $z_n \to z$ as $n \to \infty$. Thus,

$$\langle x, z\rangle = \lim_n \langle x, z_n\rangle = 0,$$

whence $x \in (\overline{\text{lin}}M)^\perp$. ■

3.3.3 Examples

Let $X = \mathbb{R}^3$. The vectors $(-3, 0, 2)$ and $(4, 1, 6)$ are orthogonal since

$$\langle(-3, 0, 2), (4, 1, 6)\rangle = (-3)(4) + 0(1) + (2)(6) = 0.$$ ■

If $M = \ell_0$, the linear subspace of ℓ_2 consisting of all scalar sequences $(x_i)_1^\infty$ with only a finite number of nonzero terms, then $M^\perp = \{0\}$. Indeed, suppose that $y = (y_i)_{i=1}^\infty \in M^\perp$. Let

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

and $e_n = (\delta_{nj})_{n,j=1}^\infty$. Then $e_n \in M$ for each $n \in \mathbb{N}$, and hence,

$$0 = \langle y, e_i \rangle = \sum_{j=1}^\infty y_j \overline{\delta_{ij}} = y_i \quad \text{for all} \quad i = 1, 2, \ldots.$$

That is, $y = 0$, whence $M^\perp = \{0\}$.

3.3.1 Theorem

(Pythagoras). *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over a field $\mathbb{F}$ and let $x, y \in X$.*

[1] If $\mathbb{F} = \mathbb{R}$, then $x \perp y$ if and only if

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

[2] If $\mathbb{F} = \mathbb{C}$, then $x \perp y$ if and only if

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 \text{ and } \|x + iy\|^2 = \|x\|^2 + \|y\|^2.$$

Proof. [1] "$\Rightarrow$". If $x \perp y$, then

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2.$$

" $\Leftarrow$". Suppose that $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Then

$$\begin{aligned} \langle x + y, x + y \rangle &= \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle &= \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad 2\langle x, y \rangle = 0 \quad &\Rightarrow \quad \langle x, y \rangle = 0. \end{aligned}$$

[2] "$\Rightarrow$". Assume that $x \perp y$. Then

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

"$\Leftarrow$". Assume that $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ and $\|x + iy\|^2 = \|x\|^2 + \|y\|^2$. Then

$$\begin{aligned} \langle x + y, x + y \rangle &= \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle &= \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad \langle x, y \rangle + \langle y, x \rangle = 0 \quad \Rightarrow \quad 2\Re e\langle x, y \rangle &= 0 \quad \Rightarrow \quad \Re e\langle x, y \rangle = 0. \end{aligned}$$

Also,

$$\begin{aligned} &\langle x + yi, x + yi \rangle = \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad &\langle x, x \rangle - i\langle x, y \rangle + i\langle y, x \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle \\ \Rightarrow \quad &-i\langle x, y \rangle + i\langle y, x \rangle = 0 \\ \Rightarrow \quad &-i\,[\langle x, y \rangle - \langle y, x \rangle] = 0 \\ \Rightarrow \quad &-i\left[\langle x, y \rangle - \overline{\langle x, y \rangle}\right] = 0 \\ \Rightarrow \quad &-i\,[2i\Im m\langle x, y \rangle] = 0 \quad \Rightarrow \quad \Im m\langle x, y \rangle = 0. \end{aligned}$$

Since $\Re e\langle x, y \rangle = 0 = \Im m\langle x, y \rangle$, we have that $\langle x, y \rangle = 0$. ■

3.3.4 Corollary
If $M = \{x_1,\ x_2, \ldots, x_n\}$ is an orthogonal set in an inner product space $(X, \langle\cdot,\cdot\rangle)$ then

$$\left\|\sum_{i=1}^{n} x_i\right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

Proof. Exercise. ■

3.4 Best Approximation in Hilbert Spaces

3.4.1 Definition
Let K be a closed subset of an inner product space $(X, \langle\cdot,\cdot\rangle)$. For a given $x \in X \setminus K$, a **best approximation** *or* **nearest point** *to x from K is any element $y_0 \in K$ such that*

$$\|x - y_0\| \leq \|x - y\| \quad \text{for all} \quad y \in K.$$

Equivalently, $y_0 \in K$ is a best approximation to x from K if

$$\|x - y_0\| = \inf_{y \in K} \|x - y\| = d(x, K).$$

The (possibly empty) set of all best approximations to x from K is denoted by $P_K(x)$. That is,

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

The (generally set-valued) map P_K which associates each x in X with its best approximations in K is called the **metric projection** *or the* **nearest point map**. *The set K is called*

[1] **proximinal** *if each $x \in X$ has a best approximation in K; i.e., $P_K(x) \neq \emptyset$ for each $x \in X$;*

[2] **Chebyshev** *if each $x \in X$ has a unique best approximation in K; i.e., the set $P_K(x)$ consists of a single point.*

The following important result asserts that if K is a complete convex subset of an inner product space $(X, \langle\cdot,\cdot\rangle)$, then each $x \in X$ has one and only one element of best approximation in K.

3.4.1 Theorem
Every nonempty complete convex subset K of an inner product space $(X, \langle\cdot,\cdot\rangle)$ is a Chebyshev set.

Proof. Existence: Without loss of generality, $x \in X \setminus K$. Let

$$\delta = \inf_{y \in K} \|x - y\|.$$

By definition of the infimum, there exists a sequence $(y_n)_1^\infty$ in K such that

$$\|x - y_n\| \to \delta \quad \text{as} \quad n \to \infty.$$

We show that $(y_n)_1^\infty$ is a Cauchy sequence. By the Parallelogram Identity (Theorem 3.1.3),

$$\begin{aligned}
\|y_m - y_n\|^2 &= \|(x - y_n) - (x - y_m)\|^2 \\
&= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - (y_n + y_m)\|^2 \\
&= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \left(\frac{y_n + y_m}{2}\right)\right\|^2 \\
&\leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2,
\end{aligned}$$

since $\dfrac{y_n + y_m}{2} \in K$ by convexity of K. Thus,

$$\|y_m - y_n\|^2 \le 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2 \to 0 \quad \text{as} \quad n, m \to \infty.$$

That is, $(y_n)_1^\infty$ is a Cauchy sequence in K. Since K is complete, there exists $y \in K$ such that $y_n \to y$ as $n \to \infty$. Since the norm is continuous,

$$\|x - y\| = \|x - \lim_{n\to\infty} y_n\| = \|\lim_{n\to\infty}(x - y_n)\| = \lim_{n\to\infty}\|x - y_n\| = \delta.$$

Thus,

$$\|x - y\| = \delta = d(x, K).$$

Uniqueness: Assume that $y,\ y_0 \in K$ are two best approximations to x from K. That is,

$$\|x - y_0\| = \|x - y\| = \delta = d(x, K).$$

By the Parallelogram Identity,

$$\begin{aligned}
0 \le \|y - y_0\|^2 &= \|(y - x) + (x - y_0)\|^2 \\
&= 2\|x - y\|^2 + 2\|x - y_0\|^2 - \|2x - (y + y_0)\|^2 \\
&= 2\delta^2 + 2\delta^2 - 4\left\|x - \left(\frac{y + y_0}{2}\right)\right\|^2 \\
&\le 4\delta^2 - 4\delta^2 = 0.
\end{aligned}$$

Thus, $y_0 = y$. ■

3.4.2 Corollary
Every nonempty closed convex subset of a Hilbert space is Chebyshev.

The following theorem characterizes best approximations from a closed convex subset of a Hilbert space.

3.4.2 Theorem
Let K be a nonempty closed convex subset of a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$, $x \in \mathcal{H} \setminus K$ and $y_0 \in K$. Then y_0 is the best approximation to x from K if and only if

$$\Re\mathfrak{e}\langle x - y_0, y - y_0\rangle \le 0 \ \text{ for all } \ y \in K.$$

Proof. The existence and uniqueness of the best approximation to x in K are guaranteed by Theorem 3.4.1. Let y_0 be the best approximation to x in K. Then, for any $y \in K$ and any $0 < \lambda < 1$, $\lambda y + (1 - \lambda)y_0 \in K$ since K is convex. Thus,

$$\begin{aligned}
\|x - y_0\|^2 &\le \|x - [\lambda y + (1 - \lambda)y_0]\|^2 = \|(x - y_0) - \lambda(y - y_0)\|^2 \\
&= \langle x - y_0) - \lambda(y - y_0), x - y_0) - \lambda(y - y_0)\rangle \\
&= \langle x - y_0, x - y_0\rangle - \lambda[\langle x - y_0, y - y_0\rangle + \langle y - y_0, x - y_0\rangle] \\
&\quad +\lambda^2\langle y - y_0, y - y_0\rangle \\
&= \|x - y_0\|^2 - 2\lambda\Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) + \lambda^2\|y - y_0\|^2 \\
\Rightarrow \quad 2\lambda\Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) &\le \lambda^2\|y - y_0\|^2 \\
\Rightarrow \quad \Re\mathfrak{e}(\langle x - y_0, y - y_0\rangle) &\le \frac{\lambda}{2}\|y - y_0\|^2.
\end{aligned}$$

As $\lambda \to 0$, $\frac{\lambda}{2}\|y - y_0\|^2 \to 0$, and consequently $\mathfrak{Re}\langle x - y_0, y - y_0\rangle \leq 0$.

Conversely, assume that for each $y \in K$, $\mathfrak{Re}\langle x - y_0, y - y_0\rangle \leq 0$. Then, for any $y \in K$,

$$\begin{aligned}
\|x - y\|^2 &= \|(x - y_0) - (y - y_0)\|^2 \\
&= \langle (x - y_0) - (y - y_0), (x - y_0) - (y - y_0)\rangle \\
&= \langle x - y_0, x - y_0\rangle - \langle x - y_0, y - y_0\rangle - \langle y - y_0, x - y_0\rangle + \langle y - y_0, y - y_0\rangle \\
&= \langle x - y_0, x - y_0\rangle - [\langle x - y_0, y - y_0\rangle + \langle y - y_0, x - y_0\rangle] + \langle y - y_0, y - y_0\rangle \\
&= \langle x - y_0, x - y_0\rangle - [\langle x - y_0, y - y_0\rangle + \overline{\langle x - y_0, y - y_0\rangle}] + \langle y - y_0, y - y_0\rangle \\
&= \|x - y_0\|^2 - 2\mathfrak{Re}\langle x - y_0, y - y_0\rangle + \|y - y_0\|^2 \\
&\geq \|x - y_0\|^2.
\end{aligned}$$

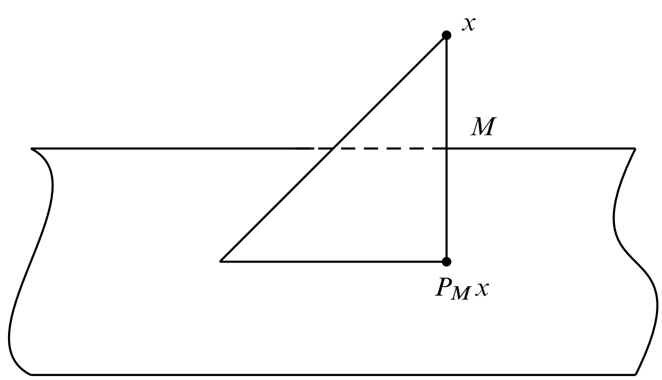
Taking the positive square root both sides, we have that $\|x - y_0\| \leq \|x - y\|$ for all $y \in K$. ■

As a corollary to Theorem 3.4.2, one gets the following characterization of best approximations from a closed linear subspace of a Hilbert space.

3.4.3 Corollary

(Characterization of Best Approximations from closed subspaces). *Let M be a closed subspace of a Hilbert space $\mathcal{H}$ and let $x \in \mathcal{H} \setminus M$. Then an element $y_0 \in M$ is the best approximation to x from M if and only if $\langle x - y_0, y\rangle = 0$ for all $y \in M$ (i.e., $x - y_0 \in M^{\perp}$).*

Corollary 3.4.3 says that if M is a closed linear subspace of a Hilbert space $\mathcal{H}$, then $y_0 = P_M(x)$ (i.e., y_0 is the best approximation to x from M) if and only if $x - P_M(x) \perp M$. That is, the unique best approximation is obtained by "dropping the perpendicular from x onto M". It is for this reason that the map $P_M : x \to P_M(x)$ is also called the **orthogonal projection** of $\mathcal{H}$ onto M.



3.4.4 Example

Let $X = \mathcal{C}_2[-1, 1]$, $M = \mathbb{P}_2 =$ lin$\{1, t, t^2\}$, and $x(t) = t^3$. Find $P_M(x)$.

Solution. Note that $\mathcal{C}_2[-1, 1]$ is an *incomplete* inner product space. Since M is finite-dimensional, it is complete, and consequently proximinal in $\mathcal{C}_2[-1, 1]$. Uniqueness of best approximations follows from the Parallelogram Identity.

Let $y_0 = \sum_{i=0}^{2} \alpha_i t^i \in M$. By Corollary 3.4.3,

$$
\begin{aligned}
y_0 = P_M(x) \quad &\iff \quad x - y_0 \in M^\perp \\
&\iff \quad \langle x - y_0, t^j \rangle = 0 \quad \text{for all } j = 0,1,2 \\
&\iff \quad \left\langle t^3 - \sum_{i=0}^{2} \alpha_i t^i, t^j \right\rangle = 0 \quad \text{for all } j = 0,1,2 \\
&\iff \quad \sum_{i=0}^{2} \alpha_i \langle t^i, t^j \rangle = \langle t^3, t^j, \quad \rangle \text{ for all } j = 0,1,2 \\
&\iff \quad \sum_{i=0}^{2} \alpha_i \int_{-1}^{1} t^i \cdot t^j \, dt = \int_{-1}^{1} t^3 \cdot t^j \, dt \quad \text{for all } j = 0,1,2 \\
&\iff \quad \sum_{i=0}^{2} \alpha_i \int_{-1}^{1} t^{i+j} \, dt = \int_{-1}^{1} t^{3+j} \, dt \quad \text{for all } j = 0,1,2 \\
&\iff \quad \sum_{i=0}^{2} \alpha_i \left. \frac{t^{i+j+1}}{i+j+1} \right]_{-1}^{1} = \left. \frac{t^{j+4}}{j+4} \right]_{-1}^{1} \quad \text{for all } j = 0,1,2 \\
&\iff \quad \sum_{i=0}^{2} \alpha_i \frac{1}{i+j+1} \left[1 - (-1)^{i+j+1} \right] = \frac{1}{j+4} \left[1 - (-1)^{j+4} \right] \\
&\qquad\qquad \text{for all } j = 0,1,2 \\
&\iff \quad \begin{cases} 2\,\alpha_0 + 0\,\alpha_1 + \frac{2}{3}\,\alpha_2 &= 0 \\ 0\,\alpha_0 + \frac{2}{3}\,\alpha_1 + 0\,\alpha_2 &= \frac{2}{5} \\ \frac{2}{3}\,\alpha_0 + 0\,\alpha_1 + \frac{2}{5}\,\alpha_2 &= 0 \end{cases} \\
&\iff \quad \alpha_0 = 0, \qquad \alpha_1 = \frac{3}{5}, \qquad \alpha_2 = 0.
\end{aligned}
$$

Thus, $P_M(x) = y_0 = \frac{3}{5} t$. ■

3.4.3 Theorem

(Projection Theorem). *Let $\mathcal{H}$ be a Hilbert space, M a closed subspace of $\mathcal{H}$. Then*

[1] $\mathcal{H} = M \oplus M^\perp$. That is, each $x \in \mathcal{H}$ can be uniquely decomposed in the form

$$x = y + z \quad \textit{with} \quad y \in M \quad \textit{and}\ z \in M^\perp.$$

[2] $M = M^{\perp\perp}$.

Proof.

[1] If $x \in M$, then $x = x + 0$, and we are done. Assume that $x \notin M$. Let $y = P_M(x)$ be the unique best approximation to x from M as advertised in Theorem 3.4.1. Then $z = x - P_M(x) \in M^\perp$, and

$$x = P_M(x) + (x - P_M(x)) = y + z$$

is the unique representation of x as a sum of an element of M and an element of $M^\perp$.

[2] Since the containment $M \subset M^{\perp\perp}$ is clear, we only show that $M^{\perp\perp} \subset M$. To that end, let $x \in M^{\perp\perp}$. Then by [1] above

$$x = y + z, \quad \text{where } \; y \in M \quad \text{and } \; z \in M^{\perp}.$$

Since $M \subset M^{\perp\perp}$ and $M^{\perp\perp}$ is a subspace, $z = x - y \in M^{\perp\perp}$. But $z \in M^{\perp}$ implies that $z \in M^{\perp} \cap M^{\perp\perp}$ which, in turn, implies that $z = 0$. Thus, $x = y \in M$.

■

3.4.5 Corollary
If M is a closed subspace of a Hilbert space $\mathcal{H}$, and if $M \neq \mathcal{H}$, then there exists $z \in \mathcal{H} \setminus \{0\}$ such that $z \perp M$.

Proof. Let $x \in \mathcal{H} \setminus M$. Then by the Projection Theorem,

$$x = y + z, \text{ where } \; y \in M \text{ and } \; z \in M^{\perp}.$$

Hence $z \neq 0$ and $z \perp M$. ■

3.4.6 Proposition
Let S be a nonempty subset of a Hilbert space $\mathcal{H}$. Then

[1] $S^{\perp\perp} = \overline{\text{lin}} S$.

[2] $S^{\perp} = \{0\}$ if and only if $\overline{\text{lin}} S = \mathcal{H}$.

Proof.

[1] Since $S^{\perp} = (\overline{\text{lin}} S)^{\perp}$ by Proposition 3.3.2, we have, by the Projection Theorem, that

$$\overline{\text{lin}} S = (\overline{\text{lin}} S)^{\perp\perp} = S^{\perp\perp}.$$

[2] If $S^{\perp} = \{0\}$, then by [1]

$$\overline{\text{lin}} S = S^{\perp\perp} = \{0\}^{\perp} = \mathcal{H}.$$

On the other hand, if $\mathcal{H} = \overline{\text{lin}} S$, then $\mathcal{H} = S^{\perp\perp}$ by [1], and so

$$S^{\perp} = S^{\perp\perp\perp} = \mathcal{H}^{\perp} = \{0\}.$$

■

3.5 Orthonormal Sets and Orthonormal Bases

In this section we extend to Hilbert spaces the finite-dimensional concept of an orthonormal basis.

3.5.1 Definition
Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space over $\mathbb{F}$. A set $S = \{x_\alpha : \alpha \in \Lambda\}$ of elements of X is called an **orthonormal set** *if*

(a) $\langle x_\alpha, x_\beta \rangle = 0$ for all $\alpha \neq \beta$ (i.e., S is an orthogonal set), and

(b) $\|x_\alpha\| = 1$ for all $\alpha \in \Lambda$.

If $S = \{x_\alpha : \alpha \in \Lambda\}$ is an orthonormal set and $x \in X$, then the numbers $\langle x, x_\alpha \rangle$ are called the **Fourier coefficients** *of x with respect to S and the formal series $\sum_{\alpha \in \Lambda} \langle x, x_\alpha \rangle x_\alpha$ the* **Fourier series** *of x.*

3.5.1 Theorem
An orthonormal set S in a separable inner product space $(X, \langle\cdot,\cdot\rangle)$ is at most countable.

Proof. If S is finite, then there is nothing to prove. Assume that S is infinite. Observe that if $x, y \in S$, then $\|x - y\| = \sqrt{2}$ (since x and y are orthonormal). Let $D = \{y_n \mid n \in \mathbb{N}\}$ be a countable dense subset of X. Then to each $x \in S$ corresponds an element $y_n \in D$ such that $\|x - y_n\| < \frac{\sqrt{2}}{4}$. This defines a map $f : S \to \mathbb{N}$ given by $f(x) = n$, where n corresponds to the y_n as indicated above. Now, if x and y are distinct elements of S, then there are distinct elements y_n and y_m in D such that

$$\|x - y_n\| < \frac{\sqrt{2}}{4} \text{ and } \|y - y_m\| < \frac{\sqrt{2}}{4}.$$

Hence,

$$\sqrt{2} = \|x - y\| \leq \|x - y_n\| + \|y_n - y_m\| + \|y_m - y\| < \frac{\sqrt{2}}{2} + \|y_n - y_m\| \iff \frac{\sqrt{2}}{2} < \|y_n - y_m\|,$$

and so $y_n \neq y_m$. In particular, $n \neq m$. Thus, we have a one-to-one correspondence between the elements of S and a subset of $\mathbb{N}$. ■

3.5.2 Definition
An orthonormal set S in an inner product space $(X, \langle\cdot,\cdot\rangle)$ is said to be **complete** *in X if $S \subset T$ and T is an orthonormal set in X, then $S = T$.*

Simply put, a complete orthonormal set S in an inner product space is an orthonormal set that is not properly contained in any other orthonormal set in X; in other words, S is complete if it is a maximal orthonormal set in X.

It is easy exercise to show that a set S is complete in an inner product $(X, \langle\cdot,\cdot\rangle)$ if and only if $S^{\perp} = \{0\}$.

3.5.3 Examples

[1] In $\mathbb{R}^3$ the set $S = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is orthonormal.

[2] In ℓ_2, let $S = \{e_n : n \in \mathbb{N}\}$, where $e_n = (\delta_{1n}, \delta_{2n}, \ldots)$ with

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Then S is an orthonormal set. Furthermore, for each $x = (x_i)_{i=1}^{\infty} \in \ell_2$, $\langle x, e_n\rangle = x_n$ for all n. Thus

$$\langle x, e_n\rangle = 0 \text{ for all } n \iff x_n = 0 \text{ for all } n \iff x = 0.$$

That is, $S^{\perp} = \{0\}$, hence, S is complete in ℓ_2.

3.5.2 Theorem
Let $(X, \langle\cdot,\cdot\rangle)$ be a separable inner product space over $\mathbb{F}$.

[1] (Best Fit). *If $\{x_1, x_2, \ldots, x_n\}$ is a finite orthonormal set in X and $M = lin\{x_1, x_2, \ldots, x_n\}$, then for each $x \in X$ there exists $y_0 \in M$ such that*

$$\|x - y_0\| = d(x, M).$$

In fact, $y_0 = \sum_{k=1}^{n} \langle x, x_k\rangle x_k$.

[2] (Bessel's Inequality). *Let $(x_n)_{n=1}^{\infty}$ be an orthonormal sequence in X. Then for any $x \in X$,*

$$\sum_{k=1}^{\infty} |\langle x, x_k\rangle|^2 \le \|x\|^2.$$

In particular, $\langle x, x_k\rangle \to 0$ as $k \to \infty$.

Proof.

[1] For any choice of scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$,

$$\begin{aligned}
\left\| x - \sum_{k=1}^{n} \lambda_k x_k \right\|^2 &= \left\langle x - \sum_{i=1}^{n} \lambda_i x_i, x - \sum_{j=1}^{n} \lambda_j x_j \right\rangle \\
&= \|x\|^2 - \sum_{i=1}^{n} \lambda_i \langle x_i, x\rangle - \sum_{j=1}^{n} \overline{\lambda_j} \langle x, x_j\rangle + \sum_{i=1}^{n} \lambda_i \overline{\lambda_i} \\
&= \|x\|^2 - \sum_{i=1}^{n} \lambda_i \overline{\langle x, x_i\rangle} - \sum_{j=1}^{n} \overline{\lambda_j} \langle x, x_j\rangle + \sum_{i=1}^{n} \lambda_i \overline{\lambda_i} \\
&= \|x\|^2 + \sum_{i=1}^{n} [\lambda_i \overline{\lambda_i} - \lambda_i \overline{\langle x, x_i\rangle} - \overline{\lambda_i} \langle x, x_i\rangle + \langle x, x_i\rangle \overline{\langle x, x_i\rangle}\,] \\
&\qquad - \sum_{i=1}^{n} \langle x, x_i\rangle \overline{\langle x, x_i\rangle} \\
&= \|x\|^2 + \sum_{i=1}^{n} [\,(\lambda_i - \langle x, x_i\rangle)(\overline{\lambda_i} - \overline{\langle x, x_i\rangle})\,] - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2 \\
&= \|x\|^2 + \sum_{i=1}^{n} [\,(\lambda_i - \langle x, x_i\rangle)\overline{(\lambda_i - \langle x, x_i\rangle)}\,] - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2 \\
&= \|x\|^2 - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2 + \sum_{i=1}^{n} |\lambda_i - \langle x, x_i\rangle|^2.
\end{aligned}$$

Therefore, $\left\| x - \sum_{k=1}^{n} \lambda_k x_k \right\|^2$ is minimal if and only if $\lambda_k = \langle x, x_k\rangle$ for each $k = 1,\ 2,\ \ldots,\ n$.

[2] For each positive integer n, and with $\lambda_k = \langle x, x_k\rangle$, the above argument shows that

$$0 \le \left\| x - \sum_{k=1}^{n} \lambda_k x_k \right\|^2 = \|x\|^2 - \sum_{i=1}^{n} |\langle x, x_i\rangle|^2.$$

Thus,

$$\sum_{k=1}^{n} |\langle x, x_k\rangle|^2 \le \|x\|^2.$$

Taking the limit as $n \to \infty$, we get

$$\sum_{k=1}^{\infty} |\langle x, x_k\rangle|^2 \le \|x\|^2. \qquad \blacksquare$$

3.5.3 Theorem
(Riesz-Fischer Theorem). *Let $(x_n)_1^\infty$ be an orthonormal sequence in a separable Hilbert space $\mathcal{H}$ and let $(c_n)_1^\infty$ be a sequence of scalars. Then the series $\sum_{k=1}^{\infty} c_k x_k$ converges in $\mathcal{H}$ if and only if $c = (c_n)_1^\infty \in \ell_2$. In this case,*

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\| = \left(\sum_{k=1}^{\infty} |c_k|^2 \right)^{\frac{1}{2}}.$$

Proof. Assume that the series $\sum_{k=1}^{\infty} c_k x_k$ converges to x. Then for each $j, n \in \mathbb{N}$,

$$\left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \sum_{k=1}^{n} c_k \langle x_k, x_j \rangle = c_j.$$

Using continuity of the inner product

$$\langle x, x_j \rangle = \left\langle \sum_{k=1}^{\infty} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} c_j = c_j.$$

By Bessel's Inequality, we have that

$$\sum_{k=1}^{\infty} |c_k|^2 = \sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \le \|x\|^2 < \infty.$$

That is, $c = (c_n)_1^\infty \in \ell_2$.

Conversely, assume that $c = (c_n)_1^\infty \in \ell_2$. Set $z_n = \sum_{k=1}^{n} c_k x_k$. Then for $1 \le n \le m$,

$$\|z_n - z_m\|^2 = \left\| \sum_{k=n+1}^{m} c_k x_k \right\|^2 = \sum_{k=n+1}^{m} |c_k|^2 \to 0 \quad \text{as } n \to \infty.$$

Hence, $(z_n)_1^\infty$ is a Cauchy sequence in $\mathcal{H}$. Since $\mathcal{H}$ is complete the sequence $(z_n)_1^\infty$ converges to some $x \in \mathcal{H}$. Hence the series $\sum_{k=1}^{\infty} c_k x_k$ converges to some element in $\mathcal{H}$.

Also,

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\|^2 = \lim_{n\to\infty} \left\| \sum_{k=1}^{n} c_k x_k \right\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} |c_k|^2,$$

whence,

$$\left\| \sum_{k=1}^{\infty} c_k x_k \right\| = \left(\sum_{k=1}^{\infty} |c_k|^2 \right)^{\frac{1}{2}}.$$

Note that Bessel's Inequality says that

$$\sum_{k=1}^{\infty} |\langle x, x_k \rangle|^2 \le \|x\|^2 < \infty.$$

That is, $(\langle x, x_n\rangle)_1^\infty \in \ell_2$. Hence, by Theorem 3.5.3, the series $\sum_{k=1}^{\infty}\langle x, x_k\rangle x_k$ converges. There is however no reason why this series should converge to x. In fact, the following example shows that this series may *not* converge to x.

3.5.4 Example
Let $(e_n) \in \ell_2$, where $e_n = (\delta_{1n}, \delta_{2n}, \ldots)$ with

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

For each $n \in \mathbb{N}$, let $f_n = e_{n+1}$. Then $(f_n)_{n=1}^\infty$ is an orthonormal sequence in ℓ_2. For any $x = (x_n)_1^\infty \in \ell_2$,

$$\sum_{k=1}^{\infty}\langle x, f_k\rangle f_k = \sum_{k=1}^{\infty}\langle x, e_{k+1}\rangle e_{k+1} = (0, x_2, x_3, \ldots) \neq (x_1, x_2, x_3, \ldots) = x.$$

3.5.5 Definition
Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{F}$. An orthonormal set $\{x_n\}$ is called an **orthonormal basis** *for X if for each $x \in X$,*

$$x = \sum_{k=1}^{\infty}\langle x, x_k\rangle x_k.$$

That is, the sequence of partial sums (s_n), where $s_n = \sum_{k=1}^{n}\langle x, x_k\rangle x_k$, converges to x.

3.5.4 Theorem
Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space and assume that $S = \{x_n\}$ is an orthonormal set in $\mathcal{H}$. Then the following statements are equivalent:

[1] S is complete in $\mathcal{H}$; i.e., $S^\perp = \{0\}$.

[2] $\overline{\text{lin}}S = \mathcal{H}$; i.e., the linear span of S is norm-dense in $\mathcal{H}$.

[3] (Fourier Series Expansion.) *For any $x \in \mathcal{H}$, we have*

$$x = \sum_{i=1}^{\infty}\langle x, x_i\rangle x_i.$$

That is, S is an orthonormal basis for $\mathcal{H}$.

[4] (Parseval's Identity.) *For all $x, y \in \mathcal{H}$,*

$$\langle x, y\rangle = \sum_{k=1}^{\infty}\langle x, x_k\rangle\overline{\langle y, x_k\rangle}.$$

[5] For any $x \in \mathcal{H}$,

$$\|x\|^2 = \sum_{k=1}^{\infty}|\langle x, x_k\rangle|^2.$$

Proof. "[1] $\iff$ [2]". This equivalence was proved in Proposition 3.4.6[2].

"[1] $\Rightarrow$ [3]". Let $x \in \mathcal{H}$ and $s_n = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$. Then for all $n > m$,

$$\|s_n - s_m\|^2 = \left\| \sum_{i=m+1}^{n} \langle x, x_i \rangle x_i \right\|^2 = \sum_{m+1}^{n} |\langle x, x_i \rangle|^2 \le \|x\|^2.$$

Thus, (s_n) is a Cauchy sequence in $\mathcal{H}$. Since $\mathcal{H}$ is complete, this sequence converges to some element which we denote by $\sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$. We show that $x = \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i$. Indeed, for each fixed $j \in \mathbb{N}$,

$$\begin{aligned}
\left\langle x - \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i, x_j \right\rangle &= \left\langle x - \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, x_i \rangle x_i, x_j \right\rangle \\
&= \lim_{n\to\infty} \left\langle x - \sum_{i=1}^{n} \langle x, x_i \rangle x_i, x_j \right\rangle \\
&= \lim_{n\to\infty} \left(\langle x, x_j \rangle - \sum_{i=1}^{n} \langle x, x_i \rangle \langle x_i, x_j \rangle \right) \\
&= \lim_{n\to\infty} (\langle x, x_j \rangle - \langle x, x_j \rangle) = 0.
\end{aligned}$$

Thus, by [1], $x - \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i = 0$, whence

$$x = \sum_{i=1}^{\infty} \langle x, x_i \rangle x_i.$$

"[3] $\Rightarrow$ [4]". Let $x, y \in \mathcal{H}$. Then

$$\begin{aligned}
\langle x, y \rangle &= \lim_{n\to\infty} \left\langle \sum_{i=1}^{n} \langle x, x_i \rangle x_i, \sum_{j=1}^{n} \langle y, x_j \rangle x_j \right\rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x, x_i \rangle \overline{\langle y, x_j \rangle} \langle x_i, x_j \rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \langle x, x_i \rangle \overline{\langle y, x_i \rangle} = \sum_{i=1}^{\infty} \langle x, x_i \rangle \overline{\langle y, x_i \rangle}.
\end{aligned}$$

"[4] $\Rightarrow$ [5]". Take $x = y$ in [4].

"[5] $\Rightarrow$ [1]". Since $\|x\|^2 = \sum_{k} \langle x, x_k \rangle \overline{\langle x, x_k \rangle}$, if $x \perp S$ then $\langle x, x_k \rangle = 0$ for all k. Thus, $\|x\|^2 = 0$, whence $x = 0$. That is, $S^{\perp} = \{0\}$. ■

3.5.6 Examples

[1] In ℓ_2, the set $S = \{e_n : n \in \mathbb{N}\}$, where $e_n = (\delta_{1n}, \delta_{2n}, \ldots)$ with

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

is an orthonormal basis for ℓ_2. ■

[2] In $L_2[-\pi,\pi]$, the set $\left\{\frac{1}{\sqrt{2\pi}}e^{int} \;:\; n \in \mathcal{Z}\right\}$ is an orthonormal basis for the *complex* $L_2[-\pi,\pi]$.

[3] The set $S = \left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nt}{\sqrt{\pi}}, \frac{\sin nt}{\sqrt{\pi}}\right\}_{n=1}^{\infty}$ is an orthonormal basis for the *real* $L_2[-\pi,\pi]$.

Hence, if $x \in L_2[-\pi,\pi]$, then by Theorem 3.5.4[3],

$$\begin{aligned}
x(t) &= \left\langle x(t), \frac{1}{\sqrt{2\pi}}\right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{n=1}^{\infty}\left[\left\langle x(t), \frac{\cos nt}{\sqrt{\pi}}\right\rangle \frac{\cos nt}{\sqrt{\pi}} + \left\langle x(t), \frac{\sin nt}{\sqrt{\pi}}\right\rangle \frac{\sin nt}{\sqrt{\pi}}\right] \\
&= \frac{1}{2\pi}\langle x(t), 1\rangle + \sum_{n=1}^{\infty}\left[\frac{1}{\pi}\langle x(t), \cos nt\rangle \cos nt + \frac{1}{\pi}\langle x(t), \sin nt\rangle \sin nt\right] \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\,dt \\
&\quad + \sum_{n=1}^{\infty}\left[\left(\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos nt\,dt\right)\cos nt + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin nt\,dt\right)\sin nt\right] \\
&= a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt),
\end{aligned}$$

where

$$\left.\begin{aligned}
a_0 &= \tfrac{1}{\sqrt{2\pi}}\left\langle x(t), \tfrac{1}{\sqrt{2\pi}}\right\rangle = \tfrac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\,dt, \\
a_n &= \tfrac{1}{\sqrt{\pi}}\left\langle x(t), \tfrac{\cos nt}{\sqrt{\pi}}\right\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos nt\,dt, \quad \text{and} \\
b_n &= \tfrac{1}{\sqrt{\pi}}\left\langle x(t), \tfrac{\sin nt}{\sqrt{\pi}}\right\rangle = \tfrac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin nt\,dt
\end{aligned}\right\} \quad n = 1, 2, \ldots.$$

That is, the Fourier series expansion of x is

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt). \tag{3.5.6.1}$$

It is clear from above that for all $n = 1, 2, \ldots$,

$$2\pi|a_0|^2 = \left|\left\langle x(t), \frac{1}{\sqrt{2\pi}}\right\rangle\right|^2, \quad \pi|a_n|^2 = \left|\left\langle x(t), \frac{\cos nt}{\sqrt{\pi}}\right\rangle\right|^2, \text{ and}$$

$$\pi|b_n|^2 = \left|\left\langle x(t), \frac{\sin nt}{\sqrt{\pi}}\right\rangle\right|^2.$$

By Theorem 3.5.4 [5] we have that

$$\begin{aligned}
\int_{-\pi}^{\pi} |x(t)|^2 dt = \|x\|^2 &= \left|\left\langle x(t), \frac{1}{\sqrt{2\pi}}\right\rangle\right|^2 \\
&+ \sum_{n=1}^{\infty}\left(\left|\left\langle x(t), \frac{\cos\ nt}{\sqrt{\pi}}\right\rangle\right|^2 + \left|\left\langle x(t), \frac{\sin\ nt}{\sqrt{\pi}}\right\rangle\right|^2\right) \\
&= 2\pi|a_0|^2 + \sum_{n=1}^{\infty}(\pi|a_n|^2 + \pi|b_n|^2) \\
&= \pi\left(2|a_0|^2 + \sum_{n=1}^{\infty}(|a_n|^2 + |b_n|^2)\right).
\end{aligned}$$

We now apply the above results to a particular function: Let $x(t) = t$. Then

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} t\ dt = 0 \quad \text{since } x(t) = t \text{ is an odd function.}$$

For $n = 1, 2, \ldots,$

$$\begin{aligned}
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos nt\ dt = 0 \quad \text{since } t\cos nt \text{ is an odd function,} \\
b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin nt\ dt = \frac{2}{\pi}\int_{0}^{\pi} t\sin nt\ dt \\
&= \frac{2}{\pi}\left[\left.\frac{-t\cos nt}{n}\right|_0^{\pi} + \frac{1}{n}\int_0^{\pi}\cos nt\ dt\right] \\
&= \frac{2}{\pi}\left[\frac{-\pi}{n}\cos n\pi\right] = \frac{2(-1)^{n+1}}{n}.
\end{aligned}$$

Hence, by Theorem 3.5.4[3],

$$x(t) = \sum_{n=1}^{\infty}\frac{2(-1)^{n+1}}{n}\ \sin nt = \sum_{n=1}^{\infty}\frac{2(-1)^{n+1}\sqrt{\pi}}{n}\ \frac{\sin nt}{\sqrt{\pi}}.$$

It now follows that

$$\frac{2(-1)^{n+1}\sqrt{\pi}}{n} = \left\langle t, \frac{\sin nt}{\sqrt{\pi}}\right\rangle.$$

Now,

$$\|x\|_2^2 = \int_{-\pi}^{\pi} t^2\ dt = 2\int_0^{\pi} t^2\ dt = \left.\frac{2}{3}t^3\right|_0^{\pi} = \frac{2\pi^3}{3}.$$

Also, by Theorem 3.5.4[5],

$$\|x\|_2^2 = \sum_{n=1}^{\infty}\left|\left\langle t, \frac{\sin nt}{\sqrt{\pi}}\right\rangle\right|^2 = \sum_{n=1}^{\infty}\left|\frac{2(-1)^{n+1}\sqrt{\pi}}{n}\right|^2 = \sum_{n=1}^{\infty}\frac{4\pi}{n^2}.$$

Thus,

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \qquad \square$$

We can express the Fourier Series Expansion (3.5.6.1) of $x \in L_2[-\pi, \pi]$ in exponential form. Recall that

$$e^{i\theta} = \cos\theta + i \sin\theta \text{ (Euler's Formula).}$$

Therefore

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ and } \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Equation (3.5.6.1) now becomes

$$\begin{aligned} x(t) &= a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt) \\ &= a_0 + \sum_{n=1}^{\infty}\left[a_n\left(\frac{e^{int} + e^{-int}}{2}\right) + b_n\left(\frac{e^{int} - e^{-int}}{2i}\right)\right] \\ &= a_0 + \sum_{n=1}^{\infty}\left[\left(\frac{a_n - ib_n}{2}\right)e^{int} + \left(\frac{a_n + ib_n}{2}\right)e^{-int}\right] \\ &= a_0 + \sum_{n=1}^{\infty}\left(\frac{a_n - ib_n}{2}\right)e^{int} + \sum_{n=1}^{\infty}\left(\frac{a_n + ib_n}{2}\right)e^{-int}. \qquad (3.5.6.2) \end{aligned}$$

For each $n = 1, 2, 3, \ldots$, let $c_n = \frac{1}{2}(a_n + ib_n)$. Then $\overline{c_n} = \frac{1}{2}(a_n - ib_n)$ for each $n = 1, 2, 3, \ldots$, and so equation (3.5.6.2) becomes

$$x(t) = a_0 + \sum_{n=1}^{\infty}\overline{c_n}\, e^{int} + \sum_{n=1}^{\infty} c_n e^{-int}. \qquad (3.5.6.3)$$

Re-index the first sum in (3.5.6.3) by letting $n = -k$. Then

$$x(t) = a_0 + \sum_{k=-1}^{-\infty}\overline{c_{-k}}\, e^{-ikt} + \sum_{n=1}^{\infty} c_n e^{-int}. \qquad (3.5.6.4)$$

For $n = -1, -2, -3, \ldots$, define

$$c_n = \overline{c_{-n}}$$

and let $c_0 = a_0$. The we can rewrite equation (3.5.6.4) as

$$x(t) = \sum_{-\infty}^{\infty} c_n e^{-int}. \qquad (3.5.6.5)$$

This is the complex exponential form of the Fourier Series of $x \in L_2[-\pi, \pi]$.

Note that,

$$c_0 = a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)dt$$

and for $n = 1, 2, 3, \ldots,$

$$\begin{aligned} c_n &= \frac{1}{2}(a_n + i b_n) = \frac{1}{2}\left[\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos nt\, dt + i\,\frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin nt\, dt\right] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\,(\cos nt + i\,\sin nt)\, dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}\, dt, \end{aligned}$$

and

$$c_{-n} = \overline{c_n} = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}\, dt} = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)\overline{e^{-int}}\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}\, dt.$$

Therefore, for all $n \in \mathbb{Z}$,

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{int}\, dt.$$

Now, for $n = 1, 2, 3, \ldots,$

$$\begin{aligned} |c_n|^2 &= c_n \cdot \overline{c_n} = \frac{1}{2}(a_n + i b_n)\cdot\frac{1}{2}(a_n - i b_n) \\ &= \frac{1}{4}\left(a_n^2 + b_n^2\right) = \frac{1}{4}\left(|a_n|^2 + |b_n|^2\right). \end{aligned}$$

Therefore

$$\sum_{n=1}^{\infty}|c_n|^2 = \frac{1}{4}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right). \tag{3.5.6.6}$$

Since for $n = 1, 2, 3, \ldots, c_{-n} = \overline{c_n}$, it follows that

$$|c_{-n}|^2 = c_{-n}\cdot\overline{c_{-n}} = \overline{c_n}\cdot\overline{\overline{c_n}} = \overline{c_n}\cdot c_n = |c_n|^2.$$

Hence, for $n = 1, 2, 3, \ldots,$

$$\sum_{n=1}^{\infty}|c_{-n}|^2 = \sum_{n=1}^{\infty}|c_n|^2 = \frac{1}{4}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right). \tag{3.5.6.7}$$

From (3.5.6.6) and (3.5.6.7), we have that

$$\begin{aligned} \sum_{n\in\mathbb{Z}}|c_n|^2 &= \sum_{n=1}^{\infty}|c_{-n}|^2 + |c_0|^2 + \sum_{n=1}^{\infty}|c_n|^2 \\ &= \frac{1}{4}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right) + |a_0|^2 + \frac{1}{4}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right) \\ &= |a_0|^2 + \frac{1}{2}\sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right) \\ &= \frac{1}{2\pi}\cdot\pi\left(2\,|a_0|^2 + \sum_{n=1}^{\infty}\left(|a_n|^2 + |b_n|^2\right)\right) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|x(t)|^2\, dt. \end{aligned}$$

That is,

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |x(t)|^2 \, dt. \tag{3.5.6.8}$$

Let $\{x_1,\ x_2,\ \ldots,\ x_n\}$ be a basis of an n-dimensional linear subspace M of an inner product space $(X, \langle\cdot,\cdot\rangle)$. We have seen in Theorem 3.5.2 that if the set $\{x_1,\ x_2,\ \ldots,\ x_n\}$ is orthonormal, then the orthogonal projection (=best approximation) of any $x \in X$ onto M is given by

$$P_M(x) = \sum_{k=1}^{n} \langle x, x_k\rangle x_k.$$

It is clearly easy to compute orthogonal projections from a linear subspace that has an orthonormal basis: the coefficients in the orthogonal projection of $x \in X$ are just the "Fourier coefficients" of x. If the basis of M is not orthogonal, it may be advantageous to find an orthonormal basis for M and express the orthogonal projection as a linear combination of the new orthonormal basis. The process of finding an orthonormal basis from a given (non-orthonormal) basis is known as the Gram-Schmidt Orthonormalisation Procedure.

3.5.5 Theorem
(Gram-Schmidt Orthonormalisation Procedure). *If $\{x_k\}_1^\infty$ is a linearly independent set in an inner product space $(X, \langle\cdot,\cdot\rangle)$ then there exists an orthonormal set $\{e_k\}_1^\infty$ in X such that*

$$\text{lin}\{x_1, x_2, \ldots, x_n\} = \text{lin}\{e_1, e_2, \ldots, e_n\} \quad \text{for all} \quad n.$$

Proof. Set $e_1 = \dfrac{x_1}{\|x_1\|}$. Then $\text{lin}\{x_1\} = \text{lin}\{e_1\}$. Next, let $y_2 = x_2 - \langle x_2, e_1\rangle e_1$. Then

$$\langle y_2, e_1\rangle = \langle x_2 - \langle x_2, e_1\rangle e_1, e_1\rangle = \langle x_2, e_1\rangle - \langle x_2, e_1\rangle\langle e_1, e_1\rangle = \langle x_2, e_1\rangle - \langle x_2, e_1\rangle = 0.$$

That is, $e_1 \perp y_2$. Set $e_2 = \dfrac{y_2}{\|y_2\|}$. Then $\{e_1, e_2\}$ is an orthonormal set with the property that $\text{lin}\{x_1, x_2\} = \text{lin}\{e_1, e_2\}$. In general, for each $k = 2, 3, \ldots$, we let

$$y_k = x_k - \sum_{i=1}^{k-1} \langle x_k, e_i\rangle e_i.$$

Then for $k = 2, 3, \ldots$

$$\langle y_k, e_1\rangle = \langle y_k, e_2\rangle = \langle y_k, e_3\rangle = \cdots = \langle y_k, e_{k-1}\rangle = 0.$$

Set $e_k = \dfrac{y_k}{\|y_k\|}$. Then $\{e_1, e_2, \ldots, e_k\}$ is an orthonormal set in X with the property that
$\text{lin}\{e_1, e_2, \ldots, e_k\} = \text{lin}\{x_1, x_2, \ldots, x_k\}$. ■

We have made the point that ℓ_2 is a Hilbert space. In this final part of this chapter we want to show that every separable infinite-dimensional Hilbert space "looks like" ℓ_2 in the sense defined below.

3.5.7 Definition
Two linear spaces X and Y over the same field $\mathbb{F}$ are said to be **isomorphic** *it there is a one-to-one map T from X onto Y such that for all $x_1,\ x_2 \in X$ and all $\alpha, \beta \in \mathbb{F}$,*

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2). \tag{3.5.7.1}$$

3.5.8 Remark
Any map that satisfies condition (3.5.7.1) of Definition 3.5.7 is called a linear operator. Chapter 4 is devoted to the study of such maps. Clearly, the linear structures of the two linear spaces X and Y are preserved under the map T.

3.5.9 Definition
Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be two normed linear spaces and $T : X \to Y$. Then T is called an **isometry** *if*

$$\|Tx\| = \|x\| \text{ for all } x \in X. \tag{3.5.9.1}$$

Simply put, an isometry is a map that preserves lengths.

3.5.10 Remark
It is implicit in the above definition that the norm on the left of equation (3.5.9.1) is in Y and that on the right is in X. In order to avoid possible confusion, we should perhaps have labelled the norms as $\|\cdot\|_X$ and $\|\cdot\|_Y$ for the norms in X and Y respectively. This notation is however cumbersome and will therefore be avoided.

Normed linear spaces that are isometrically isomorphic are essentially identical.

3.5.11 Lemma
Let $M = lin\{x_n\}$ be a linear subspace of X. Then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ which has the following properties:

(i) $lin\{x_{n_k}\} = M$;

(ii) $\{x_{n_k}\}$ is linearly independent.

Proof. We define the subsequence inductively as follows: Let x_{n_1} be the first nonzero element of the sequence $\{x_n\}$. Therefore $x_n = 0 \cdot x_{n_1}$ for all $n < n_1$. If there is an $\alpha \in \mathbb{F}$ such that $x_n = \alpha x_{n_1}$ for all $n > n_1$, then we are done. Otherwise, let x_{n_2} be the first element of the sequence $\{x_n\}_{n>n_1}$ that is not a multiple of x_{n_1}. Thus there is an $\alpha \in \mathbb{F}$ such that $x_n = \alpha x_{n_1} + 0 x_{n_2}$ for all $n < n_2$. If $x_n = \alpha x_{n_1} + \beta x_{n_2}$ for some $\alpha,\ \beta \in \mathbb{F}$ and all $n > n_2$ then we are done. Otherwise let x_{n_3} be the first element of the sequence $\{x_n\}$ which is not a linear combination of x_{n_1} and x_{n_2}. Then $x_n = \alpha x_{n_1} + \beta x_{n_2} + 0 x_{n_3}$ for all $n < n_3$. If $x_n = \alpha x_{n_1} + \beta x_{n_2} + \gamma x_{n_3}$ for all $n > n_3$, then we are done. Otherwise let x_{n_4} be the first element of the sequence $\{x_n\}$ that is not in $\text{lin}\{x_{n_1}, x_{n_2}, x_{n_3}\}$. Continue in this fashion to obtain elements $x_{n_1},\ x_{n_2},\ \ldots$. If $x \in \text{lin}\{x_1, x_2, \ldots, x_n\}$, then $x \in \text{lin}\{x_{n_1}, x_{n_2}, \ldots, x_{n_r}\}$ for r sufficiently large. That is $\text{lin}\{x_{n_k}\} = M$. The subsequence $\{x_{n_k}\}$ is, by its construction, linearly independent. ■

3.5.6 Theorem
Every separable Hilbert space $\mathcal{H}$ has a countable orthonormal basis.

Proof. By Theorem 2.7.1 there is a set $\{x_n \mid n \in \mathbb{N}\}$ such that $\overline{\text{lin}}\{x_n \mid n \in \mathbb{N}\} = \mathcal{H}$. Using Lemma 3.5.11 extract from $\{x_n \mid n \in \mathbb{N}\}$ a linearly independent subsequence $\{x_{n_k}\}$ such that $\text{lin}\{x_n\} = \text{lin}\{x_{n_k}\}$. Apply the Gram-Schmidt Orthonormalisation Procedure to the subsequence $\{x_{n_k}\}$ to obtain an orthonormal basis for $\mathcal{H}$. ■

3.5.7 Theorem
Every separable infinite-dimensional Hilbert space $\mathcal{H}$ is isometrically isomorphic to ℓ_2.

Proof. Let $\{x_n \mid n \in \mathbb{N}\}$ be an orthonormal basis for $\mathcal{H}$. Define $T : \mathcal{H} \to \ell_2$ by

$$Tx = (\langle x, x_n \rangle)_{n\in\mathbb{N}} \text{ for each } x \in \mathcal{H}.$$

It follows from Bessel's Inequality that the right hand side is in ℓ_2. We must show that T is a surjective linear isometry. (One-to-oneness follows from isometry.)

(i) <u>T is linear</u>: Let $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T(x+y) &= (\langle x+y, x_n\rangle)_{n\in\mathbb{N}} = (\langle x, x_n\rangle + \langle y, x_n\rangle)_{n\in\mathbb{N}} \\ &= (\langle x, x_n\rangle)_{n\in\mathbb{N}} + (\langle y, x_n\rangle)_{n\in\mathbb{N}} = Tx + Ty, \end{aligned}$$

and

$$T(\lambda x) = (\langle \lambda x, x_n\rangle)_{n\in\mathbb{N}} = (\lambda\langle x, x_n\rangle)_{n\in\mathbb{N}} = \lambda\,(\langle x, x_n\rangle)_{n\in\mathbb{N}}\,.$$

(ii) <u>T is surjective</u>: Let $(c_n)_{n\in\mathbb{N}} \in \ell_2$. By the Riesz-Fischer Theorem (Theorem 3.5.3), the series $\sum_{k=1}^{\infty} c_k x_k$ converges to some $x \in \mathcal{H}$. By continuity of the inner product, we have that for each $j \in \mathbb{N}$,

$$\langle x, x_j\rangle = \lim_{n\to\infty} \left\langle \sum_{k=1}^{n} c_k x_k, x_j \right\rangle = \lim_{n\to\infty} c_j = c_j.$$

Hence, $Tx = (\langle x, x_n\rangle)_{n\in\mathbb{N}} = (c_n)_{n\in\mathbb{N}}$.

(iii) <u>T is an isometry</u>: For each $x \in \mathcal{H}$,

$$\|Tx\|_2^2 = \sum_{n\in\mathbb{N}} |\langle x, x_n\rangle|^2 = \|x\|^2,$$

where the second equality follows from the fact that $\{x_n\}_{n\in\mathbb{N}}$ is an orthonormal basis and Theorem 3.5.4[5]. ■

Chapter 4

Bounded Linear Operators and Functionals

4.1 Introduction

An essential part of functional analysis is the study of continuous linear operators acting on linear spaces. This is perhaps not surprising since functional analysis arose due to the need to solve differential and integral equations, and differentiation and integration are well known linear operators. It turns out that it is advantageous to consider this type of operators in this more abstract way. It should also be mentioned that in physics, operator means a linear operator from one Hilbert space to another.

4.1.1 Definition
Let X and Y be linear spaces over the same field $\mathbb{F}$. A **linear operator** *from X into Y is a mapping $T : X \to Y$ such that*

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 \text{ for all } x_1, x_2 \in X \text{ and all } \alpha, \beta \in \mathbb{F}.$$

Simply put, a linear operator between linear spaces is a mapping that preserves the structure of the underlying linear space.
We shall denote by $\mathcal{L}(X, Y)$ the set of all linear operators from X into Y. We shall write $\mathcal{L}(X)$ for $\mathcal{L}(X, X)$.

4.1.2 Exercise
Let X and Y be linear spaces over the same field $\mathbb{F}$. Show that if T is a linear operator from X into Y, then $T(0) = 0$.

The **range** of a linear operator $T : X \to Y$ is the set

$$\operatorname{ran}(T) = \{y \in Y \mid y = Tx \text{ for some } x \in X\} = TX,$$

and the **null space** or the **kernel** of $T \in \mathcal{L}(X, Y)$ is the set

$$\mathcal{N}(T) = \ker(T) = \{x \in X : Tx = 0\} = T^{-1}(0).$$

If $T \in \mathcal{L}(X, Y)$, then $\ker(T)$ is a linear subspace of X and $\operatorname{ran}(T)$ is a linear subspace of Y.

An operator $T \in \mathcal{L}(X, Y)$ is **one-to-one (or injective)** if $\ker(T) = \{0\}$ and **onto (or surjective)** if $\operatorname{ran}(T) = Y$. If $T \in \mathcal{L}(X, Y)$ is one-to-one, then there exists a map $T^{-1} : \operatorname{ran}(T) \to \operatorname{dom}(T)$ which maps each $y \in \operatorname{ran}(T)$ onto that $x \in \operatorname{dom}(T)$ for which $Tx = y$. In this case we write $T^{-1}y = x$ and the map T^{-1} is called the **inverse** of the operator $T \in \mathcal{L}(X, Y)$. An operator $T : X \to Y$ is **invertible** if it has an inverse T^{-1}.

It is easy exercise to show that an invertible operator can have only one inverse.

4.1.3 Proposition
Let X and Y be linear spaces over $\mathbb{F}$. Suppose that $T \in \mathcal{L}(X,Y)$ is invertible. Then

(a) T^{-1} *is also invertible and* $(T^{-1})^{-1} = T$.

(b) $TT^{-1} = I_Y$ and $T^{-1}T = I_X$.

(c) T^{-1} *is a linear operator.*

Proof. We shall leave (a) and (b) as an easy exercise.
(c) Linearity of T^{-1}: Let $x, y \in Y$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} T^{-1}(\alpha x + y) &= T^{-1}(\alpha TT^{-1}x + TT^{-1}y) = T^{-1}[T(\alpha T^{-1}x + T^{-1}y)] \\ &= T^{-1}T(\alpha T^{-1}x + T^{-1}y) = \alpha T^{-1}x + T^{-1}y. \end{aligned}$$

Let X and Y be linear spaces over $\mathbb{F}$. For all $T, S \in \mathcal{L}(X,Y)$ and $\alpha \in \mathbb{F}$, define the operations of addition and scalar multiplication as follows:

$$\begin{aligned} (T+S)(x) &= Tx + Sx \quad \text{and} \\ (\alpha T)(x) &= \alpha Tx \quad \text{for each} \quad x \in X. \end{aligned}$$

Then $\mathcal{L}(X,Y)$ is a linear space over $\mathbb{F}$.

The most important class of linear operators is that of *bounded* linear operators.

4.1.4 Definition
Let X and Y be normed linear spaces over the same field $\mathbb{F}$. A linear operator $T : X \to Y$ is said to be **bounded** *if there exists a constant $M > 0$ such that*

$$\|Tx\| \le M\|x\| \quad \textit{for all} \quad x \in X.$$

(It should be emphasised that the norm on the left side is in Y and that on the right side is in X.)
An operator $T : X \to Y$ is said to be **continuous at** *$x_0 \in X$ if given any $\epsilon > 0$ there is a $\delta > 0$ such that*

$$\|Tx - Tx_0\| < \epsilon \quad \textit{whenever } \|x - x_0\| < \delta.$$

T is continuous on X if it is continuous at each point of X.

We shall denote by $\mathcal{B}(X,Y)$ the set of all bounded linear operators from X into Y. We shall write $\mathcal{B}(X)$ for $\mathcal{B}(X,X)$.

4.1.5 Definition
Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and let $T \in \mathcal{B}(X,Y)$. The **operator norm** *(or simply* **norm***) of T, denoted by $\|T\|$, is defined as*

$$\|T\| = \inf\{M : \|Tx\| \le M\|x\|, \quad \textit{for all } x \in X\}.$$

Since T is bounded, $\|T\| < \infty$. Furthermore,

$$\|Tx\| \le \|T\|\|x\| \quad \text{for all} \quad x \in X.$$

4.1.1 Theorem
Let X and Y be normed linear spaces over a field $\mathbb{F}$ and let $T \in \mathcal{B}(X,Y)$. Then

$$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \ne 0\right\} = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\|Tx\| : \|x\| \le 1\}.$$

Proof. Let $\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}$, $\beta = \sup\{\|Tx\| : \|x\| = 1\}$, and $\gamma = \sup\{\|Tx\| : \|x\| \leq 1\}$. We first show that $\|T\| = \alpha$. Now, for all $x \in X \setminus \{0\}$ we have that $\frac{\|Tx\|}{\|x\|} \leq \alpha$, and therefore $\|Tx\| \leq \alpha\|x\|$. By definition of $\|T\|$ we have that $\|T\| \leq \alpha$. On the other hand, for all $x \in X$, we have that $\|Tx\| \leq \|T\|\|x\|$. In particular, for all $x \in X \setminus \{0\}$, $\frac{\|Tx\|}{\|x\|} \leq \|T\|$, and therefore $\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} \leq \|T\|$. Thus, $\alpha = \|T\|$.

Next, we show that $\alpha = \beta = \gamma$. Now, for each $x \in X$

$$\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} = \left\{\left\|T\left(\frac{x}{\|x\|}\right)\right\| : x \neq 0\right\} \subset \{\|Tx\| : \|x\| = 1\} \subset \{\|Tx\| : \|x\| \leq 1\}.$$

Thus,

$$\alpha = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} \leq \beta = \sup\{\|Tx\| : \|x\| = 1\} \leq \gamma = \sup\{\|Tx\| : \|x\| \leq 1\}.$$

But for all $x \neq 0$

$$\frac{\|Tx\|}{\|x\|} \leq \alpha \quad \Rightarrow \quad \|Tx\| \leq \alpha\|x\| \leq \alpha \quad \text{for all } x \text{ such that } \|x\| \leq 1.$$

Therefore,

$$\gamma = \sup\{\|Tx\| : \|x\| \leq 1\} \leq \alpha.$$

That is, $\alpha \leq \beta \leq \gamma \leq \alpha$. Hence, $\alpha = \beta = \gamma$. ■

4.1.2 Theorem

Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then the function $\|\cdot\|$ defined above is a norm on $\mathcal{B}(X, Y)$.

Proof. Properties N1 and N2 of a norm are easy to verify. We prove N3 and N4. Let $T \in \mathcal{B}(X, Y)$ and $\alpha \in \mathbb{F}$.

N3. $\|\alpha T\| = \sup\{\|\alpha Tx\| : \|x\| = 1\} = |\alpha| \sup\{\|Tx\| : \|x\| = 1\} = |\alpha|\|T\|$.

N4. Let $T, S \in \mathcal{B}(X, Y)$. Then for each $x \in X$,

$$\|(T + S)(x)\| = \|Tx + Sx\| \leq \|Tx\| + \|Sx\| \leq (\|T\| + \|S\|)\|x\|.$$

Thus, $\|T + S\| \leq \|T\| + \|S\|$. ■

4.1.6 Examples

[1] Let $X = \mathbb{F}^n$ with the uniform norm $\|\cdot\|_\infty$. For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$, define $T : \mathbb{F}^n \to \mathbb{F}^n$ by

$$Tx = T(x_1, x_2, \ldots, x_n) = \left(\sum_{j=1}^{n} \alpha_{1j} x_j, \sum_{j=1}^{n} \alpha_{2j} x_j, \ldots, \sum_{j=1}^{n} \alpha_{nj} x_j\right).$$

It is easy to show that T is a linear operator on X. We show that T is bounded.

$$\begin{aligned} \|Tx\|_\infty &= \sup_{1 \leq i \leq n} \left|\sum_{j=1}^{n} \alpha_{ij} x_j\right| \leq \sup_{1 \leq i \leq n} \sum_{j=1}^{n} |\alpha_{ij}||x_j| \\ &\leq \sup_{1 \leq i \leq n} \sum_{j=1}^{n} |\alpha_{ij}| \sup_{1 \leq j \leq n} |x_j| = M\|x\|_\infty, \end{aligned}$$

where $M = \sup_{1 \le i \le n} \sum_{j=1}^{n} |\alpha_{ij}|$. Hence, $\|T\| \le M$.

We claim that $\|T\| = M$. We need to show that $\|Tx\|_\infty \ge M\|x\|_\infty$. To that end, choose an index k such that $\sum_{j=1}^{n} |\alpha_{kj}| = M = \sup_{1 \le i \le n} \sum_{j=1}^{n} |\alpha_{ij}|$ and let x be the unit vector whose j-th component is $\frac{\overline{\alpha_{kj}}}{|\alpha_{kj}|}$. Then

$$\|Tx\|_\infty = \sup_{1 \le i \le n} \left| \sum_{j=1}^{n} \alpha_{ij} x_j \right| \ge \left| \sum_{j=1}^{n} \alpha_{kj} x_j \right| = \sum_{j=1}^{n} |\alpha_{kj}| = M\|x\|_\infty.$$

Thus $\|T\| = \sup_{1 \le i \le n} \sum_{j=1}^{n} |\alpha_{ij}|$. ■

[2] Let $\{x_n \mid n \in \mathbb{N}\}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. For $(\lambda_i)_{i=1}^{\infty} \in \ell_\infty$, define $T : \mathcal{H} \to \mathcal{H}$ by

$$Tx = \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i.$$

Then T is a bounded linear operator on $\mathcal{H}$. Linearity is an immediate consequence of the inner product.

Boundedness:

$$\begin{aligned} \|Tx\|^2 &= \left\| \sum_{i=1}^{\infty} \lambda_i \langle x, x_i \rangle x_i \right\|^2 = \sum_{i=1}^{\infty} |\lambda_i|^2 |\langle x, x_i \rangle|^2 \|x_i\|^2 \\ &\le M^2 \sum_{i=1}^{\infty} |\langle x, x_i \rangle|^2, \quad \text{where } M = \sup_{i \in \mathbb{N}} |\lambda_i| \\ &\le M^2 \|x\|^2 \quad \text{by Bessel's Inequality.} \end{aligned}$$

Thus, $\|Tx\| \le M\|x\|$, and consequently $\|T\| \le M$.

We show that $\|T\| = \sup_{i \in \mathbb{N}} |\lambda_i|$. Indeed, for any $\epsilon > 0$, there exists λ_k such that $|\lambda_k| > M - \epsilon$. Hence,

$$\|T\| \ge \|Tx_k\| = \|\lambda_k x_k\| = |\lambda_k| > M - \epsilon.$$

Since ϵ is arbitrary, we have that $\|T\| \ge M$. ■

[3] Define an operator $L : \ell_2 \to \ell_2$ by

$$Lx = L((x_1, x_2, x_3, \ldots)) = (x_2, x_3, \ldots).$$

The L is a bounded linear operator on ℓ_2.

Linearity: Easy.

Boundedness: For all $x = (x_1, x_2, x_3, \ldots) \in \ell_2$,

$$\|Lx\|_2^2 = \sum_{i=2}^{\infty} |x_i|^2 \le \sum_{i=1}^{\infty} |x_i|^2 = \|x\|_2^2.$$

That is, L is a bounded linear operator and $\|L\| \leq 1$. We show that $\|L\| = 1$. To that end, consider $e_2 = (0, 1, 0, 0, \ldots) \in \ell_2$. Then

$$\|e_2\|_2 = 1 \quad \text{and} \quad Le_2 = (1, 0, 0, \ldots) \quad \text{which implies that} \quad \|Le_2\|_2 = 1.$$

Thus, $\|L\| = 1$. The operator L is called the **left-shift operator**. ■

[4] Let $\mathcal{BC}[0, \infty)$ be the linear space of all bounded continuous functions on the interval $[0, \infty)$ with the uniform norm $\|\cdot\|_\infty$. Define $T : \mathcal{BC}[0, \infty) \to \mathcal{BC}[0, \infty)$ by

$$(Tx)\,(t) = \frac{1}{t}\int_0^t x(\tau)d\tau.$$

Then T is a bounded linear operator on $\mathcal{BC}[0, \infty)$.

Linearity: For all $x, y \in \mathcal{BC}[0, \infty)$ and all $\alpha, \beta \in \mathbb{F}$,

$$(T(\alpha x + \beta y))\,(t) = \frac{1}{t}\int_0^t (x + y)(\tau)d\tau = \alpha\left(\frac{1}{t}\int_0^t x(\tau)d\tau\right) + \beta\left(\frac{1}{t}\int_0^t y(\tau)d\tau\right).$$

Boundedness: For each $x \in \mathcal{BC}[0, \infty)$,

$$\begin{aligned}\|Tx\|_\infty &= \sup_t |(Tx)(t)| = \sup_t \left|\frac{1}{t}\int_0^t x(\tau)d\tau\right| \\ &\leq \sup_t \frac{1}{t}\int_0^t |x(\tau)|d\tau \leq \left(\sup_t \frac{1}{t}\int_0^t d\tau\right)\|x\|_\infty = \|x\|_\infty.\end{aligned}$$

[5] Let M be a closed subspace of a normed linear space X and $Q_M : X \to X/M$ the qoutient map. Then Q_M is bounded and $\|Q_M\| = 1$. Indeed, since $\|Q_M(x)\| = \|x + M\| \leq \|x\|$, Q_M is bounded and $\|Q_M\| \leq 1$. But since Q_M maps the open unit ball in X onto the open unit ball in X/M, it follows that $\|Q_M\| = 1$.

[6] Let $X = mathcal P[0, 1]$ - the set of polynomials on the interval $[0, 1]$ with the uniform norm $\|x\|_\infty = \max_{0 \leq t \leq 1} |x(t)|$. For each $x \in X$, define $T : X \to X$ by

$$Tx = x'(t) = \frac{dx}{dt} \quad \text{(differentiation with respect to } t\text{)}.$$

Linearity: For $x, y \in X$ and all $\alpha, \beta \in \mathbb{F}$,

$$T(\alpha x + \beta y) = (\alpha x + \beta y)'(t) = \alpha x'(t) + \beta y'(t) = \alpha Tx + \beta Ty.$$

T is **not** bounded: Let $x_n(t) = t^n,\ n \in \mathbb{N}$. Then

$$\|x_n\| = 1, \quad Tx_n = x_n'(t) = nt^{n-1}, \quad \text{and} \quad \|Tx_n\| = \frac{\|Tx_n\|}{\|x_n\|} = n.$$

Hence T is unbounded. ■

4.1.3 Theorem

Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then $T \in \mathcal{L}(X, Y)$ is bounded if and only if T maps a bounded set into a bounded set.

Proof. Assume that T is bounded. That is, there exists a constant $M > 0$ such that $\|Tx\| \leq M\|x\|$ for all $x \in X$. If $\|x\| \leq k$ for some constant k, then $\|Tx\| \leq M\|x\| \leq kM$. That is, T maps a bounded set into a bounded set.

Now assume that T maps a bounded set into a bounded set. Then T maps the unit ball $B = \{x \in X : \|x\| \leq 1\}$ into a bounded set. That is, there exists a constant $M > 0$ such that $\|Tx\| \leq M$ for all $x \in B$. Therefore, for any nonzero $x \in X$,

$$\frac{\|Tx\|}{\|x\|} = \left\| T\left(\frac{x}{\|x\|}\right) \right\| \leq M.$$

Hence, $\|Tx\| \leq M\|x\|$. That is, T is bounded. ∎

4.1.7 Exercise

Show that the inverse of a bounded linear operator is not necessarily bounded.

4.1.8 Proposition

Let $T \in \mathcal{B}(X, Y)$. Then T^{-1} exists and is bounded if and only if there is a constant $K > 0$ such that

$$\|Tx\| \geq K\|x\| \text{ for all } x \in X.$$

Proof. Assume that there is a constant $K > 0$ such that $\|Tx\| \geq K\|x\|$ for all $x \in X$. If $x \neq 0$, then $Tx \neq 0$ and so T is one-to-one and hence T^{-1} exists. Also, given $y \in \operatorname{ran}(T)$, let $y = Tx$ for some $x \in X$. Then

$$\|T^{-1}y\| = \|T^{-1}(Tx)\| = \|x\| \leq \frac{1}{K}\|Tx\| = \frac{1}{K}\|y\|,$$

i.e., $\|T^{-1}y\| \leq \frac{1}{K}\|y\|$ for all $y \in Y$. Thus T^{-1} is bounded.

Assume that T^{-1} exists and is bounded. Then for each $x \in X$,

$$\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\|Tx\| \iff \frac{1}{\|T^{-1}\|}\|x\| \leq \|Tx\| \iff K\|x\| \leq \|Tx\|,$$

where $K = \frac{1}{\|T^{-1}\|}$. ∎

The following theorem asserts that continuity and boundedness are equivalent concepts for linear operators.

4.1.4 Theorem

Let X and Y be normed linear spaces over a field $\mathbb{F}$ and $T \in \mathcal{L}(X, Y)$. The following statements are equivalent:

(1) T is continuous on X;

(2) T is continuous at some point in X;

(3) T is bounded on X.

Proof. The implication (1) ⇒ (2) is obvious.

(2) ⇒ (3): Assume that T is continuous at $x \in X$, but T is *not* bounded on X. Then there is a sequence (x_n) in X such that $\|Tx_n\| > n\|x_n\|$ for each $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, let

$$y_n = \frac{x_n}{n\|x_n\|} + x.$$

Then

$$\|y_n - x\| = \frac{1}{n} \to 0 \text{as } n \to \infty;$$

i.e., $y_n \stackrel{n\to\infty}{\longrightarrow} x$, but

$$\|Ty_n - Tx\| = \frac{\|Tx_n\|}{n\|x_n\|} > \frac{n\|x_n\|}{n\|x_n\|} = 1.$$

That is, $Ty_n \not\to Tx$ as $n \to \infty$, contradicting (2).

(3) $\Rightarrow$ (1): Assume that T is bounded on X. Let (x_n) be a sequence in X which converges to $x \in X$. Then

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \le \|T\|\|x_n - x\| \ \to 0 \ \text{ as } \ n \to \infty.$$

Thus, T is continuous on X. ■

4.1.5 Theorem

Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be normed linear spaces with $\dim(X) < \infty$ and $T : X \to Y$ be a linear operator. Then T is continuous. That is, every linear operator on a finite-dimensional normed linear space is automatically continuous.

Proof. Define a new norm $\|\cdot\|_0$ on X by

$$\|x\|_0 = \|x\| + \|Tx\| \ \text{ for all } x \in X.$$

Since X is finte-dimensional, the norms $\|\cdot\|_0$ and $\|\cdot\|$ on X are equivalent. Hence there are constants α and β such that

$$\alpha\|x\|_0 \le \|x\| \le \beta\|x\|_0 \ \text{ for all } x \in X.$$

Hence,

$$\|Tx\| \le \|x\|_0 \le \frac{1}{\alpha}\|x\| = K\|x\|,$$

where $K = \frac{1}{\alpha}$. Therefore T is bounded. ■

4.1.9 Definition

Let X and Y be normed linear spaces over a field $\mathbb{F}$.

(1) A sequence $(T_n)_1^\infty$ in $\mathcal{B}(X, Y)$ is said to be **uniformly operator convergent** *to T if*

$$\lim_{n\to\infty} \|T_n - T\| = 0.$$

This is also referred to as **convergence in the uniform topology** *or* **convergence in the operator norm topology** *of $\mathcal{B}(X, Y)$. In this case T is called the* **uniform operator limit** *of the sequence $(T_n)_1^\infty$.*

(2) A sequence $(T_n)_1^\infty$ in $\mathcal{B}(X, Y)$ is said to be **strongly operator convergent** *to T if*

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0 \quad \text{for each } x \in X.$$

In this case T is called the **strong operator limit** *of the sequence $(T_n)_1^\infty$.*

Of course, if T is the uniform operator limit of the sequence $(T_n)_1^\infty \subset \mathcal{B}(X, Y)$, then $T \in \mathcal{B}(X, Y)$. On the other hand, the strong operator limit T of a sequence $(T_n)_1^\infty \subset \mathcal{B}(X, Y)$ need not be bounded in general.

The following proposition asserts that uniform convergence implies strong convergence.

4.1.10 Proposition

If the sequence $(T_n)_1^\infty$ in $\mathcal{B}(X, Y)$ is uniformly convergent to $T \in \mathcal{B}(X, Y)$, then it is strongly convergent to T.

Proof. Since, for each $x \in X$, $\|T_n x - Tx\| = \|(T_n - T)(x)\| \leq \|T_n - T\|\|x\|$, if $\|T_n - T\| \to 0$ as $n \to \infty$, then

$$\|(T_n - T)(x)\| \quad \to \ 0 \quad \text{as} \ \ n \to \infty.$$

The converse of Proposition 4.1.10 does not hold.

4.1.11 Example

Consider the sequence (T_n) of operators, where for each $n \in \mathbb{N}$,

$T_n : \ell_2 \to \ell_2$ is given by

$$T_n(x_1, x_2, \ldots) = (0, 0, \ldots, 0, x_{n+1}, x_{n+2}, \ldots).$$

Let $\epsilon > 0$ be given. Then for each $x = (x_i)_{i=1}^{\infty} \in \ell_2$, there exists N such that

$$\sum_{n+1}^{\infty} |x_i|^2 < \epsilon^2, \quad \text{for all} \ \ n \geq N.$$

Hence, for all $n \geq N$,

$$\|T_n x\|_2^2 = \sum_{n+1}^{\infty} |x_i|^2 < \epsilon^2.$$

That is, for each $x \in \ell_2$, $T_n x \to 0$. Hence, $T_n \to 0$ strongly.

Now, since

$$\|T_n x\|_2^2 = \sum_{n+1}^{\infty} |x_i|^2 \leq \sum_{1}^{\infty} |x_i|^2 = \|x\|_2^2$$

for $n \in \mathbb{N}$ and $x = (x_i)_{i=1}^{\infty} \in \ell_2$, it follows that $\|T_n\| \leq 1$ for each $n \in \mathbb{N}$.

But $\|T_n\| \geq 1$ for all n. To see this, take $x = (0, 0, \ldots, 0, x_{n+1}, 0, \ldots) \in \ell_2$, where $x_{n+1} \neq 0$. Then

$$T_n x = x \quad \text{and hence} \quad \|T_n x\|_2 = |x_{n+1}|, \ \text{and consequently,} \ \|T_n\| \geq 1.$$

That is, (T_n) does not converge to zero in the uniform topology. ∎

4.1.6 Theorem

Let X and Y be normed linear spaces over a field $\mathbb{F}$. Then $\mathcal{B}(X, Y)$ is a Banach space if Y is a Banach space.

Proof. We have shown that $\mathcal{B}(X, Y)$ is a normed linear space. It remains to show that it is complete if Y is complete. To that end, let $(T_n)_1^{\infty}$ be a Cauchy sequence in $\mathcal{B}(X, Y)$. Then given any $\epsilon > 0$ there exists a positive integer N such that

$$\|T_n - T_{n+r}\| < \epsilon \quad \text{for all } n > N,$$

whence,

$$\|T_n x - T_{n+r} x\| \leq \|T_n - T_{n+r}\|\|x\| < \epsilon\|x\| \quad \text{for all } x \in X. \tag{4.1.6.1}$$

Hence, $(T_n x)_1^{\infty}$ is a Cauchy sequence in Y. Since Y is complete there exists $y \in Y$ such that $T_n x \to y$ as $n \to \infty$. Set $Tx = y$. We show that $T \in \mathcal{B}(X, Y)$ and $T_n \to T$. Let $x_1, x_2 \in X$, and $\alpha, \beta \in \mathbb{F}$. Then

$$\begin{aligned} T(\alpha x_1 + \beta x_2) &= \lim_{n\to\infty} T_n(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} [\alpha T_n x_1 + \beta T_n x_2] \\ &= \alpha \lim_{n\to\infty} T_n x_1 + \beta \lim_{n\to\infty} T_n x_2 = \alpha T x_1 + \beta T x_2. \end{aligned}$$

That is, $T \in \mathcal{L}(X, Y)$. Taking the limit as $r \to \infty$ in (4.1.6.1) we get that

$$\|(T_n - T)x\| = \|T_n x - Tx\| \leq \epsilon\|x\| \quad \text{for all } n > N, \ \text{and all } x \in X.$$

That is, $T_n - T$ is a bounded operator for all $n > N$. Since $\mathcal{B}(X, Y)$ is a linear space, $T = T_n - (T_n - T) \in \mathcal{B}(X, Y)$.

Finally,

$$\|T_n - T\| = \sup\{\|T_n x - Tx\| : \|x\| \leq 1\} \leq \sup\{\|x\|\epsilon : \|x\| \leq 1\} \leq \epsilon \quad \text{for all } \; n > N.$$

That is, $T_n \to T$ as $n \to \infty$. ■

4.1.12 Definition

Let $T : X \to Y$ and $S : Y \to Z$. We define the **composition** *of T and S as the map $ST : X \to Z$ defined by*

$$(ST)(x) = (S \circ T)(x) = S(Tx).$$

4.1.7 Theorem

Let X, Y and Z be normed linear spaces over a field $\mathbb{F}$ and let $T \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$. Then $ST \in \mathcal{B}(X, Z)$ and $\|ST\| \leq \|S\|\|T\|$.

Proof. Since linearity is trivial, we only prove boundedness of ST. Let $x \in X$. Then

$$\|(ST)(x)\| = \|S(Tx)\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|.$$

Thus, $\|ST\| \leq \|S\|\|T\|$. ■

Let X be a normed linear space over $\mathbb{F}$. For $S, T_1, T_2 \in \mathcal{B}(X)$ it is easy to show that

$$\begin{aligned} (ST_1)T_2 &= S(T_1T_2) \\ S(T_1 + T_2) &= ST_1 + ST_2 \\ (T_1 + T_2)S &= T_1S + T_2S. \end{aligned}$$

The operator I defined by $Ix = x$ for all $x \in X$ belongs to $\mathcal{B}(X)$, $\|I\| = 1$, and it has the property that $IT = TI = T$ for all $T \in \mathcal{B}(X)$. We call I the **identity operator**. The set $\mathcal{B}(X)$ is therefore an algebra with an identity element. In fact, $\mathcal{B}(X)$ is a **normed algebra** with an identity element. If X is a Banach space then $\mathcal{B}(X)$ is a **Banach algebra**.

We now turn our attention to a very special and important class of bounded linear operators, namely, **bounded linear functionals**.

4.1.13 Definition

Let X be a linear space over $\mathbb{F}$. A linear operator $f : X \to \mathbb{F}$ is called a **linear functional** *on X. Of course, $\mathcal{L}(X, \mathbb{F})$ denotes the set of all linear functionals on X.*

Since every linear functional is a linear operator, all of the foregoing discussion on linear operators applies equally well to linear functionals. For example, if X is a normed linear space then we say that $f \in \mathcal{L}(X, \mathbb{F})$ is bounded if there exists a constant $M > 0$ such that $|f(x)| \leq M\|x\|$ for all $x \in X$. The **norm** of f is defined by

$$\|f\| = \sup\{|f(x)| : \|x\| \leq 1\}.$$

We shall denote by $X^* = \mathcal{B}(X, \mathbb{F})$ the set of all bounded (i.e., continuous) linear functionals on X. We call X^* the **dual** of X. It follows from Theorem 4.1.6 that X^* is always a Banach space under the above norm.

4.1.14 Examples

[1] Let $X = \mathcal{C}[a,b]$. For each $x \in X$, define $f : X \to \mathbb{R}$ by

$$f(x) = \int_a^b x(t)dt.$$

Then f is a bounded linear functional on X.

Linearity: For any $x, y \in X$ and any $\alpha, \beta \in \mathbb{R}$,

$$f(\alpha x + \beta y) = \int_a^b (\alpha x + \beta y)(t)dt = \alpha \int_a^b x(t)dt + \beta \int_a^b y(t)dt = \alpha f(x) + \beta f(y).$$

Boundedness:

$$|f(x)| = \left| \int_a^b x(t)dt \right| \le \int_a^b |x(t)|dt \le \max_{a \le t \le b} |x(t)|(b-a) = \|x\|_\infty (b-a).$$

Hence f is bounded and $\|f\| \le b - a$. We show that $\|f\| = b - a$. Take $x = 1$, the constant function 1. Then

$$f(1) = \int_a^b dt = b - a, \quad \text{i.e., } |f(1)| = b - a.$$

Hence

$$b - a = \frac{|f(1)|}{1} \le \sup \left\{ \frac{|f(x)|}{\|x\|} : x \ne 0 \right\} = \|f\| \le b - a.$$

That is, $\|f\| = b - a$. ■

[2] Let $X = \mathcal{C}[a,b]$ and let $t \in (a,b)$ be fixed. For each $x \in X$, define $\delta_t : X \to \mathbb{R}$ by

$$\delta_t(x) = x(t), \quad \text{(i.e., } \delta_t \text{ is a point evaluation at } t).$$

Then δ_t is a bounded linear functional on X. Linearity of δ_t is easy to verify.

Boundedness: For each $x \in X$,

$$|\delta_t(x)| = |x(t)| \le \max_{a \le \tau \le b} |x(\tau)| = \|x\|_\infty.$$

That is, δ_t is a bounded linear operator and $\|\delta_t\| \le 1$. We show that $\|\delta_t\| = 1$. Take $x = 1$, the constant 1 function. Then $\delta_t(1) = 1$ and so

$$1 = \frac{|\delta_t(1)|}{1} \le \sup\{|\delta_t(x)| \ : \ \|x\| = 1\} = \|\delta_t\| \le 1.$$

That is, $\|\delta_t\| = 1$. ■

[3] Let $c_1, c_2, \ldots, c_n$ be real numbers and let $X = \mathcal{C}[a,b]$. Define $f : X \to \mathbb{R}$ by

$$f(x) = \sum_{i=1}^n c_i x(t_i), \qquad \text{where} \quad t_1, t_2, \ldots, t_n \quad \text{are in} \quad [a,b].$$

Then f is a bounded linear functional on X.

Linearity: Clear.

Boundedness: For any $x \in X$,

$$|f(x)| = \left|\sum_{i=1}^{n} c_i x(t_i)\right| \leq \sum_{i=1}^{n} |c_i x(t_i)| = \sum_{i=1}^{n} |c_i||x(t_i)| \leq \|x\|_\infty \sum_{i=1}^{n} |c_i|.$$

Hence, f is bounded and

$$\|f\| \leq \sum_{i=1}^{n} |c_i|.$$

□

[4] Let X be a linear space. The norm $\|\cdot\| : X \to \mathbb{R}$ is an example of a nonlinear functional on X.

4.2 Examples of Dual Spaces

4.2.1 Definition

Let X and Y be normed linear spaces over the same field $\mathbb{F}$. Then X and Y are said to be isomorphic to each other, denoted by $X \simeq Y$, if there is a bijective linear operator T from X onto Y. If, in addition, T is an isometry, i.e., $\|Tx\| = \|x\|$ for each $x \in X$, then we say that T is an **isometric isomorphism**. *In this case, X and Y are are said to be isometrically isomorphic and we write $X \cong Y$.*

Two normed linear spaces which are isometrically isomorphic can be regarded as identical, the isometry merely amounting to a relabelling of the elements.

4.2.2 Proposition

Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and T a linear operator from X onto Y. Then T is an isometry if and only if

(i) T is one-to-one;

(ii) T is continuous on X;

(iii) T has a continuous inverse (in fact, $\|T^{-1}\| = \|T\| = 1$);

(iv) T is distance-preserving: For all $x, y \in X$, $\|Tx - Ty\| = \|x - y\|$.

Proof. If T satisfies (iv), then, taking $y = 0$, we have that $\|Tx\| = \|x\|$ for each $x \in X$; i.e., T is an isometry.

Conversely, assume that T is an isometry. If $x \neq y$, then

$$\|Tx - Ty\| = \|T(x - y)\| = \|x - y\| > 0.$$

Hence, $Tx \neq Ty$. This shows that T is one-to-one and distance-preserving. Since $\|Tx\| = \|x\|$ for each $x \in X$, it follows that T is bounded and $\|T\| = 1$. By Theorem 4.1.4, T is continuous on X.
Let $y_1,\ y_2 \in Y$ and $\alpha_1,\ \alpha_2 \in \mathbb{F}$. Then there exist $x_1,\ x_2 \in X$ such that $Tx_i = y_i$ for $i = 1,\ 2$. Therefore

$$\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 Tx_1 + \alpha_2 Tx_2 = T(\alpha_1 x_1 + \alpha_2 x_2) \text{ or}$$

$$T^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1 + \alpha_2 x_2 = \alpha_1 T^{-1} y_1 + \alpha_2 T^{-1} y_2.$$

That is, T^{-1} is linear. Furthermore, for $y \in Y$, let $x = T^{-1}y$. Then,

$$\|T^{-1}y\| = \|x\| = \|Tx\| = \|y\|.$$

Therefore T^{-1} is bounded and $\|T^{-1}\| = 1$. ■

4.2.3 Remark

It is clear from Proposition 4.2.2 that two normed linear spaces X and Y are isometrically isomorphic if and only if there is a linear isometry from X onto Y.

[1] *The dual of ℓ_1 is (isometrically isomorphic to) ℓ_∞; i.e., $\ell_1^* \cong \ell_\infty$.*

Proof. Let $y = (y_n) \in \ell_\infty$ and define $\Phi : \ell_\infty \to \ell_1^*$ by

$$(\Phi y)(x) = \sum_{j=1}^{\infty} x_j y_j \text{ for } x = (x_n) \in \ell_1.$$

Claim 1: $\Phi y \in \ell_1^*$.

Linearity of Φy: Let $x = (x_n)$, $z = (z_n) \in \ell_1$ and $\alpha \in \mathbb{F}$. Then

$$\begin{aligned}(\Phi y)(\alpha x + z) &= \sum_{j=1}^{\infty}(\alpha x_j + z_j) y_j = \sum_{j=1}^{\infty} \alpha x_j y_j + \sum_{j=1}^{\infty} z_j y_j \\ &= \alpha \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} z_j y_j \\ &= \alpha(\Phi y)(x) + (\Phi y)(z).\end{aligned}$$

Boundedness of Φy: For any $x = (x_n) \in \ell_1$,

$$|(\Phi y)(x)| = \left|\sum_{j=1}^{\infty} x_j y_j\right| \le \sum_{j=1}^{\infty} |x_j y_j| \le \|y\|_\infty \sum_{j=1}^{\infty} |x_j| = \|y\|_\infty \|x\|_1.$$

That is, $\Phi y \in \ell_1^*$ and

$$\|\Phi y\| \le \|y\|_\infty. \tag{$\star$}$$

Claim 2: Φ is a surjective linear isometry.

(i) Φ is a surjective: A basis for ℓ_1 is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n-th position and zeroes elsewhere. Let $f \in \ell_1^*$ and $x = (x_n) \in \ell_1$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n) = \sum_{n=1}^{\infty} x_n z_n,$$

where, for each $n \in \mathbb{N}$, $z_n = f(e_n)$. We show that $z = (z_n) \in \ell_\infty$. Indeed, for each $n \in \mathbb{N}$

$$|z_n| = |f(e_n)| \le \|f\| \|e_n\| = \|f\|.$$

Hence, $z = (z_n) \in \ell_\infty$. Also, for any $x = (x_n) \in \ell_1$,

$$(\Phi z)(x) = \sum_{n=1}^{\infty} x_n z_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x).$$

That is, $\Phi z = f$ and so Φ is surjective. Furthermore,

$$\|z\|_\infty = \sup_{n\in\mathbb{N}} |z_n| = \sup_{n\in\mathbb{N}} |f(e_n)| \le \|f\| = \|\Phi z\|. \tag{$\star\star$}$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_\infty$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in \ell_1$,

$$\begin{aligned}[\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j(\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x).\end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta \Phi y + \Phi$, which proves linearity of Φ.

(iii) Φ is an isometry: This follows from $(\star)$ and $(\star\star)$. □

[2] *The dual of c_0 is (isometrically isomorphic to) ℓ_1, i.e., $c_0^* \cong \ell_1$.*

Proof. Let $y = (y_n) \in \ell_1$ and define $\Phi : \ell_1 \to c_0^*$ by

$$(\Phi y)(x) = \sum_{j=1}^{\infty} x_j y_j \text{ for } x = (x_n) \in c_0.$$

Proceeding as in Example 1 above, one shows that Φy is a bounded linear functional on c_0 and

$$\|\Phi y\| \leq \|y\|_1. \qquad (\star)$$

Claim: Φ is a surjective linear isometry.

(i) Φ is a surjective: A basis for c_0 is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n-th position and zeroes elsewhere. Let $f \in c_0^*$ and $x = (x_n) \in c_0$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n) = \sum_{n=1}^{\infty} x_n w_n,$$

where, for each $n \in \mathbb{N}$, $w_n = f(e_n)$. For $n,\ k \in \mathbb{N}$, let

$$z_{nk} = \begin{cases} \dfrac{|w_k|}{w_k} & \text{if } w_k \neq 0 \text{ and } k \leq n \\ 0 & \text{if } w_k = 0 \text{ or } k > n, \end{cases}$$

and let

$$z_n = (z_{n1},\ z_{n2},\ \ldots,\ z_{nn},\ 0,\ 0,\ \ldots).$$

Then $z_n \in c_0$ and

$$\|z_n\|_\infty = \sup_{k \in \mathbb{N}} |z_{nk}| = 1.$$

Also,

$$f(z_n) = \sum_{k=1}^{\infty} z_{nk} w_k = \sum_{k=1}^{n} |w_k|.$$

Hence, for each $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} |w_k| = |f(z_n)| \leq \|f\| \|z_n\| \leq \|f\|.$$

Since the right hand side is independent of n, it follows that $\sum_{k=1}^{\infty} |w_k| \leq \|f\|$ Hence, $w = (w_n) \in \ell_1$. Also, for any $x = (x_n) \in c_0$,

$$(\Phi w)(x) = \sum_{n=1}^{\infty} x_n w_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x).$$

That is, $\Phi w = f$ and so Φ is surjective. Furthermore,

$$\|w\|_1 = \sum_{k=1}^{\infty} |w_k| \leq \|f\| = \|\Phi w\|. \qquad (\star\star)$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_1$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in c_0$,

$$\begin{aligned} [\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j(\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x). \end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta \Phi y + \Phi z$, which proves linearity of Φ.

(iii) Φ is an isometry: This follows from $(\star)$ and $(\star\star)$. □

[3] *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. *Then the dual of* ℓ_p *is (isometrically isomorphic to)* ℓ_q, *i.e.,* $\ell_p^* \cong \ell_q$.

Proof. Let $y = (y_n) \in \ell_q$ and define $\Phi : \ell_q \to \ell_p^*$ by

$$(\Phi y)(x) = \sum_{j=1}^{\infty} x_j y_j \text{ for } x = (x_n) \in \ell_p.$$

It is straightforward to show that Φy is linear. We show that Φy is bounded. By Hölder's Inequality,

$$|(\Phi y)(x)| = \left|\sum_{j=1}^{\infty} x_j y_j\right| \leq \sum_{j=1}^{\infty} |x_j y_j| \leq \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^{\infty} |y_j|^q\right)^{\frac{1}{q}} = \|x\|_p \|y\|_q.$$

That is, $\Phi y \in \ell_p^*$ and

$$\|\Phi y\| \leq \|y\|_q. \qquad (\star)$$

Claim: Φ is a surjective linear isometry.

(i) Φ is a surjective: A basis for ℓ_p is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n-th position and zeroes elsewhere. Let $f \in \ell_p^*$ and $x = (x_n) \in \ell_p$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n) = \sum_{n=1}^{\infty} x_n w_n,$$

where, for each $n \in \mathbb{N}$, $w_n = f(e_n)$. For $n,\ k \in \mathbb{N}$, let

$$z_{nk} = \begin{cases} \dfrac{|w_k|^q}{w_k} & \text{if } k \leq n \text{ and } w_k \neq 0 \\ 0 & \text{if } w_k = 0 \text{ or } k > n, \end{cases}$$

and let

$$z_n = (z_{n1},\ z_{n2},\ \ldots,\ z_{nn},\ 0,\ 0,\ \ldots).$$

Then $z_n \in \ell_p$ and

$$f(z_n) = \sum_{k=1}^{\infty} z_{nk} w_k = \sum_{k=1}^{n} |w_k|^q.$$

Hence, for each $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} |w_k|^q = |f(z_n)| \le \|f\| \|z_n\|_p.$$

Since

$$\begin{aligned}\|z_n\|_p &= \left(\sum_{k=1}^{\infty} |z_{nk}|^p\right)^{1/p} = \left(\sum_{k=1}^{n} |z_{nk}|^p\right)^{1/p} \\ &= \left(\sum_{k=1}^{n} |w_k|^{p(q-1)}\right)^{1/p} = \left(\sum_{k=1}^{n} |w_k|^q\right)^{1/p},\end{aligned}$$

it follows that , for each $n \in \mathbb{N}$,

$$\begin{aligned}\sum_{k=1}^{n} |w_k|^q \le \|f\| \|z_n\|_p \quad &\Longleftrightarrow \quad \sum_{k=1}^{n} |w_k|^q \le \|f\| \left(\sum_{k=1}^{n} |w_k|^q\right)^{1/p} \\ &\Longleftrightarrow \quad \left(\sum_{k=1}^{n} |w_k|^q\right)^{1-1/p} \le \|f\| \\ &\Longleftrightarrow \quad \left(\sum_{k=1}^{n} |w_k|^q\right)^{1/q} \le \|f\|.\end{aligned}$$

Since the right hand side is independent of n, it follows that $\left(\sum_{k=1}^{\infty} |w_k|^q\right)^{1/q} \le \|f\|$, and so $w = (w_n) \in \ell_q$. Also, for any $x = (x_n) \in \ell_p$,

$$(\Phi w)(x) = \sum_{n=1}^{\infty} x_n w_n = \sum_{n=1}^{\infty} x_n f(e_n) = f(x).$$

That is, $\Phi w = f$ and so Φ is surjective. Furthermore,

$$\|w\|_q = \left(\sum_{k=1}^{\infty} |w_k|^q\right)^{1/q} \le \|f\| = \|\Phi w\|. \qquad (\star\star)$$

(ii) Φ is linear: Let $y = (y_n), z = (z_n) \in \ell_q$ and $\beta \in \mathbb{F}$. Then, for any $x = (x_n) \in \ell_p$,

$$\begin{aligned}[\Phi(\beta y + z)](x) &= \sum_{j=1}^{\infty} x_j(\beta y_j + z_j) = \beta \sum_{j=1}^{\infty} x_j y_j + \sum_{j=1}^{\infty} x_j z_j \\ &= \beta(\Phi y)(x) + (\Phi z)(x) = [\beta \Phi y + \Phi z](x).\end{aligned}$$

Hence, $\Phi(\beta y + z) = \beta \Phi y + \Phi z$, which proves linearity of Φ.

(iii) Φ is an isometry: This follows from $(\star)$ and $(\star\star)$. □

The following result is an immediate consequence of Theorem 4.1.5.

4.2.1 Theorem
Every linear functional on a finite-dimensional normed linear space is continuous.

4.2.4 Proposition
Let X be a normed linear space over $\mathbb{F}$. If X is finite-dimensional, then X^* is also finite-dimensional and $\dim X = \dim X^*$.

Proof. Let $\{x_1, x_2, \cdots, x_n\}$ be a basis for X. For each $j = 1, 2, \ldots, n$, let x_j^* be defined by $x_j^*(x_k) = \delta_{jk}$ for $k = 1, 2, \ldots, n$. Then each x_j^* is a bounded linear functional on X. We show that $\{x_j^* \mid j = 1, 2, \ldots, n\}$ is a basis for X^*. Let x^* be an element of X^* and define $\lambda_j = x^*(x_j)$ for each $j = 1, 2, \ldots, n$. Then for any $k = 1, 2, \ldots, n$,

$$\left(\sum_{j=1}^{n} \lambda_j x_j^*\right)(x_k) = \sum_{j=1}^{n} \lambda_j \delta_{jk} = \lambda_k = x^*(x_k).$$

Hence $x^* = \sum_{j=1}^{n} \lambda_j x_j^*$; i.e., $\{x_j^* \mid j = 1, 2, \ldots, n\}$ spans X^*. It remains to show that $\{x_j^* \mid j = 1, 2, \ldots, n\}$ is linearly independent. Suppose that $\sum_{j=1}^{n} \alpha_j x_j^* = 0$. Then, for each $k = 1, 2, \ldots, n$,

$$0 = \left(\sum_{j=1}^{n} \alpha_j x_j^*\right)(x_k) = \sum_{j=1}^{n} \alpha_j \delta_{jk} = \alpha_k.$$

Hence $\{x_j^* \mid j = 1, 2, \ldots, n\}$ is a linearly independent set. ■

4.3 The Dual Space of a Hilbert Space

If $\mathcal{H}$ is a Hilbert space then bounded linear functionals on $\mathcal{H}$ assume a particularly simple form.

Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Choose and fix $y \in X \setminus \{0\}$. Define a map $f_y : X \to \mathbb{F}$ by $f_y(x) = \langle x, y\rangle$. We claim that f_y is a bounded (= continuous) linear functional on X.
Linearity: Let $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{F}$. Then

$$f_y(\alpha x_1 + \beta x_2) = \langle \alpha x_1 + \beta x_2, y\rangle = \alpha\langle x_1, y\rangle + \beta\langle x_2, y\rangle = \alpha f_y(x_1) + \beta f_y(x_2).$$

Boundedness: For any $x \in X$,

$$|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\| \quad \text{(by the CBS Inequality)}.$$

That is, f_y is bounded and $\|f_y\| \leq \|y\|$. Since

$$f_y(y) = \langle y, y\rangle = \|y\|^2 \quad \Rightarrow \quad \frac{|f_y(y)|}{\|y\|} = \|y\|,$$

we have that $\|f_y\| = \|y\|$.

The above observation simply says that each element y in an inner product space $(X, \langle\cdot,\cdot\rangle)$ determines a bounded linear functional on X.

The following theorem asserts that if $\mathcal{H}$ is a Hilbert space then the converse of this statement is true. That is, every bounded linear functional on a Hilbert space $\mathcal{H}$ is, in fact, determined by some element $y \in \mathcal{H}$.

4.3.1 Theorem
(Riesz-Fréchet Theorem). *Let $\mathcal{H}$ be a Hilbert space over $\mathbb{F}$. If $f : \mathcal{H} \to \mathbb{F}$ is a bounded linear functional on $\mathcal{H}$ (i.e., $f \in \mathcal{H}^*$) then there exists one and only one $y \in \mathcal{H}$ such that*

$$f(x) = \langle x, y\rangle \quad \text{for all } x \in \mathcal{H}.$$

Moreover, $\|f\| = \|y\|$.

Proof. Existence: If $f = 0$ then take $y = 0$. Assume that $f \neq 0$. Let $N = \{x \in \mathcal{H} \mid f(x) = 0\}$, the kernel of f. Then N is a closed proper subspace of $\mathcal{H}$. By Corollary 3.4.5 there exists $z \in N^{\perp} \setminus \{0\}$. Without loss of generality, $\|z\| = 1$. Put $u = f(x)z - f(z)x$. Then

$$f(u) = f(f(x)z - f(z)x) = f(x)f(z) - f(z)f(x) = 0, \qquad \text{i.e., } u \in N.$$

Thus,
$$0 = \langle u, z\rangle = \langle f(x)z - f(z)x, z\rangle = f(x)\langle z, z\rangle - f(z)\langle x, z\rangle = f(x) - f(z)\langle x, z\rangle,$$
whence, $f(x) = f(z)\langle x, z\rangle = \langle x, \overline{f(z)}z\rangle$. Take $y = \overline{f(z)}z$. Then $f(x) = \langle x, y\rangle$.

Uniqueness: Assume that $f(x) = \langle x, y\rangle = \langle x, y_0\rangle$ for each $x \in \mathcal{H}$. Then

$$0 = \langle x, y\rangle - \langle x, y_0\rangle = \langle x, y - y_0\rangle, \quad \text{for all} \quad x \in \mathcal{H}.$$

In particular, take $x = y - y_0$,

$$0 = \langle y - y_0, y - y_0\rangle = \|y - y_0\|^2 \quad \Rightarrow \quad y - y_0 = 0 \quad \Rightarrow \quad y = y_0.$$

Finally, for any $x \in \mathcal{H}$,

$$|f(x)| = |\langle x, y\rangle| \leq \|x\|\|y\| \quad \text{(by the CBS Inequality).}$$

That is, $\|f\| \leq \|y\|$. Since

$$f(y) = \langle y, y\rangle = \|y\|^2 \quad \Rightarrow \quad \frac{|f(y)|}{\|y\|} = \|y\|,$$

we have that $\|f\| = \|y\|$. ∎

4.3.1 Remarks

(a) The element $y \in \mathcal{H}$ as advertised in Theorem 4.3.1 is called the **representer** of the functional f.

(b) The conclusion of Theorem 4.3.1 may fail if $(X, \langle\cdot,\cdot\rangle)$ is an *incomplete* inner product space.

4.3.2 Example
Let X be the linear space of polynomials over $\mathbb{R}$ with the inner product defined by

$$\langle x, y\rangle = \int_0^1 x(t)y(t)\,dt.$$

For each $x \in X$, let $f : X \to \mathbb{R}$ be defined by

$$f(x) = x(0), \quad \text{(i.e., } f \text{ is a point evaluation at } 0).$$

Then f is a bounded linear functional on X. We show that there does not exist an element $y \in X$ such that

$$f(x) = \langle x, y\rangle \quad \text{for all } \ x \in X.$$

Assume that such an element exists. Then for each $x \in X$

$$f(x) = x(0) = \int_0^1 x(t)y(t)\,dt.$$

Since for any $x \in X$ the functional f maps the polynomial $tx(t)$ onto zero, we have that

$$\int_0^1 tx(t)y(t)\,dt = 0 \quad \text{for all } x \in X.$$

In particular, with $x(t) = ty(t)$ we have that

$$\int_0^1 t^2[y(t)]^2\,dt = 0;$$

whence $y \equiv 0$, i.e. y is the zero polynomial. Hence, for all $x \in X$,

$$f(x) = \langle x, y\rangle = \langle x, 0\rangle = 0.$$

That is, f is the zero functional, a contradiction since f maps a polynomial with a nonzero constant term to that constant term. □

4.3.2 Theorem
Let $\mathcal{H}$ be a Hilbert space.

(a) If $\mathcal{H}$ is a real Hilbert space, then $\mathcal{H} \cong \mathcal{H}^*$.

(b) If $\mathcal{H}$ is a complex Hilbert space, then $\mathcal{H}$ is isometrically embedded onto $\mathcal{H}^*$.

Proof. For each $y \in \mathcal{H}$, define $\Lambda : \mathcal{H} \to \mathcal{H}^*$ by

$$\Lambda y = f_y, \text{ where } f_y(x) = \langle x, y\rangle \text{ for each } x \in \mathcal{H}.$$

Let $y, z \in \mathcal{H}$. Then, for each $x \in \mathcal{H}$,

$$y \neq z \iff \langle x, y\rangle \neq \langle x, z\rangle \iff f_y \neq f_z \iff \Lambda y \neq \Lambda z.$$

Hence, Λ is well defined and one-to-one. Furthermore, since

$$\|\Lambda y\| = \|f_y\| = \|y\|$$

for each $y \in \mathcal{H}$, Λ is an isometry.
If $f \in \mathcal{H}^*$, then by Riesz-Fréchet Theorem (Theorem 4.3.1), there is a unique $y_f \in \mathcal{H}$ such that $f(x) = \langle x, y_f\rangle$. Hence $\Lambda y_f = f$, i.e., Λ is onto.
The inverse Λ^{-1} of Λ is given by

$$\Lambda^{-1} f = y, \text{ where } f(x) = \langle x, y\rangle \text{ for all } x \in \mathcal{H}.$$

Since $\|\Lambda^{-1} f\| = \|y\| = \|f\|$ for each $f \in \mathcal{H}^*$, Λ^{-1} is bounded (in fact an isometry).
If $\mathcal{H}$ is real, then Λ is linear. Indeed, for all $x, y, z \in \mathcal{H}$ and all $\alpha \in \mathbb{R}$, then

$$\begin{aligned} (\Lambda(\alpha y + z))(x) &= f_{(\alpha y + z)}(x) = \langle x, \alpha y + z\rangle \\ &= \langle x, \alpha y\rangle + \langle x, z\rangle = \alpha\langle x, y\rangle + \langle x, z\rangle \\ &= (\alpha\Lambda y)(x) + (\Lambda z)(x) = (\alpha\Lambda y + \Lambda z)(x). \end{aligned}$$

Hence, $\Lambda(\alpha y + z) = \alpha\Lambda y + \Lambda z$.

If $\mathcal{H}$ is complex, then Λ is conjugate-linear; i.e., $\Lambda(\alpha y + z) = \overline{\alpha}\Lambda y + \Lambda z$. ∎

4.3.3 Exercise

[1] Let X and Y be linear spaces over the same field $\mathbb{F}$ and $T \in \mathcal{L}(X, Y)$.

(a) Show that ran(T) is a linear subspace of Y and ker(T) is a linear subspace of X.

(b) T is one-to-one if and only if ker$(T) = \{0\}$.

[2] Let X and Y be normed linear spaces over the same field $\mathbb{F}$. Show that if $T \in \mathcal{B}(X, Y)$ then ker(T) is a closed linear subspace of X.

[3] Show that the mapping $R : \ell_2 \to \ell_2$ given by

$$Rx = R(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots)$$

is a bounded linear operator on ℓ_2 and find its norm. The operator R is called the **right-shift operator**.

[4] Fix $x \in \mathcal{C}[-\pi, \pi]$. Define an operator $M_x : L_2[-\pi, \pi] \to L_2[-\pi, \pi]$ by

$$M_x y = xy \quad \text{where} \quad (M_x y)(t) = x(t)y(t) \text{ for all } t \in [-\pi, \pi].$$

Show that M_x is a bounded linear operator on $L_2[-\pi, \pi]$. The operator M_x is called a **multiplication operator**. The function x is the **symbol** of M_x.

[5] Fix $x = (x_1, x_2, \ldots) \in \ell_\infty$. Define an operator $M_x : \ell_2 \to \ell_2$ by

$$M_x y = M_x(y_1, y_2, \ldots) = (x_1 y_1, x_2 y_2, \ldots).$$

Show that M_x is a bounded linear operator on ℓ_2 and $\|M_x\| = \|x\|_\infty$.

[6] Show that if S is a subset of a Hilbert space $\mathcal{H}$ that is dense in $\mathcal{H}$ and T_1 and T_2 are operators such that $T_1 x = T_2 x$ for all $x \in S$, then $T_1 = T_2$.

[7] Find the general form of a bounded linear functional on $L_2[-\pi, \pi]$.

[8] Find the general form of a bounded linear functional on ℓ_2.

[9] Define $f : \ell_2 \to \mathbb{C}$ by

$$f(x) = \sum_{n=1}^{\infty} \frac{x_n}{n^2}, \quad \text{where } x = (x_1, x_2, \ldots) \in \ell_2.$$

Show that f is a bounded linear functional on ℓ_2 and that $\|f\| = \dfrac{\pi^2}{3\sqrt{10}}$.

Chapter 5

The Hahn-Banach Theorem and its Consequences

The Hahn-Banach theorem is one of the most important results in functional analysis since it is required for many other results and also because it encapsulates the spirit of analysis. The theorem was proved independently by Hahn in 1927 and by Banach in 1929 although Helly proved a less general version much earlier in 1912. Intersetingly, the complex version was proved only in 1938 by Bohnenblust and Sobczyk. We prove the Hahn-Banach theorem using Zorn's lemma which is equivalent to the axiom of choice. It should be noted, however, that the Hahn-Banach is in fact strictly weaker than the axiom of choice. Since the publication of the original result, there have been many versions published in different settings but that is beyond the scope of this course.

5.1 Introduction

In this chapter the Hahn-Banach theorem is established along with a few of its many consequences. Before doing that, we briefly discuss Zorn's Lemma.

5.1.1 Definition
A binary relation $\preccurlyeq$ on a set P is a **partial order** *if it satisfies the following properties: For all $x,\ y,\ z \in P$,*

(i) $\preccurlyeq$ is reflexive: $x \preccurlyeq x$;

(ii) $\preccurlyeq$ is antisymmetric: if $x \preccurlyeq y$ and $y \preccurlyeq x$, then $x = y$;

(iii) $\preccurlyeq$ is transitive: if $x \preccurlyeq y$ and $y \preccurlyeq z$, then $x \preccurlyeq z$.

A partially ordered set is a pair $(P, \preccurlyeq)$, where P is a set $\preccurlyeq$ is a partial order on P.

5.1.2 Examples

[1] Let $P = \mathbb{R}$ and take $\preccurlyeq$ to be $\leq$, the usual less than or equal to relation on $\mathbb{R}$.

[2] Let $P = \mathcal{P}(X)$ the power set of a set X and take $\preccurlyeq$ to be $\subseteq$, the usual set inclusion relation.

[3] Let $P = \mathcal{C}[0, 1]$, the space of continuous real-valued functions on the interval $[0, 1]$ and take $\preccurlyeq$ to be the relation $\leq$ given by $f \leq g$ if and only if $f(x) \leq g(x)$ for each $x \in [0, 1]$.

5.1.3 Definition
Let C be a subset of a partially ordered set $(P, \preccurlyeq)$.

(i) An element $u \in P$ is an **upper bound** *of C if $x \preccurlyeq u$ for every $x \in C$;*

(ii) An element $m \in C$ is said to be **maximal** *if for any element $y \in C$, the relation $m \preccurlyeq y$ implies that $m = y$.*

5.1.4 Definition

Let $(P, \preccurlyeq)$ be a partially ordered set and $x, y \in P$. We say that x and y are **comparable** *if either $x \preccurlyeq y$ or $y \preccurlyeq x$. Otherwise, x and y are* **incomparable**.
A partial order $\preccurlyeq$ is called a **linear order** *(or a* **total order***) if any two elements of P are comparable. In this case we say that $(P, \preccurlyeq)$ is a linearly ordered (or totally ordered) set. A linearly ordered set is also called a* **chain**.

5.1.1 Theorem

(Zorn's Lemma). *Let $(P, \preccurlyeq)$ be a partially ordered set. If each linearly ordered subset of P has an upper bound, then P has a maximal element.*

5.1.5 Definition

Let M and N be linear subspaces of a linear space X with $M \subset N$ and let f be a linear functional on M. A linear functional F on N is called an **extension** *of f to N if $F|_M = f$; i.e., $F(x) = f(x)$ for each $x \in M$.*

5.1.6 Definition

Let X be a linear space. A function $p : X \to \mathbb{R}$ is called a **sublinear functional** *provided that:*

(i) $p(x + y) \leq p(x) + p(y)$ for $x, y \in X$;

(ii) $p(\lambda x) = \lambda p(x), \quad \lambda \geq 0.$

Observe that any linear functional or any norm on X is a sublinear functional. Also, every positive scalar multiple of a sublinear functional is again a sublinear functional.

5.1.7 Lemma

Let M be a proper linear subspace of a real linear space X, $x_0 \in X \setminus M$, and $N = \{m + \alpha x_0 \mid m \in M, \alpha \in \mathbb{R}\}$. Suppose that $p : X \to \mathbb{R}$ a sublinear functional defined on X, and f a linear functional defined on M such that $f(x) \leq p(x)$ for all $x \in M$. Then f can be extended to a linear functional F defined on N such that $F(x) \leq p(x)$ for all $x \in N$.

Proof. Since $x_0 \notin M$, it is readily verified that $N = M \oplus \text{lin}\{x_0\}$. Therefore each $x \in N$ has a unique representation of the form $x = m + \lambda x_0$ for some unique $m \in M$ and $\lambda \in \mathbb{R}$. Define a functional F on N by

$$F(x) = f(m) + \lambda c \quad \text{for some } c \in \mathbb{R}.$$

This functional F is well defined since each $x \in N$ is uniquely determined. Furthermore F is linear and $F(y) = f(y)$ for all $y \in M$. It remains to show that it is possible to choose a $c \in \mathbb{R}$ such that for each $x \in N$,

$$F(x) \leq p(x).$$

Let $y_1,\ y_2 \in M$. Since $f(y) \leq p(y)$ for all $y \in M$, we have that

$$\begin{aligned} f(y_1) - f(y_2) = f(y_1 - y_2) &\leq p(y_1 - y_2) = p(y_1 + x_0 - y_2 - x_0) \\ &\leq p(y_1 + x_0) + p(-y_2 - x_0) \\ \iff \quad -f(y_2) - p(-y_2 - x_0) &\leq p(y_1 + x_0) - f(y_1). \end{aligned}$$

Therefore, for fixed $y_1 \in M$, the set of real numbers $\{-f(y_2) - p(-y_2 - x_0) \mid y_2 \in M\}$ is bounded above and hence has the least upper bound. Let

$$a = \sup\{-f(y_2) - p(-y_2 - x_0) \mid y_2 \in M\}.$$

Similarly, for fixed $y_2 \in M$, the set $\{p(y_1 + x_0) - f(y_1) \mid y_1 \in M\}$ is bounded below. Let

$$b = \inf\{p(y_1 + x_0) - f(y_1) \mid y_1 \in M\}.$$

Of course, $a \le b$. Hence there is a real number c such that $a \le c \le b$. Therefore

$$-f(y) - p(-y - x_0) \le c \le p(y + x_0) - f(y)$$

for each $y \in M$.

Now, let $x = y + \lambda x_0 \in N$. If $\lambda = 0$, then $F(x) = f(x) \le p(x)$. If $\lambda > 0$, then

$$\begin{aligned} c \le p\left(\frac{y}{\lambda} + x_0\right) - f(y/\lambda) \quad &\Longleftrightarrow \quad \lambda c \le p(y + \lambda x_0) - f(y) \\ &\Longleftrightarrow \quad f(y) + \lambda c \le p(y + \lambda x_0) \\ &\Longleftrightarrow \quad F(x) \le p(x). \end{aligned}$$

Finally, if $\lambda < 0$, then

$$\begin{aligned} -f(y/\lambda) - p\,(-y/\lambda - x_0) \le c \quad &\Longleftrightarrow \quad -\frac{1}{\lambda} f(y) + \frac{1}{\lambda} p(y + \lambda x_0) \le c \\ &\Longleftrightarrow \quad f(y) - p(y + \lambda x_0) \le -\lambda c \\ &\Longleftrightarrow \quad f(y) + \lambda c \le p(y + \lambda x_0) \\ &\Longleftrightarrow \quad F(x) \le p(x). \end{aligned}$$

We now state our main result. What this theorem essentially states is that there are enough bounded (continuous) linear functionals for a rich theory and as mentioned before it is used ubiquitously thoughout functional analysis.

5.1.2 Theorem

(Hahn-Banach Extension Theorem for real linear spaces). *Let p be a sublinear functional on a real linear space X and let M be a subspace of X. If f is a linear functional on M such that $f(x) \le p(x)$ for all $x \in M$, then f has an extension F to X such that $F(x) \le p(x)$ for all $x \in X$.*

Proof. Let $\mathcal{F}$ be the set of all pairs (M_α, f_α), where M_α is a subspace of X containing M, $f_\alpha(y) = f(y)$ for each $y \in M$, i.e., f_α is an extension of f, and $f_\alpha(x) \le p(x)$ for all $x \in M_\alpha$. Clearly, $\mathcal{F} \neq \emptyset$ since $(M, f) \in \mathcal{F}$. Define a partial order on $\mathcal{F}$ by:

$$(M_\alpha, f_\alpha) \prec (M_\beta, f_\beta) \iff M_\alpha \subset M_\beta \text{ and } f_\beta|_{M_\alpha} = f_\alpha.$$

Let $\mathcal{T}$ be a totally ordered subset of $\mathcal{F}$ and let

$$X_0 = \bigcup\{M_\alpha \mid (M_\alpha, f_\alpha) \in \mathcal{T}\}.$$

Then X_0 is a linear subspace of X since $\mathcal{T}$ is totally ordered. Define a functional $f_0 : X_0 \to \mathbb{R}$ by $f_0(x) = f_\alpha(x)$ for all $x \in M_\alpha$. Then f_0 is well-defined, since if $x \in M_\alpha \cap M_\beta$, then $x \in M_\alpha$ and $x \in M_\beta$. Therefore $f_0(x) = f_\alpha(x)$ and $f_0(x) = f_\beta(x)$. By total ordering of $\mathcal{T}$, either f_α extends f_β or vice versa. Hence $f_\alpha(x) = f_\beta(x)$. It is clear that f_0 is a linear extension of f. Furthermore $f_0(x) \le p(x)$ for all $x \in X_0$ and $(M_\alpha, f_\alpha) \prec (X_0, f_0)$ for all $(M_\alpha, f_\alpha) \in \mathcal{T}$, i.e., (X_0, f_0) is an upper bound for $\mathcal{T}$. By Zorn's lemma, $\mathcal{F}$ has a maximal element (X_1, F). To complete the proof, it suffices to show that $X_1 = X$. If $X_1 \neq X$, then choose $y \in X \setminus X_1$. By Lemma 5.1.7, we can extend F to a linear functional $\widetilde{F}$ defined on $\widetilde{X} = X_1 \oplus \operatorname{lin}\{y\}$ and extending f such that $\widetilde{F}(x) \le p(x)$ for all $x \in \widetilde{X}$. Thus $(\widetilde{X}, \widetilde{F}) \in \mathcal{F}$ and $(X_1, F) \prec (\widetilde{X}, \widetilde{F})$, which contradicts the maximality of (X_1, F). ■

5.1.8 Definition

A **seminorm** *p on a (complex) linear space X is a function $p : X \to \mathbb{R}$ such that for all $x, y \in X$ and $\lambda \in \mathbb{C}$,* *(i)* $p(x) \geq 0$ *and* $p(0) = 0$;

(ii) $p(x + y) \leq p(x) + p(y)$, *and*

(iii) $p(\lambda x) = |\lambda| p(x)$.

5.1.3 Theorem
(Hahn-Banach Extension Theorem for (complex) linear spaces). *Let X be a real or complex linear space, p be a seminorm on X and f a linear functional on a linear subspace M of X such that $|f(x)| \leq p(x)$ for all $x \in M$. Then there is a linear functional F on X such that $F|_M = f$ and $|F(x)| \leq p(x)$ for all $x \in X$.*

Proof. Assume first that X is a real linear space. Then, by Theorem 5.1.2, there is an extension F of f such that $F(x) \leq p(x)$ for all $x \in X$. Since

$$-F(x) = F(-x) \leq p(-x) = p(x) \text{ for all } x \in X,$$

it follows that $-p(x) \leq F(x) \leq p(x)$, or $|F(x)| \leq p(x)$ for all $x \in X$.

Now assume that X is a complex linear space. Then we may regard X as a real linear space by restricting the scalar field to $\mathbb{R}$. We denote the resulting real linear space by $X_{\mathbb{R}}$ and the real linear subspace by $M_{\mathbb{R}}$. Write f as $f = f_1 + i f_2$, where f_1 and f_2 are real linear functionals given by $f_1(x) = \Re e[f(x)]$ and $f_2(x) = \Im m[f(x)]$. Then f_1 is a real linear functional of $M_{\mathbb{R}}$ and $f_1(x) \leq |f(x)| \leq p(x)$ for all $x \in M_{\mathbb{R}}$. Hence, by Theorem 5.1.2, f_1 has a real linear extension F_1 such that $F_1(x) \leq p(x)$ for all $x \in X_{\mathbb{R}}$. Since,

$$\begin{aligned} f(ix) = if(x) \quad &\iff \quad f_1(ix) + if_2(ix) = if_1(x) - f_2(x) \\ &\iff \quad f_2(x) = -f_1(ix) \quad \text{and} \quad f_2(ix) = f_1(x), \end{aligned}$$

we can write $f(x) = f_1(x) - if_1(ix)$. Set

$$F(x) = F_1(x) - iF_1(ix) \quad \text{for all } x \in X.$$

Then F is a real linear extension of f and, for all $x, y \in X$,

$$\begin{aligned} F(x + y) &= F_1(x + y) - iF_1(ix + iy) = F_1(x) - iF_1(ix) + F_1(y) - iF_1(iy) \\ &= F(x) + F(y). \end{aligned}$$

For all $x \in X$,

$$F(ix) = F_1(ix) - iF_1(-x) = F_1(ix) + iF_1(x) = i(F_1(x) - iF_1(ix)) = iF(x).$$

If $\alpha = a + bi$ for $a,\ b \in \mathbb{R}$, and $x \in X$, then

$$\begin{aligned} F(\alpha x) = F((a + bi)x) &= F(ax + bix) = F(ax) + F(bix) \\ &= aF(x) + bF(ix) = aF(x) + biF(x) = (a + bi)F(x) \\ &= \alpha F(x). \end{aligned}$$

Hence, F is also complex linear. Finally, for $x \in X$, write $F(x) = |F(x)|e^{i\theta}$. Then, since $\Re e F = F_1$,

$$|F(x)| = F(x)e^{-i\theta} = F(xe^{-i\theta}) = F_1(xe^{-i\theta}) \leq p(xe^{-i\theta}) = |e^{-i\theta}|p(x) = p(x). \qquad \blacksquare$$

Suppose that M is a subspace of a normed linear space X and f is a bounded linear functional on M. If F is any extension of f to X, then the norm of F is at least as large as $\|f\|$ because

$$\begin{aligned} \|F\| = \sup\{\, |F(x)| : x \in X, \|x\| \leq 1 \,\} \quad &\geq \quad \sup\{\, |F(x)| : x \in M, \|x\| \leq 1 \,\} \\ &= \quad \sup\{\, |f(x)| : x \in M, \|x\| \leq 1 \,\} = \|f\|. \end{aligned}$$

The following consequence of the Hahn-Banach theorem states that it is always possible to find a bounded extension of f to the whole space which has the *same*, i.e., smallest possible, norm.

5.1.4 Theorem
(Hahn-Banach Extension Theorem for Normed linear spaces). *Let M be a linear subspace of the normed linear space $(X, \|\cdot\|)$ and let $f \in M^*$. Then there exists an extension $x^* \in X^*$ of f such that $\|x^*\| = \|f\|$.*

Proof. Define p on X by $p(x) = \|f\|\|x\|$. Then p is a seminorm on X and $|f(x)| \le p(x)$ for all $x \in M$. By Theorem 5.1.3, f has an extension F to X such that $|F(x)| \le p(x)$ for all $x \in X$. That is, $|F(x)| \le \|f\|\|x\|$. This shows that F is bounded and $\|F\| \le \|f\|$. Since F must have norm at least as large as $\|f\|$, $\|F\| = \|f\|$ and the result follows
with $x^* = F$. ■

5.2 Consequences of the Hahn-Banach Extension Theorem

5.2.1 Theorem
Let M be a linear subspace of a normed linear space $(X, \|\cdot\|)$ and $x \in X$ such that

$$d = d(x, M) := \inf_{y \in M} \|x - y\| > 0.$$

Then there is an $x^* \in X^*$ such that

(i) $\|x^*\| = 1$

(ii) $x^*(x) = d$

(iii) $x^*(m) = 0$ *for all* $m \in M$.

Proof. Let $Y = M + \operatorname{lin}\{x\} := \{m + \alpha x,\ m \in M,\ \alpha \in \mathbb{F}\}$. Then each y in Y is uniquely expressible in the form $y = m + \alpha x$ for some $m \in M$ and some scalar α. Indeed, if

$$y = m_1 + \alpha x = m_2 + \beta x$$

for some $m_1, m_2 \in M$ and some scalars α and β, then $(\beta - \alpha)x = m_1 - m_2 \in M$.
Claim: $\alpha = \beta$. If $\alpha \neq \beta$, then since M is a subspace

$$x = \frac{1}{\alpha - \beta}(m_1 - m_2) \in M,$$

a contradiction since $x \notin M$. Hence, $\alpha = \beta$ and consequently $m_1 = m_2$.
Define $f : Y \to \mathbb{F}$ by

$$f(y) = f(m + \alpha x) = \alpha d.$$

Since the scalar α is uniquely determined, f is well defined.
Claim: f is a bounded linear functional on Y.
Linearity: Let $y_1 = m_1 + \alpha_1 x$ and $y_2 = m_2 + \alpha_2 x$ be any two elements of Y and $\lambda \in \mathbb{F}$. Then

$$f(\lambda y_1 + y_2) = f((\lambda m_1 + m_2) + (\lambda\alpha_1 + \alpha_2)x) = (\lambda\alpha_1 + \alpha_2)d = \lambda\alpha_1 d + \alpha_2 d = \lambda f(y_1) + f(y_2).$$

Boundedness: Let $y = m + \alpha x \in Y$. Then

$$\|y\| = \|m + \alpha x\| = |\alpha| \left\|\frac{m}{\alpha} + x\right\| = |\alpha| \left\|x - \left(-\frac{m}{\alpha}\right)\right\| \ge |\alpha| d = |f(y)|,$$

i.e., $|f(y)| \leq \|y\|$ for all $y \in Y$. Thus, f is bounded and $\|f\| \leq 1$.
We show next that $\|f\| = 1$. By definition of infimum, given any $\epsilon > 0$, there is an element $m_\epsilon \in M$ such that $\|x - m_\epsilon\| < d + \epsilon$. Let $z = \dfrac{x - m_\epsilon}{\|x - m_\epsilon\|}$. Then $z \in Y$, $\|z\| = 1$ and

$$|f(z)| = \frac{d}{\|x - m_\epsilon\|} > \frac{d}{d + \epsilon}.$$

Since ϵ is arbitrary, it follows that $|f(z)| \geq 1$. Thus

$$1 \leq |f(z)| \leq \|f\|\|z\| = \|f\|.$$

Thus, $\|f\| = 1$.
It is clear that and $f(m) = 0$ for all $m \in M$ and $f(x) = d$. By Theorem 5.1.4, there is an $x^* \in X^*$ such that

$$x^*(y) = f(y) \text{ for all } y \in Y \text{ and } \|x^*\| = \|f\|.$$

Hence, $\|x^*\| = 1$ and $x^*(m) = 0$ for all $m \in M$ and $x^*(x) = d$. ■

5.2.1 Corollary
Let $(X, \|\cdot\|)$ be a normed linear space and $x_0 \in X \setminus \{0\}$. Then there exists an $x^* \in X^*$, such that $x^*(x_0) = \|x_0\|$ and $\|x^*\| = 1$.

Proof. Consider $M = \{0\}$. Since $x_0 \in X \setminus \{0\}$, it follows that $x_0 \notin M$ and so $d = d(x_0, M) = \|x_0\| > 0$. By Theorem 5.2.1, there is an $x^* \in X^*$ such that $x^*(x_0) = \|x_0\|$ and $\|x^*\| = 1$. ■

The following result asserts that X^* is big enough to distinguish the points of X.

5.2.2 Corollary
Let $(X, \|\cdot\|)$ be a normed linear space and $y,\ z \in X$. If $y \neq z$, then there exists an $x^* \in X^*$, such that $x^*(y) \neq x^*(z)$.

Proof. Consider $M = \{0\}$. Since $y \neq z$, it follows that $y - z \notin M$ and consequently

$$d = d(y - z, M) = \|y - z\| > 0.$$

By Theorem 5.2.1, there is an $x^* \in X^*$ such that $x^*(y - z) = d > 0$. Hence $x^*(y) \neq x^*(z)$. ■

5.2.3 Corollary
For each x in a normed linear space $(X, \|\cdot\|)$,

$$\|x\| = \sup\{|x^*(x)| \mid x^* \in X^*,\ \|x^*\| = 1\}.$$

Proof. If $x = 0$, then the result holds vacuously. Assume $x \in X \setminus \{0\}$. For any $x^* \in X^*$ with $\|x^*\| = 1$,

$$|x^*(x)| \leq \|x^*\|\|x\| = \|x\|.$$

Hence, $\sup\{|x^*(x)| \mid x^* \in X^*,\ \|x^*\| = 1\} \leq \|x\|$.
By Corollary 5.2.1, there is a $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(x) = \|x\|$. Therefore

$$\|x\| = |x^*(x)| \leq \sup\{|x^*(x)| \mid x^* \in X^*,\ \|x^*\| = 1\},$$

whence $\|x\| = \sup\{|x^*(x)| \mid x^* \in X^*,\ \|x^*\| = 1\}$. ■

5.2.2 Theorem
If the dual X^* of a normed linear space $(X, \|\cdot\|)$ is separable, then X is also separable.

Proof. Let $S = S(X^*) = \{x^* \in X^* \mid \|x^*\| = 1\}$. Since any subset of a separable space is separable, S is separable. Let $\{x_n^* \mid n \in \mathbb{N}\}$ be a countable dense subset of S. Since $x_n^* \in S$ for each $n \in \mathbb{N}$, we have that $\|x_n^*\| = 1$. Hence, for each $n \in \mathbb{N}$ there is an element $x_n \in X$ such that $\|x_n\| = 1$ and $|x_n^*(x_n)| > \frac{1}{2}$. (Otherwise $|x_n^*(x)| \leq \frac{1}{2}$ for all $x \in X$ and so $\|x_n^*\| \leq \frac{1}{2}$, a contradiction.) Let

$$M = \overline{\mathrm{lin}}(\{x_n \mid n \in \mathbb{N}\}).$$

Then M is separable since M contains a countable dense subset comprising all linear combinations of the x_n's with coefficients whose real and imaginary parts are rational.
Claim: $M = X$. If $M \neq X$, then there is an element $x_0 \in X \setminus M$ such that $d = d(x_0, M) > 0$. By Theorem 5.2.1, there is an $x^* \in X^*$ such that $\|x^*\| = 1$, i.e. $x^* \in S$, and $x^*(y) = 0$ for all $y \in M$. In particular, $x^*(x_n) = 0$ for all $n \in \mathbb{N}$. Now, for each $n \in \mathbb{N}$,

$$\frac{1}{2} < |x_n^*(x_n)| = |x_n^*(x_n) - x^*(x_n)| = |(x_n^* - x^*)(x_n)| \leq \|x_n^* - x^*\|.$$

But this contradicts the fact that the set $\{x_n^* \mid n \in \mathbb{N}\}$ is dense in S. Hence $M = X$ and, consequently, X is separable. ∎

The converse of Theorem 5.2.2 does not hold. That is, if X separable, it does not follow that its dual X^* is also be separable. Take, for example, ℓ_1. Its dual is (isometrically isomorphic to) ℓ_∞. The space ℓ_1 is separable whereas ℓ_∞ is not. This also shows that the dual of ℓ_∞ is not (isometrically isomorphic to) ℓ_1.

5.2.4 Definition
Let M be a subset of a normed linear space X. The **annihilator** *of M, denoted by $M^\perp$, is the set*

$$M^\perp = \{x^* \in X^* \mid x^*(y) = 0 \ \textit{for all } y \in M\}.$$

It is easy to show that $M^\perp$ is a closed linear subspace of X^*.

5.2.3 Theorem
Let M be a linear subspace of a normed linear space X. Then

$$X^*/M^\perp \cong M^*.$$

Proof. Define $\Phi : X^*/M^\perp \to M^*$ by

$$\Phi(x^* + M^\perp)(m) = x^*(m) \ \text{ for all } x^* \in X^* \text{ and all } m \in M.$$

We show that Φ is well-defined. Let $x^*,\ y^* \in X^*$ such that $x^* + M^\perp = y^* + M^\perp$. Then $x^* - y^* \in M^\perp$ and so $x^*(m) = y^*(m)$ for all $m \in M$. Thus $\Phi(x^* + M^\perp) = \Phi(y^* + M^\perp)$; i.e., Φ is well-defined. Clearly, $\Phi(x^* + M^\perp)$ is a linear functional on M.
We show that Φ is linear. Let $x^*,\ y^* \in X^*$ and $\lambda \in \mathbb{F}$. Then, for all $m \in M$,

$$\begin{aligned}\Phi((x^* + M^\perp) + \lambda(y^* + M^\perp))(m) &= \Phi(x^* + \lambda y^* + M^\perp)(m) = (x^* + \lambda y^*)(m)\\ &= x^*(m) + \lambda y^*(m)\\ &= \Phi(x^* + M^\perp)(m) + \lambda\Phi(y^* + M^\perp)(m)\\ &= \left(\Phi(x^* + M^\perp) + \lambda\Phi(y^* + M^\perp)\right)(m).\end{aligned}$$

Hence, $\Phi((x^* + M^\perp) + \lambda(y^* + M^\perp)) = \Phi(x^* + M^\perp) + \lambda\Phi(y^* + M^\perp)$.
We now show that Φ is surjective. Let $y^* \in M^*$. Then, by Theorem 5.1.4, there is an $x^* \in X^*$ such that $y^*(m) = x^*(m)$ for all $m \in M$ and $\|y^*\| = \|x^*\|$. Hence, for all $m \in M$,

$$\Phi(x^* + M^\perp)(m) = x^*(m) = y^*(m).$$

Thus $\Phi(x^* + M^\perp) = y^*$. Furthermore,

$$\|x^* + M^\perp\| \le \|x^*\| = \|y^*\| = \|\Phi(x^* + M^\perp)\|.$$

But for any $y^* \in M^\perp$, $x^* + M^\perp = (x^* + y^*) + M^\perp$. Hence, for all $m \in M$,

$$|\Phi(x^* + M^\perp)(m)| = |(x^* + y^*)(m)| \le \|x^* + y^*\| \|m\|.$$

That is, $\Phi(x^* + M^\perp)$ is a bounded linear functional on M and $\|\Phi(x^* + M^\perp)\| \le \|x^* + y^*\|$ for all $y^* \in M^\perp$. Thus

$$\|\Phi(x^* + M^\perp)\| \le \inf_{y^* \in M^\perp} \|x^* + y^*\| = \|x^* + M^\perp\|.$$

It now follows that $\|\Phi(x^* + M^\perp)\| = \|x^* + M^\perp\|$. ■

5.3 Bidual of a normed linear space and Reflexivity

Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$ and $x \in X$. Define a functional $\Phi_x : X^* \to \mathbb{F}$ by

$$\Phi_x(x^*) = x^*(x) \quad \text{for all } x^* \in X^*.$$

It is easy to verify that Φ_x is linear and for each $x^* \in X^*$,

$$|\Phi_x(x^*)| = |x^*(x)| \le \|x^*\| \|x\|.$$

That is, Φ_x is bounded and $\|\Phi_x\| \le \|x\|$. By Corollary 5.2.3,

$$\|x\| = \sup\{|x^*(x)| \mid x^* \in X^*,\ \|x^*\| = 1\} = \sup\{|\Phi_x(x)| \mid x^* \in X^*,\ \|x^*\| = 1\} = \|\Phi_x\|.$$

This shows that Φ_x is a bounded linear functional on X^*, i.e., $\Phi_x \in (X^*)^* = X^{**}$ and $\|\Phi_x\| = \|x\|$. The space X^{**} is called the **second dual space** or **bidual space** of X. It now follows that we can define a map $J_X : X \to X^{**}$ by

$$J_X x = \Phi_x, \quad \text{for } x \in X, \quad \text{that is, } (J_X x)(x^*) = x^*(x) \ \text{ for } x \in X \ \text{ and } x^* \in X^*.$$

It is easy to show that J_X is linear and $\|x\| = \|\Phi_x\| = \|J_X x\|$. That is, J_X is a linear isometry of X into its bidual X^{**}. The map J_X as defined above is called the **canonical** or **natural embedding** of X into its bidual X^{**}. This shows that we can identify X with the subspace $J_X X = \{J_X x \mid x \in X\}$ of X^{**}.

5.3.1 Definition
Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$. Then X is said to be **reflexive** *if the canonical embedding $J_X : X \to X^{**}$ of X into its bidual X^{**} is surjective. In this case $X \cong X^{**}$.*

If X is reflexive, we customarily write $X = X^{**}$. The equality simply means that X is isometrically isomorphic to X^{**}. Reflexivity of X basically means that each bounded linear functional on X^* is an evaluation functional. Since dual spaces are complete, a reflexive normed linear space is necessarily a Banach space. It is therefore appropriate to speak of a reflexive Banach space rather than a reflexive normed linear space.

5.3.1 Theorem

(1) Every finite-dimensional normed linear space is reflexive.

(2) A closed linear subspace of a reflexive space is reflexive.

Proof. (1). If $\dim X < \infty$, then Proposition 4.2.4 implies that $\dim X = \dim X^* = \dim X^{**}$. Since $J_X X$ is isometrically isomorphic to X, $\dim(J_X X) = \dim X = \dim X^{**}$. Since $J_X X$ is a subspace of X^{**}, it must equal to X^{**}.

(2). Let X be reflexive and M a closed linear subspace of X. Given $y^{**} \in M^{**}$, it must be shown that there exists $y \in M$ such that $J_M y(y^*) = y^*(y) = y^{**}(y^*)$ for all $y^* \in M^*$. Define a functional ψ on X^* by

$$\psi(x^*) = y^{**}(x^*|_M), \quad x^* \in X^*.$$

Clearly, ψ is linear and

$$|\psi(x^*)| \le \|y^{**}\|\|x^*|_M\| \le \|y^{**}\|\|x^*\|$$

so $\psi \in X^{**}$. By reflexivity of X, there exists $y \in X$ such that $J_X y = \psi$. That is, $\psi(x^*) = x^*(y)$ for each $x^* \in X^*$. If $y \notin M$, then by Theorem 5.2.1, there exists an $x_0^* \in X^*$ such that $x_0^*(y) \neq 0$ and $x_0^*(m) = 0$ for all $m \in M$. Then

$$0 \neq x_0^*(y) = \psi(x_0^*) = y^{**}(x_0^*|_M) = y^{**}(0) = 0$$

which is absurd. Thus $y \in M$ and $x^*(y) = \psi(x^*) = y^{**}(x^*|_M), \quad x^* \in X^*$. By Theorem 5.1.4, every $y^* \in M^*$ is of the form $y^* = x^*|_M$ for some $x^* \in X^*$. Thus
$(J_M y)(y^*) = y^*(y) = y^{**}(y^*), \quad y^* \in M^*$, and the proof is complete. ■

5.3.2 Theorem

A Banach space X is reflexive if and only if its dual X^* is reflexive.

Proof. Assume that X is reflexive. Let $J_X : X \to X^{**}$ and $J_{X^*} : X^* \to (X^*)^{**} = X^{***}$ be the canonical embeddings of X and X^* respectively. We must show that J_{X^*} is surjective. To that end, let $x^{***} \in X^{***} = (X^{**})^*$ and consider the following diagram:

$$X \xrightarrow{J_X} X^{**} \xrightarrow{x^{***}} \mathbb{F}.$$

Define a functional x^* on X by $x^* = x^{***}J_X$. It is obvious that x^* is linear since both x^{***} and J_X are linear. Also, for each $x \in X$,

$$|x^*(x)| = |x^{***}J_X(x)| \le \|x^{***}\|\|J_X x\| = \|x^{***}\|\|x\|.$$

i.e., x^* is bounded and $\|x^*\| \le \|x^{***}\|$. Hence $x^* \in X^*$. We now show that $J_{X^*}(x^*) = x^{***}$. Let $x^{**} \in X^{**}$ be any element. Since J_X is surjective, there is an $x \in X$ such that $x^{**} = J_X x$. Hence

$$x^{***}(x^{**}) = x^{***}(J_X x) = x^*(x) = J_X x(x^*) = J_{X^*}x^*(J_X x) = J_{X^*}x^*(x^{**}),$$

and therefore $J_{X^*}x^* = x^{***}$. That is, J_{X^*} is surjective.

Assume that X^* is reflexive. Then the canonical embedding $J_{X^*} : X^* \to X^{***}$ is surjective. If $J_X X \neq X^{**}$, let $x^{**} \in X^{**} \setminus J_X X$. Since $J_X X$ is a closed linear subspace of X^{**}, it follows from Theorem 5.2.1 that there is a functional $\phi \in X^{***}$ such that $\|\phi\| = 1$, $\phi(x^{**}) = d(x^{**}, J_X X)$, and $\phi(J_X x) = 0$ for all $x \in X$. Since J_{X^*} is surjective, there is an $x^* \in X^*$ such that $J_{X^*}x^* = \phi$. Hence, for each $x \in X$,

$$0 = \phi(J_X x) = J_{X^*}x^*(J_X x) = (J_X x)(x^*) = x^*(x),$$

i.e., $x^*(x) = 0$ for all $x \in X$. This implies that $x^* = 0$. But then $0 = J_{X^*}x^* = \phi$, a contradiction since $\phi \neq 0$. Hence $J_X X = X^{**}$; i.e., J_X is surjective. ■

5.3.2 Exercise

Show that if X is a non-reflexive normed linear space, then the natural inclusions $X \subset X^{**} \subset X^{****} \subset \cdot$ and $X^* \subset X^{***} \subset X^{*****} \subset \cdot$ are all strict.

We showed earlier (Theorem 5.2.2) that if the dual space X^* of a normed linear space X is separable, then X is also separable, but not conversely. However, if X is reflexive, then the converse holds.

5.3.3 Theorem
If X is a reflexive separable Banach space, then its dual X^* is also separable.

Proof. Since X is reflexive and separable, its bidual $X^{**} = J_X X$ is also separable. Hence, by Theorem 5.2.2, X^* is separable. ■

5.3.3 Examples

(1) For $1 < p < \infty$, the sequence space ℓ_p is reflexive.

(2) The spaces c_0, c, ℓ_1, and ℓ_∞ are non-reflexive.

(3) Every Hilbert space $\mathcal{H}$ is reflexive.

5.4 The Adjoint Operator

5.4.1 Definition
Let X and Y be normed linear spaces and $T \in \mathcal{B}(X, Y)$. The **Banach space adjoint** *(or simply* **adjoint***) of T, denoted by T^*, is the operator $T^* : Y^* \to X^*$ defined by*

$$(T^* y^*)(x) = y^*(Tx) \text{ for all } y^* \in Y^* \text{ and all } x \in X.$$

The following diagram helps make sense of the above definition.

$$X \xrightarrow{T} Y$$

$$X^* \xleftarrow{T^*} Y^*.$$

It is straightforward to show that for any $y^* \in Y^*$, $T^* y^*$ is a linear functional on X. Furthermore, for any $y^* \in Y^*$ and $x \in X$

$$|T^* y^*(x)| = |y^*(Tx)| \leq \|y^*\| \|T\| \|x\|,$$

i.e., $T^* y^*$ is a bounded linear functional on X and $\|T^* y^*\| \leq \|T\| \|y^*\|$.

5.4.2 Example
Let $X = \ell_1 = Y$ and define $T : \ell_1 \to \ell_1$ by

$$Tx = T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, x_3, \dots), \text{ where } x = (x_n) \in \ell_1,$$

the right-shift operator. Then the adjoint of T is $T^* : \ell_\infty \to \ell_\infty$ is given by

$$T^* y = T^*(y_1, y_2, y_3, \dots) = (y_2, y_3, \dots), \text{ where } y = (y_n) \in \ell_\infty,$$

the left-shift operator.

5.4.1 Theorem
Let X and Y be normed linear spaces over $\mathbb{F}$ and let $T \in \mathcal{B}(X, Y)$.

(a) T^* is a bounded linear operator on Y^*.

(b) The map $\Lambda : \mathcal{B}(X, Y) \longrightarrow \mathcal{B}(Y^*, X^*)$ defined by $\Lambda T = T^*$ is an isometric isomorphism of $\mathcal{B}(X, Y)$ into $\mathcal{B}(Y^*, X^*)$.

Proof. (a) Let $y_1^*,\ y_2^* \in Y^*$ and $\alpha \in \mathbb{F}$. Then, for all $x \in X$

$$\begin{aligned} T^*(\alpha y_1^* + y_2^*)(x) &= (\alpha y_1^* + y_2^*)(Tx) = \alpha y_1^*(Tx) + y_2^*(Tx) \\ &= \alpha T^* y_1^*(x) + T^* y_2^*(x) = (\alpha T^* y_1^* + T^* y_2^*)(x). \end{aligned}$$

Hence, $T^*(\alpha y_1^* + y_2^*) = \alpha T^* y_1^* + T^* y_2^*$.
Furthermore, as shown above, $\|T^* y^*\| \leq \|T\| \|y^*\|$. Hence, $T^* \in \mathcal{B}(Y^*, X^*)$ and $\|T^*\| \leq \|T\|$.

(b) We show that $\|T^*\| = \|T\|$, whence $\|\Lambda T^*\| = \|T^*\|$. Indeed,

$$\begin{aligned} \|T\| &= \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\|=1} \left(\sup_{\|y^*\|=1} |y^*(Tx)| \right) \qquad \text{(by Corollary 5.2.3)} \\ &= \sup_{\|y^*\|=1} \left(\sup_{\|x\|=1} |y^*(Tx)| \right) = \sup_{\|y^*\|=1} \|T^* y^*\| \\ &= \|T^*\|. \end{aligned}$$

5.5 Weak Topologies

We have made the point that a norm on a linear space X induces a metric. A metric, in turn, induces a topology on X called the metric topology. It now follows that a norm on a linear space X induces a topology which we shall refer to as the *norm topology on X*. In this section we define other topologies on a linear space X that are weaker than the norm topology. We also investigate some of the properties of these weak topologies.

5.5.1 Definition
Let $(X, \|\cdot\|)$ be a normed linear space and $\mathcal{F} \subset X^*$. The **weak topology on X induced by the family $\mathcal{F}$**, *denoted by $\sigma(X, \mathcal{F})$, is the weakest topology on X with respect to which each $x^* \in \mathcal{F}$ is continuous.*

5.5.2 Remark
The weak topology on X induced by the dual space X^* is simply referred to as "*the weak topology on X*" and is denoted by $\sigma(X, X^*)$.

What do the basic open sets for the weak topology $\sigma(X, X^*)$ look like?
Unless otherwise indicated, we shall denote by $\Phi,\ \Phi_1,\ \Phi_2\ \ldots$ *finite* subsets of X^*.
Let $x_0 \in X$, Φ and $\epsilon > 0$ be given. Consider all sets of the form

$$\begin{aligned} V(x_0; \Phi; \epsilon) &:= \{x \in X \mid |x^*(x) - x^*(x_0)| < \epsilon,\ x^* \in \Phi\} \\ &= \bigcap_{x^* \in \Phi} \{x \in X \mid |x^*(x) - x * (x_0)| < \epsilon\}. \end{aligned}$$

5.5.3 Proposition

[1] $x_0 \in V(x_0; \Phi; \epsilon)$.

[2] Given $V(x_0; \Phi_1;\ \epsilon_1)$ and $V(x_0; \Phi_2;\ \epsilon_2)$, we have

$$V(x_0; \Phi_1 \cup \Phi_2;\ \min\{\epsilon_1, \epsilon_2\}) \subset V(x_0;\ \Phi_1;\ \epsilon_1) \cap V(x_0;\ \Phi_2;\ \epsilon_2).$$

[3] If $x \in V(x_0; \Phi;\ \epsilon)$, then there is a $\delta > 0$ such that $V(x; \Phi;\ \delta) \subset V(x_0; \Phi;\ \epsilon)$.

Proof. (1) It is obvious that $x_0 \in V(x_0; \Phi;\ \epsilon)$.

(2) Let $x \in V(x_0; \Phi_1 \cup \Phi_2;\ \min\{\epsilon_1, \epsilon_2\})$. Then for each $x^* \in \Phi_1$,

$$|x^*(x) - x^*(x_0)| < \min\{\epsilon_1, \epsilon_2\} \le \epsilon_1.$$

Hence $x \in V(x_0;\ \Phi_1;\ \epsilon_1)$. Similarly, $x \in V(x_0;\ \Phi_2;\ \epsilon_2)$. It now follows that $x \in V(x_0;\ \Phi_1;\ \epsilon_1) \cap V(x_0;\ \Phi_2;\ \epsilon_2)$ and, consequently

$$V(x_0; \Phi_1 \cup \Phi_2;\ \min\{\epsilon_1, \epsilon_2\}) \subset V(x_0;\ \Phi_1;\ \epsilon_1) \cap V(x_0;\ \Phi_2;\ \epsilon_2).$$

(3) Let $x \in V(x_0; \Phi;\ \epsilon)$ and $\gamma = \max\{|x^*(x) - x^*(x_0)| \mid x^* \in \Phi\}$. Then $0 \le \gamma < \epsilon$. Choose δ such that $0 < \delta < \epsilon - \gamma$. Then, for any $y \in V(x; \Phi;\ \delta)$ and any $x^* \in \Phi$, we have

$$|x^*(y) - x^*(x_0)| \le |x^*(y) - x^*(x)| + |x^*(x) - x^*(x_0)| < \delta + \gamma < \epsilon. \qquad \blacksquare$$

Recall that a collection $\mathcal{B}$ of subsets of a set X is a base for a topology on X if and only if

(i) $X = \bigcup\{B \mid B \in \mathcal{B}\}$; i.e., each $x \in X$ belongs to some $B \in \mathcal{B}$, and

(ii) if $x \in B_1 \cap B_2$ for some B_1 and B_2 in $\mathcal{B}$, then there is a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

5.5.1 Theorem
Let $\mathcal{B} = \{V(x; \Phi;\ \epsilon) \mid x \in X,\ \Phi(\text{finite}) \subset X^*,\ \epsilon > 0\}$. Then $\mathcal{B}$ is a base for a Hausdorff topology on X.

Proof. (i) It is clear that $x \in V(x; \Phi;\ \epsilon)$ for each $x \in X$.

(ii) Let $x \in V(x_1;\ \Phi_1;\ \epsilon_1) \cap V(x_2;\ \Phi_2;\ \epsilon_2)$. Then $x \in V(x_1;\ \Phi_1;\ \epsilon_1)$ and $x \in V(x_2;\ \Phi_2;\ \epsilon_2)$. By Proposition 5.5.3 (3), there are $\delta_1 > 0$ and $\delta_2 > 0$ such that $V(x;\ \Phi_1;\ \delta_1) \subset V(x_1;\ \Phi_1;\ \epsilon_1)$ and $V(x;\ \Phi_2;\ \delta_2) \subset V(x_2;\ \Phi_2;\ \epsilon_2)$. By Proposition 5.5.3 (2),

$$V(x;\ \Phi_1 \cup \Phi_2;\ \min\{\delta_1, \delta_2\}) \subset V(x;\ \Phi_1;\ \delta_1) \cap V(x;\ \Phi_2;\ \delta_2) \subset V(x_1;\ \Phi_1;\ \epsilon_1) \cap V(x_2;\ \Phi_2;\ \epsilon_2).$$

Hence, $\mathcal{B}$ is a base for a topology on X.
Finally, we show that the topology generated by $\mathcal{B}$ is Hausdorff. Let x and y be distinct elements of X. By Corollary 5.2.2, there is an $x^* \in X^*$ such that $x^*(x) \ne x^*(y)$. Let $0 < \epsilon < |x^*(x) - x^*(y)|$. Then $V(x;\ x^*;\ \frac{\epsilon}{2})$ and $V(y;\ x^*;\ \frac{\epsilon}{2})$ are disjoint neighbourhoods of x and y respectively. $\blacksquare$

It is easy to see that each $x^* \in X^*$ is continuous with respect to the topology generated by $\mathcal{B}$. Indeed, let $x_0 \in X$, $x^* \in X^*$ and $\epsilon > 0$. Since x^* is continuous with respect to the norm topology on X, there is a norm neighbourhood U of x_0 such $|x^*(x) - x^*(x_0)| < \epsilon$ for all $x \in U$. It now follows $V(x_0;\ x^*;\ \epsilon)$ is a neighbourhood of x_0 in the topology generated by $\mathcal{B}$ and $|x^*(x) - x^*(x_0)| < \epsilon$ for all $x \in V(x_0;\ x^*;\ \epsilon)$.

One shows quite easily that the topology generated by

$$\mathcal{B} = \{V(x; \Phi;\ \epsilon) \mid x \in X,\ \Phi(\text{finite}) \subset X^*,\ \epsilon > 0\}$$

is precisely $\sigma(X, X^*)$, the weak topology on X induced by X^*. Therefore, a set G is open in the topology $\sigma(X, X^*)$ if and only if for each $x \in G$ there is a finite set
$\Phi = \{x^*,\ x_1^*,\ x_2^*,\ \ldots,\ x_n^*\} \subset X^*$ and an $\epsilon > 0$ such that $V(x;\ \Phi;\ \epsilon) \subset G$.

It now follows that a normed linear space X carries two natural topologies: the norm topology induced by the norm on X and the weak topology induced by its dual space X^*. Topological concepts that are associated with the weak topology are usually preceded by the word "weak"; for example, weak compactness, weak closure, etc. Those topological concepts pertaining to the topology generated by the norm on X are usually preceded by the word "norm", e.g. norm-closure or by the word "strong", e.g. strongly open set.

5.5.4 Lemma
Let $\{x^*,\ x_1^*,\ x_2^*,\ \ldots,\ x_n^*\} \subset X^*$. Then

(1) $x^* \in \text{lin}\{x_1^*,\ x_2^*,\ \dots,\ x_n^*\}$ *if and only if* $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$.

(2) If $\{x_1^*,\ x_2^*,\ \dots,\ x_n^*\}$ *is a linearly independent set, then for any set of scalars* $\{c_1,\ c_2,\ \dots,\ c_n\}$, $\bigcap_{i=1}^n \{x \in X \mid x_i^*(x) = c_i\} \neq \emptyset$.

Proof. (1) If $x^* \in \text{lin}\{x_1^*,\ x_2^*,\ \dots,\ x_n^*\}$, then $x^* = \sum_{i=1}^n \alpha_i x_i^*$ for some scalars $\alpha_1,\ \alpha_2,\ \dots,\ \alpha_n$. Let $x \in \bigcap_{i=1}^n \ker(x_i^*)$. Then $x_i^*(x) = 0$ for each $i = 1,\ 2,\ \dots, n$. Hence, $\sum_{i=1}^n \alpha_i x_i^*(x) = 0$ and consequently, $x^*(x) = 0$; i.e., $x \in \ker(x^*)$. Therefore $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$.

Conversely, assume that $\bigcap_{i=1}^n \ker(x_i^*) \subset \ker(x^*)$. We use induction on n. Let us first show that if $\ker(x_1^*) = \ker(x^*)$, then $x^* = \alpha x_1^*$ for some nonzero scalar α. Let $K = \ker(x_1^*)$ and $z \in X \setminus K$. Then, proceeding as in Theorem 5.2.1, each $x \in X$ is uniquely expressible as $x = y + \lambda z$, where $y \in K$ and $\lambda \in \mathbb{F}$. Hence, since $x^*(y) = 0 = x_1^*(y)$,

$$x^*(x) = \lambda x^*(z) = \frac{\lambda x^*(z)}{x_1^*(z)} x_1^*(z) = \left(\frac{x^*(z)}{x_1^*(z)}\right) \lambda x_1^*(z) = \left(\frac{x^*(z)}{x_1^*(z)}\right) x_1^*(x) = \alpha x_1^*(x),$$

where $\alpha = \dfrac{x^*(z)}{x_1^*(z)}$.

Assume that the result is true for $n-1$. For each $i = 1,\ 2,\ \dots, n$, x_i^* is not a linear combination of the x_j^*'s for $j = 1,\ 2,\ \dots, n$ and $i \neq j$. Hence, $\bigcap_{j\neq i} \ker(x_j^*)$ is not contained in $\ker(x_i^*)$. Therefore there is an $x_i \in \bigcap_{j\neq i} \ker(x_j^*)$ such that $x_i^*(x_i) = 1$. Let $\alpha_i = x^*(x_i)$ for each $i = 1,\ 2,\ \dots, n$. Let $x \in X$ and $y = x - \sum_{i=1}^n x_i^*(x) x_i$. Then, for each $j = 1,\ 2,\ \dots, n$,

$$x_j^*(y) = x_j^*(x) - \sum_{i=1}^n x_i^*(x) x_j^*(x_i) = x_j^*(x) - x_j^*(x) = 0.$$

Thus, $y \in \bigcap_{i=1}^n \ker(x_i^*)$. By the assumption, $y \in \ker(x^*)$. Therefore

$$0 = x^*(y) = x^*(x) - \sum_{i=1}^n x_i^*(x) x^*(x_i) = x^*(x) - \sum_{i=1}^n \alpha_i x_i^*(x) \iff x^*(x) = \sum_{i=1}^n \alpha_i x_i^*(x),$$

whence $x^* = \sum_{i=1}^n \alpha_i x_i^*$.

(2) Let $H_i = \{x \in X \mid x_i^*(x) = c_i\}$ for each $i = 1,\ 2,\ \dots, n$. We want to show that $\bigcap_{i=1}^n H_i \neq \emptyset$. The proof is by induction on n. If $n = 1$, then, since $x_1^* \neq 0$, it follows that $H_1 \neq \emptyset$. Assume true for $n = k$ and let $H = \bigcap_{i=1}^{k+1} H_i$. By the linear independence of $\{x_1^*,\ x_2^*,\ \dots,\ x_{k+1}^*\}$, $\bigcap_{i=1}^k \ker(x_i^*) \not\subset \ker(x_{k+1}^*)$. Hence, there is an $x_0 \in \bigcap_{i=1}^k \ker(x_i^*)$ such that $x_{k+1}^*(x_0) \neq 0$. Take any $x \in \bigcap_{i=1}^k \ker(x_i^*)$ and set $y = x + \alpha x_0$, where $\alpha = c_{k+1} - \dfrac{x_{k+1}^*(x)}{x_{k+1}^*(x_0)}$. Then $x_i^*(y) = x_i^*(x) = c_i$ for each $i = 1,\ 2,\ \dots, k$ and

$x^*_{k+1}(y) = c_{k+1}$. That is, $y \in H$. ∎

5.5.2 Theorem
Let τ denote the norm topology on X. Then

(a) $\sigma(X, X^*) \subset \tau$.

(b) $\sigma(X, X^*) = \tau$ if and only if X is finite-dimensional. Thus, if X is infinite-dimensional, then the weak topology $\sigma(X, X^*)$ is strictly weaker than the norm topology.

Proof. (a) The topology $\sigma(X, X^*)$ is the weakest topology on X making each $x^* \in X^*$ continuous. Hence, $\sigma(X, X^*)$ is weaker than the norm topology τ.

(b) Assume that $\sigma(X, X^*) = \tau$ and let $x^* \in X^*$. Then, since x^* is continuous when X is equipped with the norm topology and, by the hypothesis, it is continuous in the weak topology $\sigma(X, X^*)$, it is continous at 0. Therefore there is a finite set $\Phi = \{x^*_1,\ x^*_2,\ \ldots,\ x^*_n\} \subset X^*$ and an $\epsilon > 0$ such that $|x^*(x)| < 1$ for all $x \in V(0;\ \Phi;\ \epsilon)$. Let $z \in \bigcap_{i=1}^{n} \ker(x^*_i)$. Then $x^*_i(z) = 0$ and so $|x^*_i(z)| < \epsilon$ for each $i = 1,\ 2,\ \ldots, n$. That is, $z \in V(0;\ \Phi;\ \epsilon)$. If $x \in \bigcap_{i=1}^{n} \ker(x^*_i)$, then $mx \in \bigcap_{i=1}^{n} \ker(x^*_i)$ for each $m \in \mathbb{Z}^+$ since $\bigcap_{i=1}^{n} \ker(x^*_i)$ is a linear subspace of X. It now follows that $mx \in V(0;\ \Phi;\ \epsilon)$ for each $m \in \mathbb{Z}^+$. This, in turn, implies that

$$1 > |x^*(mx)| = m|x^*(x)| \quad \Longleftrightarrow \quad |x^*(x)| < \frac{1}{m}.$$

Since m is arbitrary, $x^*(x) = 0$; i.e., $x \in \ker(x^*)$. Hence $\bigcap_{i=1}^{n} \ker(x^*_i) \subset \ker(x^*)$. By Lemma 5.5.4, $x^* \in X^*$ is expressible as $x^* = \sum_{i=1}^{n} \alpha_i x^*_i$ for some scalars $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$. Hence X^* is spanned by the set $\{x^*_1,\ x^*_2,\ \ldots,\ x^*_n\}$. This shows that X^* is finite-dimensional. By Proposition 4.2.4, X is also finite-dimensional.

Conversely, assume that X is finite-dimensional. Let $\{x_1,\ x_2,\ \ldots,\ x_n\}$ be a basis for X such that $\|x_k\| = 1$ for each $k = 1,\ 2,\ \ldots, n$. Let $U \subset X$ be open in the norm topology of X. We want to show that U is open in the weak topology of X. Let $x_0 \in U$. Then there is an $r > 0$ such that $B(x_0, r) \subset U$. For any $x \in X$, $x = \sum_{k=1}^{n} \alpha_k x_k$. Define $x^*_i : X \to \mathbb{F}$ by $x^*_i(x) = \alpha_i$ for each $i = 1,\ 2,\ \ldots, n$. Since the α_i's are uniquely determined, x^*_i is well-defined. One shows quite easily that $x^*_i \in X^*$ for each $i = 1,\ 2,\ \ldots, n$. Let $\Phi = \{x^*_1,\ x^*_2,\ \ldots,\ x^*_n\}$ and $\epsilon = \frac{r}{n}$. Then, for any $x \in V(x_0;\ \Phi;\ \epsilon)$, we have $|x^*_i(x) - x^*_i(x_0)| < \epsilon$ for each $i = 1,\ 2,\ \ldots,\ n$. Hence, if $x \in V(x_0;\ \Phi;\ \epsilon)$, then

$$\|x - x_0\| = \left\| \sum_{k=1}^{n} x^*_k(x - x_0)x_k \right\| \le \sum_{k=1}^{n} |x^*_k(x - x_0)| < n\epsilon = r.$$

That is, $x \in B(x_0, r) \subset U$. It now follows that for each $x \in U$, there is a $V(x;\ \Phi;\ \epsilon)$ such that $V(x;\ \Phi;\ \epsilon) \subset U$. Hence U is open in the weak topology $\sigma(X, X^*)$. Thus, $\sigma(X, X^*) = \tau$.
∎

The following result asserts that the weak topology and the norm topology yield exactly the same continuous linear functionals. That is, the linear functionals on X that are continuous with respect to the topology $\sigma(X, X^*)$ are those that are in X^*. Therefore weakening the topology does not affect the dual space of X.

5.5.3 Theorem
Let X be a normed linear space. Then the dual of X under the topology $\sigma(X, X^*)$ is X^*; i.e., $(X,\ \sigma(X, X^*))^* = X^*$.

Proof. By definition of the topology $\sigma(X, X^*)$, it is clear X^* is a subset of the dual of X under the topology $\sigma(X, X^*)$; i.e., $X^* \subset (X,\ \sigma(X, X^*))^*$.

Let $f \in (X,\ \sigma(X, X^*))^*$. Then, proceeding as in Theorem 5.5.2, there is a finite set $\{x_1^*,\ x_2^*,\ \ldots,\ x_n^*\} \subset X^*$ and scalars $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$ such that $f = \sum_{i=1}^{n} \alpha_i x_i^*$. Therefore $f \in X^*$. ∎

5.5.4 Theorem
Let K be a convex subset of a normed linear space X. Then the closure of K relative to the weak topology $\sigma(X, X^*)$ is the same as the norm-closure of K, i.e., $\overline{K}^{\sigma(X,X^*)} = \overline{K}$.

Proof. Since $\overline{K}^{\sigma(X,X^*)}$ is closed and $K \subset \overline{K}^{\sigma(X,X^*)}$ and $\overline{K}$ is the smallest closed set containing K, it follows that $\overline{K} \subset \overline{K}^{\sigma(X,X^*)}$.

It remains to show that $\overline{K}^{\sigma(X,X^*)} \subset \overline{K}$. Let $x_0 \in X \setminus \overline{K}$. Then, by Hahn-Banach's Theorem, there is an $x^* \in X^*$ and real numbers c_1 and c_2 such that

$$\Re e\left(x^*(x_0)\right) \le c_1 < c_2 \le \Re e\left(x^*(x)\right) \quad \text{for all } x \in K.$$

Consider $V = V(x_0;\ x^*;\ c_2 - c_1) = \{x \in X \mid |x^*(x) - x^*(x_0)| < c_2 - c_1\}$. Then V is a weak neighbourhood of x_0 and $V \cap K = \emptyset$. Thus $x_0 \notin \overline{K}^{\sigma(X,X^*)}$, and consequently $\overline{K}^{\sigma(X,X^*)} \subset \overline{K}$.
∎

Since the topology $\sigma(X, X^*)$ is weaker than the norm topology, every weakly closed set in X is closed. However, for convex sets we have the following.

5.5.5 Corollary
A convex subset K of a normed linear space X is closed if and only it is weakly closed.

We now turn our attention to the dual space X^* of a normed linear space X. X^* carries three natural topologies: the norm topology, the weak topology $\sigma(X^*, X^{**})$ induced by X^{**} and the weak* topology $\sigma(X^*, X)$ induced by X.

Let J_X be the canonical embedding of X into its bidual X^{**}. Then $X \cong J_X X \subset X^{**}$. A typical basic open set in the topology $\sigma(X^*, J_X X)$ on X^* induced by $J_X X$ is

$$\begin{aligned} V(x^*;\ \Phi;\ \epsilon) &:= \{y^* \in X^* \mid |(J_X x)(x^*) - (J_X x)(y^*)| < \epsilon,\ \epsilon > 0,\ x \in \Psi(\text{finite}) \subset X\} \\ &= \{y^* \in X^* \mid |x^*(x) - y^*(x)| < \epsilon,\ \epsilon > 0,\ x \in \Psi(\text{finite}) \subset X\}. \end{aligned}$$

It now follows that the weak* topology $\sigma(X^*, X)$ on X^* is precisely the weak topology on X^* induced by $J_X X$. That is, $\sigma(X^*, X) = \sigma(X^*, J_X X)$ – the weak topology on X^* induced by elements of X acting as continous linear functionals on X^*.

Let us observe, in passing, that X^{**} has a weak* topology $\sigma(X^{**}, X^*)$ induced by X^*. Since $X \cong J_X X \subset X^{**}$, the weak topology $\sigma(X, X^*)$ on X turns out to be the relative topology on X induced by $\sigma(X^{**}, X^*)$.

5.5.5 Theorem
Let τ^* denote the norm topology on X^*. Then

(a) $\sigma(X^*, X) \subset \sigma(X^*, X^{**}) \subset \tau^*$.

(b) $\sigma(X^*, X^{**}) = \tau^*$ *if and only if* X *is finite-dimensional.*

(c) $\sigma(X^*, X) = \sigma(X^*, X^{**})$ *if and only if* X *is reflexive. Thus, if* X *is non-reflexive, then the weak* topology* $\sigma(X^*, X)$ *is strictly weaker than the weak topology* $\sigma(X^*, X^{**})$.

Proof. (a) Since $J_X X \subset X^{**}$, it follows that $\sigma(X^*, X) = \sigma(X^*, J_X X) \subset \sigma(X^*, X^{**})$. The containment $\sigma(X^*, X^{**}) \subset \tau^*$ follows from the fact that $\sigma(X^*, X^{**})$ is the weakest topology on X^* making each $x^{**} \in X^{**}$ continuous and each $x^{**} \in X^{**}$ is continuous with respect to τ^*.

(b) An argument similar to that used in Theorem 5.5.2(b) shows that $\sigma(X^*, X^{**}) = \tau^*$ if and only if X^* is finite-dimensional. But by Proposition 4.2.4, X^* is finite-dimensional if and only if X is finite-dimensional.

(c) X is reflexive if and only if $J_X X \cong X^{**}$ if and only if $\sigma(X^*, J_X X) \subset \sigma(X^*, X^{**})$ if and only if $\sigma(X^*, X) \subset \sigma(X^*, X^{**})$. ■

5.5.6 Theorem

Let X *be a normed linear space. Then the dual of* X^* *under the weak* topology* $\sigma(X^*, X)$ *is* X*; i.e.,* $(X^*, \sigma(X^*, X))^* = X$.

Proof. Exercise. ■

Observe that $X^* \subset \mathbb{F}^X = \prod_X \mathbb{F}$ and that the weak* topology $\sigma(X^*, X)$ on X^* is the relative topology on X^* induced by the product topology on $\prod_X \mathbb{F}$.

5.5.7 Theorem

(Banach-Alaoglu-Bourbaki Theorem). *Let* X *be a normed linear space over* $\mathbb{F}$*. Then the closed unit ball in* X^* *is weak* compact; i.e., the set*

$$B^* = B(X^*) = \{x^* \in X^* \mid \|x^*\| \leq 1\}$$

is compact for the topology $\sigma(X^*, X)$.

Proof. For each $x \in X$, let $D_x = \{\lambda \in \mathbb{F} \mid |\lambda| \leq \|x\|\}$. Then, for each $x \in X$, D_x is a closed interval in $\mathbb{R}$ or a closed disk in $\mathbb{C}$ according to whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Equipped with the standard topology, D_x is compact for each $x \in X$. Let $D = \prod\{D_x \mid x \in X\}$. By Tychonoff's Theorem, D is compact.

The points of D are just functions f on X such that $f(x) \in D_x$ for each $x \in X$. If $x^* \in B(X^*)$, then

$$|x^*(x)| \leq \|x^*\|\|x\| \leq \|x\| \text{ for each } x \in X.$$

Hence $x^*(x) \in D_x$ for each $x^* \in B(X^*)$ and $x \in X$. That is, $B(X^*) \subset \prod\{D_x \mid x \in X\}$. We observe that the topology that D induces on $B(X^*)$ is precisely the weak* topology on $B(X^*)$. It remains to show that $B(X^*)$ is a closed subset of D. To this end, let $\{x_\delta\}$ be a net in $B(X^*)$ and $x^*_\delta \to x^* \in D$. Then $x_\delta(x) \to x^*(x)$ for all $x \in X$, and for all x, y in X and α, β in $\mathbb{F}$,

$$x^*(\alpha x + \beta y) = \lim_\delta x^*_\delta(\alpha x + \beta y) = \lim_\delta [\, \alpha x^*_\delta(x) + \beta x^*_\delta(y)\,] = \alpha x^*(x) + \beta x^*(x).$$

Thus x^* is linear. Since

$$|x^*(x)| = \lim_\delta |x^*_\delta(x)| \leq \|x\|$$

for all $x \in X$, x^* is bounded and $\|x^*\| \leq 1$. That is, $x^* \in B(X^*)$. Therefore $B(X^*)$ is closed in D and hence compact. ■

5.5.8 Theorem

(Helly). *Let X be a normed linear space over $\mathbb{F}$and $x^{**} \in X^{**}$. Then, for any finite-dimensional subspace Φ of X^* and any $\epsilon > 0$, there is an $x_0 \in X$ such that*

(i) $(J_X x_0)(x^*) = x^{**}(x^*) \iff x^*(x_0) = x^{**}(x^*)$ *for each* $x^* \in \Phi$, *and*

(ii) $\|x_0\| \leq \|x^{**}\| + \epsilon$.

Proof. Let $\{x_1^*, x_2^*, \ldots, x_n^*\}$ be a basis for Φ. Then (i) is equivalent to

(i') $x_i^*(x_0) = x^{**}(x_i^*)$ for each $i = 1, 2, \ldots, n$.

Let $H_i = \{x \in X \mid x_i^*(x) = x^{**}(x_i^*)\}$ for each $i = 1, 2, \ldots, n$ and $H = \bigcap_{i=1}^{n} H_i$. Then, by Lemma 5.27, $H \neq \emptyset$. Choose any $x_0 \in H$ such that $\|x_0\| < d(0, H) + \epsilon$. Obviously, x_0 satisfies (i'), hence (i). To complete the proof, it suffices to show that $d(0, H) \leq \| x^{**}|_\Phi \|$. Fix an $h \in H$ and set $h_0 = h - x_0$ and $H_0 = H - x_0$. Then $H_0 = \bigcap_{i=1}^n \ker(x_i^*)$ and $d(0, H) = d(-x_0, H_0) = d(x_0, H_0)$. By the Hahn-Banach Theorem and Lemma 5.5.4, it follows that

$$
\begin{aligned}
d(0, H) &= \max\{x^*(x_0) \mid x^* \in H_0^\perp,\ \|x^*\| \leq 1\} \\
&= \max\{\sum_{i=1}^{n} \alpha_i x_i^*(x_0) \mid \left\| \sum_{i=1}^{n} \alpha_i x_i^* \right\| \leq 1\} \\
&= \max\{\sum_{i=1}^{n} \alpha_i x_0^{**}(x_i^*) \mid \left\| \sum_{i=1}^{n} \alpha_i x_i^* \right\| \leq 1\} \\
&= \max\{x_0^{**}\left(\sum_{i=1}^{n} \alpha_i x_i^*\right) \mid \left\| \sum_{i=1}^{n} \alpha_i x_i^* \right\| \leq 1\} \\
&\leq \max\{x_0^{**}(x^*) \mid x^* \in \Phi,\ \|x^*\| \leq 1\} = \| x^{**}\big|_\Phi \|.
\end{aligned}
$$

■

5.5.9 Theorem

(Goldstine). *Let X be a normed linear space and J_X the canonical embedding of X into X^{**}. Let $B = \{x \in X \mid \|x\| \leq 1\}$ and $B^{**} = \{x^{**} \in X^{**} \mid \|x^{**}\| \leq 1\}$. Then $J_X B$ is dense in B^{**} relative to the weak* topology $\sigma(X^{**}, X^*)$ on X^{**}. That is,*

$$\overline{J_X B}^{\sigma(X^{**},X^*)} = B^{**}.$$

Proof. We must show that for each $x^{**} \in B^{**}$, each finite subset $\Phi = \{x_1^*, x_2^*, \ldots, x_n^*\} \subset X^*$ and each $\epsilon > 0$, there is an $x \in B$ such that $J_X x \in V(x^{**};\ \Phi;\ \epsilon)$; i.e.,

$$|J_X x(x_i^*) - x^{**}(x_i^*)| < \epsilon \text{ for each } i = 1, 2, \ldots, n.$$

Let $x^{**} \in B^{**}$. If $\|x^{**}\| < 1$, then, with $\epsilon = 1 - \|x^{**}\|$, we have, by Theorem 5.5.8, that there is an $x \in X$ such that $(J_X x)(x_i^*) = x^{**}(x_i^*)$ for each $i = 1, 2, \ldots, n$ and $\|x\| < \|x^{**}\| + \epsilon = 1$; i.e., $x \in B$. Hence, $x \in B$ and $0 = |J_X x(x_i^*) - x^{**}(x_i^*)| < \epsilon$ for each $i = 1, 2, \ldots, n$.

If $\|x^{**}\| = 1$, let $r = \max_{1 \leq i \leq n} \|x_i^*\|$ and $y^{**} = \left(1 - \frac{\epsilon}{2r}\right) x^{**}$. Then $\|y^{**}\| < 1$, and so by the first part, there is an $x \in B$ such that $(J_X x)(x_i^*) = y^{**}(x_i^*)$ for each $i = 1, 2, \ldots, n$. Furthermore, for each $i = 1, 2, \ldots, n$,

$$|J_X x(x_i^*) - x^{**}(x_i^*)| = |y^{**}(x_i^*) - x^{**}(x_i^*)| \leq \frac{\epsilon}{2r} \leq \frac{\epsilon}{2} < \epsilon$$

■

5.5.6 Corollary

Let X be a normed linear space over $\mathbb{F}$ and let J_X be the canonical embedding of X into X^{**}. Then $J_X X$ is dense in X^{**} relative to the weak topology $\sigma(X^{**}, X^*)$ on X^{**}. That is,*

$$\overline{J_X X}^{\sigma(X^{**},X^*)} = X^{**}.$$

Proof. Let $x^{**} \in X^{**} \setminus \{0\}$. Then

$$\frac{x^{**}}{\|x^{**}\|} \in B^{**} = \overline{J_X B}^{\sigma(X^{**},X^*)} \subset \overline{J_X X}^{\sigma(X^{**},X^*)}.$$

Since $\overline{J_X X}^{\sigma(X^{**},X^*)}$ is a linear subspace of X^{**}, it now follows that $x^{**} \in \overline{J_X X}^{\sigma(X^{**},X^*)}$. Hence $X^{**} \subset \overline{J_X X}^{\sigma(X^{**},X^*)}$. Of course, since $J_X X \subset X^{**}$, we have that $\overline{J_X X}^{\sigma(X^{**},X^*)} \subset X^{**}$, and consequently $\overline{J_X X}^{\sigma(X^{**},X^*)} = X^{**}$. ■

5.5.10 Theorem

Let X be a normed linear space over $\mathbb{F}$ and $B = \{x \in X \mid \|x\| \leq 1\}$. Then X is reflexive if and only if B is weakly compact.

Proof. Assume that X is reflexive and let J_X be the canonical embedding of X into X^{**}. Equip B (respectively, B^{**}) with the weak (respectively, weak*) topology and consider the map $f : B^{**} \to B$ defined by $f(J_X x) = x$. Now, B^{**} is weak* compact by Banach-Alaoglu-Bourbaki Theorem and $f(B^{**}) = B$. To prove weak compactness of B, it suffices to show that f is continuous. To that end, let $(J_X x_\delta)$ be a net in B^{**} that converges to $J_X x$ in the topology $\sigma(X^{**}, X^*)$ on X^{**}. Then, for each $x^* \in X^*$, we have that

$$x^*\big(f(J_X x_\delta)\big) = x^*(x_\delta) = (J_X x_\delta)(x^*) \to J_X x(x^*) = x^*(x) = x^*(f(J_X x)).$$

Thus, $f(J_X x_\delta) \to f(J_X x)$ in the weak topology on B.

Conversely, assume that B is weakly compact. Equip B (respectively, B^{**}) with the weak (respectively, weak*) topology. It follows that J_X is continuous. Hence, $J_X B$ is weak* compact in B^{**}. But $\overline{J_X B}^{\sigma(X^{**},X^*)} = B^{**}$. Hence $J_X X = X^{**}$ and so X is reflexive. ■

Chapter 6

Baire's Category Theorem and its Applications

6.1 Introduction

Recall that a subset S of a metric space (X, d) is dense in X if $\overline{S} = X$; i.e., for each $x \in X$ and each $\epsilon > 0$, there is an element $y \in S$ such that $d(x, y) < \epsilon$, or equivalently, $S \cap B(x, \epsilon) \neq \emptyset$.

6.1.1 Theorem
Let (X, d) be a complete metric space. If (G_n) is a sequence of nonempty, open and dense subsets of X then $G = \bigcap_{n \in \mathbb{N}} G_n$ is dense in X.

Proof. Let $x \in X$ and $\epsilon > 0$. Since G_1 is dense in X, there is a point x_1 in the open set $G_1 \cap B(x, \epsilon)$. Let r_1 be a number such that $0 < r_1 < \dfrac{\epsilon}{2}$ and

$$\overline{B(x_1, r_1)} \subset G_1 \cap B(x, \epsilon).$$

Since G_2 is dense in X, there is a point x_2 in the open set $G_2 \cap B(x_1, r_1)$. Let r_2 be a number such that $0 < r_2 < \dfrac{\epsilon}{2^2}$ and

$$\overline{B(x_2, r_2)} \subset G_2 \cap B(x_1, r_1).$$

Since G_3 is dense in X, there is a point x_3 in the open set $G_3 \cap B(x_2, r_2)$. Let r_3 be a number such that $0 < r_3 < \dfrac{\epsilon}{2^3}$ and

$$\overline{B(x_3, r_3)} \subset G_3 \cap B(x_2, r_2).$$

Continuing in this fashion, we obtain a sequence (x_n) in X and a sequence (r_n) of radii such that for each $n = 1, 2, 3, \ldots$,

$$0 < r_n < \frac{\epsilon}{2^n}, \quad \overline{B(x_{n+1}, r_{n+1})} \subset G_{n+1} \cap B(x_n, r_n) \text{ and } \overline{B(x_1, r_1)} \subset G_1 \cap B(x, \epsilon).$$

It is clear that

$$\overline{B(x_{n+1}, r_{n+1})} \subset B(x_n, r_n) \subset B(x_{n-1}, r_{n-1}) \subset \cdots \subset B(x_1, r_1) \subset B(x, \epsilon).$$

Let $N \in \mathbb{N}$. If $k > N$ and $\ell > N$, then both x_k and x_ℓ lie in $B(x_N, r_N)$. By the triangle inequality

$$d(x_k, x_\ell) \leq d(x_k, x_N) + d(x_N, x_\ell) < 2r_N < \frac{2\epsilon}{2^N} = \frac{\epsilon}{2^{N-1}}.$$

Hence, (x_n) is a Cauchy sequence in X. Since X is complete, there is a $y \in X$ such that $x_n \to y$ as $n \to \infty$. Since x_k lies in the closed set $\overline{B(x_n, r_n)}$ if $k > n$, it follows that y lies in each $\overline{B(x_n, r_n)}$. Hence y lies in each G_n. That is, $G = \bigcap_{n \in \mathbb{N}} G_n \neq \emptyset$. It is also clear that $y \in B(x, \epsilon)$. ■

6.1.1 Definition
A subset S of metric space (X, d) is said to be **nowhere dense** *in X if the set $X \setminus \overline{S}$ is dense in X; i.e., $\overline{X \setminus \overline{S}} = X$.*

6.1.2 Proposition
A subset S of a metric space (X, d) is nowhere dense in X if and only if the closure $\overline{S}$ of S contains no interior points.

Proof. Assume that S is nowhere dense in X and that $(\overline{S})^\circ \neq \emptyset$. Then there is an $\epsilon > 0$ and an $x \in \overline{S}$ such that $B(x, \epsilon) \subset \overline{S}$. But then $X \setminus \overline{S} \subset X \setminus B(x, \epsilon)$. Since $X \setminus B(x, \epsilon)$ is closed, $X \setminus B(x, \epsilon) = \overline{X \setminus B(x, \epsilon)}$. Therefore

$$\overline{X \setminus \overline{S}} \subset X \setminus B(x, \epsilon) \subset X,$$

where the second containment is proper. This is a contradiction. Hence, $(\overline{S})^\circ = \emptyset$.

Conversely, assume that $(\overline{S})^\circ = \emptyset$. Then, for each $x \in \overline{S}$ and each $\epsilon > 0$,

$$B(x, \epsilon) \cap X \setminus \overline{S} \neq \emptyset.$$

This means that each $x \in \overline{S}$ is a limit point of the set $X \setminus \overline{S}$. That is, $\overline{S} \subset \overline{X \setminus \overline{S}}$. Thus,

$$X = \overline{S} \cup (X \setminus \overline{S}) \subset \overline{X \setminus \overline{S}} \cup X \setminus \overline{S} = \overline{X \setminus \overline{S}} \subset X.$$

Hence $X = \overline{X \setminus \overline{S}}$ and so S is nowhere dense in X. ■

6.1.3 Example
Each finite subset of $\mathbb{R}$ is nowhere dense in $\mathbb{R}$.

6.1.4 Definition
A subset S of a metric space (X, d) is said to be

(a) of **first category** *or* **meagre** *in X if S can be written as a countable union of sets which are nowhere dense in X. Such sets are also called thin.*

(b) of **second category** *or* **nonmeagre** *in X if it is not of first category in X. Such sets are also called fat or thick.*

It is clear that a subset of a set of first category is itself a set of first category. Also, a countable union of sets of first category is again a set of first category.

6.1.5 Example
The set $\mathbb{Q}$ of rationals is of first category in $\mathbb{R}$.

6.1.2 Theorem
(Baire's Category Theorem). *A complete metric space (X, d) is of second category in itself.*

Proof. Assume that X is of first category. Then there is a sequence (G_n) of sets which are nowhere dense in X such that $X = \bigcup_n G_n$. Replacing each G_n by its closure, we get $X = \bigcup_n \overline{G_n}$. The sets $\overline{G_n}$ are closed and nowhere dense in X. It follows that the sets $U_n = X \setminus \overline{G_n}$ are open and dense in X. Since X is complete, it follows, by Theorem 6.1.1, that $U = \bigcap_n U_n$ is dense in X and therefore nonempty since X is nonempty. However $X = \bigcup_n \overline{G_n}$ implies that

$$\emptyset \neq \bigcap_n U_n = \bigcap_n (X \setminus \overline{G_n}) = X \setminus \bigcup_n \overline{G_n} = \emptyset,$$

which is absurd. ∎

6.2 Uniform Boundedness Principle

We have made the point that if X and Y are normed linear spaces, then $\mathcal{B}(X, Y)$ is a normed linear space.

6.2.1 Definition
A subset $\mathcal{F}$ of $\mathcal{B}(X, Y)$ is said to be

(a) **norm (or uniformly) bounded** *if*

$$\sup\{\|T\| \mid T \in \mathcal{F}\} < \infty.$$

(b) **pointwise bounded** *on X if*

$$\sup\{\|Tx\| \mid T \in \mathcal{F}\} < \infty$$

for each $x \in X$.

Clearly, a norm bounded set is pointwise bounded on X. Uniform Boundedness Principle (or Banach-Steinhaus Theorem) says that if X is a Banach space, then the converse also holds.

6.2.1 Theorem
(Uniform Boundedness Principle). *Let X be a Banach space, Y a normed linear space and let $\mathcal{F}$ be subset of $\mathcal{B}(X, Y)$ such that* $\sup\{\|Tx\| \mid T \in \mathcal{F}\} < \infty$ *for each $x \in X$. Then* $\sup\{\|T\| \mid T \in \mathcal{F}\} < \infty$.

Proof. For each $k \in \mathbb{N}$, let

$$A_k = \{x \in X \mid \|Tx\| \leq k \text{ for all } T \in \mathcal{F}\}.$$

Since T is continuous, A_k is closed. Indeed, let $x \in \overline{A_k}$. Then there is a sequence $(x_n) \subset A_k$ such that $\lim_{n\to\infty} x_n = x$. Since $x_n \in A_k$ for each n, $\|Tx_n\| \leq k$ for all $T \in \mathcal{F}$. Hence

$$\|Tx\| \leq \|Tx - Tx_n\| + \|Tx_n\| \leq \|T\|\|x_n - x\| + k \;\to\; k \text{ as } n \to \infty.$$

That is, $\|Tx\| \leq k$ and consequently $x \in A_k$.

By the hypothesis, $X = \bigcup_{k=1}^{\infty} A_k$. By Baire's Category Theorem, there is an index k_0 such that $(\overline{A_{k_0}})^\circ \neq \emptyset$. That is, there is an $x_0 \in \overline{A_{k_0}}$ and an $\epsilon > 0$ such that $B(x_0, \epsilon) \subset \overline{A_{k_0}} = A_{k_0}$.

Let $x \in X \setminus \{0\}$ and set $z = x_0 + \lambda x$, where $\lambda = \dfrac{\epsilon}{2\|x\|}$. Then $\|z - x_0\| = \lambda\|x\| = \dfrac{\epsilon}{2} < \epsilon$. Hence $z \in B(x_0, \epsilon) \subset A_{k_0}$ and, consequently, $\|Tz\| \leq k_0$ for all $T \in \mathcal{F}$. It now follows that

$$\|Tx\| = \frac{1}{\lambda}\|Tz - Tx_0\| \leq \frac{1}{\lambda}(\|Tz\| + \|Tx_0\|) \leq \frac{2k_0}{\lambda} = \frac{4k_0}{\epsilon}\|x\|.$$

Hence $\|T\| \leq \dfrac{4k_0}{\epsilon}$ for all $T \in \mathcal{F}$. ∎

It is essential that X be complete in Theorem 6.2.1. Consider the subset $\ell_0 \subset \ell_1$ of finitely nonzero sequences in ℓ_1. The set ℓ_0 is dense but not closed in ℓ_1. For each $n \in \mathbb{N}$, let $T_n x = n x_n$, where $x = (x_n) \in \ell_0$. For each $x \in \ell_0$, $T_n x = 0$ for sufficiently large n. Clearly, (T_n) is pointwise bounded on ℓ_0. On the other hand, for $(e_n) \in \ell_0$, $\|e_n\| = 1$ and $\|T_n\| \geq T_n e_n = n$ for all $n \in \mathbb{N}$. Thus (T_n) is not norm bounded.

6.2.2 Corollary
Let S be a subset of a normed linear space $(X, \|\cdot\|)$ such that the set $\{x^*(x) \mid x \in S\}$ is bounded for each $x^* \in X^*$. Then the set S is bounded.

Proof. Let J_X be the canonical embedding of X into X^{**}. By the hypothesis, the set $\{J_X x(x^*) \mid x \in S\}$ is bounded for each $x^* \in X^*$. Since X^* is a Banach space, it follows from the Uniform Boundedness Principle that the set $\{J_X x \mid x \in S\}$ is bounded. Since $\|J_X x\| = \|x\|$, the set S is also bounded. ■

Let X and Y be normed linear spaces. We remarked earlier that the strong operator limit T of a sequence $(T_n) \subset \mathcal{B}(X, Y)$ need not be bounded. However, if X is complete, then T is also bounded. This is a consequence of the Uniform Boundedness Principle.

6.2.3 Corollary
Let (T_n) be a sequence of bounded linear operators from a Banach space X into a normed linear space Y. If T is the strong operator limit of the sequence (T_n), then $T \in \mathcal{B}(X, Y)$.

Proof. The proof of linearity of T is straightforward.

We show that T is bounded. Since for each $x \in X$, $T_n x \to Tx$ as $n \to \infty$, the sequence $(T_n x)$ is bounded for each $x \in X$. By the Uniform Boundedness Principle, we have that the sequence $(\|T_n\|)$ is bounded. That is, there is a constant $M > 0$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Therefore

$$\|T_n x\| \leq \|T_n\|\|x\| \leq M\|x\| \text{ for all } n \in \mathbb{N}.$$

By continuity of the norm,

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| \leq \|Tx - T_n x\| + M\|x\| \to M\|x\| \text{ as } n \to \infty.$$

Hence, $\|Tx\| \leq M\|x\|$ for each $x \in X$, i.e., $T \in \mathcal{B}(X, Y)$. ■

6.3 The Open Mapping Theorem

6.3.1 Definition
Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and let $T : X \to Y$. Then we say that T is an **open mapping** *if TU is open in Y whenever U is open in X.*

6.3.2 Lemma
Let X and Y be Banach spaces over the field $\mathbb{F}$ and let T be a bounded linear operator from X onto Y. Then there is a constant $r > 0$ such that

$$B_Y(0, 2r) := \{y \in Y \mid \|y\| < 2r\} \subset \overline{TB_X(0, 1)}.$$

Proof. It is easy to see that $X = \bigcup_{n=1}^{\infty} nB_X(0, 1)$. Indeed, if $x \in X$, then there is an $n \in \mathbb{N}$ such that $\|x\| < n$. Hence, $x \in nB_X(0, 1)$. Since T is surjective,

$$Y = TX = T\left(\bigcup_{n=1}^{\infty} nB_X(0, 1)\right) = \bigcup_{n=1}^{\infty} nTB_X(0, 1) = \bigcup_{n=1}^{\infty} n\overline{TB_X(0, 1)}.$$

By Baire's Category Theorem, there is a positive integer n_0 such that $(n_0\overline{TB_X(0,1)})^\circ \neq \emptyset$. This implies that $(\overline{TB_X(0,1)})^\circ \neq \emptyset$. Hence, there is a constant $r > 0$ and an element $y_0 \in Y$ such that $B_Y(y_0, 4r) \subset \overline{TB_X(0,1)}$. Since $y_0 \in \overline{TB_X(0,1)}$, it follows, by symmetry, that $-y_0 \in \overline{TB_X(0,1)}$. Therefore

$$B_Y(0,4r) = B_Y(y_0,4r) - y_0 \subset \overline{TB_X(0,1)} + \overline{TB_X(0,1)}.$$

Since $\overline{TB_X(0,1)}$ is a convex set, $\overline{TB_X(0,1)} + \overline{TB_X(0,1)} = 2\overline{TB_X(0,1)}$. Hence, $B_Y(0,4r) \subset 2\overline{TB_X(0,1)}$ and, consequently, $B_Y(0,2r) \subset \overline{TB_X(0,1)}$. ■

6.3.3 Lemma

Let X and Y be Banach spaces over the field $\mathbb{F}$ and let T be a bounded linear operator from X onto Y. Then there is a constant $r > 0$ such that

$$B_Y(0,r) := \{y \in Y \mid \|y\| < r\} \subset TB_X(0,1).$$

Proof. By Lemma 6.3.2, there is a constant $r > 0$ such that $B_Y(0,2r) \subset \overline{TB_X(0,1)}$. Let $y \in B_Y(0,r)$, i.e., $y \in Y$ and $\|y\| < r$. Then, with $\epsilon = \frac{r}{2}$, there is an element $z_1 \in X$ such that

$$\|z_1\| < \frac{1}{2} \text{ and } \|y - Tz_1\| < \frac{r}{2}.$$

Since $y - Tz_1 \in Y$ and $\|y - Tz_1\| < \frac{r}{2} < r$, it follows that $y - Tz_1 \in B_Y(0,r)$. Therefore there is an element $z_2 \in X$ such that

$$\|z_2\| < \frac{1}{2^2} \text{ and } \|(y - Tz_1) - Tz_2\| < \frac{r}{2^2}.$$

In general, having chosen elements $z_k \in X$, $1 \le k \le n$, such that $\|z_k\| < \dfrac{1}{2^k}$ and

$$\|y - (Tz_1 + Tz_2 + \cdots + Tz_n)\| < \frac{r}{2^n},$$

pick $z_{n+1} \in X$ such that $\|z_{n+1}\| < \dfrac{1}{2^{n+1}}$ and

$$\|y - T(z_1 + z_2 + \cdots + z_n + z_{n+1})\| = \|y - (Tz_1 + Tz_2 + \cdots + Tz_n + Tz_{n+1})\| < \frac{r}{2^{n+1}}$$

Claim: The series $\displaystyle\sum_{k=1}^{\infty} z_k$ converges to a point $x \in B_X(0,1)$ and $Tx = y$.

Proof of Claim: Since X is complete, it suffices to show that $\displaystyle\sum_{k=1}^{\infty} \|z_k\| < \infty$. But this is obviously true since

$$\sum_{k=1}^{\infty} \|z_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

Hence, the series $\displaystyle\sum_{k=1}^{\infty} z_k$ converges to some $x \in X$ with $\|x\| < 1$, i.e., $x \in B_X(0,1)$. Since

$$\lim_{n\to\infty} \left\| y - T\left(\sum_{k=1}^{n} z_k\right) \right\| = \lim_{n\to\infty} \frac{r}{2^n} = 0,$$

continuity of T implies that

$$Tx = \lim_{n\to\infty} T\left(\sum_{k=1}^{n} z_k\right) = y.$$

That is, $Tx = y$. ■

6.3.1 Theorem
(Open Mapping Theorem). *Let X and Y be Banach spaces and suppose that $T \in \mathcal{B}(X, Y)$. If T maps X onto Y, then T is an open mapping.*

Proof. Let U be an open set in X. We need to show that TU is open in Y. Let $y \in TU$. Since T is surjective, there is an $x \in U$ such that $Tx = y$. Since U is open, there is an $\epsilon > 0$ such that $B_X(x, \epsilon) = x + B_X(0, \epsilon) \subset U$. But then $y + TB_X(0, \epsilon) \subset TU$. By Lemma 6.3.3, there is a constant $r > 0$ such that $B_Y(0, r) \subset TB_X(0, 1)$. Hence $B_Y(0, r\epsilon) \subset TB_X(0, \epsilon)$. Therefore

$$B(y, r\epsilon) = y + B_Y(0, r\epsilon) \subset y + TB_X(0, \epsilon) \subset TU.$$

Hence TU is open in Y. ■

6.3.4 Corollary
(Banach's Theorem). *Let X and Y be Banach spaces and assume $T \in \mathcal{B}(X, Y)$ is bijective. Then T^{-1} is a bounded linear operator from Y onto X, i.e., $T^{-1} \in \mathcal{B}(Y, X)$.*

Proof. We have shown that T^{-1} is linear. It remains to show that T^{-1} is bounded. By Theorem 4.1.4, it suffices to show that T^{-1} is continuous on Y. To that end, let U be an open set in X. By Theorem 6.3.1, $(T^{-1})^{-1}(U) = TU$ is open in Y. Hence T^{-1} is continuous on Y. ■

6.4 Closed Graph Theorem

6.4.1 Definition
Let X and Y be linear spaces over a field $\mathbb{F}$ and $T : X \to Y$. The **graph** *of T, denoted by $\mathcal{G}(T)$, is the subset of $X \times Y$ given by*

$$\mathcal{G}(T) = \{(x, Tx) \mid x \in X\}.$$

Since T is linear, $\mathcal{G}(T)$ is a linear subspace of $X \times Y$. Let $\|\cdot\|_X$ and $\|\cdot\|_Y$ be norms on X and Y respectively. Then, for $x \in X$ and $y \in Y$, $\|(x, y)\| := \|x\|_X + \|y\|_Y$ defines a norm on $X \times Y$. If X and Y are Banach spaces, then so is $X \times Y$.

6.4.2 Definition
Let X and Y be normed linear spaces over $\mathbb{F}$. A linear operator $T : X \to Y$ is **closed** *if its graph $\mathcal{G}(T)$ is a closed linear subspace of $X \times Y$.*

6.4.1 Theorem
(Closed Graph Theorem). *Let X and Y be Banach spaces and $T : X \to Y$ a closed linear operator. Then T is bounded.*

Proof. Since $X \times Y$, with the norm defined above, is a Banach space, and by the hypothesis $\mathcal{G}(T)$ is closed, it follows that $\mathcal{G}(T)$ is also a Banach space. Consider the map $P : \mathcal{G}(T) \to X$ given by $P(x, Tx) = x$. Then P is linear and bijective. It is also bounded since

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|.$$

That is, P is bounded and $\|P\| \leq 1$. By Banach's Theorem (Corollary 6.3.4), it follows that $P^{-1} : X \to \mathcal{G}(T)$ given by $P^{-1}x = (x, Tx)$ for $x \in X$, is also bounded. Hence $\|(x, Tx)\| = \|P^{-1}x\| \leq \|P^{-1}\|\|x\|$. Therefore

$$\|(x, Tx)\| = \|x\| + \|Tx\| \leq \|P^{-1}\|\|x\| \iff \|Tx\| \leq \|P^{-1}\|\|x\|.$$

That is, T is bounded and $\|T\| \leq \|P^{-1}\|$. ■