Lecture Notes
on
Functional Analysis

Kai-Seng Chou
Department of Mathematics
The Chinese University of Hong Kong
Hong Kong

May 29, 2014

Contents

若到松江呼小渡，莫驚鴛鷺，
四橋盡是、老子經行處。
蘇軾《青玉案》

Chapter 1

Normed Space: Examples

Generally speaking, in functional analysis we study infinite dimensional vector spaces of functions and the linear operators between them by analytic methods. This chapter is of preparatory nature. First, we use Zorn's lemma to prove there is always a basis for any vector space. It fills up a gap in elementary linear algebra where the proof was only given for finite dimensional vector spaces. The inadequacy of this notion of basis for infinite dimensional spaces motivates the introduction of analysis to the study of function spaces. Second, we discuss three basic inequalities, namely, Young's, Hölder's, and Minkowski's inequalities. We establish Young's inequality by elementary means, use it to deduce Hölder's inequality, and in term use Hölder's inequality to prove Minkowski's inequality. The latter will be used to introduce norms on some common vector spaces. As you will see, these spaces form our principal examples throughout this book.

1.1 Vector Spaces of Functions

Recall that a vector space is over a field $\mathbb{F}$. Throughout this book it is always assumed this field is either the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. In the following $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$.

It is true that many vector spaces can be viewed as vector spaces of functions. To describe this unified point of view, let S be a non-empty set and denote the collection of all functions from S to $\mathbb{F}$ by $F(S)$. It is routine to check that $F(S)$ forms a vector space over $\mathbb{F}$ under the obvious rules of addition and scalar multiplication for functions: For $f, g \in F(S)$ and $\alpha \in \mathbb{F}$,

$$(f+g)(p) \equiv f(p)+g(p), \qquad (\alpha f)(p) \equiv \alpha f(p).$$

In fact, these algebraic operations are inherited from the target $\mathbb{F}$.

First, take $S = \{p_1, \cdots, p_n\}$ a set consisting of n many elements. Every function $f \in F(S)$ is uniquely determined by its values at $p_1, \cdots, p_n$, so f can be identified with the n-triple $(f(p_1), \cdots, f(p_n))$. It is easy to see that $F(\{p_1, \cdots, p_n\})$ is linearly isomorphic to $\mathbb{F}^n$. More precisely, the mapping $f \mapsto (f(p_1), \cdots, f(p_n))$ is a linear bijection between $F(\{p_1, \cdots, p_n\})$ and $\mathbb{F}^n$.

Second, take $S = \{p_1, p_2, \cdots\}$. As above, any $f \in F(S)$ can be identified with the sequence $(f(p_1), f(p_2), f(p_3) \cdots)$. The vector space $F(\{p_j\}_{j=1}^\infty)$ may be called the space of sequences over $\mathbb{F}$.

Finally, taking $S = [0, 1]$, $F([0, 1])$ consists of all $\mathbb{F}$-valued functions.

The vector spaces we are going to encounter are mostly these spaces and their subspaces.

1.2 Zorn's Lemma

In linear algebra, it was pointed out that every vector space has a basis no matter it is of finite or infinite dimension, but the proof was only given in the finite dimensional case. Here we provide a proof of the general case. The proof depends critically on Zorn's lemma, an assertion equivalent to the axiom of choice.

To formulate Zorn's lemma, we need to consider a partial order on a set.

A relation $\leq$ on a non-empty set X is called a **partial order** on X if it satisfies

(PO1) $x \leq x, \ \forall x \in X$;

(PO2) $x \leq y$ and $y \leq x$ implies $x = y$.

(PO3) $x \leq y, \ y \leq z$ implies $x \leq z$.

The pair $(X, \leq)$ is called a **partially ordered set** or a **poset** for short. A non-empty subset Y of X is called a **chain** or a **totally ordered set** if for any two $y_1, y_2 \in Y$, either $y_1 \leq y_2$ or $y_2 \leq y_1$ holds. In other words, every pair of elements in Y are related. An **upper bound** of a non-empty subset Y of X is an element u, which may or may not be in Y, such that $y \leq u$ for all $y \in Y$. Finally, a **maximal element** of $(X, \leq)$ is an element z in X such that $z \leq x$ implies $z = x$.

Example 1.1. Let S be a set and consider $X = \mathscr{P}(S)$, the power set of S. It is clear that the relation "set inclusion" $A \subset B$ is a partial order on $\mathscr{P}(S)$. It has a unique maximal element given by S itself.

Example 1.2. Let $X = \mathbb{R}^2$ and define $x \prec y$ if and only if $x_1 \leq y_1$ and $x_2 \leq y_2$. For instance, $(-1, 5) \prec (0, 8)$ but $(-2, 3)$ and $(35, -1)$ are unrelated. Then $(X, \prec)$ forms a poset without any maximal element.

Zorn's Lemma. *Let $(X, \leq)$ be a poset. If every chain in X has an upper bound, then X has at least one maximal element.*

Although called a lemma by historical reason, Zorn's lemma, a constituent in the Zermelo-Fraenkel set theory, is an axiom in nature. It is equivalent to the axiom of choice as well as the Hausdorff maximality principle. You may look up Hewitt-Stromberg's "Real and Abstract Analysis" for further information. A readable account on this "lemma" can also be found in Wikipedia.

1.3 Existence of Basis

As a standard application of Zorn's lemma, we show there is a basis in any vector space. To refresh your memory, let's recall that a subset S in a vector space X is called a linearly independent set if any finite number of vectors in S are linearly independent. In other words, letting $\{x_1, \cdots, x_n\}$ be any subset of S, if $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ for some scalars $\alpha_i, \ i = 1, \cdots, n$, then $\alpha_i = 0$ for all i. On the other hand, given any subset S, denote all linear combinations of vectors from S by $\langle S \rangle$. It is easy to check that $\langle S \rangle$ forms a subspace of X called the subspace spanned by S. A subset S is called a spanning set of X if $\langle S \rangle$ is X, and it is called a basis of X if it is also a linearly independent spanning set. When X admits a finite spanning set, it has a basis consisting of finitely many vectors. Moreover, all bases have the same number of vectors and we call this number the dimension of the space X. The space X is of infinite dimension if it does not have a finite spanning set.

Theorem 1.1. *Every non-zero vector space has a basis.*

This basis is sometimes called a **Hamel basis**.

Proof. Let $\mathcal{X}$ be the set of all linearly independent subsets of a given vector space V. Since V is non-zero, $\mathcal{X}$ is a non-empty set. Clearly the set inclusion $\subset$ makes it into a poset. To apply Zorn's lemma, let's verify that every chain in it has an upper bound. Let $\mathcal{Y}$ be a chain in $\mathcal{X}$, consider the following subset of V,

$$S = \bigcup_{C \in \mathcal{Y}} C.$$

We claim that (i) $S \in \mathcal{X}$, that's, S is a linearly independent set, (ii) $C \subset S$, $\forall C \in \mathcal{Y}$, that's, S is an upper bound of $\mathcal{Y}$. Since (ii) is obvious, it is sufficient to verify (i).

To this end, pick $v_1, \cdots, v_n \in S$. By definition, we can find $C_1, \cdots, C_n$ in $\mathcal{Y}$ such that $v_1 \in C_1, \cdots, v_n \in C_n$. As $\mathcal{Y}$ is a chain, $C_1, \cdots, C_n$ satisfy $C_i \subset C_j$ or $C_j \subset C_i$ for any i, j. After rearranging the indices, one may assume $C_1 \subset C_2 \subset \cdots \subset C_n$, and so $\{v_1, \cdots, v_n\} \subset C_n$. Since C_n is a linearly independent set, $\{v_1, \cdots, v_n\}$ is linearly independent. This shows that S is a linearly independent set.

After showing that every chain in $\mathcal{X}$ has an upper bound, we appeal to Zorn's lemma to conclude that $\mathcal{X}$ has a maximal element B. We claim that B is a basis for V. For, first of all, B belonging to $\mathcal{X}$ means that B is a linearly independent set. To show that it spans V, we pick $v \in V$. Suppose v does not belong to $\langle B \rangle$, so v is independent from all vectors in B. But then the set $\widetilde{B} = B \cup \{v\}$ is a linearly independent set which contains B as its proper subset, contradicting the maximality of B. We conclude that $\langle B \rangle = V$, so B forms a basis of V. $\square$

The following example may help you in understanding the proof of Theorem 1.1.

Example 1.3. Consider the power set of $\mathbb{R}^3$ which is partially ordered by set inclusion. Let $\mathcal{X}$ be the subset of all linearly independent sets in $\mathbb{R}^3$. Then

$$\mathcal{Y}_1 \equiv \Big\{\{(1,0,0)\}, \{(1,0,0),(1,1,0)\}, \{(1,0,0),(1,1,0),(0,0,-3)\}\Big\}$$

and

$$\mathcal{Y}_2 \equiv \Big\{\{(1,3,5),(2,4,6)\}, \{(1,3,5),(2,4,6),(1,0,0)\}\Big\}$$

are chains but

$$\mathcal{Y}_3 \equiv \Big\{\{(1,0,0)\}, \{(1,0,0),(0,1,0)\}, \{(1,0,0),(0,-2,0),(0,0,1)\}\Big\}$$

is not a chain in $\mathcal{X}$.

For a finite dimensional vector space, it is relatively easy to find an explicit basis, and bases are used in many occasions such as in the determination of the dimension of the vector space and in the representation of a linear operator as a matrix. However, in contrast, the existence of a basis in infinite dimensional space is proved via a non-constructive argument. It is not easy to write down a basis. For example, consider the space of sequences $\mathcal{S} \equiv \{x = (x_1, x_2, \cdots, x_n \cdots) : \ x_i \in \mathbb{F}\}$. Letting $e_j = (0, \cdots, 1, \cdots)$ where "1" appears in the j-th place, it is tempting from the formula $x = \sum_{j=1}^{\infty} x_j e_j$ to assert that $\{e_j\}_1^{\infty}$ forms a basis for $\mathcal{S}$. But, this is not true. Why? It is because infinite sums are not linear combinations. Indeed, one cannot talk about infinite sums in a vector space as there is no means to measure convergence. According to Theorem 1.1, however, there is a rather mysterious basis. In general, a non-explicit basis is difficult to work with, and thus lessens its importance in the study of infinite dimensional spaces. To proceed further, analytical structures will be added to vector spaces. Later, we will see that for a reasonably nice infinite dimensional vector space, any basis must consist of uncountably many vectors (see Proposition 4.14). Suitable generalizations of this notion are needed. For an infinite dimensional normed space, one may introduce the so-called Schauder basis as a replacement. For a complete inner product spaces (a Hilbert space), an even more useful notion, a complete orthonormal set, will be much more useful.

Mathematics is a deductive science. A limited number of axioms is needed to build up the tower of mathematics, and Zorn's lemma is one of them. We will encounter this lemma again in later chapters. You may also google for more of its applications.

1.4 Three Inequalities

Now we come to Young's, Hölder's and Minkowski's inequalities.

Two positive numbers p and q are **conjugate** if $1/p+1/q=1$. Notice that they must be greater than one and q approaches infinity as p approaches 1. In the following paragraphs q is always conjugate to p.

Proposition 1.2 (Young's Inequality). *For any $a, b>0$ and $p>1$,*

$$ab \le \frac{a^p}{p}+\frac{b^q}{q},$$

and equality holds if and only if $a^p=b^q$.

Proof. Consider the function

$$\varphi(x)=\frac{x^p}{p}+\frac{1}{q}-x, \quad x\in(0,\infty).$$

From the sign of $\varphi'(x)=x^{p-1}-1$ we see that φ is strictly decreasing on $(0,1)$ and strictly increasing on $(1,\infty)$. It follows that $x=1$ is the strict minimum of φ on $(0,\infty)$. So, $\varphi(x)\ge\varphi(1)$ and equality holds if and only if $x=1$. In other words,

$$\frac{x^p}{p}+\frac{1}{q}-x\ge\frac{1}{p}+\frac{1}{q}-1,$$

that is ,

$$\frac{x^p}{p}+\frac{1}{q}\ge x.$$

Letting $x=ab/b^q$, we get the Young's inequality. Equality holds if and only if $ab/b^q=1$, i.e., $a^p=b^q$. □

Proposition 1.3 (Hölder's Inequality). *For $a, b\in\mathbb{R}^n$, $p>1$,*

$$\sum_{k=1}^n |a_k||b_k| \le \|a\|_p\|b\|_q,$$

where $\|a\|_p=\left(\sum_{k=1}^n|a_k|^p\right)^{\frac{1}{p}}$ and $\|b\|_q=\left(\sum_{k=1}^n|b_k|^q\right)^{\frac{1}{q}}$.

Proof. The inequality clearly holds when $a=(0,\cdots,0)$. We may assume $a\ne(0,\cdots,0)$ in the following proof. By Young's inequality, for each $\varepsilon>0$ and k,

$$|a_kb_k|=|\varepsilon a_k||\varepsilon^{-1}b_k|\le\frac{\varepsilon^p|a_k|^p}{p}+\frac{\varepsilon^{-q}|b_k|^q}{q}.$$

Thus

$$\begin{aligned}\sum_{k=1}^n|a_k||b_k| &= |a_1||b_1|+\cdots|a_n||b_n|\\ &\le \frac{\varepsilon^p}{p}\sum_{k=1}^n|a_k|^p+\frac{\varepsilon^{-q}}{q}\sum_{k=1}^n|b_k|^q\\ &= \frac{\varepsilon^p}{p}\|a\|_p^p+\frac{\varepsilon^{-q}}{q}\|b\|_q^q, \qquad (1.1)\end{aligned}$$

for any $\varepsilon>0$. To have the best choice of ε, we minimize the right hand side of this inequality. Taking derivative of the right hand side of (1.1) as a function of ε, we obtain

$$\varepsilon^{p-1}\|a\|_p^p-\varepsilon^{-q-1}\|b\|_q^q=0,$$

that is,

$$\varepsilon=\frac{\|b\|_q^{\frac{q}{p+q}}}{\|a\|_p^{\frac{p}{p+q}}}.$$

is the minimum point. (Clearly this function has only one critical point and does not have any maximum.) Plugging this choice of ε into the inequality yields the Hölder's inequality after some manipulation. □

Proposition 1.4 (Minkowski's Inequality). *For $a, b \in \mathbb{F}^n$ and $p \geq 1$,*

$$\|a+b\|_p \leq \|a\|_p + \|b\|_p.$$

Proof. The inequality clearly holds when $p = 1$ or $\|a+b\| = 0$. In the following proof we may assume $p > 1$ and $\|a+b\| > 0$. For each k,

$$\begin{aligned} |a_k + b_k|^p &= |a_k + b_k||a_k + b_k|^{p-1} \\ &\leq |a_k||a_k + b_k|^{p-1} + |b_k||a_k + b_k|^{p-1}. \end{aligned} \tag{1.2}$$

Applying Hölder's inequality to the two terms on right hand side of (1.2) separately (more precisely, to the pairs of *real* vectors $(|a_1|, \cdots, |a_n|)$ and $(|a_1 + b_1|^{p-1}, \cdots, |a_n + b_n|^{p-1})$, and $(|b_1|, \cdots, |b_n|)$ and $(|a_1 + b_1|^{p-1}, \cdots, |a_n + b_n|^{p-1})$), we have

$$\begin{aligned} \sum_{k=1}^n |a_k + b_k|^p &\leq \|a\|_p \left(\sum_{k=1}^n |a_k + b_k|^{(p-1)q} \right)^{\frac{1}{q}} + \|b\|_p \left(\sum_{k=1}^n |a_k + b_k|^{(p-1)q} \right)^{\frac{1}{q}} \\ &= (\|a\|_p + \|b\|_p) \left(\sum_{k=1}^n |a_k + b_k|^p \right)^{\frac{1}{q}}, \end{aligned}$$

and Minkowski's inequality follows. □

Look up Wikipedia for the great mathematician Hermann Minkowski (1864-1909), the best friend of David Hilbert and a teacher of Albert Einstein, who died unexpectedly at forty-five. The biography "Hilbert" by C. Reid contains an interesting account on Minkowski and Hilbert.

The last two inequalities allow the following generalization.

Hölder's Inequality for Sequences. *For any two sequences a and b in $\mathbb{F}$, and $p > 1$,*

$$\sum_{k=1}^{\infty} |a_k||b_k| \leq \|a\|_p \|b\|_q,$$

where now the summation in the sums on the right runs from 1 to ∞.

Since the norms $\|a\|_p$ and $\|b\|_q$ are allowed to be zero or infinity, we adopt the convention $0 \times \infty = 0$ in the above inequality.

Minkowski's Inequality for Sequences. *For any two sequences a and b in $\mathbb{F}$ and $p \geq 1$,*

$$\|a+b\|_p \leq \|a\|_p + \|b\|_p,$$

where now the summation in the sums runs from 1 to ∞.

Hölder's Inequality for Functions. *For $p > 1$ and Riemann integrable functions f and g on $[a, b]$, we have*

$$\int_a^b |fg| \leq \left(\int_a^b |f|^p \right)^{\frac{1}{p}} \left(\int_a^b |g|^q \right)^{\frac{1}{q}}.$$

Minkowski's Inequality for Functions. *For $p \geq 1$ and Riemann integrable functions f and g on $[a, b]$, we have*

$$\left(\int_a^b |f+g|^p\right)^{\frac{1}{p}} \le \left(\int_a^b |f|^p\right)^{\frac{1}{p}} + \left(\int_a^b |g|^p\right)^{\frac{1}{p}},$$

We leave the proofs of these generalizations as exercises.

1.5 Normed Vector Spaces

Let $(X, +, \cdot)$ be a vector space over $\mathbb{F}$. A **norm** on X is a function from X to $[0, \infty)$ satisfying the following three properties: For all $x, y \in X$ and $\alpha \in \mathbb{F}$,

(N1) $\|x\| \ge 0$ and "=" holds if and only if $x = 0$,

(N2) $\|x + y\| \le \|x\| + \|y\|$,

(N3) $\|\alpha x\| = |\alpha|\|x\|$.

The vector space with a norm, $(X, +, \cdot, \|\cdot\|)$, or $(X, \|\cdot\|)$, or even stripped to a single X when the context is clear, is called a **normed vector space** or simply a **normed space**.

Here are some normed vector spaces.

Example 1.4. $(\mathbb{F}^n, \|\cdot\|_p), 1 \le p < \infty$, where

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}.$$

Clearly, (N1) and (N3) hold. According to the Minkowski's inequality (N2) holds too. When $p = 2$ and $\mathbb{F}^n = \mathbb{R}^n$ or $\mathbb{C}^n$, the norm is called the **Euclidean norm** or the **unitary norm**.

Example 1.5. $(\mathbb{F}^n, \|\cdot\|_\infty)$ where

$$\|x\|_\infty = \max_{k=1,\cdots,n} |x_k|.$$

is called the sup-norm.

Example 1.6. Let ℓ^p, $1 \le p < \infty$, be the collection of all $\mathbb{F}$-valued sequences $x = (x_1, x_2, \cdots)$ satisfying

$$\sum_{k=1}^\infty |x_k|^p < \infty.$$

First of all, from the Minkowski's inequality for sequences the sum of two sequences in ℓ^p belongs to ℓ^p. With the other easily checked properties, ℓ^p forms a vector space. The function $\|\cdot\|_p$, i.e.,

$$\|x\|_p = \left(\sum_{k=1}^\infty |x_k|^p\right)^{\frac{1}{p}}$$

clearly satisfies (N1) and (N3). Moreover, (N2) also holds by Minkowski's inequality for sequences. Hence it defines a norm on ℓ^p.

Example 1.7. Let ℓ^∞ be the collection of all $\mathbb{F}$-valued bounded sequences. Define the sup-norm

$$\|x\|_\infty = \sup_k |x_k|.$$

Clearly ℓ^∞ forms a normed vector space over $\mathbb{F}$.

Example 1.8. Let $C[a,b]$ be the vector space of all continuous functions on the interval $[a,b]$. For $1 \le p < \infty$, define

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{\frac{1}{p}} .$$

By the Minkowski's inequality for functions, one sees that $(C[a,b], \|\cdot\|_p)$ forms a normed space under this norm.

Example 1.9. Let $B([a,b])$ be the vector space of all bounded functions on $[a,b]$. The sup-norm

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

defines a norm on $B([a,b])$.

Example 1.10. One may have already observed that the normed spaces in Examples 1.5, 1.7 and 1.9 are of the same nature. In fact, let $F_b(S)$ be the vector subspace of $F(S)$ consisting of all bounded functions from S to $\mathbb{F}$. The sub-norm can be defined on $F_b(S)$ and these examples are special cases obtained by taking different sets S.

Example 1.11. Any vector subspace of a normed vector space forms a normed vector space under the same norm. In this way we obtain many many normed vector spaces. Here are some examples: The space of all convergent sequences, $\mathcal{C}$, the space of all sequences which converges to 0, $\mathcal{C}_0$, and the space of all sequences which have finitely many non-zero terms, $\mathcal{C}_{00}$, are normed subspaces of ℓ^∞ under the sup-norm. The space of all continuous functions on $[a,b]$, $C[a,b]$, is an important normed subspace of $B([a,b])$. The spaces $\{f : f(a) = 0,\ f \in C[a,b]\}, \{f : f \text{ is differentiable},\ f \in C[a,b]\}$ and $\{f : f \text{ is the restriction of a polynomial on } [a,b]\}$ are normed subspaces of $C[a,b]$ under the sup-norm. But the set $\{f :\ f(a) = 1,\ f \in C[a,b]\}$ is not a normed space because it is not a subspace.

To accommodate more applications, one needs to replace $[a,b]$ by more general sets in the examples above. For any closed and bounded subset K in $\mathbb{R}^n$, one may define $C(K)$ to be the collection of all continuous functions in K. As any continuous function in a closed and bounded set must be bounded (with its maximum attained at some point), its sup-norm is well-defined. Thus $(C(K), \|\cdot\|_\infty)$ forms a normed space. On the other hand, let R be any rectangular box in $\mathbb{R}^n$. We know that Riemann integration makes sense for bounded, continuous functions in R. Consequently, we may introduce the normed $\|\cdot\|_p = (\int_R |f|^p)^{1/p}$ to make all bounded, continuous functions in R a normed space. However, this p-norm does not form a norm on the space of Riemann integrable functions. Which axiom of the norm is not satisfied?

In addition to Example 10 where new normed spaces are found by restricting to subspaces, there are two more general ways to obtain them. For any two given normed spaces $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$ the function $\|(x,y)\| = \|x\|_1 + \|y\|_2$ defines a norm on the product space $X \times Y$ and thus makes $X \times Y$ the product normed space. On the other hand, to each subspace of a normed space one may form a corresponding quotient space and endow it the quotient norm. We will do this in the next chapter.

These examples of normed spaces will be used throughout this book. For simplicity the norm of the space will usually be suppressed. For instance, $\mathbb{F}^n$ always stands for the normed space under the Euclidean or the unitary norm, ℓ^p and ℓ^∞ are always under the p-norms and sup-norm respectively and a single $C(K)$ refers to the space of continuous functions on the closed, bounded set K under the sup-norm.

Exercise 1

1. Find a relation which satisfies (PO1) and (PO2) but not (PO3), and one which satisfies (PO1) and (PO3) but not (PO2).

2. Let V be a vector space. Two subspaces U and W form a direct sum of V if for every $v \in V$, there exist unique $u \in U$ and $w \in W$ such that $v = u + w$. Show that for every subspace U, there exists a subspace W so that U and W forms a direct sum of V. Suggestion: Try Zorn's lemma.

3. Prove or disprove whether B is a basis for the vector space V in the followings:
   (a) $V = \ell^1$, and $B = \{e_j\}_{j=1}^{\infty}$. ($e_j$ is the $j-$th canonical vector.)
   (b) $V = \mathcal{C}_{00}$, and $B = \{e_j\}_{j=1}^{\infty}$. ($\mathcal{C}_{00}$ consists of all sequences with finitely many non-zero terms.)
   (c) $V = \{$ all continuous functions on $[0,1]\}$, and $B = \{x^k\}_{k=0}^{\infty}$.
   (d) $V = \{$ all smooth, 2π-periodic functions$\}$, and $B = \{1, \cos nx, \sin nx\}_1^{\infty}$.
4. Determine when equality in Hölder's inequality (Proposition 1.3) holds. Hint: Keep tracking the equality sign in the proof of the proposition.
5. Let φ be a strictly increasing function on $(0,\infty)$ satisfying $\varphi(0) = 0$. Denote its inverse function by ψ. Establish the following general form of Young's inequality: For $a, b > 0$,
$$ab \leq \int_0^a \varphi(x)dx + \int_0^b \psi(x)dx.$$
6. Prove Hölder's and Minkowski's inequalities for sequences stated in §1.4, Chapter 1.
7. Recall that f is Riemann integrable if $\forall \varepsilon > 0$ there exists $\delta > 0$ such that
$$\left| \int_a^b f - R(f, \dot{P}) \right| < \varepsilon, \quad \forall \|P\| < \delta,$$
where $R(f, \dot{P})$ is the Riemann sum. (Notations as in MATH2060.) Use this fact and Propositions 1.3 and 1.4 to prove Hölder's inequality and Minkowski's inquality for functions stated in §1.4, Chapter 1.
8. Apply Hölder's inequality to establish the following interpolation inequality: $\forall a \in \mathbb{R}^n$, $p, q \geq 1$, $r = (1-\lambda)p + \lambda q$, $\lambda \in [0,1]$,
$$\|a\|_r \leq \|a\|_p^{1-\lambda} \|a\|_q^{\lambda}.$$
Then extend this interpolation inequality to functions in $C[a,b]$.
9. Is $\|\cdot\|_p$ a norm on the space of all Riemann integrable functions on $[a,b]$? If not, discuss how to make all Riemann integrable functions a normed space under this norm. This problem involves the concept of sets of measure zero. Skip it if you have not learnt it.
10. Let f be a continuously differentiable function on $[a,b]$. For $p \geq 1$, define
$$\|f\|_{1,p} \equiv \|f\|_p + \|f'\|_p,$$
where f' is the derivative of f. Show that $\|\cdot\|_{1,p}$ forms a norm on the space $C^1[a,b] \equiv \{f : f \text{ and } f' \text{ are continuous on } [a,b]\}$.
11. Let $X \times Y$ be the product space of two normed spaces X and Y. Show that it is also a normed space under the product norm $\|(x,y)\| = \|x\|_X + \|y\|_Y$.
12. Give an example to show that $\|\cdot\|_p$ is not a norm on $\mathbb{F}^n$ when $n \geq 2$ and $p \in (0,1)$. Note: In fact, there are reverse Hölder's and Minkowski's inequalities when $p \in (0,1)$. Google for them.

離愁漸遠漸無窮，迢迢不斷如春水。
.....平蕪盡處是春山，行人更在春山外。
歐陽修《踏莎行》

Chapter 2

Normed Space: Analytical Aspects

When a vector space is endowed with a norm, one can talk about the distance between two vectors and consequently it makes sense to talk about limit, convergence and continuity. The underlying structure is that of a metric space. We give a brief introduction to metric space in the first section and use it to discuss three analytical properties of a normed vector space, namely, separability, completeness and Bolzano-Weierstrass property, in later sections. Emphasis is on how these properties are preserved, modified or lost when one passes from finite to infinite dimensions.

Our discussion on metric spaces is minimal in order to avoid possible overlap with a course on point set topology. Chapters 2-4 in Rudin's "Principles in Mathematical Analysis" on metric spaces contain more than enough materials for us.

2.1 Normed Space As Metric Space

Let M be a non-empty set. A function $d : M \times M \mapsto [0, \infty)$ is called a **metric** on M if $\forall p, q, r \in M$

(D1) $d(p, q) \geq 0$, and "=" holds if and only if $p = q$.

(D2) $d(p, q) = d(q, p)$.

(D3) $d(p, q) \leq d(p, r) + d(r, q)$.

The pair (M, d) is called a **metric space**.

Notions such as convergence of sequences, Cauchy sequences, continuity of functions,... which we discussed in $\mathbb{R}$ or $\mathbb{R}^n$ in Elementary Analysis and Advanced Calculus make sense naturally in a metric space. To be precise, we have

• Let $\{p_n\}$ be a sequence in (M, d). We call $p \in M$ the **limit** of $\{p_n\}$ if for any $\varepsilon > 0$, there exists n_0 such that $d(p_n, p) < \varepsilon$ for all $n \geq n_0$. Write $p = \lim_{n\to\infty} p_n$ or simply $p_n \to p$..

• The sequence $\{p_n\}$ is called a **Cauchy sequence** if for any $\varepsilon > 0$, there exists n_0 such that $d(p_n, p_m) < \varepsilon$, for all $n, m \geq n_0$.

• Let $f : (M, d) \mapsto (N, \rho)$ where (N, ρ) is another metric space be a function and $p_0 \in M$. f is **continuous** at p_0 if $f(p_0) = \lim_{n\to\infty} f(p_n)$ whenever $\lim_{n\to\infty} p_n = p$. Alternatively, for any $\varepsilon > 0$, it is required that there exists $\delta > 0$ such that $\rho(f(p), f(p_0)) < \varepsilon$ whenever $d(p, p_n) < \delta$. f is called a **continuous function** on M if it is continuous at every point.

Very often it is more convenient to use the language of topology (open and closed sets) to describe these concepts. To introduce it let's denote the metric ball centered at p, $\{q \in M : d(q, p) < r\}$, by $B_r(p)$. A non-empty subset G of M is called an **open set** if $\forall p \in G$, there exists a positive r (depending on p) such that $B_r(p) \subset G$. We define the empty set to be an open set. Also the whole M is open because it contains every ball. It is easy to see that any metric ball $B_R(p_0)$ is an open set. For, let $p \in B_R(p_0)$, we claim that $B_r(p)$, $r = R - d(p, p_0)$, is contained inside $B_R(p_0)$. This is a consequence of the triangle inequality (D3): Let $q \in B_r(p)$, then $d(q, p_0) \leq d(q, p) + d(p, p_0) < r + d(p, p_0) = R$, so $q \in B_R(p_0)$. Roughly speaking, an open set is a set without boundary. A subset E is called a **closed set** if its complement $M \setminus E$ is an open set. The empty set is a closed set as its complement is the whole space. By the same reason M is also closed. So the empty set and the whole space are both open and closed.

Proposition 2.1. *Let (M, d) be a metric space. Then the union of open sets and the intersection of finitely many open sets are open. The intersection of closed sets and the union of finitely many closed sets are closed.*

Proof. That any countable or uncountable open sets still form an open set comes from definition. As for finite intersections, let $G = \bigcap_{k=1}^{n} G_k$ where G_k is open. For $x \in G \subset G_k$, we can find a metric ball $B_{r_k}(x) \subset G_k$ for each k since G_k is open. It follows that the ball $B_r(x), r = \min\{r_1, r_2, \cdots, r_n\}$ is contained in G, so G is open.

The assertions on closed sets come from taking complements of the assertions on open sets.

□

Notice that infinite intersection of open sets may not be open. Let us consider the open intervals $I_n = (-1/n, 1 + 1/n), n \in \mathbb{N}$ in $\mathbb{R}$ under the Euclidean metric. Then $\bigcap I_n = [0, 1]$ which is not open. Similarly, let $\{a_1, a_2, a_3, \cdots, \}$ be an enumeration of all rational numbers and set $F_n = \{a_1, a_2, \cdots, a_n\}$. Then each F_n is closed, but $\bigcup F_n$ is the set of all rational numbers which is clearly not closed in $\mathbb{R}$.

To have a better picture about the closed set we introduce the notion of a the limit point of a set. We call a point $p \in M$ a **limit point** of a set E if for all $r > 0$, $B_r(x) \setminus \{x\} \bigcap E \neq \emptyset$. The limit point is related to a set, but the limit, although it is also a point, is related to a sequence. They are not the same. For example, consider the sequence $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\}$, its limit is clearly 0. If we regard $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\}$ as a set, 0 is its unique limit point. However, for the sequence $\{0, 2, 1, 1, 1, \cdots\}$, the limit is 1 but as a set it has no limit point.

Proposition 2.2. *A non-empty set E is closed if and only if it contains all its limit points.*

Proof. Let E be a closed set. By definition $M \setminus E$ is open. If $p \in M \setminus E$, there exists r such that $B_r(p) \subset M \setminus E$, ie, $B_r(p) \bigcap E = \emptyset$. It follows that p cannot be a limit point of E. This shows that any limit point of E must belong to E.

Conversely, we need to show $M \setminus E$ is open. Since E already contains all limit points, any point $p \in M \setminus E$ cannot be a limit point of E. Therefore, there is an r such that $B_r(p) \bigcap E = \emptyset$, but that means $B_r(p) \subset M \setminus E$, so $M \setminus E$ is open. □

The **closure** of E, denoted by $\overline{E}$, is defined to be the union of E and its limit points. By Proposition 2.1 it is easily shown that $\overline{E}$ is the smallest closed set containing E, that is, $\overline{E} \subset F$ whenever F is a closed set containing E.

In terms of the language of open-closed sets (or topology), a sequence $\{x_n\} \to x$ can be expressed as, for each open set G containing x, there exists an n_0 such that $x_n \in G$ for all $n \geq n_0$.

For $f : (M, d) \mapsto (N, \rho)$ where (N, ρ) is another metric space. In terms of topology, we have the following characterization of continuity:

Proposition 2.3. *$f : (M, d) \mapsto (N, \rho)$ is continuous if and only if $f^{-1}(G)$ is open for any open G in N.*

Proof. Assume on the contrary that there is an open set G in N whose preimage is not open. We can find some $p_0 \in f^{-1}(G)$ and $p_n \in M \setminus f^{-1}(G)$ with $\{p_n\} \to p_0$. By continuity, $\{f(p_n)\} \to f(p_0)$. As G is open, there exists some n_0 such that $f(p_n) \in G$ for all $n \geq n_0$. But this means that $f^{-1}(G)$ contains p_n for all $n \geq n_0$, contradiction holds. We conclude that $f^{-1}(G)$ must be open when G is open.

On the other hand, suppose f is not continuous, then there exists $\{p_n\} \to p_0$ in M but $\{f(p_n)\}$ does not converge to $f(p_0)$. Then there exists $\rho > 0$ and a subsequence $\{f(p_{n_j})\}$, $f(p_{n_j}) \notin B_\rho(f(p_0))$, $\forall n_j$. As $B_\rho(f(p_0))$ is open, $f^{-1}(B_\rho(f(p_0)))$ is open in M, so there exists n_0 such that $p_n \in f^{-1}(B_\rho(f(p_0)))$. But then $f(p_n) \in B_\rho(f(p_o))$ for all $n \geq n_0$, contradiction holds. □

Let E be any nonempty subset of (M, d). Then it is clear that (E, d) forms a metric space. It is called a metric subspace or simply a subspace. As we will see, the subspaces formed by closed subsets are particularly important since they inherit many properties of M.

Now, let us return to normed spaces. Let $(X, \|\cdot\|)$ be a normed space. Define $d(x, y) = \|x - y\|$. Using (N1)-(N3), it is easy to verify (D1)-(D3) hold for d, so (X, d) becomes a metric space. This metric is called the **induced metric** of the norm $\|\cdot\|$. Of course, there are many metrics which are not induced by norms. But in functional analysis most metrics are induced in this way. ¿From now on whenever we have a normed space, we can talk about convergence and continuity implicitly referring to this metric. The following statements show that the norm and the algebraic operations on the vector spaces interact nicely with the metric.

Proposition 2.4. *Let $(X, \|\cdot\|)$ be a normed space. Then*

(a) The norm $\|\cdot\|$ is a continuous function from X to $[0, \infty)$;

(b) Addition, as considered as a map $X \times X \mapsto X$, and scalar multiplication, a map $\mathbb{F} \times X \mapsto X$, are continuous in $X \times X$ and $\mathbb{F} \times X$ respectively.

Proof. (a) $p_n \to p$ means $d(p_n, p) \to 0$. But then $\|p_n - p\| = d(p_n, p) \to 0$, $|\|p_n\| - \|p\|| \leq \|p_n - p\| \to 0$.

(b) We need to show $p_n \to p$ and $q_n \to q$ implies $p_n + q_n \to p + q$. This is clear from $d(p_n + q_n, p + q) = \|(p_n + q_n) - (p + q)\| \leq \|p_n - p\| + \|q_n - q\| = d(p_n, p) + d(q_n, q)$.

For scalar multiplication, need to show $\alpha_n \to \alpha$ and $p_n \to p$ implies $\alpha_n p_n \to \alpha p$. By (a), we have $\|p_n\| \to \|p\|$. Hence for $\varepsilon = 1$, there exists some n_0 such that $|\|p_n\| - \|p\|| < 1$, or $\|p_n\| \leq 1 + \|p\|$, for all $n \geq n_0$.

$$\begin{aligned} d(\alpha_n p_n, \alpha p) &= \|\alpha_n p_n - \alpha p\| = \|\alpha_n p_n - \alpha p_n + \alpha(p_n - p)\| \\ &\leq |\alpha_n - \alpha| \|p_n\| + |\alpha| \|p_n - p\| \\ &\leq |\alpha_n - \alpha|(\|p\| + 1) + |\alpha| \|p_n - p\| \\ &= |\alpha_n - \alpha|(\|p\| + 1) + |\alpha| d(p_n, p) \to 0 \end{aligned}$$

as $n \to \infty$.

□

As an interesting application of the continuity of norm, we study the equivalence problem for norms. Consider two norms defined on the same space $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$. We call $\|\cdot\|_2$ is **stronger** than $\|\cdot\|_1$ if there exists $C > 0$ such that

$$\|x\|_1 \leq C\|x\|_2, \ \forall x \in X.$$

In particular, it means $x_n \to x$ in $\|\cdot\|_2$ implies $x_n \to x$ in $\|\cdot\|_1$. Two norms are **equivalent** if $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$ and $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$. In other words, there exists $C_1, C_2 > 0$ such that

$$C_1\|x\|_2 \leq \|x\|_1 \leq C_2\|x\|_2, \ \forall x \in X.$$

Example 2.1. On $\mathbb{F}^n$ consider the p-metric $d_p(x, y) = \|x - y\|_p$ induced from the p-norm $(1 \leq p \leq \infty)$. In a previous exercise we were asked to show all these metrics are equivalent. In fact, this is a general fact as established by the following result.

Theorem 2.5. *Any two norms on a finite dimensional space are equivalent.*

Proof. In the following proof we assume the space is over $\mathbb{R}$. The same arguments work for spaces over $\mathbb{C}$.

Step 1: Take $X = \mathbb{R}^n$ first. It suffices to show that any norm on $\mathbb{R}^n$ is equivalent to the Euclidean norm. Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. For $x = \sum \alpha_j e_j$, recalling that $\|x\|_2 = \sqrt{\sum |\alpha_j|^2}$, we have

$$\|x\| \leq \sum |\alpha_j| \|e_j\| \leq \sqrt{\sum |\alpha_j|^2} \sqrt{\sum \|e_j\|^2} = C\|x\|_2,$$

where $C = (\sum_j \|e_j\|^2)^{1/2}$. This shows that $\|\cdot\|_2$ is stronger than $\|\cdot\|$. To establish the other inequality, letting $\varphi(x) = \|x\|$, from the triangle inequality $|\varphi(x) - \varphi(y)| \leq \|x-y\| \leq C\|x-y\|_2$ φ is a continuous function with respect to the Euclidean norm. Consider

$$\alpha \equiv \inf\{\varphi(x) : x \in \mathbb{R}^n, \|x\|_2 = 1\}.$$

As the function φ is positive on the unit sphere of $\|\cdot\|_2$, α is a nonnegative number. The second inequality will come out easily if α is positive. To see this we observe that for every nonzero $x \in \mathbb{R}^n$,

$$0 < \alpha \leq \varphi(\frac{x}{\|x\|_2}) = \frac{\|x\|}{\|x\|_2},$$

i.e.,

$$\alpha\|x\|_2 \leq \|x\|, \ \forall x.$$

To show that α is positive, we use the fact that every continuous function on a closed and bounded subset of $\mathbb{R}^n$ must attain its minimum. Applying it to φ and the unit sphere $\{\|x\|_2 = 1\}$, the infimum α is attained at some point x_0 and so in particular $\alpha = \varphi(x_0) > 0$.

Step 2: For any n dimensional space X, fix a basis $\{x_1, x_2, \cdots, x_n\}$. For any $x \in X$, we have a unique representation $x = \sum_{k=1}^n \alpha_k x_k$. The map $x \mapsto (\alpha_1, \cdots, \alpha_n)$ is a linear isomorphism from X to $\mathbb{R}^n$. Any norm $\|\cdot\|$ on X induces a norm $\|\cdot\widetilde{\|}$ on $\mathbb{R}^n$ by

$$\|(\alpha_1, \cdots, \alpha_n)\widetilde{\|} = \|\sum \alpha_k x_k\|.$$

Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X and let $\|\cdot\widetilde{\|}_1$ and $\|\cdot\widetilde{\|}_2$ be the corresponding norms on $\mathbb{R}^n$. From Step 1, there exist $C_1, C_2 > 0$ such that

$$C_1\|x\|_2 = C_1\|\alpha\widetilde{\|}_2 \leq \|\alpha\widetilde{\|}_1 = \|x\|_1 \leq C_2\|\alpha\widetilde{\|}_2 = C_2\|x\|_2.$$

□

Example 2.2. Consider the norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$ on $C[a,b]$. On one hand, from the obvious estimate

$$\|f-g\|_1 = \int_a^b |f-g|(x)dx \leq (b-a)\|f-g\|_\infty$$

we see that $\|\cdot\|_\infty$ is stronger than $\|\cdot\|_1$. But they are not equivalent. It is easy to find a sequence of functions $\{f_n\}$ in $C[a,b]$ which satisfies $\|f_n\|_\infty = 1$ but $\|f_n\|_1 \to 0$. Consequently, it is impossible to find a constant C such that $\|f\|_\infty \leq C\|f\|_1$ for all f. In other words, $\|\cdot\|_1$ cannot be stronger than $\|\cdot\|_\infty$.

2.2 Separability

There are some important and basic properties of the space of all real numbers which we would like to study in a general normed space. They are

- Separability
- Completeness
- Bolzano-Weierstrass property.

We study the first item in this section.

As we all know, the rational numbers are dense in the space of all real numbers. The notion of a dense set makes perfect sense in a metric space. A subset E of (M, d) is a **dense set** if its closure is the whole M, or equivalently, for every $p \in M$, there exists $\{p_n\}$ in E, $p_n \to p$. A metric space is called **separable** if it has a countable dense subset.

Thus $\mathbb{R}$ is separable because it contains the countable dense subset $\mathbb{Q}$. The following two results show that there are many separable normed spaces.

Proposition 2.6. *The following normed spaces are separable:*

(a) $(\mathbb{F}^n, \|\cdot\|_p)$ $(1 \le p \le \infty)$,

(b) ℓ^p $(1 \le p < \infty)$,

(c) $(C[a,b], \|\cdot\|_p)$ $(1 \le p \le \infty)$.

Proof. We only prove (c) and leave (a) and (b) to you. For any continuous, real-valued f, given any $\varepsilon > 0$, by Weierstrass approximation theorem there exists a polynomial p such that $\|f - p\|_\infty < \varepsilon$. The coefficients of p are real numbers in general, but we can approximate them by rational numbers, so without loss of generality we may assume its coefficients are rational. The set

$$E = \{p \in C[a,b] : p \text{ is a polynomial with rational coefficients}\}$$

forms a countable, dense subset of $(C[a,b], \|\cdot\|_\infty)$.

For any finite p, we observe that

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx\right)^{\frac{1}{p}} \le (b-a)^{\frac{1}{p}} \|f\|_\infty .$$

As for every f, there exists $p_n \in E$, $\|p_n - f\|_\infty \to 0$, we also have $\|p_n - f\|_p \to 0$, so E is also dense in $(C[a,b], \|\cdot\|_p)$.

When the function is complex-valued, simply apply the above result to its real and imaginary parts. □

Proposition 2.7. *Any subset of a separable metric space is again separable.*

Proof. Let $Y \subset X$ and E a countable, dense subset of X. Write $E = \{p_k\}_{k=1}^\infty$. For each m, $B_{\frac{1}{m}}(p_k) \cap Y$ may or may not be empty. Pick a point $p_{m,k}$ if it is not empty. The collection of all these $p_{m,k}$ points forms a countable subset S of Y. We claim that it is dense in Y. For, any $p \in Y$, and $m > 0$, there exists $p_k \in B_{\frac{1}{m}}(p)$, $p_k \in E$ by assumption. But then $p \in B_{\frac{1}{m}}(p_k)$ which means $B_{\frac{1}{m}}(p_k) \cap Y \neq \emptyset$. Then we have $p_{m,k} \in B_{\frac{1}{m}}(p_k)$ and so $d(p, p_{m,k}) \le d(p, p_k) + d(p_k, p_{m,k}) < 2/m$. □

Now we give an example of a non-separable space.

Proposition 2.8. ℓ^∞ *is not separable.*

Proof. Consider the subset F of ℓ^∞ consisting of all sequences of the form $(a_1, a_2, a_3, \cdots)$ where $a_k = 1$ or 0. In view of Proposition 2.7, it suffices to show that F is not separable. First of all, it is an uncountable set as easily seen from the correspondence

$$(a_1, a_2, a_3, \cdots) \leftrightarrow 0.a_1 a_2 a_3 \cdots \text{ in binary representation of a real number}$$

which maps F onto $[0, 1]$. For each $x, y \in F$, we have $d(x, y) = \|x - y\|_\infty = \sup_k |x_k - y_k| = 1$. It follows that the balls $B_{1/2}(x),\ x \in F$, are mutually disjoint. Let E be a dense set in F. By definition there exists some $p_x \in B_{1/2}(x) \cap E$. Since these balls are disjoint, all p_x are distinct, so $\{p_x\}$ forms an uncountable subset of E. Thus E is also uncountable. We have shown that there are no countable dense subsets in F, that is, F is not separable. □

Can you find more non-separable normed spaces?

2.3 Completeness

A metric space (M, d) is **complete** if every Cauchy sequence converges. As we all know, $\mathbb{R}$ is a complete metric space.

Proposition 2.9. *The following spaces are complete:*

(a) $(\mathbb{F}^n, \|\cdot\|_p)$ $(1 \le p \le \infty)$,

(b) ℓ^p $(1 \le p \le \infty)$,

(c) $(C[a, b], \|\cdot\|_\infty)$.

Proof. (a) Let $\{p^k\}$ be a Cauchy sequence in $\mathbb{F}^n$. For $p^k = (p_1^k, \cdots, p_n^k)$, from

$$|p_j^k - p_j^l| \le \|p^k - p^l\|_p$$

we see that $\{p_j^k\}$ is a Cauchy sequence in $\mathbb{F}$ for each $j = 1, 2, \cdots, n$. By the completeness of $\mathbb{F}$ there exists p_j such that $p_j^k \to p_j$ as $k \to \infty$ for each j. Given $\varepsilon > 0$, there exists k_0 such that

$$|p_j^k - p_j| \le \varepsilon, \quad \forall k \ge k_0.$$

Summing up over j, $\|p^k - p\|_p < n^{1/p} \max_j |p_j^k - p_j| < n^{1/p}\varepsilon$, $\forall k \ge k_0$, which shows that $p^k \to p \equiv (p_1, \cdots, p_n)$. We leave the proofs of (b) and (c) to the reader. Note that (c) was a theorem on uniform convergence in elementary analysis. □

But $C[a, b]$ is not complete in the L^p-norm $(1 \le p < \infty)$. To find a divergent Cauchy sequence we let

$$\varphi_n(x) = \begin{cases} 1, & x \in [-1, 0] \\ -nx + 1, & x \in [0, \frac{1}{n}] \\ 0, & x \in [\frac{1}{n}, 1]. \end{cases}$$

and

$$\varphi(x) = \begin{cases} 1, & x \in [-1, 0] \\ 0, & x \in (0, 1]. \end{cases}$$

It is easy to see that $\|\varphi_n - \varphi\|_p \to 0$. Therefore,

$$\|f_n - f_m\|_p \le \|f_n - f\|_p + \|f_m - f\|_p \to 0,$$

as $n, m \to \infty$, that is, $\{f_n\}$ is a Cauchy sequence in p-norm. To show that it does not converge to a continuous function let's assume on the contrary it converges to some continuous f. From

$$\Big(\int_{-1}^{0} |f - \varphi|^p\Big)^{1/p} \le \Big(\int_{-1}^{0} |f - \varphi_n|^p\Big)^{1/p} + \Big(\int_{-1}^{1} |\varphi_n - \varphi|^p\Big)^{1/p} \to 0,$$

as $n \to \infty$, we see that f is identical to φ on $[-1, 0]$ since both functions are continuous on $[-1, 0]$. In particular, $f(0) = 1$, so by continuity $f > 0$ on $[0, \delta]$ for some positive δ. However, since f and φ are continuous on $[\delta, 1]$ by a similar argument as above f is identical to φ on $[\delta, 1]$, but then $g(\delta) = \varphi(\delta) = 0$, contradiction holds.

Fortunately, one can make any metric space complete by putting in ideal points. In general, a map $f : (M, d) \longrightarrow (N, \rho)$ is called a **metric preserving map** if $\rho(f(x), f(y)) = d(x, y)$ for all x, y in M. Note that a metric preserving map is necessarily injective. In some texts the name an **isometry** is used instead of a metric preserving map. We prefer to use the former and reserve the latter for a metric preserving and surjective map. A complete metric space $(\widetilde{M}, \widetilde{d})$ is called the **completion** of a metric space (M, d) if there exists an metric preserving map Φ of M into $\widetilde{M}$ such that $\Phi(M)$ is dense in $\widetilde{M}$.

Theorem 2.10. *Every metric space has a completion.*

Proof. Let $\mathscr{C}$ be the collection of all Cauchy sequences in (M, d). We introduce a relation $\sim$ on $\mathscr{C}$ by $x \sim y$ if and only if $d(x_n, y_n) \to 0$ as $n \to \infty$. It is routine to verify that $\sim$ is an equivalence relation on $\mathscr{C}$. Let $\widetilde{M} = \mathscr{C}/\sim$ and define a map: $\widetilde{M} \times \widetilde{M} \mapsto [0, \infty)$ by

$$\widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} d(x_n, y_n)$$

where $x = (x_1, x_2, x_3, \cdots)$ and $y = (y_1, y_2, y_3, \cdots)$ are respective representatives of $\widetilde{x}$ and $\widetilde{y}$. We note that the limit in the definition always exists: For

$$d(x_n, y_n) \le d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n)$$

and, after switching m and n,

$$|d(x_n, y_n) - d(x_m, y_m)| \le d(x_n, x_m) + d(y_m, y_n).$$

As x and y are Cauchy sequences, $d(x_n, x_m)$ and $d(y_m, y_n) \to 0$ as $n, m \to \infty$, so $\{d(x_n, y_n)\}$ is a Cauchy sequence of real numbers.

Step 1. Well-definedness of $\widetilde{d}$. To show that $\widetilde{d}(\widetilde{x}, \widetilde{y})$ is independent of their representatives let $x \sim x'$ and $y \sim y'$. We have

$$d(x_n, y_n) \le d(x_n, x_n') + d(x_n', y_n') + d(y_n', y_n).$$

After switching x and x', and y and y',

$$|d(x_n, y_n) - d(x_n', y_n')| \le d(x_n, x_n') + d(y_n, y_n').$$

As $x \sim x'$ and $y \sim y'$, the right hand side of this inequality tends to 0 as $n \to \infty$. Hence $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x_n', y_n')$.

Step 2. $\widetilde{d}$ is a metric. This is straightforward and is left as an exercise.

Step 3. Φ is metric preserving and has a dense image in $\widetilde{M}$. More precisely, we need to show that there is a map $\Phi : M \mapsto \widetilde{M}$ so that $\widetilde{d}(\Phi(x), \Phi(y)) = d(x, y)$ and $\Phi(M)$ is dense in $\widetilde{M}$.

Given any x in M, the "constant sequence" $(x, x, x, \cdots)$ is clearly a Cauchy sequence. Let $\widetilde{x}$ be its equivalence class in $\mathscr{C}$. Then $\Phi x = \widetilde{x}$ defines a map from M to $\widetilde{M}$. Clearly

$$\widetilde{d}(\Phi(x), \Phi(y)) = \lim_{n\to\infty} d(x_n, y_n) = d(x, y)$$

since $x_n = x$ and $y_n = y$ for all n, so Φ is metric preserving and it is injective in particular.

To show that $\Phi(M)$ is dense in $\widetilde{M}$ we observe that any $\widetilde{x}$ in $\widetilde{M}$ is represented by a Cauchy sequence $x = (x_1, x_2, x_3, \cdots)$. Consider the constant sequence $x^n = (x_n, x_n, x_n, \cdots) \in \Phi(M)$. We have

$$\widetilde{d}(\widetilde{x}, \widetilde{x}_n) = \lim_{m\to\infty} d(x_m, x_n).$$

Given $\varepsilon > 0$, there exists n_0 such that $d(x_m, x_n) < \varepsilon/2$ for all $m, n \geq n_0$. Hence $\widetilde{d}(\widetilde{x}, \widetilde{x}_n) = \lim_{m\to\infty} d(x_m, x_n) < \varepsilon$ for $n \geq n_0$. That is $\widetilde{x}^n \to \widetilde{x}$ as $n \to \infty$, so $\Phi(M)$ is dense in M.

Step 4. $\widetilde{d}$ is a complete metric on $\widetilde{M}$. Let $\{\widetilde{x}^n\}$ be a Cauchy sequence in $\widetilde{M}$. As $\Phi(M)$ is dense in $\widetilde{M}$, for each n we can find a $\widetilde{y}^n$ in $\Phi(M)$ such that

$$\widetilde{d}(\widetilde{x}^n, \widetilde{y}^n) < \frac{1}{n}.$$

So $\{\widetilde{y^n}\}$ is Cauchy in $\widetilde{d}$. Let y_n be the point in M so that $y^n = (y_n, y_n, y_n, \cdots)$ represents $\widetilde{y}^n$. Since Φ is metric preserving and $\{\widetilde{y}^n\}$ is Cauchy in $\widetilde{d}$, $\{y_n\}$ is a Cauchy sequence in M. Let $(y_1, y_2, y_3, \cdots) \in \widetilde{y}$ in $\widetilde{M}$. We claim that $\widetilde{y} = \lim_{n\to\infty} \widetilde{x}^n$ in $\widetilde{M}$. For, we have

$$\begin{aligned}\widetilde{d}(\widetilde{x}^n, \widetilde{y}) &\leq \widetilde{d}(\widetilde{x}^n, \widetilde{y}^n) + \widetilde{d}(\widetilde{y}^n, \widetilde{y}) \\ &\leq \frac{1}{n} + \lim_{m\to\infty} d(y_n, y_m) \to 0\end{aligned}$$

as $n \to \infty$.

□

The idea of this proof is due to Cantor, who used equivalence classes of Cauchy sequences of rational numbers to construct real numbers. Another popular approach for the real number system is by "Dedekind cut". See the old book by E. Landau "Foundations of Analysis" for more.

The uniqueness of the completion could be formulated as follows. Let $\Psi_i : (M, d) \to (M_i, d_i), i = 1, 2,$ be two metric preserving maps with dense images. Then the map $\Psi_2\Psi_1^{-1} : \Psi_1(M_1) \to M_2$ can be extended to be an isometry between M_1 and M_2. I leave the proof of this fact to you.

When a given metric space is induced from a normed space, it is naturally to ask it is possible to make the completion into a normed space so that the complete metric is induced by the norm on the completion. To this question we have a satisfactory answer as stated in the theorem below. You may also formulate a uniqueness result for the completion of a normed space.

Theorem 2.11. *Let $(X, \|\cdot\|)$ be a normed space and $\widetilde{X}$ its completion under the induced metric of X. There is a unique normed space structure on $\widetilde{X}$ so that the quotient map $x \longmapsto \widetilde{x}$ becomes linear and norm-preserving. Moreover, the metric induced by this norm $\widetilde{X}$ is identical to the completion metric.*

Proof. We only give the outline of the proof and leave the patient reader to provide all details.

Step 1. Let $\Phi : X \longrightarrow \widetilde{X}$ be the quotient map. As $\Phi(X)$ is dense in $\widetilde{X}$, for any $\widetilde{x}, \widetilde{y}$ in $\widetilde{X}$, we can find sequences $\{\widetilde{x_n}\}$, $\{\widetilde{y_n}\}$ converging to $\widetilde{x}, \widetilde{y}$ respectively. We define an addition and a scalar multiplication on $\widetilde{X}$ by

$$\widetilde{x} + \widetilde{y} \equiv \lim_{n\to\infty} \Phi(x_n + y_n),$$

and

$$\alpha\widetilde{x} \equiv \lim_{n\to\infty} \Phi(\alpha x_n),$$

where x_n and y_n are representatives of $\widetilde{x_n}$ and $\widetilde{y_n}$ respectively. You need to establish three things. First, these operations are well-defined, that's, they are independent of the representatives. Second, they make

$\widetilde{X}$ into a vector space. Third, the map Φ is linear from X to $\widetilde{X}$. (In fact, this follows immediately from the definitions.)

Step 2. Introduce a map on $\widetilde{X}$ by
$$\|\widetilde{x}\| \equiv \widetilde{d}(\widetilde{x}, 0).$$
Then verify the following three facts: First, translational invariance: $\widetilde{d}(\widetilde{x}+\widetilde{y}, \widetilde{y}) = \widetilde{d}(\widetilde{x}, 0)$ for all $\widetilde{x}$ and $\widetilde{y}$. Second, use translational invariance to show that this map really defines a norm on $\widetilde{X}$. Third, show that the metric induced by this norm coincides with the completion metric. This in fact follows from the definition of the norm.

Step 3. Show that if there is another normed space structure on $\widetilde{X}$ so that the quotient map Φ becomes linear and norm-preserving, then this normed space structure is identical to the one given by Steps 1 and 2. Essentially this follows from the fact that $\Phi(X)$ is dense in $\widetilde{X}$. □

As an immediate application of these results, we let $L^p(a,b)$ be the completion of $C[a,b]$ under the L^p-norm. We shall call an element in $L^p(a,b)$ an L^p-function, although it makes sense only when the element is really in $C[a,b]$. Such terminology is based on another construction in real analysis where we really identify $L^p(a,b)$ as the function space consisting of L^p-integrable functions. We do not need this fact in this course.

A complete normed space is called a **Banach space**. Banach space is one of the fundamental concepts in functional analysis. Now we know that even a space is not complete, we can make it into a Banach space. The following nice properties of Banach spaces hold:

- Any closed subspace of a Banach space is a Banach space.
- The product space of two Banach spaces is a Banach space under the product norm.
- For any closed subspace Z of a Banach space X, the quotient space X/Z is a Banach space under the quotient norm.

We have shown that the spaces $\mathbb{F}^n, \ell^p (1 \le p \le \infty), C[a,b]$ and $L^p[a,b], p \in [1,\infty)$, are Banach spaces. In fact, for any metric space X, the space $C_b(X) = \{f : f \text{ is bounded and continuous in } X\}$ forms a Banach space under the sup-norm. For any measure space (X,μ), the space $L^p(X,\mu) = \{f : f \text{ is } L^p\text{-integrable}\}$ forms a Banach space under the L^p-norm. Finally, without requiring any topology or integrability on the functions, the space $L^\infty(X)$ consisting of all bounded functions in a nonempty set X is a Banach space under the sup-norm.

It is an exercise to show that every subspace in a finite dimensional normed space is closed. Can you find a non-closed subspace in $C[a,b]$?

So far we have encountered three types of mathematical structure, namely, those of a vector space, a metric space and a normed space. How do we identify two spaces from the same structure? Well, first of all, we view two vector spaces the same if there exists a bijective linear map between them. A bijective linear map is also called a linear isomorphism. Next, two metric spaces are the same if there exists a metric preserving bijective map, that is, an isometry, from one to the other. Finally, two normed spaces are the same if there exists a norm-preserving linear isomorphism from one to the other.

2.4 Sequential Compactness

In the space of real numbers, any bounded sequence has a convergent subsequence. This property is called the Bolzano-Weierstrass property. In a general setting, it is more convenient to put this concept in another way. Let E be a subset of the metric space (M,d). E is called **sequentially compact** if every sequence in E enjoys the Bolzano-Weierstrass property, that is, it contains a convergent subsequence, in E. Any sequentially compact set is necessarily a closed set. It is clear that the Bolzano-Weierstrass property essentially refers to the fact that the interval $[a, b]$ is sequentially compact in $\mathbb{R}$. The same as in the case of $\mathbb{R}$, one can show that every closed and bounded set in $\mathbb{R}^n$ is sequentially compact. Surprisingly, this property is a characterization of finite dimensionality.

Theorem 2.12. *Any closed ball in a normed space is sequentially compact if and only if the space is of finite dimension.*

Lemma 2.13. *Let Y be any proper finite dimensional subspace of the normed space $(X, \|\cdot\|)$. Then for any $x \in X \setminus Y$, there exists $y_0 \in Y$ such that*

$$d \equiv dist(x, Y) \equiv \inf_{y \in Y} \|x - y\| > 0.$$

is realized at y_0.

The distance d is positive because Y is closed due to finite dimensionality and x stays outside Y.

Proof. Let $\{y_k\}$ be a minimizing sequence of the distance, that is, $d = \lim_{k\to\infty} \|x - y_k\|$. We may assume $\|x - y_k\| \le d + 1$, for all k. Then

$$\|y_k\| \le \|x\| + \|y_k - x\| \le \|x\| + d + 1,$$

which means that $\{y_k\}$ is a bounded sequence in Y. Since Y is finite dimensional, it is closed and Bolzano-Weierstrass property holds in it, there exists a subsequence $\{y_{n_j}\}$ converging to some y_0 in Y. We have $d = \lim_{n_j\to\infty} \|x - y_{n_j}\| = \|x - y_0\|$, hence y_0 realizes the distance between x and Y. □

Proof of Theorem 2.12. It suffices to show that the closed unit ball $\{x \in X : \|x\| \le 1\}$ is not sequentially compact when X is of infinite dimension.

Let $\{x_1, x_2, x_3, \cdots\}$ be a sequence of linearly independent vectors in X. We are going to construct a sequence $\{z_n\}$, $z_n \in \langle x_1, x_2, \cdots, x_n\rangle$, $\|z_n\| = 1$ satisfying that $\|z_n - x\| \ge 1$, for all $x \in \langle x_1, x_2, \cdots, x_{n-1}\rangle$, $n \ge 2$.

Set $z_1 = x_1/\|x_1\|$. For $x_n \notin \langle x_1, x_2, \cdots, x_{n-1}\rangle$, $n \ge 2$, let y_{n-1} be the point in $\langle x_1, x_2, \cdots, x_{n-1}\rangle$ realizing $dist(x_n, \langle x_1, \cdots, x_{n-1}\rangle)$. Let

$$z_n = \frac{x_n - y_{n-1}}{\|x_n - y_{n-1}\|}.$$

Then $\|z_n\| = 1$ and, for all $y \in \langle x_1, \cdots, x_{n-1}\rangle$,

$$\|z_n - y\| = \left\| \frac{x_n - y_{n-1}}{\|x_n - y_{n-1}\|} - y \right\| = \frac{\|x_n - y'\|}{\|x_n - y_{n-1}\|} \ge 1,$$

where $y' = y_{n-1} + \|x_n - y_{n-1}\|\, y \in \langle x_1, \cdots, x_{n-1}\rangle$, since $\|x_n - y_{n-1}\| \le \|x_n - y'\|$.

We claim that the bounded sequence $\{z_n\}$ does not have a convergent subsequence. For, if it has, this subsequence is a Cauchy sequence. Taking $\varepsilon = 1$, we have

$$\|z_{n_k} - z_{n_j}\| < 1, \quad k, j \text{ sufficiently large.}$$

Taking $n_k > n_j$, as $\|z_{n_k} - x\| \ge 1$, for all $x \in \langle x_1, \cdots, x_{n_k-1}\rangle$ and $z_{n_j} \in \langle x_1, \cdots, x_{n_k-1}\rangle$, we have

$$\|z_{n_k} - z_{n_j}\| \ge 1,$$

contradiction holds. We conclude that the closed unit ball is not sequentially compact in an infinite dimensional normed space. □

Digressing a bit, let x be a point lying outside Y, a proper subspace of the normed space X. A point in Y realizing the distance from x to Y is called the **best approximation** from x to Y. It always exists when Y is a finite dimensional subspace. However, things change dramatically when the subspace has infinite dimension. For instance, let Y be the closed subspace of $C[-1,1]$ given by

$$\int_{-1}^{0} f(x)dx = 0, \quad \int_{0}^{1} f(x)dx = 0, \quad \forall f \in Y,$$

and h a continuous function satisfying

$$\int_{-1}^{0} h(x)dx = 1, \quad \int_{0}^{1} h(x)dx = -1.$$

One can show that the distance from h to Y is equal to 1, but it is not realized at any point on Y. You may try to prove this fact or consult chapter 5 of [L]. Later we will see that the best approximation problem has always a solution when the space X is reflexive.

2.5 Arzela-Ascoli Theorem

From the last section, we know that not all bounded sequences in an infinite dimensional normed space have convergent subsequences. It is natural to ask what additional conditions are needed to ensure this property. For the space $C[a,b]$, a complete answer is provided by the Arzela-Ascoli theorem. This theorem gives a necessary and sufficient condition when a closed and bounded set in $C[a,b]$ is sequentially compact. In order to have wider applications, we shall work on a more general space $C(K)$, where K is a closed, bounded subset of $\mathbb{R}^n$, instead of $C[a,b]$. As every continuous function in K attains its maximum and minimum, its sup-norm is always finite. It can be shown that $C(K)$ is a separable Banach space under the sup-norm.

The crux for sequential compactness for continuous functions lies on the notion of equicontinuity. Let E be a subset of $\mathbb{R}^n$. A subset $\mathcal{F}$ of $C(E)$ is **equicontinuous** if for every $\varepsilon > 0$, there exists some δ such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } f \in \mathcal{F}, \quad \text{and} \quad |x - y| < \delta, \ x, y \in E.$$

Recall that a function is uniformly continuous in E if for each $\varepsilon > 0$, there exists some δ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$, $x, y \in E$. So, equicontinuity means that δ can further be chosen independent of individual functions in $\mathcal{F}$.

There are various ways to show that a family of functions is equicontinuous. A function f defined in a subset E of $\mathbb{R}^n$ is called **Hölder continuous** if there exists some $\alpha \in (0,1)$ such that

$$|f(x) - f(y)| \le L|x - y|^{\alpha}, \quad \text{for all } x, y \in E, \tag{2.1}$$

for some constant L. The number α is called the Hölder exponent. The function is called **Lipschitz continuous** if (2.1) holds for α equals to 1. A family of functions $\mathcal{F}$ in $C(E)$ is said to satisfy a **uniform Hölder** or **Lipschitz condition** if all members in $\mathcal{F}$ are Hölder continuous with the same α or Lipschitz continuous and (2.1) holds for the same constant L. Clearly, such $\mathcal{F}$ is equicontinuous. The following situation is commonly encountered in the study of differential equations. The philosophy is that equicontinuity can be obtained if there is a good, uniform control on the derivatives of functions in $\mathcal{F}$.

Proposition 2.14. *Let $\mathcal{F}$ be a subset of $C(G)$ where G is a convex set in $\mathbb{R}^n$. Suppose that each member in $\mathcal{F}$ is differentiable and there is a uniform bound on the partial derivatives of the functions in $\mathcal{F}$. Then $\mathcal{F}$ is equicontinuous.*

Proof. For, x and y in G, $(1-t)x + ty$, $t \in [0,1]$, belongs to G by convexity. Let $\psi(t) \equiv f((1-t)x + ty)$. From the mean-value theorem

$$\psi(1) - \psi(0) = \psi'(t^*)(1-0), \ t^* \in [0,1],$$

and the chain rule

$$\psi'(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}((1-t)x + ty)(y_j - x_j),$$

we have

$$|f(y) - f(x)| \le \sum_j |\frac{\partial f}{\partial x_j}||y_j - x_j| \le \sqrt{n}M|y - x|,$$

where $M = \sup\{|\partial f/\partial x_j(x)| : x \in G, j = 1, \ldots, n, f \in \mathcal{F}\}$ after using Cauchy-Schwarz inequality. So $\mathcal{F}$ satisfies a uniform Lipschitz condition with the Lipschitz constant $n^{1/2}M$. □

Theorem 2.15 (Arzela-Ascoli). *Let $\mathcal{F}$ be a closed set in $C(K)$ where K is a closed and bounded set in $\mathbb{R}^n$. Then $\mathcal{F}$ is sequentially compact if and only if it is bounded and equicontinuous.*

A set E is called bounded if there exists $M > 0$ such that

$$|f(x)| \le M, \quad \text{for all } f \in E \text{ and } x \in K.$$

In other words, it is a bounded set in the metric induced by the sup-norm. This theorem was proved for $C[a, b]$ in the end of the nineteenth century by two italian mathematicians, the sufficient part by Ascoli and the necessary part by Arzela respectively.

We shall need the following useful fact.

Lemma 2.16. *Let E be a set in the metric space (X, d). Then*

(a) That E is sequentially compact implies that for any $\varepsilon > 0$, there exist finitely many ε-balls covering E.

(b) Assuming that E is closed and (X, d) is complete, the converse of (a) is true.

Proof. (a) Suppose on the contrary that there exists some $\varepsilon_0 >$ such that no finite collection of ε_0-balls covers E. For a fixed x_1, the ball $B_{\varepsilon_0}(x_1)$ does not cover E, so we can pick $x_2 \notin B_{\varepsilon_0}(x_1)$. As $B_{\varepsilon_0}(x_1)$ and $B_{\varepsilon_0}(x_2)$ together do not cover E, there is $x_3 \notin B_{\varepsilon_0}(x_1) \cup B_{\varepsilon_0}(x_2)$. Continuing this way, we find a sequence $\{x_n\}$ satisfying $x_n \notin B_{\varepsilon_0}(x_1) \cup \cdots \cup B_{\varepsilon_0}(x_{n-1})$. In particular, $d(x_i, x_j) \ge \varepsilon_0$ for distinct i, j, which shows that $\{x_n\}$ cannot have any convergent subsequence, a contradiction the sequential compactness of E.

(b). Let $\{x_n\}$ be a sequence in E. We may assume that it has infinitely many distinct elements, otherwise the conclusion is trivial. Since E can be covered by finitely many balls of radius 1, there exists one, say B^1, which contains infinitely many elements of E. Next, as E can be covered by finitely many balls of radius $1/2$, there exists B^2 of radius $1/2$ so that $B^1 \cap B^2$ contains infinitely many elements of E. Continuing this we get B^n of radius $1/n$ such that $B^1 \cap B^2 \cdots \cap B^n \cap E$ is non-empty for all n. Pick $x_{n_j} \in B^1 \cap B^2 \cdots \cap B^j \cap E$ with $n_{j-1} < n_j$. Then $\{x_{n_j}\}$ is a subsequence of $\{x_n\}$ which is also a Cauchy sequence. As X is complete and E is closed, it is convergent in E. We consider that E is sequentially compact. □

We shall also use the following lemma from elementary analysis.

Lemma 2.17. *Let $\{f_n\}$ be a bounded sequence of functions from the countable set $\{z_1, z_2, \cdots,\}$ to $\mathbb{F}$. There is a subsequence of $\{f_n\}$, $\{g_n\}$, such that $\{g_n(z_j)\}$ is convergent for every z_j.*

Proof. Since $\{f_n(z_1)\}$ is a bounded sequence in $\mathbb{F}$, we can extract a subsequence $\{f_n^1\}$ such that $\{f_n^1(z_1)\}$ is convergent. Next, as $\{f_n^1\}$ is bounded, it has a subsequence $\{f_n^2\}$ such that $\{f_n^2(z_2)\}$ is convergent. Keep doing in this way, we obtain sequences $\{f_n^j\}$ satisfying (i) $\{f_n^{j+1}\}$ is a subsequence of $\{f_n^j\}$ and (ii) $\{f_n^j(z_1)\}, \{f_n^j(z_2)\}, \cdots, \{f_n^j(z_j)\}$ are convergent. Then the diagonal sequence $\{g_n\}$, $g_n = f_n^n$, for all $n \ge 1$, is a subsequence of $\{f_n\}$ which converges at every z_j. □

The subsequence selected in this way is sometimes called to Cantor's diagonal sequence.

Proof of Arzela-Ascoli Theorem.

Assuming boundedness and equicontinuity of $\mathcal{F}$, we would like to show that $\mathcal{F}$ is sequentially compact.

Since K is sequentially compact in $\mathbb{R}^n$, by Lemma 2.16, for each $j \geq 1$, we can cover K by finitely many balls $D_{1/j}(x_1^j), \cdots, D_{1/j}(x_K^j)$ where the number K depends on j. For any sequence $\{f_n\}$ in $\mathcal{F}$, by Lemma 2.17, we can pick a subsequence from $\{f_n\}$, denoted by $\{g_n\}$, such that $\{g_n(x_k^j)\}$ is convergent for each x_k^j. We claim that $\{g_n\}$ is Cauchy in $C(K)$. For, due to the equicontinuity of $\mathcal{F}$, for every $\varepsilon > 0$, there exists a δ such that $|g_n(x) - g_n(y)| < \varepsilon$, whenever $|x - y| < \delta$. Pick j_0, $1/j_0 < \delta$. Then for $x \in K$, there exists $x_k^{j_0}$ such that $|x - x_k^{j_0}| < 1/j_0 < \delta$,

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_k^{j_0})| + |g_n(x_k^{j_0}) - g_m(x_k^{j_0})| + |g_m(x_k^{j_0}) - g_m(x)| \\ &< \varepsilon + |g_n(x_k^{j_0}) - g_m(x_k^{j_0})| + \varepsilon. \end{aligned}$$

As $\{g_n(x_k^{j_0})\}$ converges, there exists n_0 such that

$$|g_n(x_k^{j_0}) - g_m(x_k^{j_0})| < \varepsilon, \quad \text{for all } n, m \geq n_0. \tag{2.2}$$

Here n_0 depends on $x_k^{j_0}$. As there are finitely many $x_k^{j_0}$'s, we can choose some N_0 such that (2.2) holds for all $x_k^{j_0}$ and $n, m \geq N_0$. It follows that

$$|g_n(x) - g_m(x)| < 3\varepsilon, \quad \text{for all } n, m \geq N_0,$$

i.e., $\{g_n\}$ is Cauchy in $C(K)$. By the completeness of $C(K)$ and the closedness of $\mathcal{F}$, $\{g_n\}$ converges to some function in $\mathcal{F}$.

Conversely, by Lemma 2.16, for each $\varepsilon > 0$, there exist $f_1, \cdots, f_N \in \mathcal{F}$ such that $\mathcal{F} \subset \bigcup_{j=1}^N B_\varepsilon(f_j)$ where N depends on ε. So for any $f \in \mathcal{F}$, there exists f_j such that

$$|f(x) - f_j(x)| < \varepsilon, \quad \text{for all } x \in K.$$

As each f_j is continuous, there exists δ_j such that $|f_j(x) - f_j(y)| < \varepsilon$ whenever $|x - y| < \delta_j$. Letting $\delta = \min\{\delta_1, \cdots, \delta_N\}$, then

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < 3\varepsilon,$$

for $|x - y| < \delta$, so S is equicontinuous. As $\mathcal{F}$ can be covered by finitely many 1-balls, it is also bounded.

We have completed the proof of Arzela-Ascoli theorem.

We note the following useful corollary of the theorem, sometimes called Ascoli's theorem.

Corollary 2.18. *A sequence in $C(K)$ where K is a closed, bounded set in $\mathbb{R}^n$ has a convergent subsequence if it is uniformly bounded and equicontinuous.*

Proof. Let $\mathcal{F}$ be the closure of the sequence $\{f_j\}$. As this sequence is uniformly bounded, there exists some M such that

$$|f_j(x)| \leq M, \quad \forall x \in K, \ j \geq 1.$$

Consequently, any limit point of $\{f_j\}$ also satisfies this estimate, that is, $\mathcal{F}$ is bounded in $C(K)$. Similarly, by equicontinuity, for every $\varepsilon > 0$, there exists some δ such that

$$|f_j(x) - f_j(y)| < \frac{\varepsilon}{2}, \quad \forall x, y \in K, \ |x - y| < \delta.$$

As a result, any limit point f of $\{f_j\}$ satisfies

$$|f(x) - f(y)| \le \frac{\varepsilon}{2} < \varepsilon, \quad \forall x, y \in K, \ |x - y| < \delta,$$

so $\mathcal{F}$ is also equicontinuous. Now the conclusion follows from the Arzela-Ascoli theorem.

□

We present an application of Arzela-Ascoli theorem to ordinary differential equations. Consider the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(0) = 0,$$

where f is a continuous function defined in $[-a, a] \times [-b, b]$. We are asked to find a differentiable function $y(x)$ so that this equation is satisfied for x in some interval containing the origin. Under the further assumption that f satisfies the "Lipschitz condition": For some constant L

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2|, \text{ for all } x \in [-a, a], \ y_1, y_2 \in [-b, b],$$

we learn from a course on ordinary differential equations that there exists a *unique* solution to this initial value problem defined on the interval $I = (-a', a')$, where $a' = \min\{a, b/M\}$, where $M = \max\{|f(x, y)| : (x, y) \in [-a, a] \times [-b, b]\}$.

Now, let us show that the Lipschitz condition can be removed as far as *existence* is in concern. First of all, by Weierstrass approximation theorem, there exists a sequence of polynomials $\{f_n\}$ approaching f in $C([-a, a] \times [-b, b])$ uniformly. In particular, it means that $M_n \to M$, where $M_n = \max\{|f_n(x, y)| : (x, y) \in [-a, a] \times [-b, b]$. As each f_n satisfies the Lipschitz condition (why?), there is a unique solution y_n defined on $I_n = (-a_n, a_n), a_n = \min\{a, b/M_n\}$ for the initial value problem

$$\frac{dy}{dx} = f_n(x, y), \quad y(0) = 0.$$

¿From $|dy_n/dx| \le M_n$ and $\lim_{n\to\infty} M_n = M$, we know from Proposition 2.14 that $\{y_n\}$ forms an equicontinuous family. Clearly, it is also bounded. By Arzela-Ascoli theorem, it contains a subsequence $\{y_{n_j}\}$ converging uniformly to a continuous function $y \in I$ on every subinterval $[\alpha, \beta]$ of I and $y(0) = 0$ holds. It remains to check that y solves the differential equation for f.

Indeed, each y_n satisfies the integral equation

$$y_n(x) = \int_0^x f(t, y_n(t))dt, \quad x \in I_n.$$

As $\{y_{n_j}\} \to y$ uniformly, $\{f(x, y_{n_j}(x))\}$ also tends to $f(x, y(x))$ uniformly. By passing to limit in the formula above, we conclude that

$$y(x) = \int_0^x f(t, y(t))dt, \quad x \in I$$

holds. By the fundamental theorem of calculus, y is differentiable and a solution to our initial value problem.

The solution may not be unique without the Lipschitz condition. Indeed, the function $y_1(x) \equiv 0$ solves the initial value problem $y'(x) = y^{1/2}, \ y(0) = 0$, and yet there is another solution given by $y_2(x) = x^2/4, \ x \ge 0$ and vanishes for $x < 0$.

You may google for more applications of Arzela-Ascoli theorem.

Exercise 2

In Problems 1-3 f is a function from $(M, d) \to (N, \rho)$.

1. Let $p \in M$. Show that f is continuous at p if and only if for every open set G in N containing $f(p)$, there exists an open set U in M containing p such that $U \subset f^{-1}(G)$.

2. Show that f is continuous in M if and only if for every closed set F in N, $f^{-1}(F)$ is a closed set in M.

3. A function $f : (M, d) \to \mathbb{R}$ is lower semicontinuous at p if for every $\varepsilon > 0$, there exists an open set U containing p such that $f(q) > f(p) - \varepsilon$ for all $q \in U$. Show that f is lower semicontinuous at p if and only $\liminf_n f(p_n) \geq f(p)$ whenever $\{p_n\} \to p$.

   You may try to define upper semicontinuity so that a function which is lower and upper semicontinuous at a point means it is continuous at this point.

4. Show that in a normed space the closed metric ball with radius R centered at x, $\{y : d(y, x) \leq R\}$, is the closure of the open metric ball $B_R(x)$.

5. Show that any finite dimensional subspace of a normed space is closed. Can you find a subspace which is not closed, say, in ℓ^1? How about in $C[0, 1]$?

6. Show that any proper subspace of a normed space cannot contain any metric ball.

7. Show that $\mathcal{C}_{00}$ is not a closed subspace in ℓ^∞ and ℓ^2.

8. Show that a metric d induced from a norm on the vector space X satisfies (i) $d(x+z, y+z) = d(x, y)$ and (ii) $d(\alpha x, \alpha y)) = |\alpha| d(x, y)$. Use these properties to find two examples of metrics (on vector spaces) which are not induced by norms.

9. (a) Prove that ℓ^1 a proper vector subspace of ℓ^p for $p > 1$.

   (b) Now both the 1-norm and p-norm are norms on ℓ^1. Are they equivalent?

10. Let $\| \cdot \|_1$ and $\| \cdot \|_2$ be two norms on the vector space X.

    (a) Show that $\|x\|_M = \max\{\|x\|_1, \|x\|_2\}$ is again a norm on X.

    (b) Is this true for $\|x\|_m = \min\{\|x\|_1, \|x\|_2\}$?

11. (a) Establish the estimates
$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p} \|x\|_\infty, \quad p \geq 1.$$
    It shows that all p-norms are equivalent on $\mathbb{R}^n$.

    (b) Let M_p be the closed unit metric ball $\{x : \|x\|_p \leq 1\}$ in $\mathbb{R}^2$. Draw M_p for $p = 1, 2, 8$ and ∞.

12. Consider the vector space of polynomials P. For $p(x) = \sum_1^n a_j x^j$, define a norm by
$$\| \cdot \|_s \equiv \sum_1^n |a_j|.$$
    Compare it with the sup-norm.

13. Let $C^1[a, b]$ be the vector space of all continuously differentiable functions on $[a, b]$. Define two norms on this space by
$$\|f\|_1 \equiv \|f\|_\infty + \|f'\|_\infty,$$
    where f' is the derivative of f, and
$$\|f\|_a \equiv \|f\|_\infty + |f(a)|.$$
    Show that these two norms are equivalent.

14. Let $C^\infty[a,b]$ be the vector space of all smooth functions on $[a,b]$. Show that

$$d(f,g) \equiv \sum_1^\infty \frac{\|f^{(j)} - g^{(j)}\|_\infty}{1 + \|f^{(j)} - g^{(j)}\|_\infty},$$

defines a metric on $C^\infty[a,b]$ so that $\{f_k\}$ converges to f in this metric means $\{f_k^{(j)}\}$ converges to $f^{(j)}$ uniformly on $[a,b]$ for each j.

A subset C in a vector space X is **convex** if $\forall x, y \in C$ and $\lambda \in [0,1]$, $(1-\lambda)x + \lambda y \in C$. It is **balanced** if $x \in C$ implies $-x \in C$. It is "solid" if for any non-zero $x \in X$, there exist $0 < t_1 < t_2$ such that $tx \in C, \forall t \in (0, t_1]$ and $tx \notin C, \forall t \in [t_2, \infty)$. In the following two problems we relate a norm to a convex, balanced and "solid" set. It answers in particular the question: How many norms can be defined in a space?

15. Show that the closed unit metric ball $\{x \in X : \|x\| \le 1\}$ in any normed space $(X, \|\cdot\|)$ is convex, balanced and "solid".

16. Let C be a convex, balanced and "solid" set in a vector space X. Define its gauge function $p_C : X \mapsto [0,\infty)$ by

$$p_C(x) = \inf\{\alpha : \frac{x}{\alpha} \in C, \alpha > 0\}, \quad x \neq 0$$

and

$$p_C(0) = 0.$$

Prove that p_C is a norm on X.

17. The function

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}, \qquad p \in (0,1),$$

satisfies (N1) and (N2) in the axioms of a norm, but not (N3) on $\mathbb{R}^n$. Explain it.

18. Show that ℓ^p, $1 \le p < \infty$, is separable under the p-norm.

19. Show that $B[a,b]$ (bounded functions on $[a,b]$) is not separable under the sup-norm.

20. We know ℓ^1 is a proper subset of ℓ^p, $p > 1$. Now consider $(\ell^1, \|\cdot\|_1)$ and $(\ell^1, \|\cdot\|_p)$.

    (a) Prove that $\|x\|_p \le \|x\|_1$, $\forall x \in \ell^1$.

    (b) Are $\|\cdot\|_1$ and $\|\cdot\|_p$ equivalent on ℓ^1?

21. Let $X \times Y$ be the product space of two Banach spaces X and Y. Show that it is also a Banach space under the product norm. Recall that the product norm is given by $\|(x,y)\| = \|x\|_X + \|y\|_Y$.

22. Let Z be a closed subspace of the normed space X. On the quotient space X/Z define

$$\|\tilde{x}\tilde{\|} \equiv \inf_{z \in Z} \|x + z\|, \quad x \in \tilde{x}.$$

Verify that (a) $\|\cdot\tilde{\|}$ is a norm on X/Z and (b) X/Z is complete if X is complete. Here the quotient space X/Z is formed by the equivalence relation $x \sim y$ if and only if $x - y \in Z$.

23. Let (X,d) be a metric space and $C_b(X)$ be the vector space of all bounded, continuous functions on X. Prove that $C_b(X)$ forms a Banach space under the sup-norm.

24. Any metric space (X,d) can be mapped to $C_b(X)$ by $\phi(x) = f_x$ where $f_x(y) = d(x,y) - d(x_0,y)$ and x_0 is a prescribed point in X. (a) Verify that $f_x \in C_b(X)$. (b) Verify that ϕ is metric-preserving. (c) Use (b) to give another proof of Theorem 2.10 on metric completion.

25. Determine which of the following subspaces are closed:

    (a) {all polynomials on $[a, b]$} in $C[a, b]$.

    (b) {all continuous f with $\int_a^b f(x)\phi(x)dx = 0$} in $C[a, b]$ where ϕ is a given integrable function on $[a, b]$.

    (c) $\{x = (x_1, x_2, x_3, \cdots)$ which has finitely many non-zero entries$\}$ in ℓ^∞.

    (d) $\mathcal{C} \equiv \{x = (x_1, x_2, x_3, \cdots), x_n \to 0\}$ in ℓ^∞.

26. A sequence $\{x_k\}_{k=1}^\infty$ is called a **Schauder basis** for a Banach space X if every $x \in X$ can be uniquely expressed as $\sum_{k=1}^\infty \alpha_k x_k$ for $\alpha_k \in \mathbb{F}$.

    (a) Prove that a Banach space possessing a Schauder basis must be separable.

    (b) Show that $\{e_k\}_{k=1}^\infty$ forms a Schauder basis for ℓ^p, $1 \le p < \infty$.

    (c) Is it true that all polynomials with rational coefficients form a Schauder basis in $C[a, b]$?

    It was a problem of Banach to prove every separable Banach space has a Schauder basis. There were no exception among all familiar spaces. In 1973 Enflo surprised everyone by constructing a separable Banach space without Schauder bases. I still remember how excited Prof K.F. Ng was when he told us that the construction had to rely on the distribution of prime numbers in 1979. Even though they exist for many well-known Banach spaces, Schauder bases are difficult to find explicitly. You may google for a Schauder basis for $C[a, b]$.

27. Establish the following lemma of Riesz: For any closed, proper subspace Y of $(X, \|\cdot\|)$, there is some $x_0 \in X$, $\|x_0\| = 1$, satisfying $\|x_0 - y\| > \frac{1}{2}$, for all $y \in Y$.

28. Use Exercise 7 to give an alternative proof of Theorem 2.12.

29. Give an example to show that the point y^* realizing

    $$\|y^* - x_0\| = \min\{\|y - x_0\| : y \in Y\}$$

    where x_0 is a point not in Y, a closed subspace in $(X, \|\cdot\|)$, may not be unique. Suggestion: Work on $(\mathbb{R}^2, \|\cdot\|_\infty)$.

30. Produce a bounded sequence in $C[0, 1]$ which does not have any convergent subsequence in $C[0, 1]$.

31. Show that $\{\sin nx : x \in [0, \pi]\}$ is not equicontinuous in $C[0, \pi]$.

32. Let E be the set $\{f \in C[0, 1] : \ f(0) = 0, \ \ |f(x) - f(y)| \le |x - y|, \text{ for all } x, y \in [-1, 1]\}$. Show that there exists a function g in E such that

    $$\int_0^1 g(x)dx \ge \int_0^1 f(x)dx,$$

    for all f in E. Can you find it explicitly?

33. Let f be a continuous function on the real line. Suppose that $\{f_n(x)\} \equiv \{f(nx)\}$ forms an equicontinuous family in $C[0, 1]$. Is it necessarily that f is a constant?

34. Let $\{f_n\}$ be equicontinuous in $C(K)$ where K is a closed, bounded subset of $\mathbb{R}^n$. Show that $\{f_n\}$ uniformly converges to some continuous function f provided $\{f_n(x)\}$ converges to $f(x)$ for every $x \in K$.

    In the following three optional exercises we discuss the concept of a compact set in a metric space. Some student may have known this concept from point set topology. These exercises are served to link compactness to sequential compactness.

    A set E in a metric space (M, d) is called **compact** if every open cover has a finite subcover. In other words, suppose $\{G_\alpha\}, \alpha \in A$, is a collection of open sets in M satisfying $E \subset \bigcup_\alpha G_\alpha$, then there exist $G_1, \cdots, G_N$ from this collection such that $E \subset \bigcup_j G_j$. Any compact set is a closed set.

35. Show that a subset K in $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

36. Prove that a set C in a metric space (M, d) is compact if and only if it is sequentially compact.

    Therefore, compactness and sequential compactness are equivalent in a metric space. However, this may not be true in a topological space. Since we are only concerned with metric spaces in this course, we are happy to be free from this delicate matter.

37. The Ascoli-Arzela theorem characterizes compact sets in $C(K)$ where K is compact in some Euclidean space. The same result in fact holds when K is replaced by any compact metric space (M, d) (that is, M is compact). Prove it.

素月分輝，明河共影，表裡俱澄澈。
悠然心會，妙處難與君說。
張孝祥《念奴嬌》

Chapter 3

Dual Space

In this chapter we further our study of Banach spaces by examining continuous linear functionals on them. Each of these functionals gives very limited information on the space, but as a whole they become enormously helpful. The fundamental Hahn-Banach theorem guarantees there are sufficiently many such functionals for various purposes. They form a normed space called the dual space of the original space. We identify the dual spaces of $\mathbb{F}^n$, $\ell^p, 1 \le p < \infty$, and $C[a,b]$ in Sections 4 and 5, and introduce reflexive space in Section 6. Reflexive spaces arise naturally when we study the dual of dual spaces.

3.1 Linear Functionals

Any linear function from a vector space X to its scalar field $\mathbb{F}$ is called a **linear functional**. It is clear that the collection of all linear functionals from X to $\mathbb{F}$, denoted by $L(X,\mathbb{F})$, forms a vector space over $\mathbb{F}$ under pointwise addition and scalar multiplication of functions. Linear functionals play a crucial role in the study of the structure of vector spaces. There are two subspaces associated to a linear functional, namely, its image and its kernel, and the latter is more relevant. Indeed, let $\Lambda \in L(X,\mathbb{F})$, the **null space** (or **kernel**) of Λ is given by $N(\Lambda)$ (or $\ker \Lambda$) the set $\{x \in X : \Lambda x = 0\}$. It is clear that the kernel $N(\Lambda)$ forms a subspace of X and it is proper if and only if Λ is not identically zero.

Proposition 3.1. *Let X be a vector space over $\mathbb{F}$. Then*

(a) $L(X,\mathbb{F})$ is a vector space over $\mathbb{F}$,

(b) $N(\Lambda)$ is a subspace of X for any $\Lambda \in L(X,\mathbb{F})$, and

(c) if Λ is non-zero, then for any $x_0 \in X \setminus N(\Lambda)$, $X = N(\Lambda) \oplus \langle x_0 \rangle$.

Proof. (a) and (b) can be verified directly. It suffices to prove (c). Let x_0 be a point satisfying $\Lambda x_0 \neq 0$. For any $x \in X$, the vector $y = x - \lambda x_0$ where $\lambda = \Lambda x / \Lambda x_0$ belongs to $N(\Lambda)$:

$$\Lambda(x - \lambda x_0) = \Lambda x - \lambda \Lambda x_0 = 0.$$

Therefore, $x = y + \lambda x_0$, that is, $X = N(\Lambda) + \langle x_0 \rangle$. To show this is a direct sum, suppose that $x = y_1 + \lambda x_0 = y_2 + \mu x_0$. Then $y_1 - y_2 = (\mu - \lambda)x_0$, so $(\mu - \lambda)\Lambda x_0 = \Lambda(y_1 - y_2) = 0$ implies that $\mu = \lambda$ and $y_1 = y_2$. □

The meaning of (c) can be understood better by looking at the finite dimensional situation. Any linear functional Λ on $\mathbb{F}^n$ is completely determined by its values at a basis. For instance, consider the canonical basis $e_1, \cdots, e_n$ and let $\alpha_j = \Lambda e_j$, $j = 1, \cdots, n$, for $x \in \mathbb{F}^n$, $x = \sum_1^n x_j e_j$. Then

$$\Lambda x = \Lambda(\sum_1^n x_j e_j) = \sum_1^n \alpha_j x_j = \alpha_1 x_1 + \cdots + \alpha_n x_n,$$

gives the general formula for a linear functional on $\mathbb{F}^n$. What is $N(\Lambda)$ for a nonzero Λ? Apparently it is composed of the set $\{x \in \mathbb{F}^N : \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}$. When $\mathbb{F}^n = \mathbb{R}^n$, this is precisely the equation for a hyperplane passing through the origin whose normal direction is given by $(\alpha_1, \cdots, \alpha_n)/(\alpha_1^2 + \cdots + \alpha_n^2)^{1/2}$. In general, a hyperplane is one dimension lower than its ambient space. Thus (c) tells us that in infinite dimensional situation this is still true: After adjoining a single dimension (spanned by x_0) to it, $N(\Lambda) \oplus \langle x_0 \rangle$ is the entire space.

The abundance of linear functionals can be seen by the following abstract consideration. Let $\mathcal{B}$ be a Hamel basis for X. For each x in this basis we define a functional Λ_x by setting $\Lambda_x(\alpha x) = \alpha$ and $\Lambda_x y = 0$ for any y in $\mathcal{B}$ distinct from x where α is a fixed scalar. As every vector can be written as a finite linear combination of elements from $\mathcal{B}$, it is easy to see that Λ_x extends to become a linear functional on X. Moreover, one readily verifies that all these Λ_x's form a linearly independent set, so $L(X, \mathbb{F})$ is of infinite dimension.

When it comes to a normed space $(X, \|\cdot\|)$, a linear functional may not be related to the norm structure of the space. To have good interaction with the norm structure it is more desirable to look at linear functionals which are also continuous with respect to the norm. By linearity it is easy to show that a continuous linear functional is continuous everywhere once it is so at a single point. A related notion of continuity of a linear functional is its boundedness. We used to call a function bounded if its image is a bounded set. For linear functionals boundedness has a different meaning. We call a linear functional Λ **bounded** if it maps any bounded set in X to a bounded set in $\mathbb{F}$. In other words, for any bounded S in X, there exists a constant C such that $|\Lambda x| \le C$, for all $x \in S$. By linearity for Λ to be bounded it suffices that it maps a ball to a bounded set in the scalar field, or equivalently in the form of an estimate, there exists a constant C' such that $|\Lambda x| \le C'\|x\|$ for all $x \in X$. It turns out that for a linear functional continuity and boundedness are equivalent. We put all these in the following proposition.

Proposition 3.2. *Let $\Lambda \in L(X, \mathbb{F})$ where X is a normed space. We have*

(a) Λ is continuous if and only if Λ is continuous at one point.

(b) Λ is bounded if and only if there exists $C > 0$ such that $|\Lambda x| \le C\|x\|$ for all x.

(c) Λ is continuous if and only if Λ is bounded.

Proof. (a) It suffices to show the "if" part. Suppose Λ is continuous at x_0, that's, $\Lambda x_n \to \Lambda x_0$ whenever $x_n \to x_0$ in X. For any $x_1 \in X$ and $x_n \to x_1$, we have $x_n - x_1 + x_0 \to x_0$, so $\Lambda(x_n - x_1 + x_0) \to \Lambda(x_0)$. By linearity, $\Lambda x_n - \Lambda x_1 + \Lambda x_0 \to \Lambda x_0$ which means $\Lambda x_n \to \Lambda x_1$.

(b) Let Λ be bounded and take S to be the closed unit ball $\overline{B_1}(0)$. Then we can find a constant C_1 such that $|\Lambda x| \le C_1$. For any nonzero x, $x/\|x\| \in \overline{B_1}(0)$, we have $|\Lambda(x/\|x\|)| \le C_1$, i.e., $|\Lambda x| \le C_1\|x\|$.

Conversely, let S be a bounded set. We can find a large ball $B_R(0)$ to contain S. Then for any x in S, $|\Lambda x| \le C\|x\| \le CR$.

(c) If Λ is not bounded, there exists $\|x_n\| \le M$ but $|\Lambda x_n| \to \infty$. Then the sequence $\{y_n\}$, $y_n = x_n/|\Lambda x_n|$, satisfies $\|y_n\| \to 0$ but $|\Lambda y_n| = 1$, so Λ cannot be continuous.

On the other hand, let $x_n \to x_0$, that's, $\|x_n - x_0\| \to 0$. When Λ is bounded, by (b) $|\Lambda x_n - \Lambda x_0| = |\Lambda(x_n - x_0)| \le C\|x_n - x_0\| \to 0$, so Λ is continuous.

□

We also note the following useful characterization of a continuous linear functional. Prove it as an exercise.

Proposition 3.3. *A linear functional on a normed space is bounded if and only if its kernel is closed.*

We use X' to denote all bounded linear functionals on X. It is clear that X' is a subspace of $L(X, \mathbb{F})$. When X is of finite dimension, we have seen that every linear functional is of the form

$$\Lambda x = \sum_1^n \alpha_j x_j,$$

hence it is continuous. Thus $X' = L(X, \mathbb{F})$ when X is of finite dimension. However, this is no longer true when X is of infinite dimension. Let $\mathcal{B}$ be a Hamel basis for an infinite dimensional space X. We may pick a countably infinite set $\{x_1, x_2, x_3, \cdots\}$, $\|x_k\| = 1$, $\forall k$, from $\mathcal{B}$ and define T by assigning $Tx_k = k$, $k = 1, 2, \cdots$ and $Tx = 0$ for the remaining vectors in $\mathcal{B}$. As $\mathcal{B}$ is a basis, T can be extended to become a linear functional on X. Clearly it cannot be bounded.

Now we come to the norm structure on X' inherited from X.

Proposition 3.4. *Let X be a normed space and Λ a bounded linear functional on X and define*

$$\|\Lambda\| \equiv \sup_{x \neq 0} \frac{|\Lambda x|}{\|x\|}.$$

Then $\|\cdot\|$ is a norm on X'.

Before the proof of this proposition we point out a few things. First, the operator norm $\|\Lambda\|$ is also given by

$$\|\Lambda\| = \sup_{\|x\|=1} |\Lambda x|.$$

or

$$\|\Lambda\| = \sup_{\|x\| \leq 1} |\Lambda x|.$$

Second, we always have the useful inequality

$$|\Lambda x| \leq \|\Lambda\| \|x\|, \quad \text{for all } x \in X.$$

Third, the definition of the operator norm is basically the sup-norm for continuous functions. However, as the supremum is always infinity for any nonzero linear functional, we modify it by taking the supremum over the unit ball $\{\|x\| \leq 1\}$. Thanks to boundedness of the functional this supremum is always a finite number, and thanks to linearity it satisfies (N1).

Proof. Clearly (N1) and (N2) hold. To verify (N3), for $\Lambda_1, \Lambda_2 \in X'$ and $x \in X$, $\|x\| = 1$,

$$|(\Lambda_1 + \Lambda_2)(x)| = |\Lambda_1 x + \Lambda_2 x| \leq |\Lambda_1 x| + |\Lambda_2 x| \leq \|\Lambda_1\| + \|\Lambda_2\|.$$

Therefore,

$$\|\Lambda_1 + \Lambda_2\| = \sup_{\|x\|=1} |(\Lambda_1 + \Lambda_2)(x)| \leq \|\Lambda_1\| + \|\Lambda_2\|.$$

□

From Proposition 3.4, $(X', \|\cdot\|)$ forms a normed space called the **dual space** of $(X, \|\cdot\|)$. The norm on X' is called the **operator norm** sometimes. It is a bit surprising that X' behaves better than X as implicated by the following proposition.

Proposition 3.5. *The dual space X' of a normed space X is a Banach space.*

Proof. Let $\{\Lambda_k\}$ be a Cauchy sequence in X', that's, for every $\varepsilon > 0$, there is some k_0 such that $\|\Lambda_k - \Lambda_l\| < \varepsilon$ for all $k, l \geq k_0$. For each $x \in X$,

$$|\Lambda_k x - \Lambda_l x| \leq \|\Lambda_k - \Lambda_l\| \|x\| \leq \varepsilon \|x\|, \tag{3.1}$$

which shows that $\{\Lambda_k x\}$ is a Cauchy sequence in $\mathbb{F}$. By the completeness of $\mathbb{F}$, $\lim_{k\to\infty} \Lambda_k x$ exists for every $x \in X$. Setting $\Lambda x \equiv \lim_{k\to\infty} \Lambda_k x$, it is routine to check that Λ is linear. Moreover, by letting $l \to \infty$ in (3.1), we have

$$|\Lambda_k x - \Lambda x| \leq \varepsilon \|x\|, \ k \geq k_0. \tag{3.2}$$

It follows that

$$|\Lambda x| \le |\Lambda_{k_0} x - \Lambda x| + |\Lambda_{k_0} x| \le (\varepsilon + \|\Lambda_{k_0}\|)\|x\|,$$

so Λ is also bounded. From (3.2) we have

$$|\Lambda_k x - \Lambda x| \le \varepsilon, \quad k \ge k_0,$$

for all $x, \|x\| = 1$, so $\Lambda_k \to \Lambda$ in operator norm.

□

3.2 Concrete Dual Spaces

We determine the dual spaces of $\mathbb{F}^n$ and $\ell^p, 1 \le p < \infty$ in this section.

Recall that we identify two normed spaces $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ if there exists a norm-preserving linear isomorphism from X_1 to X_2.

Proposition 3.6. *For $1 \le p < \infty$, the dual space of ℓ^p is given by ℓ^q where p and q are conjugate.*

Proof. We only consider $p > 1$ and leave the case $p = 1$ to you.

We define a map from $(\ell^p)'$ to ℓ^q as follows. For each $\Lambda \in (\ell^p)'$, let $\Phi\Lambda$ be the sequence $\alpha = (\Lambda e_1, \Lambda e_2, \cdots)$ where $\{e_j\}$ is the "canonical sequence". This is a linear map from $(\ell^p)'$ to the space of sequences. We claim that its image belongs to ℓ^q.

Letting $\alpha^N = (\text{sgn}(\alpha_1)|\alpha_1|^{q-1}, \cdots, \text{sgn}(\alpha_N)|\alpha_N|^{q-1}, 0, 0, \cdots)$ and using the inequality $|\Lambda x| \le \|\Lambda\|\|x\|_p$ for $x = \alpha^N$ we have, as $|\Lambda\alpha^N| = \sum_1^N |\alpha_j|^q$ and $\|\alpha^N\|_p = \left(\sum_1^N |\alpha_j|^q\right)^{\frac{1}{p}}$,

$$\left(\sum_1^N |\alpha_j|^q\right)^{\frac{1}{q}} \le \|\Lambda\|.$$

Letting $N \to \infty$, we have

$$\|\Phi(\Lambda)\|_q = \|\alpha\|_q \le \|\Lambda\|. \tag{3.3}$$

We have shown that Φ maps $(\ell^p)'$ into ℓ^q.

To show that Φ is onto we construct its inverse. For each α in ℓ^q we define $\Psi\alpha = \Lambda_\alpha$ where Λ_α is given by $\Lambda_\alpha x = \sum_j \alpha_j x_j$. By Hölder inequality, this map is well-defined and

$$|\Lambda_\alpha x| \le \|\alpha\|_q \|x\|_p, \quad \text{for all } x \in \ell^p.$$

It follows that

$$\|\Psi\alpha\| = \|\Lambda_\alpha\| \le \|\alpha\|_q. \tag{3.4}$$

Next we claim that

$$\Phi\Psi\alpha = \alpha, \quad \forall \alpha \in \ell^q.$$

Indeed, for $\alpha \in \ell^q$, $\Phi\Psi\alpha = ((\Psi\alpha)e_1, (\Psi\alpha)e_2, \cdots) = (\alpha_1, \alpha_2, \cdots) = \alpha$, so the claim holds. This claim shows in particular that Φ is onto. By combining it with (3.3) and (3.4), we have

$$\|\Lambda\| = \|\Phi\Psi\Lambda\| \le \|\Phi\Lambda\|_q \le \|\Lambda\|,$$

whence $\|\Phi\Lambda\|_q = \|\Lambda\|$, that is, Φ is norm-preserving. We have succeeded in constructing a norm-preserving linear isomorphism from $(\ell_p)'$ to ℓ^q, so these two spaces are the same.

□

A similar but simpler proof shows that the dual of $\mathbb{F}^n$ under the $p-$norm is itself under the q-norm for $p \in [1, \infty]$.

We will determine the dual space of $C[a, b]$ as an application of the Hahn-Banach theorem in Section 3.5. It is a standard result in real analysis that the dual space of $L^p(a, b)$, $1 \le p < \infty$, is $L^q(a, b)$ where q is conjugate to p. But it is not true when p is infinity. See Rudin's "Real and Complex Analysis" and Hewitt-Stromberg's "Abstract Analysis" (especially for $p = \infty$) for details.

3.3 Hahn-Banach Theorem

In the last section we identified the dual space of $\mathbb{F}^n$ and $\ell^p, 1 \le p < \infty$. In particular, it shows that there are many non-trivial bounded linear functionals in these spaces. However, in a general normed space it is not clear how to find even one which distinguishes two points. The theorem of Hahn-Banach ensures that we can always do this. This extremely useful theorem, which is formulated as a statement on extension, is one of the most fundamental results in functional analysis.

Considering its applications in later chapters, it is necessary to formulate the theorem not only in the setting of a normed space but in a vector space. We call a function p defined in a vector space X to $(-\infty, \infty]$ **subadditive** if for all x, y in X,

$$p(x + y) \le p(x) + p(y),$$

and **positive homogeneous** if for all x in X and $\alpha \ge 0$,

$$p(\alpha x) = \alpha p(x).$$

Note that the norm is a subadditive, positive homogenous function due to (N2) and (N3). A non-negative, subadditive, positive homogeneous function on a vector space is sometimes called a **gauge** or a **Minkowski functional**.

Any positive multiple of the norm is a gauge on a normed space. Other gauges can be found as follows. Let C be a non-empty convex set containing 0 in a vector space X. Define

$$p_C(x) = \inf\{\alpha : \quad x \in \alpha C, \ \alpha > 0\}$$

and set $p_C(x) = \infty$ if no such α exists. We claim that p_C is a gauge. Clearly, $p_C(\alpha x) = \alpha p_C(x)$ for every positive α. On the other hand, consider x, y in X where $p_C(x)$ and $p_C(y)$ are finite (subadditivity holds trivially if they are not). According to the definition of a gauge, for every $\varepsilon > 0$, there exist positive α, β satisfying $p_C(x) \ge \alpha - \varepsilon$, $p_C(y) \ge \beta - \varepsilon$ and $x/\alpha,\ y/\beta \in C$. Therefore, by the convexity of C,

$$\frac{x+y}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta}\frac{x}{\alpha} + \frac{\beta}{\alpha+\beta}\frac{y}{\beta} \in C,$$

which implies that

$$p_C(x + y) \le \alpha + \beta \le p_C(x) + p_C(y) + 2\varepsilon.$$

So p_C is subadditive. Gauges of convex sets form an important class of subadditive, positive homogeneous functions.

Now we state and prove a general form of the Hahn-Banach theorem.

Theorem 3.7. *Let X be a vector space and p a subadditive, positive homogeneous function in X. Suppose $\Lambda \in L(Y, \mathbb{F})$ where Y is a proper subspace of X satisfies*

$$Re\Lambda x \le p(x), \quad \text{for all } x \in Y.$$

Then there is an extension of Λ to $L(X, \mathbb{F})$, $\widetilde{\Lambda}$, such that

$$Re\widetilde{\Lambda} x \le p(x), \quad \text{for all } x \in X.$$

We will treat the real case first. The complex case can be deduced from the real case. The technical part of the proof of this theorem is contained in the following lemma.

Lemma 3.8 (One-Step Extension). *Let $\mathbb{F} = \mathbb{R}$, $\Lambda \in L(Y, \mathbb{R})$ and $x_0 \in X \setminus Y$. There exists an extension Λ_1 of Λ on $\langle Y, x_0 \rangle$ such that $\Lambda_1 x \le p(x)$.*

Proof. Let $Y_1 = \langle Y, x_0 \rangle$. Every element in Y_1 is of the form $x = y + cx_0$, $y \in Y$, $c \in \mathbb{R}$. Any linear functional Λ_1 extending Λ satisfies $\Lambda_1 x = \Lambda_1(y + cx_0) = \Lambda_1 y + c\Lambda_1 x_0 = \Lambda y + c\Lambda_1 x_0$. Conversely, by assignment any value to $\Lambda_1 x_0$ one obtains an extension of Λ in this way. Nevertheless, the point is to determine the value of $\Lambda_1 x_0$ so that $\Lambda_1 x \le p(x)$ on Y_1. To show such choice is possible, let's focus at $c = \pm 1$. For an admissible extension, one should have

$$\Lambda y \pm \Lambda_1 x_0 \le p(y \pm x_0), \tag{3.5}$$

or,

$$\Lambda_1 x_0 \le p(y + x_0) - \Lambda y$$

and

$$\Lambda y - p(y - x_0) \le \Lambda_1 x_0.$$

It implies that for all $y, z \in Y$,

$$\Lambda z - p(z - x_0) \le \Lambda_1 x_0 \le p(y + x_0) - \Lambda y.$$

Therefore, if

$$\alpha \equiv \sup_{z \in Y} (\Lambda z - p(z - x_0)) \le \beta \equiv \inf_{y \in Y} (p(y + x_0) - \Lambda y) \tag{3.6}$$

holds, we can pick any $\gamma \in [\alpha, \beta]$ and set $\Lambda_1 x_0 = \gamma$, so that (3.5) holds. Before verifying this, let's show that it implies Λ_1 is our desired extension. In fact, by linearity, for any $c > 0$,

$$\begin{aligned} \Lambda_1(y \pm cx_0) &= \Lambda y \pm c\Lambda_1 x_0 = c(\Lambda(\frac{y}{c}) \pm \Lambda_1 x_0) \\ &\le cp(\frac{y}{c} \pm x_0) \\ &= p(y \pm cx_0), \end{aligned}$$

so $\Lambda_1 x \le p(x)$ for all x in Y_1. It remains to verify (3.6). But this is easy. We write (3.6) as

$$\Lambda z - p(z - x_0) \le p(y + x_0) - \Lambda y, \quad \forall y, z \in Y,$$

and it holds if and only if

$$\Lambda(y + z) \le p(z - x_0) + p(y + x_0).$$

Certainly this is true by the subaddivity of p

$$\Lambda(y + z) \le p(y + z) = p(y + x_0 - x_0 + z) \le p(y + x_0) + p(z - x_0).$$

□

Proof of Theorem 3.6. Let $\mathbb{F} = \mathbb{R}$ first. Set $\mathscr{D} = \{(Z, T)$: Z is a subspace of X containing Y and $T \in L(Z, \mathbb{R})$ is an extension of Λ satisfying $Tx \le p(x)$ on $Z.\}$. $\mathscr{D}$ is non-empty since $(Y, \Lambda) \in \mathscr{D}$. A relation "$\le$" is defined on $\mathscr{D}$, $(Z_1, T_1) \le (Z_2, T_2)$ if and only if (a) Z_1 is a subspace of Z_2 and (b) T_2 extends T_1. We check easily that $(\mathscr{D}, \le)$ is a poset.

We claim that each chain $\mathscr{C}$ in $(\mathscr{D}, \le)$ has an upper bound. Indeed, for all Z in $\mathscr{C}$, let

$$Z^* = \bigcup_{Z_\alpha \in \mathscr{C}} Z_\alpha \text{ and } T^* z = T_\alpha z \text{ for } z \in Z_\alpha.$$

We show that Z^* is a subspace of X. Let $z_1, z_2 \in Z^*$. Then $z_1 \in Z_\alpha$ and $z_2 \in Z_\beta$ for some α, β. As $\mathscr{C}$ is a chain, either Z_α is a subspace of Z_β or the other way around. Let's assume the latter, so $z_1, z_2 \in Z_\alpha$ and

$\lambda z_1 + \mu z_2 \in Z_\alpha \subset Z^*$. Z^* is a subspace. By a similar reason we can show that $T_\alpha z = T_\beta z$ if $z \in Z_\alpha \cap Z_\beta$, so T^* is well-defined. For any $z \in Z^*$, there exists some Z_α containing z, so $T^* z = T_\alpha z \le p(z)$. We have shown that (Z^*, T^*) is an upper bound for $\mathscr{C}$.

Now we apply Zorn's lemma to conclude that there is a maximal element (Z_{max}, T_{max}) in $\mathscr{D}$. We claim that $Z_{max} = X$. For, if this is not true, we can find $x_0 \in X \setminus Z_{max}$. Using the one-step extension lemma, we find T_1 on $Z_1 = \langle Z_{max}, x_0 \rangle$ extending T_{max} and $T_1 x \le p(x),\ x \in Z_1$. So $(Z_1, T_1) \in \mathscr{D}$, that is to say, (Z_{max}, T_{max}) cannot be a maximal element. This contradiction shows that $Z_{max} = X$, and $\widetilde{\Lambda} \equiv T_{max}$ is our desired extension of Λ. This completes the proof of the general Hahn-Banach theorem for the real case. □

To treat the complex case, we need the following lemma. It asserts that any complex linear functional is uniquely determined by its real or imaginary part.

Lemma 3.9. *(a) Let Λ be in $L(X, \mathbb{C})$ where X is a complex vector space. Then its real and imaginary parts are in $L(X, \mathbb{R})$ when X is regarded as a real vector space. Furthermore,*

$$\Lambda x = Re \Lambda x - i Re \Lambda(ix), \text{ for all } x \in X. \tag{3.7}$$

(b) Conversely, for any Λ_1 in $L(X, \mathbb{R})$, there exists a unique element in $L(X, \mathbb{C})$ taking Λ_1 as its real part so that the above formula holds.

Proof. It is clear that both the real and imaginary parts of a complex linear functional are linear functionals over the reals. Write

$$\Lambda x = Re(\Lambda x) + i\ Im(\Lambda x) \equiv \Lambda_r x + i\ \Lambda_i x.$$

We claim that they are related by

$$\Lambda_r(ix) = -\Lambda_i(x),\ \Lambda_i(ix) = \Lambda_r x, \tag{3.8}$$

so (3.7) holds. To see this, simply use the linearity of Λ over $\mathbb{C}$ to get

$$\Lambda(ix) = i\ \Lambda(x),$$

so,

$$\Lambda_r(ix) + i\ \Lambda_i(ix) = i\ \Lambda_r x - \Lambda_i x,$$

and (3.8) holds.

Now, given a real linear functional Λ_1 on X, define

$$\Lambda x = \Lambda_1 x - i \Lambda(ix).$$

It is clear that the real part of Λ is equal to Λ_1. It remains to check that Λ is linear. For $x_1, x_2 \in X$, and $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} \Lambda(x_1 + x_2) &= \Lambda_1(x_1 + x_2) - i\Lambda_1(i(x_1 + x_2)) \\ &= \Lambda_1 x_1 - i\ \Lambda_1(ix_1) + \Lambda_1 x_2 - i\ \Lambda_1(ix_2) \\ &= \Lambda x_1 + \Lambda x_2; \end{aligned}$$

$$\begin{aligned} \Lambda((\alpha + i\beta)x) &= \Lambda(\alpha x + i\beta x) = \Lambda(\alpha x) + \Lambda(\beta i x) \\ &= \Lambda_1(\alpha x) - i\Lambda_1(\alpha i x) + \Lambda_1(\beta i x) - i\Lambda_1(i\beta i x) \\ &= \alpha\Lambda_1 x - i\alpha\Lambda_1(ix) + \beta\Lambda_1(ix) + i\beta\Lambda_1 x \\ &= (\alpha + i\beta)\Lambda x; \end{aligned}$$

□

We complete the proof of the general Hahn-Banach theorem as follows. We first obtain a real extension Λ_1 of the real part of the complex linear functional Λ satisfying $\Lambda_1 x \le p(x)$ on X. By the lemma above, we find a complex linear functional $\tilde{\Lambda}$ on X whose real part is given by Λ_1 extending Λ. $\tilde{\Lambda}$ is our desired extension.

We have the following version of Hahn-Banach theorem on normed spaces.

Theorem 3.10. *Let $(X, \|\cdot\|)$ be a normed space and Y a proper subspace of X. Then any $\Lambda \in Y'$ admits an extension to some $\widetilde{\Lambda} \in X'$ with $\|\widetilde{\Lambda}\| = \|\Lambda\|$.*

Notice that from definition, in general, we have

$$\|\widetilde{\Lambda}\| = \sup_{x \in X \setminus \{0\}} \frac{|\widetilde{\Lambda} x|}{\|x\|} \ge \sup_{x \in Y \setminus \{0\}} \frac{|\Lambda x|}{\|x\|} = \|\Lambda\|.$$

It suffices to establish the inequality from the other direction.

Proof. Consider first the real case. Taking $p(x) = \|\Lambda\|\|x\|$, we apply the general Hahn-Banach theorem to obtain an extension of Λ, $\tilde{\Lambda}$, which satisfies $\tilde{\Lambda} x \le \|\Lambda\|\|x\|$. Replacing x by $-x$, we get $-\tilde{\Lambda} x \le \|\Lambda\|\|-x\| = \|\Lambda\|\|x\|$, so $|\tilde{\Lambda} x| \le \|\Lambda\|\|x\|$ which implies $\|\tilde{\Lambda}\| \le \|\Lambda\|$ on X.

For the complex case, let $\tilde{\Lambda}$ be an extension of Λ satisfying $Re\tilde{\Lambda} x \le \|\Lambda\|\|x\|$ on X. Replacing x by $-x$, we have $|Re\tilde{\Lambda} x| \le \|\Lambda\|\|x\|$. For any x , there is a complex number $e^{i\theta}$ such that $\tilde{\Lambda} x = |\tilde{\Lambda} x| e^{i\theta}$. It follows that $|\tilde{\Lambda} x| = \tilde{\Lambda}(e^{-i\theta} x) = Re\tilde{\Lambda}(e^{-i\theta x}) \le \|\Lambda\|\|e^{-i\theta} x\| = \|\Lambda\|\|x\|$, that is, $\|\tilde{\Lambda}\| \le \|\Lambda\|$. The proof is completed. □

We will be concerned only with Hahn-Banach theorem in the remaining sections. The general Hahn-Banach theorem can be used to establish some separation theorems, see exercises. Its importance will become evident in Chapter 7.

3.4 Consequences of Hahn-Banach Theorem

Theorem 3.11. *Let $(X, \|\cdot\|)$ be a normed space and Y a closed subspace of X. For any $x_0 \in X \setminus Y$, there exists $\Lambda \in X'$, $\|\Lambda\| = 1$, satisfying*

$$\Lambda x_0 = dist(x_0, Y),$$

and

$$\Lambda y = 0, \text{ for all } y \in Y.$$

Proof. Let $d = dist(x_0, Y)$. It is positive because Y is closed and x_0 stays outside Y. In the subspace $Y_1 = \langle Y, x_0 \rangle$, every vector can be written uniquely in the form $y + \alpha x_0$. We define Λ_0 on Y_1 by setting

$$\Lambda_0(y + \alpha x_0) = \alpha \|x_0\|.$$

Then Λ_0 is linear and vanishes on Y. Moreover, using

$$0 < d = \inf_{z \in Y} \|x_0 + z\| \le \frac{1}{|\alpha|} \|\alpha x_0 + y\|, \ \forall y \in Y,$$

We have

$$|\Lambda_0(y + \alpha x_0)| \le |\alpha|\|x_0\| \le \frac{\|x_0\|}{d} \|y + \alpha x_0\|,$$

in other words, $\Lambda_0 \in Y_1'$ and

$$\|\Lambda_0\| \le \frac{\|x_0\|}{d}.$$

We claim that $\|\Lambda_0\| = \|x_0\|/d$. For, taking $y_n \in Y$, $\|y_n + x_0\| \to d$,

$$\Lambda_0(y_n + x_0) = \|x_0\| \le \|\Lambda_0\|\|y_n + x_0\| \to \|\Lambda_0\|d,$$

hence $\|x_0\|/d \le \|\Lambda_0\|$.

Now, we apply Hahn-Banach theorem to obtain an extension $\widetilde{\Lambda}$ of Λ_0 in X' with $\|\widetilde{\Lambda}\| = \|\Lambda_0\| = \|x_0\|/d$. Then, a constant multiple of $\tilde{\Lambda}$, $d/\|x_0\|\widetilde{\Lambda}$, is our desired functional.

□

Corollary 3.12. *For any non-zero $x_0 \in X$, there exists $\Lambda \in X'$ such that $\Lambda x_0 = \|x_0\|$ and $\|\Lambda\| = 1$.*

Proof. Apply Theorem 3.9 by taking $Y = \{0\}$. □

A bounded linear functional with the properties described in this corollary may be called a "dual point" of x_0. It may not be unique. For instance, consider $(\mathbb{R}^2, \|\cdot\|_1)$ and the vector $x_0 = (1, 0)$. It is readily checked that the two linear functionals $\Lambda_1(x, y) = x$ and $\Lambda_2(x, y) = x + y$ are dual points of $(1, 0)$.

Corollary 3.13. *For any $x \in X$,*

$$\|x\| = \sup_{\Lambda \in X'\ \Lambda \ne 0} \frac{|\Lambda x|}{\|\Lambda\|}.$$

Proof. From $\|\Lambda x\| \le \|\Lambda\|\|x\|$ we obtain

$$\|x\| \ge \sup_{\Lambda \in X'\ \Lambda \ne 0} \frac{|\Lambda x|}{\|\Lambda\|}.$$

On the other hand, for a given non-zero x, pick Λ^* such that $\|\Lambda^*\| = 1$ and $\Lambda^* x_0 = \|x_0\|$. We have

$$\|x\| = \frac{|\Lambda^* x|}{\|\Lambda^*\|} \le \sup_{\Lambda \in X'\ \Lambda \ne 0} \frac{|\Lambda x|}{\|\Lambda\|}.$$

□

This corollary tells us that there are sufficiently many bounded linear functionals to determine the norm of any vector. Furthermore, the "sup" in the above expression can be strengthened to "max" as it is attained by Λ^*.

3.5 The Dual Space of Continuous Functions

The dual space of $C[a, b]$ is described essentially by a representation theorem of Riesz. To formulate it we need to introduce two new concepts: Riemann-Stieltjes integral and functions of bounded variations. Since our focus is on the application of the Hahn-Banach theorem, we simply state basic results (Facts 1 to 4) on these new concepts and leave them as exercises. You may consult Rudin's "Principles of Mathematical Analysis" or [Hewitt-Stromberg] for more on Riemann-Stieltjes integrals.

First of all, for any two complex-valued functions f and g on $[a, b]$ we define its **Riemann-Stieltjes sum** $R(f, g, P)$ with respect to a tagged partition $\dot{P}$ by

$$R(f, g, P) = \sum_1^n f(z_j)\big(g(x_j) - g(x_{j-1})\big)$$

where $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ is the partition and $z_j \in [x_{j-1}, x_j]$ is a tag. We call f is **Riemann-Stieltjes integrable** with respect to g if there exists $I \in \mathbb{F}$ satisfying: For each $\varepsilon > 0$, there exists a δ such that

$$|R(f,g,P) - I| < \varepsilon, \quad \forall P, \ \|P\| < \delta.$$

Recall that the length of the partition P, $\|P\|$, is given by $\max_{j=1}\{x_j - x_{j-1}\}$. Write $I = \int_a^b f(x)dg(x)$ or simply $\int f dg$ and denote the class of all Riemann-Stieltjes integrable functions by $R_g[a,b]$. Using the definition one can establish the following facts.

Fact 1.

(a) For f_1 and f_2 in $R_g[a,b]$ and $\alpha_1, \alpha_2 \in \mathbb{F}$, we have $\alpha_1 f_1 + \alpha_2 f_2$ belongs to $R_g[a,b]$, and

$$\int (\alpha_1 f_1 + \alpha_2 f_2) dg = \alpha_1 \int f_1 dg + \alpha_2 \int f_2 dg;$$

(b) For $f \in R_{g_1}[a,b] \cap R_{g_2}[a,b]$ and $\alpha_1, \alpha_2 \in \mathbb{F}$, f belongs to $R_{\alpha_1 g_1 + \alpha_2 g_2}[a,b]$ and

$$\int f d(\alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 \int f dg_1 + \alpha_2 \int f dg_2;$$

We will single out a class of g's so that all continuous functions are Riemann-Stieltjes integrable with respect to each of them. A function g on $[a,b]$ is called a **function of bounded variation** (a BV-function for short) if there exists a constant M such that

$$\sum_{j=1}^{n} \left| g(x_j) - g(x_{j-1}) \right| \le M$$

for all partitions P on $[a,b]$. For a function g of bounded variation, set its **total variation** to be

$$\|g\|_{BV} \equiv \sup\{\sum_{j=1}^{n} \left| g(x_j) - g(x_{j-1}) \right| : \text{ all partitions } P\}.$$

It is easy to see that $\|g\|_{BV}$ satisfies (N2) and (N3) but not (N1), which must be replaced by: "$\|g\|_{BV} = 0$ implies g is a constant". Nevertheless, we can remove this unpleasant situation by restricting to the subset, $BV_0[a,b]$, consisting of all BV-functions which vanish at a. Then $BV_0[a,b]$ forms a normed vector space under $\|\cdot\|_{BV}$. A by-now routine check shows that it is complete.

Let's look at some examples of BV-functions.

Example 3.1. Every monotone function on $[a,b]$ is of bounded variation. In fact, for any P,

$$\sum_{1}^{n} \left| g(x_j) - g(x_{j-1}) \right| = \sum_{1}^{n} \big(g(x_j) - g(x_{j-1}) \big) = g(b) - g(a)$$

when g is increasing, so $\|g\|_{BV} = g(b) - g(a)$. When g is decreasing, $\|g\|_{BV} = g(a) - g(b)$.

Example 3.2. If g is continuously differentiable on $[a,b]$, then $\|g\|_{BV} \le \|g'\|_\infty (b-a)$. For,

$$\sum_{1}^{n} \left| g(x_j) - g(x_{j-1}) \right| = \sum_{1}^{n} \left| g'(z_j)(x_j - x_{j-1}) \right| \le \|g'\|_\infty (b-a)$$

where $z_j \in [x_{j-1}, x_j]$.

Example 3.3. Not every continuous function is of bounded variation. Consider the continuous function on $[0, 2/\pi]$ given by

$$h(x) = \begin{cases} x \sin \dfrac{1}{x}, & x \ne 0 \\ 0, & x = 0 \end{cases}$$

Taking the partition P^N to be $\{\big[(n+1/2)\pi\big]^{-1}:\ n=1,\cdots,N\}$ together with the endpoints, one shows that $Vh=\infty$ after letting $N\to\infty$.

Fact 2. *Every real-valued BV-function can be expressed as the difference of two increasing functions.*

This is known as Jordan decomposition theorem.

Fact 3. *Every continuous function on $[a,b]$ is Riemann-Stieltjes integrable with respect to a BV-function in $[a,b]$.*

In other words, $\int f dg$ is well-defined when $f\in C[a,b]$ and $g\in BV[a,b]$.

Example 3.4. Taking $g(x)=x$, the Riemann-Stieltjes integral reduces to the Riemann integral.

Example 3.5. Taking g to be continuously differentiable, $\int f dg=\int f g' dx$ and $\|\Lambda_g\|\le\|g'\|_1$.

Example 3.6. Taking $g=\chi_{[c,b]}$, $a<c<b$, in

$$\sum f(z_j)\big(g(x_j)-g(x_{j-1})\big)$$

all terms vanish except the subinterval $[x_{j-1},x_j]$ containing c in its interior (we may take a partition in which c is not an endpoint of any subinterval.), so, as $\|P\|\to 0$,

$$\int f d\chi_{[c,b]}=f(c).$$

Note that we also have

$$\int f d\chi_{(c,b]}=f(c).$$

Now we come to bounded linear functionals on $C[a,b]$. Let's consider two examples. First, fix a point $c\in[a,b]$ and let $\Lambda_1 f=f(c)$. This "evaluation map" is clearly a linear functional with operator norm equal to 1. Next, fix an arbitrary continuous function ϕ and define

$$\Lambda_2(f)=\int_a^b f(x)\phi(x)dx,\quad f\in C[a,b].$$

From

$$\big|\Lambda_2 f\big|\le\int_a^b|\phi(x)|dx\|f\|_\infty,$$

Λ_2 is also in the dual of $C[a,b]$. Both functionals can be unified in the setting of Riemann-Stieltjes integrals. Indeed, the first functional corresponds to taking $g=\chi_{(c,b]}$ and the second one to taking g to be a primitive function of ϕ. In view of this, to every BV-function g we associate it with the functional

$$\Lambda_g f=\int_a^b f dg.$$

It is not hard to verify that Λ_g belongs to the dual space of $C[a,b]$ (see below). Our goal is to show that such association is a norm-preserving linear isomorphism. However, Example 3.6 is an obvious obstruction; as both functions $\chi_{(c,b]}$ and $\chi_{[c,b]}$ give the same evaluation map $f(c)$, this association cannot be injective. This difficulty turns out to be minor, and we can overcome it by further restricting the space $BV_0[a,b]$.

A function g is called right continuous if $\lim_{h\downarrow 0} g(x+h)=g(x)$. Let

$$V[a,b]=\{g\in BV_0[a,b]:\quad g \text{ is right continuous on } [a,b).\}$$

Notice that $\chi_{[c,b]}$ is right continuous but $\chi_{(c,b]}$ is not. It is clear that $V[a,b]$ is a subspace of $BV_0[a,b]$.

Fact 4.

(a) *Every BV_0-function g is equal to a unique V-function $\widetilde{g}$ except possibly at countably many points.*

(b) $\int f dg = \int f d\widetilde{g}$, *for all* $f \in C[a,b]$.

(c) For $g_1, g_2 \in V[a,b]$, $\int f dg_1 = \int f dg_2$ *implies* $g_1 = g_2$.

Fact 4 (a) can be deduced from a known result in Elementary Analysis, namely, the discontinuity points of a monotone function consist of jump discontinuity and there are at most countably many of them. Since a BV-function is the difference of two increasing functions, the same property holds for it.

From Facts 1 and 4 we see that the map $g \mapsto \Lambda_g$ defines a linear injective map Φ from $V[a,b]$ to $C[a,b]'$. In fact, we have

$$|\Phi(g)f| = |\Lambda_g f| \le \|g\|_{BV} \ \|f\|_\infty,$$

hence

$$\|\Phi(g)\| \le \|g\|_{BV}, \quad \forall g \in V[a,b]. \tag{3.9}$$

Here is a version of the Riesz representation theorem.

Theorem 3.14. *There is a norm-preserving linear isomorphism from $C[a,b]'$ to $V[a,b]$.*

Proof. The norm-preserving linear isomorphism is, of course, Φ. Let's find its inverse. Let $\Lambda \in C[a,b]'$. Observing that $C[a,b]$ is a subspace in the normed space $B[a,b]$ of bounded functions, we can use Hahn-Banach theorem to find an extension $\widetilde{\Lambda} \in B[a,b]'$ with $\|\widetilde{\Lambda}\| = \|\Lambda\|$. This is crucial!

Our desired inverse g should satisfy $\Lambda f = \Lambda_g f$ for $f \in C[a,b]$. Formally,

$$\begin{aligned} \widetilde{\Lambda}(\chi_{[a,c]}) &= \int_a^b \chi_{[a,c]} dg = \int_a^c dg = \int_a^c g'(x)dx \\ &= g(c) - g(a) = g(c), \end{aligned}$$

as g vanishes at a. Motivated by this, we define

$$g(x) = \widetilde{\Lambda}(\chi_{[a,x]}), \quad x \in (a,b],$$

and $g(0) = 0$. We claim that $g \in BV_0[a,b]$ and $\|g\|_{BV} \le \|\Lambda\|$.

For, with respect to an arbitrary partition P,

$$\begin{aligned} \sum_1^n |g(x_j) - g(x_{j-1})| &= \sum_1^n e^{i\theta_j}(g(x_j) - g(x_{j-1})) \\ &\quad \text{(for any } z \in \mathbb{C}, \text{ there exists } e^{i\theta} \text{ such that } |z| = e^{i\theta} z) \\ &= e^{i\theta_1} g(x_1) + \sum_2^n e^{i\theta_j}(g(x_j) - g(x_{j-1})) \\ &= e^{i\theta_1} \tilde{\Lambda}\chi_{[a,x_1]} + \sum_2^n e^{i\theta_j}(\tilde{\Lambda}\chi_{[a,x_j]} - \tilde{\Lambda}\chi_{[a,x_{j-1}]}) \\ &= \tilde{\Lambda}\big(e^{i\theta_1}\chi_{[a,x_1]}\big) + \sum_2^n \tilde{\Lambda}\big(e^{i\theta_j}\chi_{(x_{j-1},x_j]}\big) \\ &= \tilde{\Lambda}\big(e^{i\theta_1}\chi_{[a,x_1]} + \sum_2^n e^{i\theta_j}\chi_{(x_{j-1},x_j]}\big). \end{aligned}$$

Let's denote the function inside the above bracket by h. Noting that for all $x \in [a,b]$, there exists a unique subinterval $[a,x_1]$ or $(x_{j-1},x_j]$ containing x, so $h(x) = e^{i\theta_1}$ or $e^{i\theta_{j_0}}$ for some j_0. In any case $|h(x)| = 1$.

It follows that $\|h\|_\infty = 1$ and that

$$\sum_1^n |g(x_j) - g(x_{j-1})| \le \|\widetilde{\Lambda}\| \|h\|_\infty = \|\Lambda\|$$

for the partition P, so $g \in BV_0[a,b]$ and $\|g\|_{BV} \le \|\Lambda\|$.

We define $\Psi : C[a,b]' \to V[a,b]$ by $\Psi(\Lambda) = \tilde{g}$ where $\tilde{g}$ is the right continuous modification of g satisfying $\tilde{g}(a) = 0$ defined above. The estimate $\|\tilde{g}\|_{BV} = \|g\|_{BV} \le \|\Lambda\|$ can be written as

$$\|\Psi(\Lambda)\|_{BV} \le \|\Lambda\|, \quad \forall \Lambda \in C[a,b]'. \tag{3.10}$$

To complete the proof, we claim that $\Lambda f = \Lambda_{\tilde{g}} f$, for all $f \in C[a,b]$. It means $\Phi(\Psi(\Lambda)) = \Lambda$ on $C[a,b]'$. In particular, Φ is surjective. Moreover, from (3.10) and (3.9) we have

$$\|g\|_{BV} \le \|\Phi(g)\| \le \|g\|_{BV},$$

so Φ is norm-preserving.

It remains to verify $\Lambda f = \Lambda_g f$. Given $\varepsilon > 0$, since f is Riemann-Stieltjes integrable with respect to g, there is some δ_1 such that

$$\left| \int_a^b f dg - \sum_1^n f(x_j)(g(x_j) - g(x_{j-1})) \right| < \varepsilon,$$

for P, $\|P\| < \delta_1$. Using

$$\begin{aligned}
\sum_1^n f(x_j)(g(x_j) - g(x_{j-1})) &= f(x_1)g(x_1) + \sum_2^n f(x_j)(g(x_j) - g(x_{j-1})) \\
&= f(x_1)\tilde{\Lambda}\chi_{[a,x_1]} + \sum_2^n f(x_j)\widetilde{\Lambda}\chi_{(x_{j-1},x_j]} \\
&= \widetilde{\Lambda}\big(f(x_1)\chi_{[a,x_1]} + \sum_2^n f(x_j)\chi_{(x_{j-1},x_j]}\big) \\
&\equiv \widetilde{\Lambda}(f'),
\end{aligned}$$

where

$$f'(x) = f(x_1)\chi_{[a,x_1]} + \sum_2^n f(x_j)\chi_{(x_{j-1},x_j]}(x),$$

we get

$$\left| \int_a^b f dg - \widetilde{\Lambda}(f') \right| < \varepsilon. \tag{3.11}$$

As f is uniformly continuous on $[a,b]$, for every $\varepsilon > 0$, there is some δ_2 such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y, \ |x - y| < \delta_2.$$

We take $\delta = \min\{\delta_1, \delta_2\}$ and $\|P\| < \delta$. Then

$$\begin{aligned}
f(x) - f'(x) &= f(x)\big[\chi_{[a,x_1]}(x) + \sum_2^n \chi_{(x_{j-1},x_j]}(x)\big] - f'(x) \\
&= (f(x) - f(x_1))\chi_{[a,x_1]}(x) + \sum_2^n (f(x) - f(x_j))\chi_{(x_{j-1},x_j]}(x).
\end{aligned}$$

As each x belongs to exactly one subinterval, say, the j_0-th,

$$|f(x) - f'(x)| = |f(x) - f(x_{j_0})| < \varepsilon$$

if $\|P\| < \delta$. That means, $\|f - f'\|_\infty < \varepsilon$, so

$$|\widetilde{\Lambda} f - \widetilde{\Lambda} f'| \leq \|\widetilde{\Lambda}\| \|f - f'\|_\infty < \varepsilon \|\widetilde{\Lambda}\|.$$

Combining with (3.11),

$$\left| \int_a^b f dg - \widetilde{\Lambda} f \right| \leq \left| \int_a^b f dg - \widetilde{\Lambda} f' \right| + \left| \widetilde{\Lambda} f' - \widetilde{\Lambda} f \right| \leq (1 + \|\Lambda\|)\varepsilon.$$

Since ε is arbitrary and $\tilde{\Lambda}$ extends Λ, $\int f dg = \widetilde{\Lambda} f = \Lambda f$.

The proof of Theorem 3.14 is completed. □

Perhaps you have met other "Riesz representation theorems" in Real Analysis. All these theorems relate the space of continuous linear functionals to integrals with respect to the space of certain measures in different contexts. A reasonably general version, which may be regarded as a higher dimensional generalization of Theorem 3.14, can be found in chapter 2 of Rudin's "Real and Complex Analysis". Riesz representation theorem is significant because it links up real analysis and functional analysis. By the way, there are two famous mathematicians named Riesz, Frigyes and Marcel, both left footprints in analysis. In these lectures, the Riesz always refers to the elder brother.

3.6 Reflexive Spaces

To any normed space X there associates another normed space, namely its dual X'. Since the dual space X' is again a normed space, one may consider the double dual space $(X')'$ or simply X''. It is interesting to observe that any vector in X can be viewed as a vector in X''.

Proposition 3.15. *For $x_0 \in X$, define a functional $\tilde{x_0}$ on X' by*

$$\tilde{x_0}(\Lambda) = \Lambda x_0, \ \forall \Lambda \in X'.$$

Then $\tilde{x_0} \in X''$ and $\|\tilde{x_0}\| = \|x_0\|$. The mapping J (called ***canonical identification*** *or* ***canonical embedding****): $x_0 \mapsto \tilde{x_0}$ is a norm-preserving, linear map from X to X''.*

Notice here $\|x_0\|$ is the norm of x_0 in X and $\|\tilde{x_0}\|$ stands for the operator norm in X''.

Proof. Clearly, $J : x_0 \mapsto \tilde{x_0}$ is linear. From

$$|\tilde{x_0}(\Lambda)| = |\Lambda x_0| \leq \|\Lambda\| \|x_0\|$$

we also have $\tilde{x_0} \in X''$ with operator norm $\|\tilde{x_0}\| \leq \|x_0\|$. By Corollary 3.10 we pick Λ_0 satisfying $\|\Lambda_0\| = 1$ and $\Lambda_0 x_0 = \|x_0\|$. Then

$$\|x_0\| = \Lambda_0 x_0 = \tilde{x_0}(\Lambda_0) \leq \|\tilde{x_0}\| \|\Lambda_0\| = \|\tilde{x_0}\|,$$

so J is norm-preserving. □

A normed space is called a **reflexive space** if the canonical identification is a norm-preserving linear isomorphism. By Proposition 3.13 surjectivity of the canonical map is sufficient for the space to be reflexive. Interestingly there are non-reflexible Banach spaces with the property that there exists a norm-preserving linear isomorphism from X to X''. Of course, this isomorphism cannot be the canonical identification.

It is an easy exercise to show that all finite dimensional normed spaces are reflexive.

Proposition 3.16. *ℓ^p $(1 < p < \infty)$ is a reflexive space.*

Proof. For every $T \in (\ell^p)'$, there exists a unique $y^T \in \ell^q$ such that $Tx = \sum_j y_j^T x_j$ for all x in ℓ^p. Given $\Lambda \in (\ell^p)''$, the linear functional given by $\Lambda_1 y^T = \Lambda T$ is bounded in ℓ^q. By Proposition 3.5, there exists some $z \in \ell^p$ such that $\Lambda T = \Lambda_1 y^T = \sum_j y_j^T z_j$ for all y^T in ℓ^q. Recalling from the definition the canonical identification of z, $z^*(T) = Tz = \sum_j y_j^T z_j$. By comparison we see that $\Lambda = z^*$, that is to say, the canonical identification is surjective, so ℓ^p is reflexive.

□

Likewise, the L^p-space $L^p(X, \mu)$ where (X, μ) is a measure space and $p \in (1, \infty)$ is reflexive. This result, also known as Riesz representation theorem, is a standard one in real analysis, see, for instance, [Hewitt-Stromberg] or [Rudin].

Before giving some non-reflexive spaces, we note two results which may be viewed as necessary conditions for reflexivity.

Proposition 3.17. *A reflexive space is a Banach space.*

Proof. According to Proposition 3.4, the dual space of a normed space is a Banach space. As now $X = (X')'$ is the dual of the normed space X', it must be complete. □

From this result, we see that $(C[a, b], \|\cdot\|_p)$ is not reflexive for $p \in [1, \infty)$ since $C[a, b]$ is not complete under the L^p-norm.

Proposition 3.18. *If X' is separable, then X is also separable.*

Proof. As X' is separable, the subset $\{\Lambda \in X' : \|\Lambda\| = 1\}$ is also separable. Pick a countable dense set $\{\Lambda_k\}$ in this subset. Using the definition of the operator norm, for each Λ_k we can find x_k, $\|x_k\| = 1$, such that $\Lambda_k x_k \geq 1/2$.

Let E be the closure of the span of $\{x_k\}_1^\infty$. E is separable because all linear combinations of x_k's with coefficients in $\mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$ form a countable dense subset in E. We shall finish the proof by showing $E = X$.

For, if $X \setminus E \neq \varnothing$ we pick $x_0 \in X \setminus E$. By Corollary 3.10 we can find some $\Lambda_0 \in X'$ such that $\Lambda_0 = 0$ on E, $\|\Lambda_0\| = 1$. On the other hand, as $\{\Lambda_k\}$ is dense, for any $\varepsilon < 1/2$, there is some k_0 such that $\|\Lambda_0 - \Lambda_{k_0}\| < \varepsilon$. It follows that for all $x \in E$, $\|x\| = 1$.

$$\begin{aligned} |\Lambda_{k_0} x| &\leq |(\Lambda_{k_0} - \Lambda_0)x| + |\Lambda_0 x| \\ &= |(\Lambda_{k_0} - \Lambda_0)x| \\ &\leq \|\Lambda_{k_0} - \Lambda_0\| < \varepsilon. \end{aligned}$$

Taking $x = x_{k_0}$,

$$\frac{1}{2} \leq |\Lambda_{k_0} x_{k_0}| < \varepsilon,$$

contradiction holds. □

Using Proposition 3.16, we see that ℓ^1 is not reflexive. For, if it is, then $(\ell^\infty)' = (\ell^1)'' = \ell^1$. As ℓ^1 is separable, ℓ^∞ must be separable. However, this is in conflict with Proposition 2.7. Similarly it is not hard to show that the dual of $C[a, b]$ is not separable, so $C[a, b]$ is not reflexive.

Reflexive spaces have many nice properties. They arise from many contexts, for instance, the Sobolev spaces $W^{k,p}(\mathbb{R}^n), 1 < p < \infty$, are an indispensable tool in the modern study of partial differential equations. They reflexive and separable. The interested reader may google for it.

To end this section, we point out three further properties of a reflexive space:

First, any closed subspace of a reflexive space is also a reflexive space. Second, a Banach space is reflexive if and only if its dual is reflexive. The proofs of these two results are elementary and left as exercises. Third, the best approximation problem (see Section 2.4) always has an affirmative answer in a reflexive space. More precisely, let C be a closed, convex subset in this space and x_0 a point lying outside C. Then there exists a point z_0 in C such that $\|x - z_0\| \leq \|x - z\|$ for all $z \in C$. We will defer the proof until Chapter 7 where more properties of reflexive spaces can be found.

Exercise 3

1. (a) Let $\Lambda_i,\ i = 1, 2$ be two linear functionals on a vector space such that $N(\Lambda_1) = N(\Lambda_2)$. Show that $\Lambda_1 = c\Lambda_2$ for some non-zero scalar c.

   (b) Let $\Lambda_i,\ i = 1, \cdots, n$ be linear functionals on a vector space. For some linear functional Λ, suppose that $\Lambda x = 0$ for all those x in $\bigcap_i N(\Lambda_i)$. Show that Λ is a linear combination of Λ_i's.

2. Show that a linear functional Λ on a normed space is bounded if and only if its kernel is closed. Hint: The proof is short and elementary.

3. Let $(X, || \cdot ||))$ be an infinite dimensional normed space. Given any $\Lambda_1, \cdots, \Lambda_n$ in X'. Show that there exists a non-zero point $x \in X$ satisfying $\Lambda_j x = 0$ for all $j = 1, \cdots, n$.

4. Verify the following linear functionals are bounded and determine their operator norms (without rigorous proofs):

   (a) $\Lambda x \equiv a_1 x_1 + \cdots + a_n x_n,\ x = (x_1, \cdots, x_n)$, on $\mathbb{R}^n$ under the Euclidean norm.

   (b) $\delta \in L(C_b(\mathbb{R}), \mathbb{R})$, $\delta f \equiv f(x_0)$ where x_0 is a given point.

   (c) $T \in L(C[0, 1], \mathbb{R})$,
   $$Tf \equiv \int_0^1 f(x)g(x)dx,$$
   where g is a given continuous $\mathbb{R}$-valued function on $[0, 1]$.

   (d) The same as in (b) except now $\mathbb{R}$ is replaced by $\mathbb{C}$.

   (e) $S \in L(l^p, \mathbb{R})$, $Sx \equiv x_5$ where $x = (x_1, x_2, \cdots)$.

5. Show that $(\ell^1)' = \ell^\infty$.

6. Provide a detailed proof that the map Φ in the proof of Proposition 3.5 is onto ℓ^q and its inverse is given by Ψ.

7. Show that $(c_0)' = \ell^1$ where c_0 is the space of all sequences converging to 0 ("null sequences") under the sup-norm.

8. Let p be a gauge on the vector space X. Assume the set $C = \{x \in X :\ p(x) < 1\}$ is non-empty. Show that C is convex, and, for any $\alpha \in (0, 1),\ \alpha x \in C$.

9. Let C be a non-empty convex set containing 0 in the vector space X. (a) Show that $p_C(x) \leq 1$ for all x in C. (b) Under what conditions on C do we have $C = \{x :\ p_C(x) < 1\}$?

   The following two exercises, which will be used in Chapter 7, are concerned with separating two convex sets in a vector space.

10. Let A and B be two non-empty, disjoint, convex sets in the real vector space X. Show that there exists a non-zero linear functional Λ such that $\Lambda x \leq \Lambda y$ for all $x \in A$ and $y \in B$. Hint: Define $C = A - B + z$ where z is a point from $B - A$. Then 0 belongs to C and z lies outside. Define some linear functional Λ_0 on the one-dimensional space spanned by z so that it is dominated by the gauge of C, and apply the general Hahn-Banach theorem.

11. Let A and B as above where now X is normed. (a) Let A be open. Show that there is a bounded linear functional Λ such that $\Lambda x < \Lambda y$ for all $x \in A$ and $y \in B$. This is the weak separation form of the geometric Hahn-Banach theorem.

    (b) Let A be sequentially compact and B closed. Show that there exist a bounded linear functional Λ and two constants α, β, such that
    $$\Lambda x < \alpha < \beta < \Lambda y, \text{ for all } x \in A,\ y \in B.$$
    This is the strong separation form of the geometric Hahn-Banach theorem.

12. In the proof of Theorem 3.6 in class I assumed that the vector space is real. Read Lemma 3.8 for the complex version.

13. Let p_1 and p_2 be two subadditive and positive homogeneous functions on X over $\mathbb{R}$. Let $\Lambda \in L(X, \mathbb{R})$ satisfying
    $$\Lambda x \leq p_1(x) + p_2(x), \quad \forall x \in X.$$
    Show that there exist Λ_1 and Λ_2 in $L(X, \mathbb{R})$ such that $\Lambda = \Lambda_1 + \Lambda_2$ and $\Lambda_1 x \leq p_1(x)$, $\Lambda_2 x \leq p_2(x)$. Hint: Consider $p(x, y) = p_1(x) + p_2(y)$ on $X \times X$ which is subadditive and positive homogeneous and $\Lambda_0(x, x) \equiv \Lambda x$.

14. Define the shift operator on ℓ^∞ by $S(x_1, x_2, x_3, \cdots) = (x_2, x_3, \cdots)$. Show that there exists a bounded linear functional B on ℓ^∞ over reals satisfying

    (a) $Bx = \lim_{k\to\infty} x_k$ provided the limit exists;

    (b) $\underline{\lim}_{k\to\infty} x_k \leq Bx \leq \overline{\lim}_{k\to\infty} x_k$, and

    (c) $Bx = B(Sx), \quad \forall x \in \ell^\infty$.

    (d) Show that $B(1, 0, 1, 0, \cdots) = (1/2, 1/2, \cdots)$.

    (e) Show Bx cannot be expressed as $\sum_1^\infty k_j x_j$ for some $(k_1, k_2, \cdots) \in \ell^1$.

    Bx is called the **Banach limit** of the sequence x. Hint: Set $\Lambda^N x = (x_1 + \cdots + x_N)/N$, $Y = \{x \in \ell^\infty : \lim_{N\to\infty} \Lambda^N x \text{ exists}\}$. Then define $B_0 x = \lim_{N\to\infty} \Lambda^N x$ on Y and consider $p(x) = \overline{\lim}_{N\to\infty} \Lambda^N x$.)

15. (a) Prove that if Λ_1 and Λ_2 are both dual points of x in a normed space (that's, $\Lambda_i x = \|x\|$, $\|\Lambda_i\| = 1$, $i = 1, 2$), then $(\Lambda_1 + \Lambda_2)/2$ is also a dual point of x.

    (b) Show that the dual point in $\ell^p, 1 < p < \infty$, is unique. You need to use the equality case in Young's inequality.

16. Show that for a separable normed space X, there exists a countable set $\{\Lambda_j\} \subset X'$ such that
    $$\|x\| = \sup_j \frac{|\Lambda_j x|}{\|\Lambda_j\|}.$$

17. Optional readings: Chapters 3 and 4 of Lax's " Functional Analysis" contain more results and applications on Hahn-Banach theorem. Also look up Wikipedia on this item.

18. Prove Fact 1 in Section 3.5.

19. Let f be bounded and g be increasing on $[a, b]$. The upper and lower Darboux sums of f with respect to g and the partition P are given respectively by
    $$U(f, g, P) = \sum_j \sup_{x \in I_j} f(x)(g(x_{j+1}) - g(x_j)),$$
    and
    $$L(f, g, P) = \sum_j \inf_{x \in I_j} f(x)(g(x_{j+1}) - g(x_j)).$$
    Show that

    (a) $U(f, g, P)$ decreases and $L(f, g, P)$ increases when P is refined.

    (b) $L(f, g, P) \leq U(f, g, P')$ for any two partitions.

    (c) If $f \in C[a, b]$ then $\forall \varepsilon > 0$, there exists P such that $U(f, g, P) - L(f, g, P) < \varepsilon$.

20. Show that $C[a,b] \subset R_g[a,b]$ for any monotone function g on $[a,b]$.

21. Prove that any BV-function g can be written as the difference of two increasing functions. This is Jordan decomposition theorem. At each $x \in [a,b]$, define the increasing function N_g by

$$N_g(x) = \sup \sum_j |g(x_{j+1}) - g(x_j)|$$

where the supremum is over all partitions of $[a,x]$. Show that $N_g(y) \geq N_g(x) + |g(x) - g(y)|$, for $y > x$.

22. Let $f \in C[a,b]$ and $g \in BV[a,b]$, Then for any sequence of partitions with length tending to 0, the corresponding R-S sums tends to the RS-integral.

23. Let g be an increasing function and $\tilde{g}$ the right continuous function obtained from g. Prove that

$$\int f \, d\tilde{g} = \int f \, dg, \quad \forall f \in C[a,b].$$

24. Let g_1 and g_2 be right continuous functions in $BV_0[a,b]$ satisfying

$$\int f \, dg_1 = \int f \, dg_2, \quad \forall f \in C[a,b].$$

Show that $g_1 = g_2$ on $[a,b]$. Hint: Let $c \in [a,b)$ and take $\delta > 0$ small. Plug in

$$f(x) = \begin{cases} 1, & x \in [a,c] \\ 0, & x \in [c+\delta, b] \end{cases}$$

and f is linear in $[c, c+\delta]$. Then use

$$0 = g(c) + \int_c^{c+\delta} f \, dg,$$

where $g = g_2 - g_1$. For the second term we have the formula

$$\int_c^{c+\delta} f \, dg = -g(c) - \frac{1}{\delta} \int_c^{c+\delta} g \, dx$$

and then pass $\delta \to 0$.

25. Prove that $C[a,b]'$ is not separable. Suggestion: Consider the evaluation maps.

26. Consider the space consisting of all polynomials as a subspace of $L^2[a,b]$. Is it reflexive?

27. Use Proposition 3.14 to provide a short proof of the following completion theorem: To every normed space X there exists a Banach space $\widetilde{X}$ and a norm-preserving linear injection Φ from X to $\widetilde{X}$ so that $\Phi(X)$ is dense in $\widetilde{X}$.

28. Prove that any closed subspace of a reflexive space is also reflexive. Suggestion: By Hahn-Banach theorem there is a "restriction map" from X' to Y' where Y is a closed subspace of X. It induces a linear map from Y'' to X''. Since X is reflexive, the image of this linear map comes from some vector in X. Then show that in fact this vector belongs to Y by the spanning criterion.

29. Show that a Banach space is reflexive if and only if its dual space is reflexive.

水是眼波橫，山是眉峰聚。
王觀《卜算子》

Chapter 4

Bounded Linear Operator

We studied normed spaces in the previous three chapters. Now we come to bounded linear operators on these spaces. A bounded linear operator is the infinite dimensional analog of a matrix. The norm-preserving linear isomorphism and the canonical identification studied the previous chapters are special cases of bounded linear operators. They are very special ones. Due to the complexity of the structure of infinite dimensional spaces, bounded linear operators are much more diverse and difficult to investigate than matrices, and yet there are many applications. After introducing basic definitions and properties in Section 1 and examining some examples in Section 2, we turn to two theorems, namely, the uniform boundedness principle and the open mapping theorem. Together with Hahn-Banach theorem, they form the cornerstone of the subject. Nevertheless, unlike the Hahn-Banach theorem, both theorems depend critically on completeness. The careful reader should keep tracking how completeness is involved in the proofs of these theorems. We end this chapter with a brief discussion on the spectrum of a bounded linear operator. Being the infinite dimensional counterpart of the eigenvalues of a matrix, spectra play an important role in analyzing bounded linear operators.

4.1 Bounded Linear Operators

Let X and Y be two vector spaces over $\mathbb{F}$. Recall that a map $T : X \to Y$ is a linear operator (usually called a linear transformation in linear algebra) if for all $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{F}$,

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2).$$

The null space (or kernel) of T, $N(T)$, is the set $\{x \in X : Tx = 0\}$ and the range of T is denoted by $R(T)$. Both $N(T)$ and $R(T)$ are subspaces of X and Y respectively.

The collection of all linear operators from X to Y forms a vector space $L(X, Y)$ under pointwise addition and scalar multiplication of functions.

When $X = \mathbb{F}^n$ and $Y = \mathbb{F}^m$, any linear operator (or called linear transformation) can be represented by an $m \times n$ matrix with entries in $\mathbb{F}$. The vector space $L(\mathbb{F}^n, \mathbb{F}^m)$ is of dimension mn.

When X and Y are normed, one prefers to study continuous linear operators. $T \in L(X, Y)$ is continuous means it is continuous as a mapping from the metric space X to the metric space Y. It is called a **bounded linear operator** if it maps any bounded set in X to a bounded set in Y. By linearity, it suffices to map a ball to a bounded set. We encountered the same situation when the target space is the scalar field in Chapter 3.

Parallel to Proposition 3.2, we have

Proposition 4.1. *Let $T \in L(X, Y)$ where X and Y are normed spaces. We have*

(a) T is continuous if and only if it is continuous at a point.

(b) T is bounded if and only if there exists a constant $C > 0$ such that

$$\|Tx\| \le C\|x\|, \quad \text{for all } x.$$

(c) T is continuous if and only if T is bounded.

We leave the proof of this proposition to the reader.

We denote the collection of all bounded linear operators from X to Y by $B(X, Y)$. It is a subspace of $L(X, Y)$. They coincide when X and Y are of finite dimension, of course.

We observe that $X' = B(X, \mathbb{F})$.

The space $B(X, Y)$ not only inherits a vector space structure from X and Y but also a norm structure. For $T \in B(X, Y)$, define its **operator norm** by

$$\|T\| \equiv \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

It is immediate to check that $\|\cdot\|$ makes $B(X, Y)$ into a normed space. Furthermore, for $T \in B(X, Y)$ and $S \in B(Y, Z)$, the composite operator $ST \in B(X, Z)$ and

$$\|ST\| \leq \|S\|\|T\|.$$

Taking $X = Y = Z$, it means we have a multiplication structure on $B(X)$ which makes $B(X)$ a Banach algebra. Banach algebra is an advanced topic which has many applications in abstract harmonic analysis. We will not discuss it further. Nevertheless, the multiplicative property still has some interesting implications, where the reader can find one in Theorem 4.4.

The following proposition is useful in determining the operator norm.

Proposition 4.2. *Let $T \in B(X, Y)$. Suppose M is a positive number satisfying*

(a) $\|Tx\| \leq M\|x\|$, for all $x \in D$ where D is a dense set in X, and

(b) there exists a nonzero sequence $\{x_k\} \subset D$ such that $\|Tx_k\|/\|x_k\| \to M$.

Then $M = \|T\|$.

Proof. For any $x \in X$, pick a sequence $y_k \to x$, $y_k \in D$. Then $\|Tx\| = \lim_{k\to\infty} \|Ty_k\| \leq M \lim_{k\to\infty} \|y_k\| = M\|x\|$ shows that $\|Tx\| \leq M\|x\|$, for all $x \in X$. By the definition of the operator norm,

$$\|T\| \leq \sup_{\|x\|=1} \|Tx\| \leq M.$$

On the other hand, for the sequence $\{x_k\}$ given in (b),

$$M = \lim_{k\to\infty} \frac{\|Tx_k\|}{\|x_k\|} \leq \|T\|,$$

so $M = \|T\|$. □

The following result, which generalizes Proposition 3.4, can be established in a similar way.

Proposition 4.3. *$B(X, Y)$ is a Banach space if Y is a Banach space.*

Let $T \in B(X, Y)$ where X and Y are normed spaces. Then T is called **invertible** if it is bijective with the inverse in $B(Y, X)$. When X and Y are finite dimensional, every linear bijective map is automatically bounded, so it is always invertible. However, this is no longer true in the infinite dimensional setting. In many applications, some problem can be rephrased to solving the equation $Tx = y$ in some spaces for some linear operator T. The invertibility of T means the problem has a unique solution for every y.

Furthermore, for two solutions $Tx_i = y_i, i = 1, 2$, the continuity of T^{-1} implies the estimate $\|x_2 - x_1\| \leq C\|y_2 - y_1\|,\ C = \|T^{-1}\|$, from which we see that the solution depends continuously on the given data. This is related to the concept of well-posedness in partial differential equations.

The following general result is interesting.

Theorem 4.4. *Let $T \in B(X, Y)$ be invertible where X is a Banach space. Then $S \in B(X, Y)$ is invertible whenever S satisfies $\|I - T^{-1}S\|, \|I - ST^{-1}\| < 1$.*

The conditions $\|I - T^{-1}S\|$ and $\|I - ST^{-1}\| < 1$ should be understood as a measurement on how S is close to T. The idea behind this theorem as follows. We would like to solve $Sx = y$ for a given y. Rewriting the equation in the form $Tx + (S - T)x = y$ and applying the inverse operator to get $(I - E)x = T^{-1}y$ where I is the identity operator on $B(X, X)$ and $E = T^{-1}(T - S) \in B(X, X)$ is small in operator norm. So the solution x should be given by $(\sum_{j=0}^{\infty} E^j)T^{-1}y$ as suggested by the formula $(1 - x)^{-1} = \sum_j x^j$ for $|x| < 1$.

Our proof involves infinite series in $B(X, X)$. As parallel to what is done in elementary analysis, an infinite series $\sum_k x_k,\ x_k \in (X, \|\cdot\|)$, is **convergent** if its partial sums $s_n = \sum_1^n x_k$ form a convergent sequence in $(X, \|\cdot\|)$. We note the following criterion, "M-Test", for convergence.

Proposition 4.5. *An infinite series $\sum_k x_k$ in the Banach space X is convergent if there exist $a_k \geq 0$ such that $\|x_k\| \leq a_k$ for all k and $\sum_k a_k$ is convergent.*

Proof. We have

$$\|s_n - s_m\| = \|\sum_{m+1}^{n} x_k\| \leq \sum_{m+1}^{n} \|x_k\| \leq \sum_{m+1}^{n} a_k,$$

and the result follows from the convergence of $\sum_k a_k$ and the completeness of X. □

In particular, the series is convergent if there exists some $\rho \in (0, 1)$ such that $\|x_k\| \leq \rho^k$ for all k.

Corollary 4.6. *Let $L \in B(X, X)$ where X is a Banach space with $\|L\| < 1$. Then $I - L$ is invertible with inverse given by*

$$(I - L)^{-1} = \sum_{k=0}^{\infty} L^k.$$

Proof. By assumption, there exists some $\rho \in (0, 1)$ such that $\|L\| \leq \rho$. From $\|L^k\| \leq \|L\|^k \leq \rho^k$ and Proposition 4.5 that $\sum_{k=0}^{\infty} L^k$ converges in $B(X, X)$. Moreover,

$$(I - L)\sum_{k=0}^{\infty} L^k = \sum_{k=0}^{\infty}(I - L)L^k = \lim_{n\to\infty}\sum_{k=0}^{n}(I - L)L^k = \lim_{n\to\infty}(I - L^{n+1}) = I.$$

Similarly, $\sum_{k=0}^{\infty} L^k(I - L) = I$. □

Proof of Theorem 4.4. We adopt the notations in the above paragraph. As $\|E\| < 1$ by assumption, Corollary 4.6 implies that $\sum_{j=0}^{\infty} E^j$ is the inverse of $I - E$. Letting $x = (\sum_{j=0}^{\infty} E^j)T^{-1}y$, then $(I - E)x = T^{-1}y$, that is, $Sx = y$. We have shown that S is onto. Also it is bounded. On the other hand, from $\|(S - T)x\| = \|(ST^{-1} - I)Tx\| \leq \|ST^{-1} - I\|\|Tx\|$, we have

$$\begin{aligned} \|Sx\| &\geq \big|\|Tx\| - \|(S - T)x\|\big| \\ &\geq (1 - \|I - ST^{-1}\|)\|Tx\| \\ &\geq \frac{(1 - \|I - ST^{-1}\|)}{\|T^{-1}\|}\|x\|, \end{aligned}$$

So S has a bounded inverse. We have completed the proof of this theorem.

As an application let us show that all invertible linear operators form an open set in $B(X,Y)$ when X is complete. Let T_0 be invertible. Then for each T satisfying $\|T - T_0\| < \rho \equiv 1/\|T_0^{-1}\|$, we have $\|I - T_0^{-1}T\| \leq \|T_0^{-1}\|\|T_0 - T\| < 1$, so by this theorem T is invertible. That means the ball $B_\rho(T_0)$ is contained in the set of all invertible linear operators, and consequently it is open. For an $n \times n$-matrix, its corresponding linear transformation is invertible if and only if it is nonsingular. Again a matrix is nonsingular if and only if its determinant is non-zero. As the determinant is a continuous function on matrices (as the space $\mathbb{F}^{n^2}$), for all matrices close to a nonsingular matrix their determinants are non-zero, so all nonsingular matrices form an open set in the vector space of all $n \times n$-matrices. Theorem 4.4 shows that this result holds in general.

A main theme in linear algebra is to solve the nonhomogeneous linear system

$$Ax = b,$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$ are given. The Fredholm alternative states that either this linear system is uniquely solvable, or the homogeneous system

$$A'y = 0,$$

has nonzero solutions y, where A' is the transpose matrix of A. Moreover, when this happens, the nonhomogeneous system is solvable if and only if b is perpendicular to all solutions y of the homogeneous system. Can we extend this beautiful result to linear operators in Banach spaces? We need to answer the following question before we can proceed, namely, how do we define the transpose of a linear operator?

For a bounded linear operator T from the normed space X to another normed space Y there associates with a linear operator T' from Y' to X' called the **transpose** of T. Indeed, we define T' by

$$T'y'(x) \equiv y'(Tx), \quad \text{for all } y' \in Y', \quad x \in X.$$

It is straightforward to prove the following result.

Proposition 4.7. *Let T' be defined as above. Then*

(a) T' is a bounded linear operator from Y' to X'. Furthermore, $\|T'\| = \|T\|$.

(b) The correspondence $T \to T'$ is linear from $B(X,Y)$ to $B(Y',X')$.

(c) If $S \in B(Y,Z)$ where Z is a normed space, then $(ST)' = T'S'$.

We examine the finite dimensional situation. Let T be a linear operator from $\mathbb{F}^n$ to $\mathbb{F}^m$. Let $\{e_j\}$ and $\{f_j\}$ be the canonical bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. We have $Tx = \sum a_{kj}\alpha_j f_k$ where $x = \sum_j \alpha_j e_j$, so T is represented by the matrix $m \times n$-matrix (a_{kj}). On the other hand, we represent T' as a matrix with respect to the dual canonical bases $\{f_j'\}$ and $\{e_j'\}$ as $T'y' = \sum b_{kj}\beta_j e_k'$ where $y' = \sum_j \beta_j f_j'$. From the relation $T'y'(e_j) = y'(Te_j)$ for all j we have $b_{kj} = a_{jk}$. Thus the matrix of T' is the transpose of the matrix of T. This justifies the terminology. In some books it is called the adjoint of T. Here we shall reserve this terminology for a later occasion.

There are close relations between the ranges and kernels of T and those of its transpose which now we explore. Recall that the kernel of $T \in B(X,Y)$ is given by $N(T) = \{x \in X : Tx = 0.\}$ and its range is $R(T) \equiv T(X)$. The null space is always a closed subspace of X and $R(T)$ is a subspace of Y, but it may not be closed.

For a subspace Y of the normed space X, we define its **annihilator** to be

$$Y^\perp = \{x' \in X' : x'(y) = 0, \quad \text{for all } y \in Y\}.$$

Similarly, for a subspace G of X', its **annihilator** is given by

$$^\perp G = \{x \in X : x'(x) = 0, \quad \text{for all } x' \in G\}.$$

It is clear that the annihilators in both cases are closed subspaces, and the following inclusions hold:

$$Y \subset^{\perp} (Y^{\perp}),$$

and

$$G \subset (^{\perp}G)^{\perp}.$$

Lemma 4.8. *Let X be a normed space, Y a closed subspace of X and G a closed subspace of X'. Then*

(a)

$$Y =^{\perp} (Y^{\perp});$$

(b) in addition, if X is reflexive,

$$G = (^{\perp}G)^{\perp}.$$

Proof. (a) It suffices to show $^{\perp}(Y^{\perp}) \subset Y$. Any $x_0 \in^{\perp} (Y^{\perp})$ satisfies $\Lambda x_0 = 0$ whenever Λ vanishes on Y. By the spanning criterion (or Theorem 3.9), x_0 belongs to Y.

(b) It suffices to show $(^{\perp}G)^{\perp} \subset G$. Any $\Lambda_1 \in (^{\perp}G)^{\perp}$ satisfies $\Lambda_1 x = 0$ for all $x \in^{\perp} G$. If Λ_1 does not belong to G, as G is closed and the space is reflexive, there is some $x_1 \in X$ such that $\Lambda_1 x_1 \neq 0$ and $x_1 \in^{\perp} G$ according to Theorem 3.9, contradiction holds. □

Proposition 4.9. *Let X and Y be two normed spaces and $T \in B(X,Y)$. Then we have*

$$N(T') = \overline{R(T)}^{\perp},$$

$$N(T) =^{\perp} \overline{R(T')},$$

$$^{\perp}N(T') = \overline{R(T)},$$

$$N(T)^{\perp} = (^{\perp}\overline{R(T')})^{\perp}.$$

Proof. $T'y_0' = 0$ means $T'y_0'(x) = 0$ for all $x \in X$. By the definition of the transpose of T we have $y_0'(Tx) = 0$ for all x. Since T is continuous, $y_0' \in \overline{R(T)}^{\perp}$. We conclude that $N(T') \subset \overline{R(T)}^{\perp}$. By reversing this reasoning we obtain the other inclusion, so the first identity holds.

The second identity can be proved in a similar manner.

The third and the fourth identities are derived from the first and the second after using the previous lemma. □

It is clear that we have

Corollary 4.10. *Let X and Y be normed and $T \in B(X,Y)$. Then $R(T)$ is dense in Y if and only if T' is injective.*

The significance of this result is evident. It shows that in order to prove the solvability of the equation $Tx = y$ for any given $y \in Y$, it suffices to show that the only solution to $T'y' = 0$ is $y' = 0$. This sets up a relation between the solvability of the equation $Tx = y$ and the uniqueness of the transposed equation $T'x = 0$.

Fredholm alternative can be established for linear operators with more structure. For instance, in Chapter 6 we will show that it holds for $T = Id + K$ where K is a compact operator on a Hilbert space.

4.2 Examples of Linear Operators

There are plenty linear operators in analysis. Here we discuss some examples.

Linear operators on sequence spaces are direct generalization of linear transformations on $\mathbb{F}^n$. Let $x = (x_1, x_2, \cdots)$ be a sequence. Then $Tx = (y_1, y_2, \cdots)$ is again a sequence whose entries y_k depends linearly on x. You may write it formally as

$$y_k = \sum_{j=1}^{\infty} c_{jk} x_j .$$

Depending on which sequence space and the growth on the coefficients c_{jk}, T defines a bounded linear operator or an unbounded one.

Let us consider two cases. First, let $\{a_j\}$ be a null sequence with nonzero terms and define a linear operator from ℓ^p to itself by $Tx \equiv (a_1x_1, a_2x_2, \cdots,)$. It is clear that $|Tx|_p \leq \|a\|_\infty \|x\|_p$ so T is bounded. However, it is not invertible because its inverse is not bounded. To see this, assume T^{-1} exists. But then $T^{-1}e_j = a_j^{-1}e_j$ which implies that $|a_j^{-1}|$'s are uniformly bounded, contradicting that $\{a_j\}$ is null.

Second, the shift (to the right) operator $S_R : \ell^p \mapsto \ell^p (1 \leq p \leq \infty)$ given by

$$S_R(x_1, x_2, x_3, \cdots) = (0, x_1, x_2, x_3, \cdots).$$

It is easily checked that $S_R \in B(\ell^p, \ell^p)$ and $\|S_R\| = 1$. Obviously S_R is not onto, so it is not invertible.

Now, we consider integral operators. These operators arise as the inverse operators for differential operator as well as convolution operators. We restrict our attention on the one dimensional situation. Fix a continuous function $K \in C([a, b] \times [a, b])$ (usually called the integral kernel) and define

$$\mathscr{I} f(x) = \int_a^b K(x, y) f(y) dy, \quad \text{for all } f \in C[a, b].$$

Clearly $\mathscr{I}$ is linear. Let's show that it is also bounded on $C[a, b]$. In fact,

$$|\mathscr{I} f(x)| \leq \int_a^b |K(x, y)||f(y)| dy \leq M \|f\|_\infty$$

where $M = \sup_x \int_a^b |K(x, y)| dy$. So $\|\mathscr{I}\| \leq M$. By some careful work, one can show that $\|\mathscr{I}\|$ is equal to M precisely.

Integral operators can also be defined in other spaces. To do this we note the following lemma.

Lemma 4.11. *Let X and Y be Banach spaces and T is a linear operator from X_1 to Y where X_1 is a dense subspace of X. Suppose that there is a constant C such that*

$$\|Tx\| \leq C\|x\|,$$

for all $x \in X_1$. Then T can be uniquely extended to a bounded linear operator from X to Y whereas the above estimate holds on X.

We leave the proof of this lemma as an exercise. We shall make no difference between T and its extension.

We apply this lemma to the L^p-spaces. From the estimate

$$\begin{aligned} \int_a^b \Big| \int_a^b K(x, y) f(y) dy \Big|^p dx &\leq (b-a) \max |K|^p \Big(\int_a^b |f(y)| dy \Big)^p \\ &\leq (b-a)^p \max |K|^p \Big(\int_a^b |f(y)|^p dy \Big), \end{aligned}$$

we see that $\mathscr{I}$ can be extended to become a bounded linear operator on $L^p[a,b]$ by the lemma. Although the integral

$$\int_a^b K(x,y)f(y)dy$$

may not make sense for the "ideal points" in $L^p[a,b]$, it is customary to denote it by the same expression for all points in this space.

In passing one should note that the abuse of notation $\mathscr{I}$; $\mathscr{I}f$ first stands for $f \in C[a,b]$ but then for its extension in $L^p(a,b)$ for all p.

What is the transpose of $\mathscr{I}$? Let us determine it on $L^2(a,b)$. From real analysis we know that this space is self-dual, that is, any bounded linear functional on $L^2(a,b)$ is given by

$$\Lambda_g(f) = \int_a^b f(x)g(x)dx,$$

for some $g \in L^2(a,b)$, that is, the map $\Phi : L^2(a,b) \to L^2(a,b)'$ given by $g \mapsto \Lambda_g$ is a norm-preserving linear isomorphism. Now, from the definition for the transpose,

$$(\mathscr{I}'\Lambda_g)(f) = \Lambda_g(\mathscr{I}f) = \int_a^b \big(\int_a^b K(x,y)f(y)dy\big)g(x)dx = \int_a^b h(x)f(x)dy,$$

where

$$h(x) = \int_a^b K(y,x)g(y)dy.$$

Hence $\mathscr{I}'\Lambda_g = \Lambda_h$. Since $L^2(a,b)'$ may be identified with $L^2(a,b)$ via Φ, the transpose of $\mathscr{I}$ may be viewed as a map on $L^2(a,b)$ to itself given by

$$\mathscr{I}'f(x) = \int_a^b K(y,x)f(y)dy.$$

In particular, $\mathscr{I} = \mathscr{I}'$ when $K(x,y)$ is symmetric in x and y.

Now given a continuous function f in $C(S^1)$ (the space of all continuous, $2\pi-$periodic functions) we can define a sequence of complex numbers by its Fourier coefficients

$$c_n = \int_0^{2\pi} f(x)e^{-inx}dx, \quad \text{for } n \in \mathbb{Z}.$$

The Parseval identity

$$\sum_{n\in\mathbb{Z}} |c_n|^2 = \int_0^{2\pi} |f(x)|^2 dx$$

shows that the linear operator $\mathcal{F} : (C(S^1), \|\cdot\|_2) \to \ell^2(\mathbb{Z})$ assigning f to $\{c_n\}$ can be extended to become a norm preserving linear isomorphism from $L^2(S^1)$ to ℓ^2. In particular, $\mathcal{F}$ is invertible. This result partly justifies the assertion that a function is determined by its Fourier series.

In the study of the well-posedness of solutions of partial differential equations we encounter numerous linear operators. This provides opportunity to apply the soft method of functional analysis to partial differential equations. Very often it is crucial to find the most appropriate spaces for the differential operator or its inverse operator to act on.

Finally, let's consider the differential operator. Let X be the subspace of $C[0,1]$ consisting of continuous differentiable functions. The differential operator d/dx maps X to $C[0,1]$. It is linear but unbounded. Taking $f_k = \sin kx$, $\|df_k/dx\|_\infty = \|k\cos kx\|_\infty = k \to \infty$, but $\|f_k\|_\infty = 1$. Beyond this introductory course, unbounded operators play an important role in quantum physics.

4.3 Baire Theorem

In the next two sections we shall discuss the uniform boundedness principle and the open mapping theorem both due to Banach. The underlying idea of the proofs of these theorems is the Baire theorem for complete metric spaces.

The motivation is somehow a bit strange at first glance. It is concerned with the decomposition of a space as a union of subsets. For instance, we can decompose the plane $\mathbb{R}^2$ as the union of strips $\mathbb{R}^2 = \bigcup_{k\in\mathbb{Z}} S_k$ where $S_k = (k, k+1] \times \mathbb{R}$. In this decomposition each S_k is not so sharply different from $\mathbb{R}^2$. Aside from the boundary, the interior of each S_k is just like the interior of $\mathbb{R}^2$. On the other hand, one can make the more extreme decomposition: $\mathbb{R}^2 = \bigcup_{\alpha\in\mathbb{R}} l_\alpha$ where $l_\alpha = \{\alpha\} \times \mathbb{R}$. Each l_α is a vertical straight line and is very different from $\mathbb{R}^2$. It is simpler in the sense that it is one-dimensional and has no area. The sacrifice is now we need an uncountable union. The question is: Can we represent $\mathbb{R}^2$ as a countable union of these sets (or sets with lower dimension)? It turns out that the answer is no. The obstruction comes from the completeness of the ambient space.

We need one definition. Let (X, d) be a metric space. A subset E of X is called **nowhere dense** if its closure does not contain any metric ball. Equivalently, E is nowhere dense if $X \setminus \overline{E}$ is dense in X. Note that a set is nowhere dense if and only if its closure is nowhere dense. The following result is called **Baire theorem**.

Theorem 4.12. *Let $\{E_k\}_1^\infty$ be a sequence of nowhere dense subsets of (X, d) where (X, d) is complete. Then $X \setminus \bigcup \overline{E_k}$ is dense in X.*

In particular, this theorem asserts that it is impossible to express a complete metric space as a countable union of nowhere sets. In applications, we often use it in the following form: Suppose $X = \bigcup_1^\infty E_k$. Then at least the closure of one of the E_k's has non-empty interior. An equivalent formulation is also useful: The intersection of countably many open dense sets in a complete metric space is again a dense set (though not necessarily open.)

Lemma 4.13. *Let $\{\overline{B}_j\}$ be a sequence of closed balls in the complete metric space X which satisfies $\overline{B}_{j+1} \subset \overline{B}_j$ and diam $\overline{B}_j \to 0$. Then $\bigcap_{j=1}^\infty \overline{B}_j$ consists of a single point.*

Proof. Pick x_j from $\overline{B_j}$ to form a sequence $\{x_j\}$. As the diameters of the balls tend to zero, $\{x_j\}$ is a Cauchy sequence. By the completeness of X, $\{x_j\}$ converges to some x^*. Clearly x^* belongs to all $\overline{B}_j$ and is unique. □

Proof of Theorem 4.10. By replacing E_j by its closure if necessary, we may assume all E_j's are closed sets. Let B_0 be any ball. We want to show that $B_0 \bigcap (X \setminus \bigcup_j E_j) \neq \emptyset$.

As E_1 is nowhere dense and closed, we can find a closed ball $\overline{B_1} \subset B_0$ such that $\overline{B_1} \cap E_1 = \emptyset$ and its diameter $d_1 \leq d_0/2$, the diameter of B_0. Next, as E_2 is nowhere dense and closed, by the same reason there is a closed ball $\overline{B_2} \subset B_1$ such that $\overline{B_2} \cap E_2 = \emptyset$ and $d_2 \leq d_1/2$. Repeating this process, we obtain a sequence of closed balls $\overline{B_j}$ satisfying (1) $\overline{B_{j+1}} \subset \overline{B_j}$, (2) $d_j \leq d_0/2^j$, and (c) $\overline{B_j}$ is disjoint from $E_1, \cdots, E_j$. By Lemma 4.13 there is a point x^* in the common intersection of all E_j's. As $x^* \in \overline{B_j}$ for all j, $x^* \in B_0 \setminus \bigcup_j E_j$.

□

Baire theorem has many interesting applications. We include a short one here. You can find more in the next sections and the exercises.

Proposition 4.14. *Any basis of an infinite dimensional Banach space contains uncountably many vectors.*

Proof. First we claim any finite dimensional subspace of an infinite dimensional normed space is nowhere dense. Let E be such a subspace. As it is finite dimensional, it is closed. (Why?) Pick $x_0 \in X \setminus E, \|x_0\| = 1$

(such x_0 exists because X is of infinite dimensional). For any $x \in E$ and $\varepsilon > 0$, the point $x_\varepsilon = x + \varepsilon x_0 \in X \setminus E$ and $\|x - x_\varepsilon\| < \varepsilon$, so $E = \overline{E}$ does not contain any ball.

Let B be a countable basis of X, $B = \{x_k\}_{k=1}^{\infty}$. By the definition of a basis,

$$X = \bigcup_{n=1}^{\infty} E_n, \ E_n = \langle x_1, \cdots, x_n \rangle.$$

But this is impossible according to Baire theorem! □

4.4 Uniform Boundedness Principle

The following **uniform boundedness principle** is also called Banach-Steinhaus theorem as a tribute to its discoverers. Steinhaus was the teacher of Banach.

Theorem 4.15. *Let $\mathscr{T}$ be a family of bounded linear operators from a Banach space X to a normed space Y. Suppose that $\mathscr{T}$ is pointwisely bounded in the sense that for all x, there exists a constant C_x such that $\|Tx\| \le C_x$ for all $T \in \mathscr{T}$. Then we can find a constant M such that $\|T\| \le M$, for all $T \in \mathscr{T}$.*

Proof. Let $E_k = \{x \in X : \|Tx\| \le k, \text{ for all } T \in \mathscr{T}\}$. We observe that

$$X = \bigcup_{k=1}^{\infty} E_k.$$

This is simply because for any $x \in X$, $\|Tx\| \le C_x$ by assumption. Hence $x \in E_k$ for all $k \ge C_x$. Clearly each E_k is closed. By Baire theorem there is some E_{k_0} which contains a ball B. It follows from the lemma below that $\|T\| \le M$, for all $T \in \mathscr{T}$. □

Lemma 4.16. *Let $T \in L(X, Y)$ where X and Y are normed spaces. Suppose that $\|TB_\rho(x_0)\| \le C$. Then $\|T\| \le 2C/\rho^{-1}$.*

Proof. As $B_\rho(0) = B_\rho(x_0) - x_0$, by linearity, we have

$$\|TB_\rho(0)\| \le \|TB_\rho(x_0)\| + \|Tx_0\| \le C + \|Tx_0\|,$$

so

$$\|T\| = \sup \|TB_1(0)\| \le \frac{C + \|Tx_0\|}{\rho} \le \frac{2C}{\rho}.$$

□

Uniform boundedness principle does not hold when the completeness of X is removed, see exercise.

An alternative formulation of this principle is sometimes quite useful. A vector x_0 is called a **resonance point** for a family of bounded linear operators $\mathscr{T}$ if $\sup_{T \in \mathscr{T}} \|Tx_0\| = \infty$.

Theorem 4.17. *Let $\mathscr{T}$ be a family of bounded linear operators from X to Y where X is a Banach space and Y is a normed space. Suppose that $\sup_{T \in \mathscr{T}} \|T\| = \infty$. Then the resonance points of $\mathscr{T}$ forms a dense set in X.*

Proof. Suppose resonance points are not dense. There exists a ball $B_\rho(x_0)$ on which $\mathscr{T}$ is pointwisely bounded, that's, for all $x \in B_\rho(x_0)$, $\|Tx\| \le C_x$, for all $T \in \mathscr{T}$. For any $x \in X$, $z = \rho x/\|x\| + x_0 \in B_\rho(x_0)$,

$$\|T(\rho \frac{x}{\|x\|} + x_0)\| = \|Tz\| \le C_z, \quad \text{for all } T \in \mathscr{T}$$

implies

$$\|Tx\| \leq \frac{(C_z + \|Tx_0\|)}{\rho}\|x\|, \quad \text{for all } T \in \mathscr{T}.$$

So $\mathscr{T}$ is pointwisely bounded on the whole X. By Banach-Steinhaus theorem, $\|T\| \leq M$ for all T. But this is impossible by assumption. □

We give an application of Theorem 4.17 to Fourier series. Recall that for any Riemann integrable function f of period 2π, its Fourier series is given by

$$\frac{a_0}{2} + \sum_1^\infty (a_n \cos nx + b_n \sin nx),$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos nydy, \ n \geq 0,$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\sin nydy, \ n \geq 1.$$

We list the following facts (see, for instance, Stein and Shakarchi "Fourier Analysis"):

(1) The n-th partial sum $S_n f$ of the Fourier series

$$(S_n f)(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

has a closed form

$$(S_n f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin((n+\frac{1}{2})(y-x))}{\sin\frac{y-x}{2}} f(y)dy.$$

(2) It is believed that the formula

$$f(x) = \frac{a_0}{2} + \sum_1^\infty (a_k \cos kx + b_k \sin kx)$$

should hold for "sufficiently nice functions".

(3) Taking $f_0(x) = 1$ $(0 \leq x \leq \pi)$ and $f_0(x) = 0$ $(-\pi \leq x < 0)$ and extend it periodically in $\mathbb{R}$. The Fourier series of f_0 is

$$\frac{1}{2} + \sum_{k \text{ odd}} \frac{2}{k\pi}\sin kx.$$

We have $f_0(0) = 0$ but the value of the Fourier series at 0 is $\frac{1}{2}$. This shows that "sufficiently nice functions" should exclude discontinuous ones.

(4) For any Lipschitz continuous, 2π-periodic function f, its Fourier series converges uniformly to f everywhere.

For continuous 2π-periodic functions it took some time to produce an example, see [Stein-Shakarchi] for an explicit construction.

Here we present a soft proof of a stronger result. Denote by $C(S^1)$ the vector space of all continuous, 2π-periodic functions. It can be identified with $\{f \in C[-\pi,\pi] : f(-\pi) = f(\pi)\}$, which is a closed subspace of $C[-\pi,\pi]$ under the sup-norm.

Theorem 4.18. *The subset $\{f \in C(S^1) :$ The Fourier series of f diverges at 0$\}$ is dense in $C(S^1)$. In particular, f is not equal to its Fourier series at 0.*

Proof. We note that each partial sum $f \mapsto S_n f$ may be regarded as a linear operator, so composing with the evaluation at 0, $\Lambda_n f = (S_n f)(0)$, or

$$\Lambda_n f = \frac{a_0}{2} + \sum_1^n a_k,$$

forms a bounded linear functional on $C(S^1)$.

From the closed form of the partial sums we have

$$\Lambda_n f = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sin((n+\frac{1}{2})y)}{\sin \frac{y}{2}} f(y)dy.$$

The integral kernel $K(x) = \sin((n+1/2)x)/\sin x/2$ is continuous provided we set $K(0) = 2n+1$. Its operator norm is equal to

$$\|\Lambda_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|\sin((n+\frac{1}{2})y)|}{|\sin \frac{y}{2}|} dy$$

by the lemma below. We claim that $\sup_n \|\Lambda_n\| = \infty$. This is done by a direct computation:

$$\begin{aligned}
\|\Lambda_n\| &= \frac{1}{\pi} \int_0^{\pi} \frac{|\sin((n+\frac{1}{2})y)|}{\sin \frac{y}{2}} dy \\
&\geq \frac{2}{\pi} \int_0^{\pi} \frac{|\sin((n+\frac{1}{2})y)|}{y} dy \quad (\because 0 \leq \sin\theta \leq \theta) \\
&= \frac{2}{\pi} \int_0^{(n+\frac{1}{2})\pi} \frac{|\sin x|}{x} dx \\
&\geq \frac{2}{\pi} \sum_{j=1}^{n} \int_{(j-1)\pi}^{j\pi} \frac{|\sin x|}{x} dx \\
&> \frac{2}{\pi} \sum_{j=1}^{n} \frac{1}{j\pi} \int_{(j-1)\pi}^{j\pi} |\sin x| dx \quad \left(\because \frac{1}{x} > \frac{1}{j\pi} \text{ for } x \in [(j-1)\pi, j\pi]\right) \\
&= \frac{4}{\pi^2} \sum_{j=1}^{n} \frac{1}{j} \to \infty
\end{aligned}$$

as $n \to \infty$. By Theorem 4.17, resonance points are dense in $C(S^1)$. However, resonance points are precisely those functions in $C(S^1)$ whose Fourier series diverges at 0. □

Lemma 4.19. *Let*

$$\Lambda f = \int_a^b f(x)g(x)dx$$

where $g \in C[a,b]$. Then $\Lambda \in C[a,b]'$ with

$$\|\Lambda\| = \int_a^b |g(x)|dx.$$

Proof. This lemma could be proved by Riesz representation theorem. However, a direct proof is preferred.

Clearly we have

$$|\Lambda f| \leq \int_a^b |g(x)|dx,$$

for all $f, \|f\|_\infty \le 1$. We need to establish the reverse inequality.

First assume that g is a polynomial p. Let I_k and J_k be open subintervals on which p is positive and negative respectively. For each small $\varepsilon > 0$, let I_k' and J_k' be subintervals of I_k and J_k respectively so that the distance between I_k' (resp. J_k') and the endpoints of I_k (resp. J_k) is equal to ε. Define a function $f_\varepsilon \in C[a,b]$ by setting it to be 1 on I_k', -1 on J_k', 0 at endpoints of I_k and J_k and linear in between. Then $f_\varepsilon \in C[a,b]$ with $\|f_\varepsilon\|_\infty = 1$.

$$|\Lambda f_\varepsilon| = \Big| \int_a^b (f_\varepsilon(x) - \text{sgn} p(x)) p(x) dx + \int_a^b |p(x)| dx \Big| \ge \int_a^b |p(x)| dx - C\varepsilon,$$

for some constant C independent of ε. By Proposition 4.2,

$$\|\Lambda\| \ge \int_a^b |g(x)| dx,$$

when g is the polynomial p. By an approximation argument, this inequality also holds for every continuous g. □

We end this section with a discussion on soft and hard methods in analysis. Very often a theorem can be proved by two methods of very different nature. As a first example, consider the existence of transcendental numbers. Recall that an algebra number is a number that is a root for some polynomial in rational coefficients and a number is a transcendental number if it is not algebraic. In history the first transcendental number was found by Liouville (1844) who proved that $\sum_j 10^{-j!}$ is transcendental. The transcendentality of e and π were established by Hermite (1873) and Lindemann (1882) respectively. However, using Cantor's theory of cardinality, it is easily shown that all algebraic numbers form a countable set. Since $\mathbb{R}$ is uncountable, the set of all transcendental numbers is equal to $\mathbb{R}$ minus all algebra numbers and therefore is uncountable. The soft method shows there are infinitely many transcendental numbers, but it cannot pinpoint which one is transcendental. In the previous section we discussed Baire theorem. As an application of this theorem, in the exercise you are asked to show that all continuous, nowhere differentiable functions are dense in $C[0,1]$. In 1872, Weierstrass caused a sensation in math community by constructing such functiond explicitly. This class of functions are given by

$$\sum_j a^j \cos(b^n \pi x),$$

where $a \in (0,1)$, b an odd integer, $ab > 1 + 3\pi/2$. Again the soft method cannot give you any explicit example. Finally, in the above discussion we proved that the collection of all periodic, continuous functions whose Fourier series are divergent at 0 is a dense subset of $C(S^1)$, but again we cannot tell which one belongs to this collection. You need to find it in a hard way.

4.5 Open Mapping Theorem

The open mapping theorem asserts that a surjective bounded linear operator from a Banach space to another Banach space must be an open map. This result is uninteresting in the finite dimensional situation, but turns out to be very important for infinite dimensional spaces. From history there were several concrete, relevant results in various areas, Banach had the insight to single out the property as a theorem.

A map $f : (X,d) \mapsto (Y,\rho)$ between two metric spaces is called an **open map** if $f(G)$ is open in Y for any open set G in X. This should not be confused with continuity of a map, namely, f is continuous if $f^{-1}(E)$ is open in X for any open set E in Y. As an example, let us show that every non-zero linear functional on a normed space X is an open map. Indeed, pick $z_0 \in X$ with $\Lambda z_0 = 1$. Such point always exists when the functional Λ is non-zero. For any open set G in X, we claim that ΛG is open. Letting $\Lambda x_0 \in \Lambda G$, as $x_0 \in G$ and G is open, there exists some $R > 0$ such that $B_R(x_0)$ is contained in G. Then $x_0 + rz_0 \in B_R(x_0)$ for all $r \in (-R,R)$ and $\Lambda(x_0 + rz_0) = \Lambda x_0 + r$ imply that $(\Lambda x_0 + R, \Lambda x_0 - R) \in \Lambda G$, so ΛG is open.

Before stating the theorem, let's state a necessary and sufficient condition for a linear operator to be open.

Lemma 4.20. *Let $T \in L(X,Y)$ when X and Y are normed spaces. T is an open map if the image of a ball under T contains a ball.*

Roughly speaking, a linear operator either has "fat" image or it collapses everywhere.

Proof. We use "D" instead of "B" to denote a ball in Y. Suppose there exists $D_{r_0}(Tx_1) \subset TB_{R_0}(x_0)$ for some $x_1 \in B_{R_0}(x_0)$. By linearity, $D_{r_0}(Tx_1) = D_{r_0}(0) + Tx_1 \subset TB_{R_0}(x_0)$ implies

$$\begin{aligned} D_{r_0}(0) &\subset TB_{R_0}(x_0) - Tx_1 \\ &= TB_{R_0}(x_0 - x_1) \\ &\subset TB_{R_1}(0), \quad R_1 = R_0 + \|x_0 - x_1\|. \end{aligned}$$

Let G be an open set in X. We want to show that TG is open. So, for $Tx_0 \in TG$, $x_0 \in G$, as G is open, we can find a small $\rho > 0$ such that $B_\rho(x_0) \subset G$. From the above inclusion,

$$D_\varepsilon(0) \subset TB_\rho(0), \qquad \varepsilon = \rho \frac{r_0}{R_1},$$

or

$$D_\varepsilon(Tx_0) \subset TB_\rho(x_0)$$

which shows that the ball $D_\varepsilon(Tx_0)$ is contained in TG, so TG is open. □

Now we state and prove the **open mapping theorem**.

Theorem 4.21. *Any surjective bounded linear operator from a Banach space to another Banach space is an open map.*

Unlike the uniform boundedness principle here we require both the domain and target of the linear operator be complete.

Proof. Step 1: We claim that there exists $r > 0$ such that

$$D_r(0) \subset \overline{TB_1(0)}.$$

For, as T is onto, we have

$$Y = \bigcup_1^\infty TB_j(0) = \bigcup_1^\infty \overline{TB_j(0)}.$$

By assumption Y is complete, so we may apply Baire theorem to conclude that $\overline{TB_{j_0}(0)}$ contains a ball for some j_0, i.e.,

$$D_\rho(y_0) \subset \overline{TB_{j_0}(0)}.$$

Since $TB_{j_0}(0)$ is dense in $\overline{TB_{j_0}(0)}$, by replacing $D_\rho(y_0)$ by a smaller ball if necessary, we may assume $y_0 = Tx_0$, for some $x_0 \in B_{j_0}(0)$. Then

$$D_\rho(y_0) \subset \overline{TB_{j_0}(0)} \subset \overline{TB_R(x_0)}, \ R = j_0 + \|x_0\|,$$

so

$$D_\rho(0) \subset \overline{TB_R(0)},$$

or

$$D_r(0) \subset \overline{TB_1(0)}, \qquad r = \frac{\rho}{R}.$$

Step 2: $D_r(0) \subset TB_3(0)$.

First, note by scaling,

$$D_{\frac{r}{2^n}}(0) \subset \overline{TB_{\frac{1}{2^n}}(0)}, \quad \text{for all } n \geq 0 \tag{4.1}$$

Letting $y \in D_r(0)$, we want to find $x^* \in B_3(0), Tx^* = y$. We will do this by constructing an approximating sequence.

For $\varepsilon = \frac{r}{2}$, from (4.1) with $n = 0$, there exists $x_1 \in B_1(0)$ such that

$$\|y - Tx_1\| < \frac{r}{2}.$$

As $y - Tx_1 \in D_{\frac{r}{2}}(0)$, for $\varepsilon = \frac{r}{2^2}$, from (4.1) with $n = 1$, there exists $x_2 \in B_{\frac{1}{2}}(0)$ such that

$$\|y - Tx_1 - Tx_2\| < \frac{r}{2^2}.$$

Keep doing this we get $\{x_n\}$, $x_n \in B_{\frac{1}{2^{n-1}}}(0)$ such that

$$\|y - Tx_1 - Tx_2 - \cdots - Tx_n\| < \frac{r}{2^n}.$$

Setting $z_n = \sum_1^n x_j$, we have

$$\|y - Tz_n\| < \frac{r}{2^n}.$$

Let's verify that $\{z_n\}$ is a Cauchy sequence in X. $\forall n, m,\ m < n$,

$$\begin{aligned} \|z_n - z_m\| &= \|x_{m+1} + \cdots + x_n\| \\ &\le \|x_{m+1}\| + \cdots + \|x_n\| \\ &< \frac{1}{2^m} + \cdots + \frac{1}{2^n} \le \frac{1}{2^{m-1}} \to 0 \end{aligned}$$

as $m \to \infty$. From the completeness of X we may set $z^* = \lim_{n\to\infty} z_n$. Let's check that $z^* \in B_3(0)$ and $Tz^* = y$. For,

$$\|z_n\| \le \sum_1^n \|x_j\| \le \sum_1^n \frac{1}{2^{j-1}} \le 2 < 3.$$

So z^* belongs to the closure of $B_2(0)$, or, in $B_3(0)$. Next,

$$\begin{aligned} \|y - Tz^*\| &\le \|y - Tz_n\| + \|Tz_n - Tz^*\| \\ &\le \frac{r}{2^n} + \|T\|\|z_n - z^*\| \to 0 \end{aligned}$$

as $n \to \infty$, so $y = Tz^*$.

We have shown that the image of the ball $B_3(0)$ under T contains the ball $D_r(0)$, and the desired conclusion follows from Lemma 4.17. □

Recall that a linear operator is invertible if it is bounded, bijective and with a bounded inverse. The following theorem shows that the boundedness of the inverse comes as a consequence of boundedness and surjectivity of the operator when working on Banach spaces. This is called the **Banach inverse mapping theorem**.

Corollary 4.22. *Let $T \in B(X, Y)$ be a bijection where X and Y are Banach spaces. Then T is invertible.*

Proof. It suffices to show that the inverse map T^{-1} is bounded. From the above proof $D_r(0) \subset TB_3(0)$ holds. As T is bijective, $T^{-1}(D_r(0)) \subset B_3(0)$. In other words, T^{-1} maps a ball in Y to a bounded set in X, so T^{-1} is bounded. □

A theorem in general topology asserts that a continuous bijection from $\mathbb{R}^n$ to $\mathbb{R}^n$ must have a continuous inverse, that is, it is a homeomorphism. This property does not hold for continuous maps in a general Banach space. However, it remains to be valid when the map is a bounded linear operator.

A standard application of the open mapping theorem is the closed graph theorem. By definition a linear operator T between normed spaces X and Y is called a **closed map** or **of closed graph** if the graph of T,

$$G(T) \equiv \{(x, Tx) : \quad x \in X\} \subset X \times Y,$$

is a closed set in the product normed space $X \times Y$. Observe that $X \times Y$ is also a subspace of $X \times Y$. An alternative definition is, T is closed if whenever $x_n \to x$ and $Tx_n \to y$, we have $y = Tx$. From the definition one sees immediately that any bounded linear operator is a closed map. But the converse is not always true. As an exercise you may check that the differential operator is a closed map; but we already showed that it is unbounded. The following **closed graph theorem** provides an efficient way to verify the boundedness of a linear operator.

Theorem 4.23. *Any closed map from a Banach space to another Banach space is bounded.*

Proof. Let T be a closed map in $L(X, Y)$ where X and Y are Banach spaces. Since $X \times Y$ is complete and $G(T)$ is closed in $X \times Y$ by assumption, $G(T)$ is also a Banach space. We consider the linear map P which is simply the projection of $G(T)$ to X: $P(x, Tx) = x$. Clearly P is bijective. From the relation

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\|,$$

we see that P belongs to $B(G(T), X)$. By the above corollary we conclude that P^{-1} is bounded. There exists some constant C such that

$$\|x\| + \|Tx\| = \|P^{-1}x\| \leq C\|x\|.$$

In particular, we have $\|Tx\| \leq C\|x\|$. □

With more effort, one can deduce the open mapping theorem from the closed graph theorem. So these two results are in fact equivalent.

4.6 The Spectrum

Denote by $B(X) = B(X, X)$ the vector space of all bounded linear operators from the normed space X to itself. It is a normed space under the operator norm, and it is a Banach space when X is a Banach space. An additional algebraic operation, namely, the composition of two linear operators, makes sense in $B(X)$. In fact, we note that

(i) the identity map I is well-defined in $B(X)$;

(ii) for all $T, S \in B(X)$, $TS \in B(X)$ and $\|TS\| \leq \|T\|\|S\|$.

$B(X)$ is the prototype for Banach algebras. When the space is of finite dimension, one may regard it as $\mathbb{F}^n$, so any linear operator is essentially a square matrix. In the theory of square matrices eigenvalues and eigenvectors are of central importance. In general, it is possible to define the same notion for linear operators in $B(X)$. A scalar λ is called an **eigenvalue** for T in $L(X, X)$ if there exists a nonzero x, called an **eigenvector**, such that

$$Tx = \lambda x.$$

As before, it is readily checked that all eigenvectors form a subspace together with 0. It is closed when X is normed and T is bounded.

In the following we let X be a Banach space and the linear operator T bounded for simplicity, although much of the discussion could be extended to more general settings.

Recall that a bounded linear operator S in $B(X)$ is called invertible if it is bijective and S^{-1} belongs to $B(X)$. According to the open mapping theorem, S is invertible if and only if it is bijective. A scalar $\lambda \in \mathbb{F}$ is called a **regular value** for a bounded linear operator T if $T - \lambda I$ is invertible. The set of all regular values of T forms the **resolvent set** of T, denoted by $\rho(T)$, and we define its complement, that's, $\mathbb{F} \setminus \rho(T)$, the **spectrum** of T and denote it by $\sigma(T)$. Both terminologies are motivated by their connection with physics.

We point out that any eigenvalue λ of the bounded linear operator T must belong to the spectrum of T. Indeed, the existence of an eigenvector shows that $T-\lambda I$ is not injective, hence cannot be invertible. When the space is finite dimensional, a linear operator is injective if and only if it is surjective. Consequently, the spectrum of any linear operator consists exactly of eigenvalues. However, this is no longer the case for infinite dimensional spaces. For T on a Banach space, $T - \lambda I$ fails to be invertible for two reasons; either it is not injective or not surjective. The scalar λ is an eigenvalue when the former holds.

An example may be helpful in illustrating the situation. Let $X = C[0,1]$ over the real field and consider the linear operator T given by

$$(Tf)(x) = xf(x).$$

Clearly, $T \in B(C[0,1])$ with $\|T\| \leq 1$. If λ is an eigenvalue of T, and φ its eigenfunction, $x\varphi(x) = \lambda\varphi(x)$ will hold for all $x \in [0,1]$. This is clearly impossible, so T does not have any eigenvalue. However, for any λ not in $[0,1]$, the inverse of $T - \lambda I$ is given by the map

$$(Sf)(x) = \frac{f(x)}{x-\lambda}.$$

It is easy to check that $S \in B(C[0,1])$. When $\lambda \in [0,1]$, the inverse of $T-\lambda I$ does not exist. We conclude that although T has no eigenvalues, its spectrum is given by $\sigma(T) = [0,1]$ and resolvent set by $\mathbb{C}/[0,1]$.

As a consequence of Theorem 4.4, we have

Proposition 4.24. *Let $T \in B(X)$ where X is a Banach space. Then $I - T$ is invertible when $\|T\| < 1$.*

Corollary 4.25. *The spectrum of $T \in B(X)$ where X is a Banach space forms a closed and bounded set in $\mathbb{F}$. In fact, $|\lambda| \leq \|T\|$ for any $\lambda \in \sigma(T)$.*

Proof. If λ does not belong to $\sigma(T)$, that is, $T - \lambda I$ is invertible. There exists $\rho > 0$ such that all linear operators in $B_\rho(T - \lambda I)$ are invertible. In particular, it means $T - \mu I$, $|\lambda - \mu| < \rho$, is invertible. This shows that the complement of $\sigma(T)$ is open, hence $\sigma(T)$ is closed.

Next, if $|\lambda| > \|T\|$, then $I - \lambda^{-1}T$ and hence $T - \lambda I$ are invertible by Proposition 4.21. Hence λ cannot be in the spectrum. □

Evidently there is a natural question: Is the spectrum nonempty for any bounded linear operator in $B(X)$? After all, there are n many eigenvalues (including multiplicity) for any $n \times n$- matrix with complex entries. Remember that the proof of this fact depends on the fundamental theorem of algebra which is most easily established by using the Liouville theorem in complex analysis. It is not surprising we need to use complex analysis to establish the following two results over $\mathbb{C}$:

First, $\sigma(T)$ is always nonempty for any $T \in B(X)$;

Second, we have the formula for the "spectral radius":

$$\sup\{|\lambda| : \lambda \in \sigma(T)\} = \lim_{k\to\infty} \|T^k\|^{1/k}.$$

Theorem 4.26. *Let $T \in B(X)$ where X is a complex Banach space. Then*

(a) $\rho(T)$ is open in $\mathbb{C}$. More precisely, for any $\lambda_0 \in \rho(T)$, $\lambda \in \rho(T)$ for $|\lambda - \lambda_0| < 1/\|(\lambda_0 I - T)^{-1}\|$.

(b) For each $\Lambda \in B(X)'$, the function $\varphi(\lambda) = \Lambda(\lambda I - T)^{-1}$ is analytic in $\rho(T)$.

(c) $\sigma(T)$ is a non-empty compact set in the plane.

Proof. (a) follows immediately from Theorem 4.4 after taking $T = \lambda_0 I - T$ and $S = \lambda I - T$ in that theorem.

(b). To show analyticity we represent $\varphi(\lambda)$ as a power series around every λ_0 in $\rho(T)$. Formally, we have $\lambda I - T = (\lambda_0 I - T)[1 + (\lambda - \lambda_0)(\lambda_0 I - T)^{-1}]$, so define

$$(\lambda I - T)^{-1} = (\lambda_0 I - T)^{-1} \sum_{k=0}^{\infty} (-1)^k (\lambda_0 I - T)^{-k} (\lambda - \lambda_0)^k.$$

When $|\lambda - \lambda_0| < 1/\|(\lambda_0 I - T)^{-1}\|$, there exists some $\sigma \in (0,1)$ such that $\|(\lambda_0 I - T)^{-1}\| |\lambda - \lambda_0| < 1 - \sigma$, therefore, this power series converges and one can easily check that it converges to $(\lambda I - T)^{-1}$. Hence the above formal expression holds rigorous. For $\Lambda \in B(X)'$, we have

$$\varphi(\lambda) = \sum_{k=0}^{\infty} (-1)^k \Lambda((\lambda_0 I - T)^{-k-1})(\lambda - \lambda_0)^k$$

also converges for λ, $|\lambda - \lambda_0| < 1/\|(\lambda_0 I - T)^{-1}\|$.

(c). We first show that $\varphi(\lambda) \to 0$ as $|\lambda| \to \infty$ for any $\Lambda \in B(X)'$. In this time we expand φ at ∞. Formally

$$(\lambda I - T)^{-1} = \frac{1}{\lambda} \left(I - \frac{T}{\lambda} \right)^{-1} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{T^k}{\lambda^k}.$$

For $\lambda > \|T\|$, this can be made rigorously and so

$$\begin{aligned} |\varphi(\lambda)| &= |\Lambda(\lambda I - T)^{-1}| \\ &\le \frac{1}{|\lambda|} \sum_{k=0}^{\infty} \frac{|\Lambda T^k|}{|\lambda^k|} \\ &\le \frac{C}{|\lambda|} \|\Lambda\| \to 0 \end{aligned}$$

as $|\lambda| \to \infty$.

If $\sigma(T)$ is empty, that means φ is an entire function. As it tends to 0 at ∞, it is bounded on $\mathbb{C}$. By Liouville theorem we conclude that φ is identically zero for every $\Lambda \in B(X)'$. By Hahn-Banach theorem, $\lambda I - T = 0$ for all λ, contradiction holds. Hence the spectrum is always non-empty. In Corollary 4.22 we know that it is compact. ☐

Define the **spectral radius** of $T \in B(X)$ by

$$r_T = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

From Corollay 4.22 we know that $0 \le r_T \le \|T\|$. We have a precise formula.

Theorem 4.27. *For any $T \in B(X)$ where X is a Banach space,*

$$r_T = \overline{\lim_{n\to\infty}} \sqrt[n]{\|T^n\|}.$$

Proof. For $|\lambda| > \overline{\lim}_{n\to\infty} \sqrt[n]{\|T^n\|}$, there exists $\delta \in (0,1)$ and n_0 such that

$$|\lambda|(1-\delta) > \sqrt[n]{\|T^n\|}, \ \forall n \ge n_0.$$

So,

$$1 - \delta \ge \frac{\sqrt[n]{\|T^n\|}}{|\lambda|},$$

and $(\lambda I - T)^{-1} = 1/\lambda \sum_{k=0}^{\infty} T^k/\lambda^k$ exists, that is, $\lambda \in \rho(T)$. So $r_T \leq \overline{\lim}_{n\to\infty} \sqrt[n]{\|T^n\|}$.

On the other hand, for λ, $|\lambda| > \|T\|$, we know that $\lambda \in \rho(T)$ and, for $\Lambda \in B(X)'$,

$$\varphi(\lambda) = \sum_{k=0}^{\infty} \frac{\Lambda(T^k)}{\lambda^{k+1}}, \ |\lambda| > \|T\|$$

holds. As φ is analytic in λ, $|\lambda| > r_T$, this relation holds for all $|\lambda| > r_T$. At each λ, $|\lambda| > r_T$,

$$\left\{ \Lambda \left(\frac{T^k}{\lambda^{k+1}} \right) \right\}$$

is bounded for each Λ. Uniform boundedness principle asserts that

$$\frac{\|T^k\|}{|\lambda|^{k+1}} \leq M, \ k = 1, 2, \cdots$$

So

$$\|T^k\|^{\frac{1}{k}} \leq M^{\frac{1}{k}} |\lambda|^{\frac{k+1}{k}}$$

and

$$\overline{\lim_{n\to\infty}} \|T^k\|^{\frac{1}{k}} \leq |\lambda|.$$

As $|\lambda| > r_T$, we conclude

$$\overline{\lim_{n\to\infty}} \|T^k\|^{\frac{1}{k}} \leq r_T.$$

□

We remark that it is a good exercise to show

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \inf_n \|T^n\|^{1/n},$$

hence the "$\overline{\lim}$" in this theorem can be replaced by "lim".

Exercise 4

1. Provide a proof of Proposition 4.1.
2. Prove that $B(X, Y)$ is a Banach space when Y is a Banach space.
3. Let $T \in B(X, Y)$ where X is a Banach space and Y is normed. Suppose there exists $C > 0$ such that
$$\|Tx\| \geq C\|x\|, \ \forall x \in X.$$
Show that
(a) $R(T)$ is a complete subspace of Y, and
(b) $T \in B(X, R(T))$ is invertible.
4. Find a formula for the operator norm of an $n \times n-$matrix A in $\mathbb{F}^n$. Suggestion: Consider $\|Ax\|^2 = \langle Ax, Ax \rangle$ in the dot product.
5. Consider $T : \ell^2 \to \ell^2$ given by $Tx = (0, 4x_1, x_2, 4x_3, x_4, \cdots)$ where $x = (x_1, x_2, x_3, \cdots)$. Compare $\|T\|^2$ and $\|T^2\|$.
6. Let $S_R : \ell^2 \to \ell^2$ be $S_R(x) = (0, x_1, x_2, \cdots)$, the right shift operator.
(a) Find S_1 to satisfy $S_1 S_R = I$ on ℓ^2.
(b) Is there some S_2 satisfying $S_R S_2 = I$ on ℓ^2?

7. Let $\Phi : \ell^q \to (\ell^p)'$ be $(\Phi y)(x) = \sum_j y_j x_j$. We know that it is a norm-preserving linear isomorphism. Define $\tilde{T} = (\Phi)^{-1} T' \Phi : \ \ell^q \to \ell^q$. Show that
$$\sum_j (\tilde{T} y)_j x_j = \sum_j y_j (Tx)_j, \ \forall x \in \ell^p, \ y \in \ell^q.$$
$\tilde{T}$ may also be called the transpose of T.

8. Determine $\tilde{S_R}$ and $\tilde{S_L}$, see Problem 7 for notations.

9. Consider the integral equation
$$g(x) = f(x) + \int_0^1 A \sin(x-y) g(y) dy$$
where $f \in C[0,1]$ is given. Show that it has a unique solution $g \in C[0,1]$ if the constant $A \in (-1,1)$.

10. Let $K \in C([a,b]^2)$ and define
$$\mathscr{I}_K f(x) = \int_a^b K(x,y) f(y) dy, \quad \forall f \in C[a,b].$$
Show that $\mathscr{I}_K$ can be extended to a bounded linear operator from $L^p(a,b)$ to $L^q(a,b)$ where the operator norm
$$\|\mathscr{I}_K\| \le \left(\int_a^b \int_a^b |K(x,y)|^q dx dy \right)^{\frac{1}{q}},$$
where q is conjugate to $p \in [1, \infty)$. Hint: Use Lemma 4.9.

11. (a) Let X be a Banach space and $x : \ (a,b) \mapsto X$ be a "curve". Propose a definition of the derivative of x at $t \in (a,b)$ using the notation $x'(t)$.

    (b) Show that the initial value problem for the ODE $x' = Ax, \ x(0) = x_0$, where $A \in B(X)$ and $x_0 \in X$ are given, has a unique solution for all time. Hint: Try to define $\exp T \in B(X)$ for any $T \in B(X)$ and show the solution is given by $x(t) = \exp tA \ x_0, \ t \in \mathbb{R}$.

12. (a) Show that in a complete metric space (X,d), the intersection of countably many open, dense subsets is still a dense set.

    (b) Give an example to show that this intersection may not be open though.

13. A nowhere monotonic function in $[a,b]$ is a function which is not monotone on any subinterval. Show that all nowhere monotonic, continuous functions form a dense subset in $C[a,b]$. Suggestion: Consider the sets $E_j = \{\pm f : \text{there is some } x, (f(y)-f(x))(y-x) \ge 0 \text{ for } \ |y-x| \le 1/j\}$.

14. A function $f \in C[0,1]$ is called α-Lipschitz continuous at x_0 if
$$|f(x) - f(x_0)| \le \alpha |x - x_0|, \ \forall x \in [0,1].$$
(a) Prove that $\forall k > 0$, the set
$$L_k = \{f \in C[0,1] : f \text{ is } k\text{-Lipschitz continuous at some } x_0\}$$
is a closed and nowhere dense set in $C[0,1]$. Suggestion: First approximate f by a piecewise linear function g and then consider $g + \varepsilon \sin(Nx)$ where N is large.

    (b) Use (a) to deduce that the collection of all continuous, nowhere differentiable functions forms a dense set in $C[0,1]$. Hint: Differentiability implies Lipschitz continuity.

15. Let $\mathcal{C}_{00} = \{x \in \ell^\infty : x \text{ has finitely many non-zero terms}\}$ and define $T_n x = \sum_{j=1}^n x_j$. Show that $|T_n x| \le C\|x\|$ for all n but $\|T_n\| \to \infty$. Is it contradictory to the uniform boundedness principle?

16. Let $\{b_j\}$ be a sequence which satisfies $\sum_j a_j b_j < \infty$ whenever $\{a_j\}$ is a sequence converging to 0. Show that $\{b_j\} \in \ell^1$.

17. Let $\{T_n\} \subset B(X,Y)$ where X and Y are complete. Suppose that $\forall x$, $\lim_{n\to\infty} T_n x$ exists. Show that
(a) Define $Tx \equiv \lim_{n\to\infty} T_n x$. Then $T \in B(X,Y)$.
(b) $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$.

18. Let $L(x,y)$ be a functional on $X \times Y$ where X and Y are Banach spaces so that for each fixed x, $y \mapsto L(x,y) \in Y'$ and for each fixed y, $x \mapsto L(x,y) \in X'$. Show that L is continuous, that is, whenever $x_n \to x$ and $y_n \to y$, $L(x_n, y_n) \to L(x,y)$ as $n \to \infty$. Hint: Reduce to $x = y = 0$ and then use Banach-Steinhaus.

19. Let P be the vector space of all polynomials in $(-\infty, \infty)$ endowed with the norm
$$\|p\| = \max |a_k|, \quad p(x) = a_0 + a_1 x + \cdots + a_n x^n.$$
Show that it is not complete. Suggestion: Consider the linear functionals given by $T_k p = a_0 + \cdots + a_k,\ k \in \mathbb{N}$, and show that $|T_k p| \leq C_p$ but $\|T_k\|$ is not uniformly bounded.

20. Show without using the open mapping theorem that every non-zero linear operator from a finite dimensional normed spaces X onto another Y is an open map.

21. Let $X = \{x \in \ell^\infty : x \text{ has finitely many non-zero entries } \}$ as a subspace of ℓ^∞. Define $T : X \mapsto X$ by $TX = (x_1, \frac{x_2}{2}, \frac{x_3}{3}, \cdots)$. Verify that T is a linear bijection, but not invertible. What is wrong with the Banach inverse mapping theorem?

22. Use the Banach inverse mapping theorem to give a short proof of the following old result: Any two norms on $\mathbb{F}^n$ are equivalent.

23. Let X_1 and X_2 be two closed subspaces of the Banach space X such that $X_1 \oplus X_2 = X$. Show that there exists $M > 0$ such that $\|x_1\| + \|x_2\| \leq M\|x\|$ where $x = x_1 + x_2$, $x_i \in X_i$, $i = 1, 2$.

24. Verify the differential operator $d/dx : C^1[0,1] \mapsto C[0,1]$, where $C^1[0,1]$ is viewed as a subspace in $C[0,1]$, is closed but not bounded. What is wrong with the closed graph theorem?

25. Let $T \in L(X,Y)$ where X and Y are Banach spaces. Suppose that whenever $x_j \to 0$ in X, $\Lambda T x_j \to 0$ in Y for every $\Lambda \in Y'$. Prove that $T \in B(X,Y)$.

26. Let Z be a closed subspace of the Banach space X. Show that the projection map $\pi : X \mapsto X/Z$ where X/Z is the quotient Banach space is an open map. We shall use this problem in the next one.

27. Deduce the open mapping theorem from the closed graph theorem. Suggestion: Consider $X/N(T)$ and the closed map $\widetilde{T}$ given by $\widetilde{T}(\tilde{x}) = Tx$ on $X/N(T)$.

28. Let $T \in B(X)$ where X is normed. Show that $\sigma(T) = \sigma(T')$. Hint: Show that S is invertible if and only if S' is invertible.

29. Let S_R and S_L be the shift operators on ℓ^2 defined before. Prove the followings:
(a) All λ, $|\lambda| < 1$, are eigenvalues for S_L.
(b) $\sigma(S_L) = \{\lambda : |\lambda| \leq 1\}$.
(c) $\sigma(S_R) = \{\lambda : |\lambda| \leq 1\}$. Hint: Use Problem 1.
(d) Find the eigenvalues for S_R.

30. Let $T \in B(\ell^2)$ given by $Tx = (x_1, -x_2, x_3, -x_4, \cdots)$. Show that $\sigma(T) = \{-1, 1\}$ and both ± 1 are eigenvalues.

空城曉角， 吹入垂楊陌。 馬上單衣寒惻惻，
看盡鵝黃嫩綠， 都是江南舊相識。
姜夔《淡黃柳》

Chapter 5

Hilbert Space

The Euclidean norm is special among all norms defined in $\mathbb{R}^n$ for being induced by the Euclidean inner product (the dot product). A Hilbert space is a Banach space whose norm is induced by an inner product. An inner product enables one to define orthogonality, which in turns leads to far reaching conclusion on the structure of a Hilbert space. In particular, we show that there is always a complete orthonormal set, a substitute for a Hamel basis, in a Hilbert space. It is a natural, infinite dimensional analog of an orthonormal basis in a finite dimensional vector space. We conclude with a theorem which asserts that any infinite dimensional separable Hilbert space is "isometric" to ℓ^2. Thus once again the cardinality of the basis alone is sufficient to distinguish separable Hilbert spaces.

David Hilbert was old and partly deaf in the nineteen thirties. Yet being a diligent man, he still attended seminars, usually accompanied by his assistant Richard Courant. One day a visitor was talking on his new findings in linear operators on Hilbert spaces. The professor was puzzled first. Soon he grew impatient and finally he turned to Courant. "Richard, what is a Hilbert space?" he asked loudly.

5.1 Inner Product

An **inner product** is a map: $X \times X \mapsto \mathbb{F}$ for a vector space X over $\mathbb{F}$ satisfying

(P1) $\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle$,

(P2) $\overline{\langle x, y\rangle} = \langle y, x\rangle$,

(P3) $\langle x, x\rangle \geq 0$ and "=" if and only if $x = 0$.

The pair $(X, \langle\cdot,\cdot,\rangle)$ is called an **inner product space**. Note that (P1) and (P2) imply

$$\langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle = \overline{\alpha_1}\langle x, y_1\rangle + \overline{\alpha_2}\langle x, y_2\rangle.$$

Example 5.1 In $\mathbb{F}^n$ define the product

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \bar{y_k}.$$

It makes $(\mathbb{F}^n, \langle\cdot,\cdot\rangle)$ into an inner product space. It is called the Euclidean space when $\mathbb{F}^n = \mathbb{R}^n$ and the unitary space when $\mathbb{F}^n = \mathbb{C}^n$.

Example 5.2 $\ell^2 = \{x = (x_1, x_2, x_3, \cdots) : x_k \in \mathbb{F}, \quad \sum_{k=1}^{\infty} |x_k|^2 < \infty\}$ becomes an inner product space under the product

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

We should keep in mind that this product is finite is a consequence of Cauchy-Schwarz inequality in Chapter 1. A variant of this space is the space of bi-sequences:

$$\ell^2(\mathbb{Z}) = \{x = (\cdots, x_{-1}, x_0, x_1, x_2, \cdots) : \quad \sum_{-\infty}^{\infty} |x_k|^2 < \infty\}$$

under the product $\langle x, y\rangle = \sum_{-\infty}^{\infty} x_k \bar{y_k}$.

Example 5.3 Recall that $L^2(a, b)$ is the completion of $C[a, b]$ under the L^2-norm. On $C[a, b]$ the L^2-product

$$\langle f, g\rangle = \int f(x)\overline{g(x)}x,$$

defines an inner product in $C[a, b]$. It is not hard to show that it has an extension to $L^2(a, b)$. See next section for more details.

Example 5.4 Any subspace of an inner product space is an inner product space under the same product.

On $\mathbb{F}^2$ and the ℓ^2-space there are Cauchy-Schwarz inequality. In fact, the most general setting for the Cauchy-Schwarz inequality is an inner product space. In the following we establish this inequality and use it to introduce the angle between two non-zero vectors and the concept of orthogonality.

Proposition 5.1. *For any x and y in an inner product space $(X, \langle\cdot,\cdot\rangle)$,*

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

Moreover, equality holds in this inequality if and only if x and y are linearly dependent.

Proof. The inequality is trivial when x or y is a zero-vector, so let's assume both x and y are non-zero. When the field is complex, let θ be a number satisfying $\langle x, y\rangle = e^{i\theta}|\langle x, y\rangle|$. Then $\langle x, z\rangle = |\langle x, y\rangle|$ where $z = e^{-i\theta}y$ is a nonnegative number. When the field is real, no need to do this as $\langle x, y\rangle$ is already real. Just take z to be y. By (P3),

$$0 \le \langle x - \alpha z, x - \alpha z\rangle = \langle x, x\rangle - 2\alpha\langle x, z\rangle + \alpha^2\langle z, z\rangle.$$

This is a quadratic equation with real coefficients in α. Since it is always nonnegative, its discriminant is non-positive. In other words,

$$\langle x, z\rangle^2 - \langle x, x\rangle\langle z, z\rangle \le 0,$$

the inequality follows.

When x and y are linearly dependent, there is some $\alpha \in \mathbb{F}$, $x - \alpha y = 0$. Plugging in the inequality we readily see that equality holds. On the other hand, when $\langle x, x\rangle\langle y, y\rangle = |\langle x, y\rangle|^2$, we can take $\alpha = \langle x, y\rangle/\langle y, y\rangle$ in $\langle x - \alpha y, x - \alpha y\rangle$. By a direct computation, $0 = \langle x - \alpha y, x - \alpha y\rangle$. By (P3), $x = \alpha y$. □

It follows from this inequality that

$$\frac{|\langle x, y\rangle|}{\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}} \le 1.$$

For any two nonzero vectors x and y there is a unique $\theta \in [0, \pi]$ (the "angle" between x and y) satisfying

$$\theta = \cos^{-1}\frac{\langle x, y\rangle}{\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}} \in [0, \pi].$$

Any two vectors x and y are **orthogonal** if $\langle x, y\rangle = 0$. The zero vector is orthogonal to all vectors.

5.2 Inner Product and Norm

There is a norm canonically associated to an inner product. Indeed, the function $\|\cdot\| : (X, \langle\cdot,\cdot\rangle) \mapsto [0,\infty)$ given by

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

defines a norm on X. To verify this, we only need to prove the triangle inequality since it is evident for the other two axioms. For $x, y \in X$,

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \quad \text{(Cauchy-Schwarz inequality)} \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Notions such as limits, convergence, open/closed sets and continuity in an inner product space will be referred to this induced norm. In particular, we have

Proposition 5.2. *The inner product $X \times X \mapsto [0,\infty)$ is a continuous function.*

Proof. For $x_n \to x$ and $y_n \to y$, we want to show that $\langle x_n, y_n\rangle \to \langle x, y\rangle$. First of all, let n_0 be a positive number satisfying $\|y_n - y\| \leq 1, \ \forall n \geq n_0$. Then $\|y_n\| \leq 1 + \|y\|$ and we have

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x,y\rangle| &= |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\ &\leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\| \\ &\leq \|x_n - x\|(1 + \|y\|) + \|x\|\|y_n - y\| \\ &\longrightarrow 0 \end{aligned}$$

as $n \to \infty$. □

A complete inner product space is called a **Hilbert space**. As any closed subspace of a Banach space is complete, any closed subspace of a Hilbert space is also a Hilbert space. Products of Hilbert spaces are Hilbert spaces. Moreover, the quotient space of a Hilbert space over a closed subspace is again a Hilbert space. For completion of an inner product space we have the following rather evident result, whose proof is left to the reader.

Proposition 5.3. *Let $(\widetilde{X}, \widetilde{\|\cdot\|})$ be the completion of $(X, \|\cdot\|)$ where $\|\cdot\|$ is induced from an inner product $\langle\cdot,\cdot\rangle$. Then there exists a complete inner product on $\widetilde{\langle\cdot,\cdot\rangle}$ which extends $\langle\cdot,\cdot\rangle$ and induces $\widetilde{\|\cdot\|}$.*

When the reader runs through his/her list of normed spaces, he/she will find that there are far more Banach spaces than Hilbert spaces. However, one may wonder these Banach spaces are also Hilbert spaces, whose inner products are just too obscure to write down. A natural question arises: How can we decide which norm is induced by an inner product and which one is not? The answer to this question lies on a simple property—the parallelogram identity.

Proposition 5.4 (Parallelogram Identity). *For any x, y in $(X, \langle\cdot,\cdot\rangle)$,*

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

Proof. Expand

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

and add up. □

This rule, which involves only the norm but not the inner product, gives a necessary condition for a norm to be induced by an inner product.

As application we first show that the $\|\cdot\|_p$-norm on $\mathbb{F}^n$ $(n \geq 2)$ is induced from an inner product if and only if $p = 2$. Take $x = (1, 1, 0, \cdots, 0)$ and $y = (1, -1, 0, \cdots, 0)$ in $\mathbb{F}^n$. We have $\|x\|_p = \|y\|_p = 2^{\frac{1}{p}}$ and $\|x + y\|_p = \|x - y\|_p = 2$. If $\|\cdot\|_p$ is induced from an inner product, Proposition 5.4 asserts

$$\|x + y\|_p^2 + \|x - y\|_p^2 = 8 = 2(\|x\|_p^2 + \|y\|_p^2) = 2^{\frac{2}{p}+2}$$

which holds only if $p = 2$.

Similarly, $C[0, 1]$ does not come from an inner product. We take $f(x) = 1$ and $g(x) = x$. Then $\|f\|_\infty = \|g\|_\infty = 1$, $\|f+g\|_\infty = 2$ and $\|f-g\|_\infty = 1$. Then $\|f+g\|_\infty^2+\|f-g\|_\infty^2 = 5 \neq 2(\|f\|_\infty^2+\|g\|_\infty^2) = 4$.

Proposition 5.5. *(a) For every x, y in a real inner product space X, we have*

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

(b) On a real normed space $(X, \|\cdot\|)$, the above identity defines an inner product on X if and only if the parallelogram identity holds.

The identity in (a) is called the **polarization identity**. It shows how one can recover the inner product from the norm in an inner product space.

Proof. (a) We have

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

By subtracting, we get the polarization formula.

(b) In view of Proposition 5.4, it remain to verify that the polarization identity defines an inner product under the validity of the parallelogram identity. In fact, (P2) and (P3) are immediate. We only need to prove (P1). By the parallelogram identity,

$$2(\|x \pm z\|^2 + \|y\|^2) = \|(x \pm z) + y\|^2 + \|(x \pm z) - y\|^2.$$

By subtracting

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

In terms of $\langle\cdot, \cdot\rangle$ we have

$$\langle x + y, z\rangle + \langle x - y, z\rangle = 2\langle x, z\rangle, \quad \text{for all } x, y, z \in X.$$

Replacing x, y by $(x + y)/2$ and $(x - y)/2$ respectively, we have

$$\langle x, z\rangle + \langle y, z\rangle = 2\langle \frac{x + y}{2}, z\rangle, \quad \text{for all } x, y, z \in X.$$

Letting $y = 0$, $\langle x, z\rangle = 2\langle x/2, z\rangle$ for all x, z. It follows that

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

that is, $\langle\cdot, \cdot\rangle$ is additive in the first component. Next, we show that $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$, for all $\alpha \in \mathbb{R}$. We observe that by induction and $2\langle x, y\rangle = \langle 2x, y\rangle$ we can show $n\langle x, y\rangle = \langle nx, y\rangle$ for all $n \in \mathbb{N}$. Using $\langle x, y\rangle + \langle -x, y\rangle = \langle 0, y\rangle = 0$ we deduce $n\langle x, y\rangle = \langle nx, y\rangle$ for all $n \in \mathbb{Z}$. Then $m\langle x/m, y\rangle = \langle x/m, y\rangle + \cdots + \langle x/m, y\rangle$ (m times)$= \langle x, y\rangle$ implies $1/m\langle x/m, y\rangle = \langle x, y\rangle$. So, $\langle nx/m, y\rangle = 1/m\langle nx, y\rangle = n/m\langle x, y\rangle$ for any rational n/m. By continuity of the norm, $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$.

□

We have the following corresponding result when $\mathbb{F} = \mathbb{C}$, whose proof may be deduced from the real case. We leave it as an exercise.

Proposition 5.6. *(a) For any x, y in a complex inner product space X, we have the polarization identities*

$$Re\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2),$$

and

$$Im\langle x, y\rangle = \frac{1}{4}(\|x+iy\|^2 - \|x-iy\|^2).$$

(b) On a complex normed space X the polarization identities define an inner product on X which induces its norm if and only if the parallelogram identity holds.

5.3 Orthogonal Decomposition

In Chapter 2 we discussed the best approximation problem for closed subspaces in a Banach space. Aside from finite dimensional subspaces (see Lemma 2.13), the problem does not always have a positive solution. One may consult p.46 in [Lax] for an example in $C[-1, 1]$. Nevertheless, with the help of orthogonality, we show in this section that for a Hilbert space this problem always has a unique solution. An immediate consequence is the existence of complementary subspaces, a property which is not necessarily valid for Banach spaces.

In fact, our result extends from subspaces to convex subsets.

Theorem 5.7. *Let K be a closed, convex subset in the Hilbert space X and $x_0 \in X \setminus K$. There exists a unique point $x^* \in K$ such that*

$$\|x_0 - x^*\| = \inf_{x\in K} \|x_0 - x\|$$

Proof. Let $\{x_n\}$ be a minimizing sequence in K, in other words,

$$\|x_0 - x_n\| \to d = \inf_{x\in K} \|x_0 - x\|.$$

We claim that $\{x_n\}$ is a Cauchy sequence. For, by parallelogram identity,

$$\begin{aligned}
\|x_n - x_m\|^2 &= \|x_n - x_0 - (x_m - x_0)\|^2 \\
&= -\|x_n - x_0 + x_m - x_0\|^2 + 2(\|x_n - x_0\|^2 + \|x_m - x_0\|^2) \\
&= -4\|\frac{x_n + x_m}{2} - x_0\|^2 + 2(\|x_n - x_0\|^2 + \|x_m - x_0\|^2) \\
&\leq -4d^2 + 2(\|x_n - x_0\|^2 + \|x_m - x_0\|^2) \\
&\to 0
\end{aligned}$$

as $n, m \to \infty$. Note that $(x_n + x_m)/2 \in K$ by convexity.

By the completeness of X and the closedness of K, $x^* = \lim_{n\to\infty} x_n \in K$. As the norm is a continuous function, we have $d = \|x_0 - x^*\|$.

Suppose $x' \in K$ also minimizes the distance. Then

$$\begin{aligned}
\|x^* - x'\|^2 &\leq -4\|\frac{x^* + x'}{2} - x_0\|^2 + 2(\|x^* - x_0\|^2 + \|x' - x_0\|^2) \\
&\leq -4d^2 + 4d^2 = 0,
\end{aligned}$$

that's, $x' = x^*$. □

This theorem plays a fundamental role in convex analysis. But here we only consider the special case when K is a closed subspace. More can be said about the best approximation in this case.

Theorem 5.8 (**Best Approximation**). *Let Y be a closed subspace of a Hilbert space X and x_0 a point lying outside Y. The point y_0 which minimizes the distance between x_0 and Y satisfies*

$$\langle x_0 - z, y\rangle = 0, \qquad \forall y \in Y,$$

at $z = y_0$. Conversely, if $z \in Y$ satisfies this condition, then z must be y_0. When this holds,

$$\|x_0 - y_0\|^2 + \|y_0\|^2 = \|x_0\|^2, \tag{5.1}$$

holds.

Proof. When y_0 minimizes $\|y - x_0\|$ among all y, it also minimizes $\|y - x_0\|^2$. For any $y \in Y$, $y_0 + \varepsilon y \in Y$, so the function

$$\varphi(\varepsilon) = \|x_0 - y_0 - \varepsilon y\|^2$$

attains a minimum at $\varepsilon = 0$. By expanding, we have

$$\varphi(\varepsilon) = \|x_0 - y_0\|^2 - \varepsilon\langle x_0 - y_0, y\rangle - \varepsilon\langle y, x_0 - y_0\rangle + \varepsilon^2\|y\|^2.$$

Clearly $0 = \varphi'(0)$ implies

$$Re\langle x_0 - y_0, y\rangle = 0.$$

Replacing y by iy, $Im\langle x_0 - y_0, y\rangle = 0$.

Conversely, if $\langle x_0 - y_0, y\rangle = 0$ for all y in Y, we have

$$\begin{aligned} \|x_0 - y\|^2 &= \|x_0 - y_0 + y - y_0\|^2 \\ &= \|x_0 - y_0\|^2 + \|y - y_0\|^2 \\ &\geq \|x_0 - y_0\|^2, \end{aligned}$$

which shows that y_0 minimizes $d(x_0, Y)$.

Finally, let y_1 also minimize the distance. By the above characterization, $\langle x_0 - y_1, y\rangle = 0$ on Y. It follows that $\langle y_0 - y_1, y\rangle = \langle x_0 - y_1, y\rangle - \langle x_0 - y_0, y\rangle = 0$. Taking $y = y_0 - y_1$ we conclude $y_0 = y_1$.

□

This theorem has the following geometric meaning. For x_0 outside Y, the projection point y_0 is the unique point on Y so that $\Delta O x_0 y_0$ forms a perpendicular triangle.

For any closed subspace Y in a Hilbert space, the **projection operator** of X onto Y is given by

$$Px_0 = \begin{cases} y_0, & \text{if } x_0 \in X \setminus Y \\ x_0, & \text{if } x_0 \in Y \end{cases}$$

We have been calling Px the best approximation of x in Y. Now we may also it the **orthogonal projection** of x on Y. It is easy to check that $P \in B(X, Y)$, $P^2 = P$ and $\|P\| = 1$. For instance, to show that P is linear, we just have to verify the obvious identity $\langle \alpha x_1 + \beta x_2 - (\alpha P x_1 + \beta P x_2), y\rangle = 0$ on Y. For then it follows from the above characterization that $P(\alpha x_1 + \beta x_2) = \alpha P x_1 + \beta P x_2$.

We also note that a more general characterization holds: For any x in X, not only those in Y, $z = Px$ if and only if z satisfies $\langle x - z, y\rangle = 0$ on Y.

We will discuss two consequences of the best approximation theorem. The first is the self-duality property of the Hilbert space.

To each z in the Hilbert space X we associate a bounded linear functional Λ_z given by $\Lambda_z x = \langle x, z\rangle$. It is routine to verify that Λ_z belongs to X' with operator norm $\|\Lambda_z\| = \|z\|$. The mapping Φ defined by mapping z to Λ_z sets up a sesquilinear map from X to X'. A map T is **sesquilinear** if $T(\alpha x_1 + \beta x_2) = \overline{\alpha} T x_1 + \overline{\beta} T x_2$. Sesquilinear and linear are the same when the field is real, and they are different when the field is complex. The following Frechét-Riesz theorem shows that Φ is surjective, so it is a norm-preserving, sesquilinear isomorphism between X and X'. A Hilbert space is self-dual in the sense that every bounded linear functional on it can be identified with a unique point in the space itself.

Theorem 5.9. *Let X be a Hilbert space. For every Λ in X', there exists a unique z in X such that $\Lambda_z = \Lambda$ and $\|z\| = \|\Lambda\|$.*

We leave the proof of this theorem to you.

Next, we consider direct sum decomposition in a Hilbert space. Recall that a direct sum decomposition of a vector space, $X = X_1 \bigoplus X_2$, where X_1 and X_2 are two subspaces, means every vector x can be expressed as the sum of one vector x_1 from X_1 and the other x_2 from X_2 in a unique way. From the uniqueness of the representation one can show that the maps $x \mapsto x_1$ and $x \mapsto x_2$ are linear maps from X onto X_1 and X_2 respectively. They are called projection maps associated to the direct sum $X_1 \bigoplus X_2$.

Direct sum decomposition is clearly useful in the study of vector spaces since it breaks down the space to two smaller (and hence simpler) subspaces. When the space is normed, it is desirable to ensure that such decomposition respects the topology in some sense. Thus we may introduce the definition that the direct sum is a "topological direct sum" if the projection maps: $x \mapsto x_1$ and $x \mapsto x_2$ are bounded from X to X_i, $i = 1, 2$. When the space is complete, certainly we would like our decomposition to break into Banach spaces. We prefer X_i, $i = 1, 2$, to be closed subspaces. An advantage of this assumption is that the projections are automatically bounded, as a direct consequence of the closed graph theorem, so any direct sum decomposition of a Banach space into closed subspaces is topological.

We now are left with question: Given any closed subspace X_1 of a Banach space X, can we find a closed subspace X_2 such that $X = X_1 \oplus X_2$? Unfortunately, except when X_1 is of finite dimension, a complementary closed space X_2 does not always exist. However, this is always true for Hilbert spaces. In fact, a deep theorem asserts that if a Banach space possesses the property that any closed subspace has a topological complement, then its norm must be equivalent to a norm induced by a complete inner product.

In fact, for any closed, proper subspace X_1, we define its "orthogonal subspace" $X_1^\perp$ to be

$$X_1^\perp = \{x \in X : \langle x, x_1 \rangle = 0, \text{ for all } x_1 \in X_1\}$$

It is clear that $X_1^\perp$ is a closed subspace. (According to Riesz-Frechet theorem, $X_1^\perp$ is the annihilator of X_1, so the notation is consistent with the one we used in Chapter 4.) Thus we have the decomposition $x = Px + (x - Px) \in X_1 + X_1^\perp$ where P is the orthogonal projection operator on X_1. We claim that this is a direct sum. For, if $x_0 \in X_1 \cap X_1^\perp$, then $\langle x_0, x_1 \rangle = 0$, for all $x_1 \in X_1$. As x_0 also belongs to X_1, taking $x_1 = x_0$, we get $\|x_0\|^2 = \langle x_0, x_0 \rangle = 0$. Hence $X_1 \cap X_1^\perp = \{0\}$. Moreover, we observe that the bounded linear operator P and $I - P$ are precisely the projection maps of the direct sum $X_1 \bigoplus X_1^\perp$. We have proved the following theorem on the orthogonal decomposition in a Hilbert space.

Theorem 5.10. *For every closed subspace X_1 of a Hilbert space X, $X = X_1 \oplus X_1^\perp$. Moreover, the projection operator $P : X \mapsto X_1$ maps x to Px which is the unique point in X_1 satisfying $\|x - Px\| = d(x, X_1)$ and the projection $Q : X \mapsto X_1^\perp$ is given by $Qx = x - Px$.*

5.4 Complete Orthonormal Sets

We start by considering the following question: How can we determine Px_0 when x_0 and the subspace Y are given? It is helpful to examine this question when X is the n-dimensional Euclidean space. Let $\{x_1, \cdots, x_m\}$ be a basis of Y. Any projection y_0 has the expression $y_0 = \sum_{k=1}^m \alpha_k x_k$. From Theorem 5.8 we have $\langle y_0 - x_0, x_k \rangle = 0$ for each $k = 1, \cdots, m$. It amounts to a linear system for the unknown $\alpha_j' s$:

$$\sum_k \langle x_j, x_k \rangle \alpha_k = \langle x_j, x_0 \rangle.$$

The system assumes a simple form when $\{x_k\}$ forms an orthonormal set. We immediately solve this system to get $y_0 = \sum_k \langle x_0, x_k \rangle x_k$. This example suggests it is better to consider orthonormal spanning sets in Y.

Lemma 5.11 (**Bessel's Inequality**). *Let S be an orthonormal set in the Hilbert space X. Then for each $x \in X$, $\langle x, x_\alpha \rangle = 0$ except for at most countably many $x_\alpha \in B$. Moreover, for any sequence $\{\alpha_k\}$ from the index set B*

$$\sum_k |\langle x, x_{\alpha_k} \rangle|^2 \le \|x\|^2. \tag{5.2}$$

Proof. Step 1: Let $\{x_k\}_1^N$ be a finite orthonormal set. For $x \in X$, we claim

$$\sum_{k=1}^N |\langle x, x_k \rangle|^2 \le \|x\|^2. \tag{5.3}$$

For, let $y = \sum_{k=1}^N \langle x, x_k \rangle x_k$. Then $\langle x - y, x_k \rangle = 0$, for all $k = 1, \cdots, N$, implies that y is the orthogonal projection of x onto the space $\langle x_1, \cdots, x_N \rangle$. As $\|y\|^2 = \sum_{k=1}^N |\langle x, x_k \rangle|^2$,

$$\sum_{k=1}^N |\langle x, x_k \rangle|^2 = \|y\|^2 = \|x\|^2 - \|x - y\|^2 \le \|x\|^2,$$

Step 2: Let $x \in X$ and l a natural number. We claim that the subset S_l of S

$$S_l = \{x : |\langle x, x_\alpha \rangle| \ge \frac{1}{l}\}$$

is a finite set. For, picking $x_{\alpha_1}, \cdots, x_{\alpha_N}$ many vectors from S_l and applying Step 1, we get

$$\|x\|^2 \ge \sum_{k=1}^N |\langle x, x_{\alpha_k} \rangle|^2 \ge \frac{N}{l^2}.$$

It follows that the cardinality of S_l cannot exceed $\|x\|^2 l^2$.

Step 3: Only countably many terms $\langle x, x_\alpha \rangle$ are non-zero. Let S_x be the subset of S consisting of all x_α's such that $\langle x, x_\alpha \rangle$ is non-zero. We have the decomposition

$$S_x = \bigcup_1^\infty S_l.$$

Since each S_l is a finite set, S_x is countable.

Now the Bessel's inequality follows from passing to infinity in (5.3). □

Now we can give an answer to the question posed in the beginning of this section.

Theorem 5.12. *Let Y be a closed subspace in the Hilbert space X. Suppose that S is an orthonormal subset of Y whose linear span is dense in Y. Then for each x, its orthogonal projection on Y is given by $\sum_k \langle x, x_k \rangle x_k$ where $\{x_k\}$ is any ordering of all those x_α in S with non-zero $\langle x, x_\alpha \rangle$.*

Proof. First of all, we need to verify that the sum $\sum_k \langle x, x_k \rangle x_k$ is convergent. By the completeness of Y it suffices to show that $\{y_n\} \equiv \{\sum_k^n \langle x, x_k \rangle x_k\}$ is a Cauchy sequence. Indeed, we have

$$\begin{aligned} \|y_n - y_m\|^2 &= \|\sum_{m+1}^n \langle x, x_k \rangle x_k\|^2 \\ &= \sum_{m+1}^n |\langle x, x_k \rangle|^2. \end{aligned}$$

By the Bessel's inequality it is clear that the right hand side of this inequality tends to zero as $n, m \to \infty$. Therefore, $\{y_n\}$ is a Cauchy sequence. Now using the characterization of the orthogonal projection stated in Theorem 5.8 and the continuity of the inner product, we conclude the proof of this theorem. □

After proving that the sum $\sum_k \langle x, x_k\rangle x_k$ is the projection of x on Y, we see that this summation is independent of the ordering of those non-zero $\langle x, x_\alpha\rangle$. (You may also deduce this result by recalling rearrangement does not change the limit of an absolutely convergent series.) Therefore we can comfortably write it in the form $\sum_\alpha \langle x, x_\alpha\rangle x_\alpha$ without causing any confusion.

The discussion so far motivates us to introduce a more natural replacement of the Hamel basis for Hilbert spaces. A subset B in a Hilbert space is called a **complete orthonormal set** if it satisfies (a) it is an orthonormal set, that is, for all $x \neq y \in B, \langle x, y\rangle = 0$, and $\|x\| = 1$, and (b) $\overline{\langle B\rangle} = X$. The conditions are different from those for a basis. In contrast, for a basis B we require (a)' all vectors in B is linearly independent, and (b)' $\langle B\rangle = X$. It is an exercise to show that (a) implies (a)', but (b) is weaker than (b)'. Some authors use the terminology "an orthonormal basis" instead of "a complete orthonormal set". We prefer to use the latter.

Theorem 5.13. *Every non-zero Hilbert space admits a complete orthonormal set.*

Proof. Let $\mathscr{F}$ be the collection of all orthonormal sets in X. Clearly $\mathcal{F}$ is non-empty and has a partial order by set inclusion. For any chain $\mathscr{C}$ in $\mathscr{F}$, clearly

$$S^* = \bigcup_{S \in \mathscr{C}} S$$

is an upper bound of $\mathscr{C}$. By Zorn's lemma, $\mathscr{F}$ has a maximal element B. We claim that B is a complete orthonormal set. First of all, B consists of normalized vectors orthogonal to each other, so (a) holds. To prove (b), let's assume there is some z not in $\overline{\langle B\rangle}$. By orthogonal decomposition, $z' = (z - Pz)/\|z - Pz\|$ where P is the projection operator on $\overline{\langle B\rangle}$ is a unit vector perpendicular to $\overline{\langle B\rangle}$. It follows that $B \cup \{z'\} \in \mathscr{F}$, contradicting the maximality of B. □

We end this section with some criteria for a complete orthonormal set.

Theorem 5.14. *Let B be an orthonormal set in the Hilbert space X. The followings are equivalent:*

(a) B is a complete orthonormal set,

(b) $x = \sum_{x_\alpha \in B} \langle x, x_\alpha\rangle x_\alpha$ holds for all $x \in X$,

(c) $\|x\|^2 = \sum_{x_\alpha \in B} |\langle x, x_\alpha\rangle|^2$ holds,

(d) $\langle x, x_\alpha\rangle = 0$, for all $x_\alpha \in B$ implies that $x = 0$.

(c) is called the (**Parseval's identity**). In other words, the Bessel's inequality holds on every orthonormal set, but the Parseval's identity holds only when the set is complete.

Proof. (a)⇒(b): When $\overline{\langle B\rangle} = X$, the orthogonal projection beomes the identity map, so (b) holds by Theorem 5.12.

(b)⇒ (c):

$$\|x\|^2 - \sum_{k=1}^n |\langle x, x_k\rangle|^2 = \|x - \sum_{k=1}^n \langle x, x_k\rangle x_k\|^2 \to 0,$$

as $n \to \infty$, by Theorem 5.12.

(c)⇒ (d): Obvious.

(d)⇒ (a): Suppose on the contrary $\overline{\langle B\rangle}$ is strictly contained in X. We can find a non-zero $x_0 \in X \setminus \overline{\langle B\rangle}$ such that $\langle x_0, x_\alpha\rangle = 0$, for all $x_\alpha \in B$. However, this is impossible by (d). □

5.5 A Structure Theorem

Recall that in linear algebra we showed that two finite dimensional vector spaces are linearly isomorphic if and only if they have the same dimension. That means there is only one invariant, the dimension, to distinguish vector spaces. A similar result holds in a (separable) Hilbert space. We prove in below that every separable Hilbert space has a countable complete orthonormal set. Consequently separable Hilbert spaces are distinguished by their cardinality.

Proposition 5.15. *A Hilbert space has a countable complete orthonormal set if and only if it is separable in its induced metric.*

Proof. Let $B = \{x_k\}_1^\infty$ be a complete orthonormal set of X. By definition, the set $\langle B\rangle$ is dense in X. However, consider the subset $S = \{x \in \langle B\rangle : x$ is a linear combination of B with coefficients in $\mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$ depending on $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}\}$. It is clear that $\overline{S} = \overline{\langle B\rangle} = X$.

On the other hand, let C be a countable, dense subset of X. We can write it as a sequence $\{x_1, x_2, x_3, \cdots\}$. Step by step we can throw away vectors which are linearly dependent of the previous ones to get a subset $\{y_1, y_2, y_3, \cdots\}$ which consists of linearly independent vectors and yet still spans X. Now, apply the Gram-Schmidt process to this subset to obtain an orthonormal set $\{z_1, z_2, z_3, \cdots\}$. From construction we have that $\langle\{z_1, z_2, z_3, \cdots\}\rangle = \langle\{y_1, y_2, y_3, \cdots\}\rangle$ so $\{z_1, z_2, z_3, \cdots\}$ is a complete orthonormal set. □

Theorem 5.16. *Every infinite dimensional separable Hilbert space X is the same as ℓ^2. More precisely, there exists an inner-product preserving linear isomorphism Φ from X to ℓ^2.*

Proof. Pick a complete orthonormal set $\{x_k\}_1^\infty$ of X whose existence is guaranteed by Proposition 5.15. Then for every $x \in X$, we have $x = \sum\langle x, x_k\rangle x_k$. Define the map $\Phi : X \mapsto \ell^2$ by $\Phi(x) = (a_1, a_2, \cdots)$ where $a_k = \langle x, x_k\rangle$. By Theorem 5.14 we know that Φ is a norm-preserving linear map from X to ℓ^2. It is also onto. For, let $\{a_k\}$ be an l^2-sequence. Define $y_n = \sum_{k=1}^n a_k x_k$. Using $\|y_n - y_m\|^2 = \sum_{m+1}^n a_k^2 \to 0$ as $n, m \to \infty$, y_n converges in X to $\sum_1^\infty a_k x_k$. Clearly, $\langle x, x_k\rangle = a_k$, so Φ is onto. Finally, it is also inner-product preserving by polarization. □

We end this section with a famous example of a complete orthonormal set.

First, Let $L^2((-\pi,\pi))$ be the completion of $C([-\pi,\pi])$ under the L^2-product. For $f, g \in C([-\pi,\pi])$ over the complex field, the product is given by

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx.$$

The set $B = \{\frac{1}{\sqrt{2\pi}}e^{inx}, n \in \mathbb{Z}\}$ is a countable set consisting of orthonormal vectors:

$$\begin{aligned}\langle\frac{1}{\sqrt{2\pi}}e^{inx}, \frac{1}{\sqrt{2\pi}}e^{imx}\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}dx \\ &= \frac{1}{2\pi}\frac{1}{i(m-n)}e^{i(n-m)x}|_{-\pi}^{\pi} \\ &= 0 \text{ if } n \neq m;\end{aligned}$$

$$\langle\frac{1}{\sqrt{2\pi}}e^{inx}, \frac{1}{\sqrt{2\pi}}e^{imx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-inx} = 1.$$

For $f \in L^2((-\pi,\pi))$, we define its Fourier series to be

$$\sum_n\langle f, \frac{1}{\sqrt{2\pi}}e^{iny}\rangle\frac{1}{\sqrt{2\pi}}e^{inx} = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\,dy,$$

and write $f \sim \sum_n c_n e^{inx}$. We know from previous discussions that this series is a well-defined function in $L^2((-\pi,\pi))$. In fact, it is the orthogonal projection of f onto the closed subspace $\overline{\langle B \rangle}$. The completeness of B is a standard result in Fourier Analysis. Here we give a quick proof by using Weierstrass' approximation theorem in the plane. That is, for any continuous function f in the unit disk there exists $\{p_n(z)\}$, where each p_n is a polynomial so that $\{p_n\}$ tends to f in supnorm.

Observe any 2π-periodic function f in $[-\pi,\pi]$ induces a function $g \in C(S^1)$ where $S^1 = \{e^{i\theta} : \theta \in [-\pi,\pi]\}$ is the unit circle in the plane by $g(e^{i\theta}) = f(\theta)$. Extend g as a continuous function in the closed disc $\overline{D} = \{(x_1,x_2) : x_1^2 + x_2^2 \le 1\}$ and denote it by the same g. For $\varepsilon > 0$, by Weierstrass' theorem there exists a polynomial $p(x_1,x_2)$ such that $\|p-g\|_{\infty,\overline{D}} < \varepsilon$. When restricted to S^1 we obtain

$$\|p(x_1,x_2) - g(x_1,x_2)\|_{\infty,S^1} < \varepsilon,$$

where

$$x_1 = \frac{1}{2}(e^{ix} + e^{-ix}),\ x_2 = \frac{1}{2i}(e^{ix} - e^{-ix}),\ x \in [-\pi,\pi].$$

As $p(x_1,x_2)$ (regarded as a function of x in $[-\pi,\pi]$) is just a linear combination of functions in B, this shows that $\langle B \rangle$ is dense in the subspace of periodic functions in $C([-\pi,\pi])$ under the sup-norm. As the sup-norm is stronger than the L^2-norm, $\langle B \rangle$ is also dense in this subspace in L^2-norm. Now, for any L^2-function f, we can find a continuous function $f_1 \in C[-\pi,\pi]$ such that $\|f - f_1\|_2 < \varepsilon$. By modifying the value of f near endpoints we can find another continuous f_2, which is now periodic, $\|f_1 - f_2\| < \varepsilon$. Finally, there exists a trigonometric polynomial p such that $\|f_2 - p\|_2 < \varepsilon$. All together we obtain $\|f-p\|_2 \le \|f-f_1\|_2 + \|f_1 - f_2\|_2 + \|f_2 - p\|_2 < 3\varepsilon$. We conclude that B forms a complete orthonormal set in $L^2((-\pi,\pi))$. In particular, for every L^2-functions, its Fourier series converges to f in L^2-norm, and the Parseval's identity

$$\int_{-\pi}^{\pi} |f(x)|^2 dx = 2\pi \sum_{n=-\infty}^{\infty} |c_n|^2,$$

holds.

For a real function f, the Fourier series is usually expressed in real form,

$$f \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos nx + b_n \sin nx\big),$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos ny\ dy,\ n \ge 0, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\sin ny\ dy,\ n \ge 1,$$

and you can write down the corresponding Parseval's identity.

Other examples of orthonormal sets can be found in the exercises.

Exercise 5

1. Establish the identity

$$\|z-x\|^2 + \|z-y\|^2 = \frac{1}{2}\|x-y\|^2 + 2\|z - \frac{1}{2}(x+y)\|$$

for x, y, and z in an inner product space.

2. Prove Proposition 5.6.

3. Show that $(C[0,1], \|\cdot\|_p)$ is not induced from any inner product for $p \in [1,\infty] \setminus \{2\}$.

4. (a) Let X be a Hilbert space and $\Lambda \in X'$. Show that there exists a unique $z \in X$ such that $\Lambda x = \langle x, z\rangle$ and $\|\Lambda\| = \|z\|$. Combining with what we have done in class, the map $\Lambda \mapsto z$ is a norm-preserving linear isomorphism when X is real and a norm-preserving sesquilinear isomorphism when X is complex. Hint: The kernel of Λ is a closed subspace of codimension one. Consider $z = x - Px$ where x lies outside this subspace and P is the projection on this subspace.

   (b) Use (a) to show that a Hilbert space is always reflexive.

5. Let X be a Hilbert space. Let $B : X \times X \to \mathbb{F}$ such that $B(x, y)$ is (a) linear in x, (b) sesquilinear in y, and (c) $|B(x, y)| \leq M\|x\|\|y\|$ for some M. Prove that there exists $T \in B(X)$ satisfying $B(x, y) = \langle Tx, y\rangle$ and $\|T\| = \sup\{B(x, y) : \|x\|, \|y\| \leq 1\}$.

6. Let X be a Hilbert space. Show that if $\{x_n\}$ *weakly converges* to x, that's, $\Lambda x_n \to \Lambda x$, $\forall \Lambda \in X'$, then $x_n \to x$ provided $\|x_n\| \to \|x\|$.

7. Let X be a vector space and X_1 a closed subspace.

   (a) Show that there exists another subspace X_2 such that $X = X_1 \bigoplus X_2$. Suggestion: Apply Zorn's lemma as in the proof of existence of basis in the first chapter.

   (b) Let P_1 be the projection map which sends x to x_1 in the above direct sum decomposition. Show that P_1 is linear and satisfies $P_1^2 = P_1$.

   (c) Similarly P_2 is defined. Explain why $P_2 = I - P_1$ and $P_1P_2 = 0$.

   The following two problems refresh your memory on materials about the Gram-Schmidt process which will be used in the next lecture.

8. Let $\{x_1, x_2, \cdots, x_n\}$ be in the Hilbert space X. Show that it is a linearly independent set if and only if its "Gram determinant"
$$\det(\langle x_i, x_j\rangle)$$
   is nonzero.

9. Let $S = \{x_1, x_2, \cdots, x_n\}$ be a finite set in the inner product space X. Review the Gram-Schmidt process how to obtain an orthonormal set S' from S so that they span the same space.

10. Let K be a closed, convex set in the real Hilbert space X and $x_0 \in X/K$. In Theorem 5.7 we proved that there exists a point $y_0 \in K$ which minimizes the distance from x_0 to K. Show that the following characterization of y_0 holds: z minimizes the distance if and only if for all $y \in K$,
$$\langle x_0 - z, y\rangle \leq \langle x - z, z\rangle.$$
    This result reduces to Theorem 5.8 when K is a subspace.

11. (a) A linear map P on a vector space X is a projection if it satisfies $P^2 = P$. Show that $X = RanP \oplus N(P)$.

    (b) If now $P \in B(X)$ where X is a Hilbert space. A projection is called an orthogonal projection if it further satisfies $\langle Px, y\rangle = \langle x, Py\rangle$, $\forall x, y$. Show that there is a one-to-one correspondence between orthogonal projections and closed subspaces of X.

12. Find
$$\min\{\int_{-1}^{1} \left|x^3 - a - bx - cx^2\right|^2 dx : a, b, c \in \mathbb{R}\}.$$

13. Find the orthogonal projection of (or the best approximation to) the function f onto the subspace spanned by $\cos x$ and $1 - 2x$ in $L^2(-\pi, \pi)$ where f is (a) e^x and (b) $\cos 6x$.

14. Find the function $g \in L^2(-\pi, \pi)$ which satisfies
$$\int_{-\pi}^{\pi} f(x)x dx = 1, \quad \int_{-\pi}^{\pi} f(x) \sin x dx = 2$$
    with the minimal L^2-norm. Show that it is unique.

15. Compare the concept of a Schauder basis and a complete orthonormal basis in a Hilbert space. Is a complete orthonormal set necessarily a Schauder basis?

16. Read pg 176-184 in [Kreyszig] for two examples of complete orthonormal sets: The Legendre polynomials for $L^2[0,1]$ and Hermite functions for $L^2(\mathbb{R})$.

17. Let H be a Hilbert space.

    (a) Show that $\{x_j\} \to x_0$ weakly if and only if for every $x \in H$, $\langle x_j, x\rangle \to \langle x_0, x\rangle$ as $j \to \infty$. (See last assignment for the concept of weak convergence.)

    (b) Show that $\{x_j\} \to x_0$ weakly if for each fixed x_k, $\langle x_j, x_k\rangle \to \langle x_0, x_k\rangle$. Suggestion: Let Y be the subspace $\{x \in H : \langle x_j, x\rangle \to \langle x_0, x\rangle \text{ as } j \to \infty\}$. Observe it contains all x_k's and is closed. Then consider the orthogonal decomposition $H = Y \oplus Y^\perp$ and apply (a).

18. (a) Write down the Parseval's identities for the Fourier series in the real case.

    (b) Evaluate the Parseval's identity for the following two functions in $L^2((-\pi,\pi))$: (a) $f(x) = 1, x \in [0,\pi], = 0, x \in [-\pi, 0)$, and (b) $g(x) = \pi - x, x \in [0,\pi], = x + \pi, x \in [-\pi, 0)$.

回首向來蕭瑟處，歸去、也無風雨也無晴。
蘇軾 《定風波》

Chapter 6

Compact, Self-Adjoint Operator in Hilbert Space

In Chapter 4 we discussed general properties of bounded linear operators. To have a taste of the richness of operator theory, here we cast our attention on a special class of bounded linear operators, namely, compact, self-adjoint ones in Hilbert spaces. Our main result is a structural theorem stating that the eigenvectors of a compact, self-adjoint operators form a complete orthonormal set. This is an infinite dimensional generalization of the theorem of reduction to principal axes for self-adjoint matrices in linear algebra. Compact, self-adjoint operators come up naturally in differential and integral equations. In the last section we show how it is applied to the boundary value problems of second order ordinary differential equations.

6.1 Adjoint Operators

Let $T \in B(X_1, X_2)$ where X_1 and X_2 are Hilbert spaces over the same field. We construct a bounded linear operator called the adjoint of T, T^*, from X_2 to X_1 as follows. For any $y \in X_2$, the map $x \mapsto \langle Tx, y\rangle_{X_2}$ is linear and bounded, and hence defines an element in X_1'. By self-duality there exists a unique x^* in X_1 such that $\langle Tx, y\rangle_{X_2} = \langle x, x^*\rangle_{X_1}$. We define the **adjoint** of T to be the map $T^*y = x^*$. Then

$$\langle Tx, y\rangle_{X_2} = \langle x, T^*y\rangle_{X_1}, \text{ for all } x, y, \tag{6.1}$$

holds. We shall drop the subscripts in the inner products.

Proposition 6.1. *Let T be in $B(X_1, X_2)$ where X_1 and X_2 are Hilbert spaces. Then*

(1) $(T^*)^* = T$,

(2) $T^* \in B(X_2, X_1)$, and

(3) $\|T^*\| = \|T\|$.

Proof. (1) is straightforward from definition. Next we verify linearity. For any y_1, y_2 and scalars α and β, by (6.1),

$$\begin{aligned}
\langle x, T^*(\alpha y_1 + \beta y_2)\rangle &= \langle Tx, \alpha y_1 + \beta y_2\rangle \\
&= \overline{\alpha}\langle Tx, y_1\rangle + \overline{\beta}\langle Tx, y_2\rangle \\
&= \overline{\alpha}\langle x, T^*y_1\rangle + \overline{\beta}\langle x, T^*y_2\rangle \\
&= \langle x, \alpha T^*y_1 + \beta T^*y_2\rangle,
\end{aligned}$$

so $T^*(\alpha y_1 + \beta y_2) = \alpha T^*y_1 + \beta T^*y_2$.

Finally, by self-duality,

$$\|T^*y\| = \sup_{x\neq 0}\frac{|\langle x, T^*y\rangle|}{\|x\|} = \sup_{x\neq 0}\frac{|\langle Tx, y\rangle|}{\|x\|} \le \|T\|\|y\|,$$

so $\|T^*\| \le \|T\|$. The reverse inequality follows from (1). □

Other elementary properties of T^* are contained in the following proposition, whose proof is left to you.

Proposition 6.2. *Let T, T_1, and $T_2 \in B(X_1, X_2)$ where X_1 and X_2 are Hilbert spaces.*

(1) $(\alpha T_1 + \beta T_2)^* = \bar{\alpha}T_1^* + \bar{\beta}T_2^*$,

(2) $(ST)^* = T^*S^*$ *where* $S \in B(X_2, X_3)$ *and* X_3 *is a Hilbert space,*

(3) $(T^{-1})^* = (T^*)^{-1}$ *if* $T \in B(X_1)$ *is invertible.*

Consider $T \in L(\mathbb{F}^n, \mathbb{F}^m)$ where $\{e_1, \cdots, e_n\}$ and $\{f_1, \cdots, f_m\}$ denote the canonical bases in $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Then $Te_j = \sum_1^m a_{kj}f_k$ where (a_{kj}) is the matrix associated with T, and $T^*f_k = \sum_1^n b_{jk}e_j$ where (b_{jk}) is the matrix associated with T^*. Letting $x = \sum_1^n \alpha_j e_j$ and $y = \sum_1^m \beta_k f_k$, then

$$\langle Tx, y\rangle = \langle \sum \alpha_j Te_j, \sum \beta_k f_k\rangle = \sum \alpha_j \overline{\beta_k} a_{kj},$$

$$\langle x, T^*y\rangle = \langle \sum \alpha_j e_j, \sum \beta_k T^* f_k\rangle = \sum \alpha_j \overline{\beta_k b_{jk}},$$

From $\langle Tx, y\rangle = \langle x, T^*y\rangle$ we conclude

$$(b_{ij}) = \overline{(a_{ji})}.$$

So the matrix associated with T^* is the adjoint matrix of the matrix associated with T. This justifies the terminology of the adjoint of a linear operator.

Let X be a Hilbert space. A bounded linear operator on X to itself is called **self-adjoint** if $T^* = T$. For $T \in B(\mathbb{F}^n)$ its associated matrix satisfies $(a_{jk}) = (\overline{a_{kj}})$. That is to say, it is a self-adjoint matrix. When the scalar field is real, the matrix is called symmetric. In some texts, the terminology a "symmetric operator" is used instead of a "self-adjoint operator", and a self-adjoint operator is reserved for a densely defined unbounded operator whose adjoint is equal to itself. We never touch upon unbounded operators, so this definition will not come up; there is no chance to mess things up.

A basic property of a self-adjoint operator is that its eigenvalues must be real. Recall that λ is an eigenvalue of a linear operator T if there exists a non-zero vector x satisfying $Tx = \lambda x$. The eigenspace $\Phi_\lambda = \{x \in X : Tx = \lambda x\}$ forms a subspace of X and it is closed when T is bounded.

Proposition 6.3. *Let $T \in B(X)$ be self-adjoint where X is a Hilbert space.*

(1) All eigenvalues of T are real; and

(2) Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof. (1). If x is an eigenvector for the eigenvalue λ, then

$$\langle Tx, x\rangle = \langle \lambda x, x\rangle = \lambda\langle x, x\rangle,$$

$$\langle x, Tx\rangle = \langle x, \lambda x\rangle = \bar{\lambda}\langle x, x\rangle.$$

By self-adjointness, $\lambda\langle x, x\rangle = \bar{\lambda}\langle x, x\rangle$ which implies λ is real.

(2). Let $Tx_1 = \lambda_1 x_1$ and $Tx_2 = \lambda_2 x_2$, where λ_1 and λ_2 are distinct. We have

$$\langle Tx_1, x_2\rangle = \lambda_1\langle x_1, x_2\rangle,$$

$$\langle x_1, Tx_2\rangle = \lambda_2\langle x_1, x_2\rangle.$$

By self-adjointness, $\lambda_1\langle x_1, x_2\rangle = \lambda_2\langle x_1, x_2\rangle$ and, as the eigenvalues are distinct, $\langle x_1, x_2\rangle = 0$. □

Proposition 6.4. *Let $T \in B(X)$ be self-adjoint where X is a Hilbert space. Then*

$$\|T\| = \sup\{|\langle Tx, x\rangle| : x \in X, \|x\| = 1.\}$$

Proof. Denote the right hand side of the above formula by M. As

$$|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2,$$

taking supremum over all unit x shows that $M \leq \|T\|$.

On the other hand, from $\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$ we know that $\langle Tx, x\rangle \in \mathbb{R}$, for all x. By a direct expansion

$$\begin{aligned}\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle &= 2\langle Tx, y\rangle + 2\langle Ty, x\rangle \\ &= 4Re\langle Tx, y\rangle,\end{aligned}$$

because T is self-adjoint. As a result,

$$\begin{aligned}Re\langle Tx, y\rangle &= \frac{1}{4}(\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle) \\ &\leq \frac{M}{4}(\|x+y\|^2 + \|x-y\|^2) \\ &= \frac{M}{2}(\|x\|^2 + \|y\|^2),\end{aligned}$$

by the parallelogram identity. Taking x, $\|x\| = 1$, and $y = Tx/\|Tx\|(\neq 0)$,

$$\|Tx\| = Re\langle Tx, y\rangle \leq \frac{M}{2}(1+1) = M,$$

whence $\|T\| = M$. □

Remark 6.1 We shall use the following remarks in next section:

(a) $\sup\{|\langle Tx, x\rangle| : \|x\| = 1\}$ may be expressed as

$$\max\{\sup_{\|x\|=1} \langle Tx, x\rangle, -\inf_{\|x\|=1} \langle Tx, x\rangle\}.$$

(b) From this proposition, we know that $T \equiv 0$ if $\langle Tx, x\rangle = 0$ for all x.

6.2 Compact, Self-Adjoint Operators

A linear operator $T \in L(X_1, X_2)$ where X_1 and X_2 are normed spaces is called **compact** if whenever $\{x_n\}$, $\|x_k\| \leq M$ for some M, $\{Tx_k\}$ has a convergent subsequence. In other words, the image of a bounded sequence under a compact operator has the Bolzano-Weierstrass property. A compact operator is necessarily bounded. It is like a "regulator" which produces finite dimensional behavior.

All compact operators form a closed subspace of $B(X)$ where X is a Banach space under the operator norm. Furthermore, it is a two-sided ideal in the sense that ST and TS are compact if T is compact and S is bounded. The transpose (or the adjoint when the space is Hilbert) of a compact operator is again compact. It is a good exercise to prove all these facts.

A common class of compact operators is provided by integral operators. Letting $K \in C([a,b]^2)$ and considering the operator

$$\mathcal{I}f(x) = \int_a^b K(x,y)f(y)dy,$$

we saw in Chapter 4 that $\mathcal{I}$ is bounded on $C[a,b]$ as well as $L^p((a,b))$, $p \in [1,\infty)$. (Forgive me for abusing the same notation.) We claim that it is compact on any one of these spaces. Let's take it to be $L^p((a,b))$, $p \geq 1$. Let $\{f_j\}$ be a sequence in $L^p((a,b))$, $\|f_j\|_p \leq M$, say. By the definition of L^p-space, we can find $g_j \in C([a,b])$ such that $\|f_j - g_j\|_p < 1/j$. Then

$$\mathcal{I}g_j(x) = \int_a^b K(x,y)g_j(y)dy$$

makes sense. As K is uniformly continuous, for any $\varepsilon > 0$ there exists δ such that $|K(x,y)-K(x',y')| < \varepsilon$ for $\sqrt{(x-x')^2+(y-y')^2} < \delta$. So,

$$\begin{aligned} \left|\mathcal{I}g_j(x) - \mathcal{I}g_j(x')\right| &\leq \int_a^b \left|K(x,y) - K(x',y)\right||g_j(y)|dy \\ &\leq \varepsilon(b-a)^{1/q}\|g_j\|_p \\ &\leq \varepsilon(b-a)^{1/q}(1+M), \end{aligned}$$

where q is conjugate to p. We conclude that $\{\mathcal{I}g_j\}$ is equicontinuous in $[a,b]$. Similarly we can show that it is also uniformly bounded. Hence by Arzela-Ascoli theorem there exists $\{\mathcal{I}g_{j_k}\}$ converging uniformly to some $h \in C[a,b]$. As uniform convergence is stronger than L^p-convergence, we have

$$\|\mathcal{I}f_{j_k} - h\|_p \leq \|\mathcal{I}f_{j_k} - \mathcal{I}g_{j_k}\|_p + \|\mathcal{I}g_{j_k} - h\|_p \to 0,$$

as $k \to \infty$, so $\mathcal{I}$ is compact.

Another subclass of compact operators is provided by operators of finite rank. A bounded linear operator T is an **operator of finite rank** if its image is a finite dimensional subspace. Since $\{Tx_k\}$ is a bounded subset in a finite dimensional space whenever $\{x_k\}$ is bounded, clearly the Bolzano-Weierstrass property holds for it. In practise most compact operators are limits of operators of finite rank. For operators on Hilbert spaces, this can be established without much difficulty. For many years it was conjectured that this be true on Banach spaces, but now people have found sophisticated counterexamples even in a separable Banach space.

Here we consider linear operators in a Hilbert space to itself which is self-adjoint and compact simultaneously. The study of self-adjoint, compact operators was due to Hilbert and is an early success of functional analysis. There is a lot of information one can retrieve.

Proposition 6.5. *Let T be compact, self-adjoint in $B(X)$ where X is a Hilbert space X. Then*

(1) For any non-zero eigenvalue λ, the eigenspace of λ, Φ_λ, is a finite dimensional subspace.

(2) If eigenvalues $\{\lambda_k\}$, where all λ_k's are all distinct, converges to λ^*, then $\lambda^* = 0$.

Proof. The following proof works for (1) and (2). Assume on the contrary that there are infinitely many distinct eigenvectors. Let λ_j be a sequence of eigenvalues of T, $\lambda_j \to \lambda^* \neq 0$ and $Tx_j = \lambda_j x_j$, $\|x_j\| = 1$ where $\{x_j\}$ forms an orthonormal set. According to Proposition 6.3 (b), $\|x_j - x_k\| = \sqrt{2}$. On the other hand, by compactness there exists $Tx_{j_k} \to x_0$. That is to say, $\lambda_{j_k}x_{j_k} \to x_0$. By assumption, $\lambda_{j_k} \to \lambda^*$. It follows that $x_{j_k} = \frac{1}{\lambda_{j_k}}\lambda_{j_k}x_{j_k} \to x_0/\lambda^*$. So $\{x_{j_k}\}$ is a Cauchy sequence and $\|x_{j_k} - x_{j_l}\| \to 0$ as $j_k \neq j_l \to \infty$, contradicting $\|x_{j_k} - x_{j_l}\| = \sqrt{2}$. □

Lemma 6.6. *Let T be compact, self-adjoint in $B(X)$ where X is a Hilbert space X. Then*

$$M = \sup_{x\neq 0} \frac{\langle Tx, x\rangle}{\|x\|^2},$$

is an eigenvalue of T provided it is positive. Similarly,

$$m = \inf_{x \neq 0} \frac{\langle Tx, x \rangle}{\|x\|^2}$$

is an eigenvalue of T provided it is negative.

Proof. It suffices to consider the first case. Let $\{x_k\}$ be a sequence satisfying $\|x_k\| = 1$ and $\langle Tx_k, x_k \rangle \to M$. By compactness, there exists $Tx_{k_j} \to x_0$ in X.

Consider the self-adjoint operator $T - mI$. By Remark (a) after Proposition 6.4, $\|T - mI\| = \max\{M - m, m - m\} = M - m$. We have

$$\begin{aligned}
\|Tx_k - Mx_k\|^2 &= \|(T - mI)x_k - (M - m)x_k\|^2 \\
&= \|(T - mI)x_k\|^2 + (M - m)^2\|x_k\|^2 - \langle (T - mI)x_k, (M - m)x_k \rangle \\
&\quad -\langle (M - m)x_k, (T - mI)x_k \rangle \\
&\leq \|T - mI\|^2 + (M - m)^2 - 2(M - m)(\langle Tx_k, x_k \rangle - m) \\
&= 2(M - m)^2 - 2(M - m)(\langle Tx_k, x_k \rangle - m) \\
&\to 0
\end{aligned}$$

as $k \to \infty$. Restricting to the subsequence $\{x_{k_j}\}$ to get

$$Tx_{k_j} - Mx_{k_j} \to 0.$$

As $Tx_{k_j} \to x_0$, $x_{k_j} \to x_0/M$, $x_0 \neq 0$, and so by continuity $Tx_{k_j} \to Tx_0/M$. We conclude that x_0 is an eigenvector to to the eigenvalue M. □

The following is the main result of this chapter. It is an infinite dimensional version of the reduction to principal axes for a self-adjoint matrix. It is also called the spectral theorem for compact, self-adjoint operators.

Theorem 6.7. *Let T be compact, self-adjoint in $B(X)$ where X is Hilbert space.*

(1) Suppose $\langle Tx, x \rangle > 0$ for some $x \in X$. Then

$$\lambda_1 = \sup_{x \neq 0} \frac{\langle Tx, x \rangle}{\|x\|^2}$$

is a positive eigenvalue of T.

(2) Recursively define, for $n \geq 2$,

$$\lambda_n = \sup\{\frac{\langle Tx, x \rangle}{\|x\|^2} : x \neq 0, x \perp \langle x_1, \cdots, x_{n-1} \rangle\}$$

where x_j satisfies $Tx_j = \lambda_j x_j$, $\|x_j\| = 1$, $j = 1, 2, \cdots, n-1$. Then λ_n is a positive eigenvalue of T as long as the supremum is positive. The collection is finite when there exists some N such that

$$\langle Tx, x \rangle \leq 0, \quad \textit{for all } x \perp \langle x_1, \cdots, x_N \rangle.$$

Otherwise, there are infinitely many λ_j's and $\lambda_1 \geq \lambda_2 \geq \cdots \to 0$.

(3) For any "eigenpair" (λ, z) where $\lambda > 0$, λ must equal to λ_j for some j and z belong to the subspace spanned by all x_j. (We note that to the same λ_j there could be more than one corresponding eigenvectors by the above construction.)

(4) Similarly, all negative eigenvalues are given by

$$\lambda_1' = \inf_{x\neq 0} \frac{\langle Tx, x\rangle}{\|x\|^2},$$

and, for $n \geq 2$,

$$\lambda_n' = \inf\{\frac{\langle Tx, x\rangle}{\|x\|^2} : x \neq 0, x \perp \langle x_1', \cdots, x_{n-1}'\rangle\}$$

if $\langle Tx, x\rangle < 0$ for some x. Here x_j' is the normalized eigenvector of λ_j'.

(5) Let $\langle x_k, x_k'\rangle$ be the span of all normalized eigenvectors. Then

$$X = \overline{\langle x_k, x_k'\rangle} \oplus X_0,$$

where X_0 is the zero-eigenspace of T.

Proof. (1) follows directly from Lemma 6.6.

(2) Consider the closed subspace $X_1 = \langle x_1\rangle^\perp$. We check that $T : X_1 \mapsto X_1$. For, if $x \perp x_1$, then $0 = \langle Tx, x_1\rangle = \langle x, Tx_1\rangle = \lambda_1\langle x, x_1\rangle$, so $Tx \perp x_1$. It is routine to check that $T : X_1 \mapsto X_1$ is still a compact, self-adjoint operator. By applying Lemma 6.6 again

$$\lambda_2 = \sup\{\frac{\langle Tx, x\rangle}{\|x\|^2} : x \neq 0, x \in X_1\}$$

is an eigenvalue provided the supremum is positive. We may repeat this process to obtain the other eigenvalues until there exists an N such that $\langle Tx, x\rangle \leq 0$, for all $x \perp \langle x_1, \cdots, x_N\rangle$. Otherwise, we have an infinite sequence of decreasing eigenvalues. By Proposition 6.5, this sequence must converge to zero.

(3) Suppose λ is a positive eigenvalue with eigenvector $\widetilde{x}$. If $(\lambda, \widetilde{x})$ does not come from the above construction, we must have $\lambda \leq \lambda_1$ and there exists some n such that $\lambda \in (\lambda_{n+1}, \lambda_n]$ or $(0, \lambda_N]$ (when $N < \infty$). Consider the former first. When λ is not equal to λ_n, $\widetilde{x}$ is orthogonal to all x_n by Proposition 6.3 (b). However, by the construction of λ_{n+1}, we have

$$\lambda = \frac{\langle T\widetilde{x}, \widetilde{x}\rangle}{\|\widetilde{x}\|^2} \leq \sup\{\frac{\langle Tx, x\rangle}{\|x\|^2} : x \neq 0, x \perp \langle x_1, \cdots, x_n\rangle\} = \lambda_{n+1},$$

contradiction holds. When $\lambda = \lambda_n$, let us assume $\lambda_{n-K} > \lambda_{n-K+1} = \lambda_{n-K+2} = \cdots = \lambda_n > \lambda_{n+1}$ because of finite multiplicity. The modified vector $\overline{x} \equiv \widetilde{x} - P\widetilde{x}$ where P is the orthogonal projection of $\widetilde{x}$ to the subspace spanned by $\{x_{n-K+1}, \cdots, x_n\}$ is orthogonal to all $\{x_1, x_2, \cdots, x_n\}$ and still satisfies $T\overline{x} = \lambda\overline{x}$. Without loss of generality we may assume $\overline{x} = \widetilde{x}$. Then the above argument still produces a contradiction. A similar argument works for $\lambda \in (0, \lambda_N]$.

(4) The proof is left to the reader.

(5) By the above construction we see that for all x in $Z \equiv \overline{\langle x_k, x_k'\rangle}^\perp$, $\langle Tx, x\rangle = 0$. It is readily checked that T maps Z to itself. By Remark (b) after Proposition 6.4 we conclude that so $T \equiv 0$ on Z. In other words, $\overline{\langle x_k, x_k'\rangle}^\perp$ is the 0-eigenspace. By the theorem on orthogonal decomposition

$$X = \overline{\langle x_k, x_k'\rangle} \oplus \overline{\langle x_k, x_k'\rangle}^\perp = \overline{\langle x_k, x_k'\rangle} \oplus X_0.$$

□

This theorem may be viewed as the statement: Any matrix representation of a compact, self-adjoint operator can be diagonalized by a "rotation". Let us take the field to be real and consider T a symmetric linear transformation on the Euclidean space $\mathbb{R}^n$. For any orthonormal basis $\{x_1, \cdots, x_n\}$, the matrix $A \equiv (a_{jk})$, $Tx_k = \sum_j a_{jk}x_j$, is the matrix representation of T with respect to this basis. From linear algebra we know there exists an orthogonal matrix R such that R^*AR is equal to a diagonal matrix Λ. (An orthogonal matrix R satisfies $R^*R = I$ by definition.) Letting $y = Rx$, the matrix representation of T with respect to the new orthonormal basis $\{y_1, \cdots, y_n\}$ is the matrix $\Lambda \equiv (\lambda_j \delta_{jk})$ where y_j is the eigenvector of the eigenvalue λ_j. Now, for a compact, symmetric operator on an infinite dimensional Hilbert space the same thing happens. Let us assume for simplicity that zero is not an eigenvalue. Then for any complete orthonormal set $\{x_j\}$ the operator T is represented by an infinite matrix $A \equiv (a_{jk}), j, k \geq 1$, defined similarly as above. Let $\{\lambda_1, \lambda_2, \ldots,\}$ be an ordering of all eigenvalues of T according to $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \ldots$ and z_j the corresponding (orthonormal) eigenvectors. According to the theorem, $\{z_j\}$ forms a complete orthonormal set and the mapping defined by $z_j = \sum_k r_{kj} x_k$ defines an "orthogonal matrix" $R \equiv (r_{jk})$ which satisfies $R^*R = I$. We have $R^*AR = \Lambda$ where Λ is the diagonal matrix consisting of eigenvalues.

Consider the set $\{x : \sum a_{jk} x_j x_k = 1\}$ where (a_{jk}) is positive definite. The discussion above shows that in the new coordinates given by y_j's, as a result of a rotation of x_j's, this set becomes $\{y : \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 = 1,\}$, an ellipsoid in principal axes. In the infinite dimensional setting one may still call the eigenvectors z_j the principal axes of T, and the theorem guarantees such reduction to principal axes by a rotation is always possible.

6.3 An Application

It is a basic fact in Fourier series that the sine and cosine functions form a complete orthonormal basis in the L^2-space of 2π-periodic functions. For L^2-functions on the interval $[0, \pi]$ the functions $\sqrt{\pi/2} \sin kx\ k \geq 1$, form a complete orthonormal set. By considering the eigenvalue problem for an ordinary differential equation we shall show there are many complete orthonormal sets like the sine family.

Consider the eigenvalue problem for the second order ordinary differential equation

$$-y'' + q(x)y = \mu y, \quad y(a) = y(b) = 0, \tag{6.2}$$

where the "potential" q is a given continuous function on $[a, b]$. We want to find μ such that this problem has a non-trivial solution y. Then μ is an eigenvalue and y an "eigenfunction" corresponding to μ. We denote the differential operator $-d^2/dx^2 + q(x)$ by L.

First of all, look at the simplest case $q \equiv 0$. (6.1) becomes

$$\begin{cases} -y'' = \mu y \\ y(0) = y(\pi) = 0. \end{cases}$$

We have taken the interval to be $[0, \pi]$. By a scaling this can be achieved always. By solving the problem directly all eigenvalues are found; they are given by $j^2, j \geq 1$, and $\varphi_j = \sqrt{\frac{2}{\pi}} \sin jx$ is the corresponding normalized eigenfunction. All eigenvalues are simple (multiplicity one).

We want to extend the result to a general $q(x)$. To take advantage of the theorem on reduction to principal axes, we first convert the problem into one for an integral operator since L is unbounded. In general, differential operators behave badly but integral operators usually have some regularizing effect. We shall see shortly how nice the "inverse operator" of L will be.

To simplify our discussion, we observe that for any constant C_0, the eigenvalue problem

$$-y'' + (q(x) + C_0)y = \mu' y, \quad y(a) = y(b) = 0,$$

is equivalent to (6.2) in the sense that μ is an eigenvalue of (6.2) if and only if $\mu' = \mu + C_0$ above. By choosing C_0 a large number such $q(x) + C_0 > 0$, $\forall x$, and replace q in (6.2) by $q + C_0$ we may assume that q is positive on $[a, b]$. An advantage of this assumption is that all eigenvalues of (6.2) must be positive. To see this, letting (μ, ϕ) be an eigen-pair of (6.2) where ϕ is, say, positive somewhere, there is some x_0 at which ϕ attains maximum. Then $\phi''(x_0) \leq 0$, whence

$$\mu\phi(x_0) = -\phi''(x_0) + q(x_0)\phi(x_0) > 0,$$

implies $\mu > 0$.

In the following we assume q is positive so all eigenvalues of (6.2) are positive.

The convert of L into an integral operator is realized by using the Green's function. Although a bit obscure at first sight, this method is widely used in theory of differential equations. Pick any two non-trivial solutions $h_a(x)$ and $h_b(x)$ of

$$-y'' + q(x)y = 0,$$

which satisfies $h_a(b) = h_b(a) = 0$. These two solutions must be linearly independent. For if there exists some C such that $h_a = Ch_b$, then $h_a(a) = Ch_b(a) = 0$ means that $h_a(a) = h_a(b) = 0$ and h_a is a non-trivial eigenfunction for $\mu = 0$, contradicting to our assumption on q.

As h_a and h_b are linearly independent, from (6.2) we know that its Wronskian $W(x) = (h_a h_b' - h_a' h_b)(x)$ is equal to a non-zero constant c.

Define the Green's operator $\mathscr{G}$ by

$$(\mathscr{G}f)(x) = \int_a^b G(x,y)f(y)dy, \quad f \in C[a,b]$$

where the Green's function G is

$$G(x,y) = \begin{cases} \frac{1}{c}h_a(x)h_b(y), & x \geq y \\ \frac{1}{c}h_a(y)h_b(x), & x \leq y \end{cases}$$

It is readily check that $G \in C([a,b] \times [a,b])$ and $G(x,y) = G(y,x)$. The first fact shows that $\mathscr{G}$ is an integral operator with a continuous kernel. According to the general discussion in Section 4.2 $\mathscr{G}$ extends to be a bounded linear operator $\overline{\mathscr{G}}$ on $L^2((a,b))$. Moreover, from Section 6.2 we know that it is compact. Here we would like to use the spectral theorem to solve the eigenvalue problem (6.2) for the differential operator L. We need to derive more properties of $\mathscr{G}$. Let $E = \{\phi \in C^1([a,b]) : \phi(a) = \phi(b) = 0\}$. The following proposition clarifies in what sense $\mathscr{G}$ is the inverse of L.

Proposition 6.8. *(a)* $\mathscr{G} : C([a,b]) \to C^2([a,b]) \cap E$. *Moreover,* $L\mathscr{G}f = f$ *for every* $f \in C([a,b])$.

(b) $L : C^2([a,b]) \cap E \to C([a,b])$. *Moreover,* $\mathscr{G}L\phi = \phi$ *for every* $\phi \in C^2([a,b]) \cap E$.

Proof. We write

$$(\mathscr{G}f)(x) = \int_a^x \frac{1}{c}h_a(x)h_b(y)f(y)dy + \int_x^b \frac{1}{c}h_a(y)h_b(x)f(y)dy.$$

Clearly,

$$(\mathscr{G}f)(a) = (\mathscr{G}f)(b) = 0.$$

We compute

$$\begin{aligned}
(\mathscr{G}f)'(x) &= \int_a^x \frac{1}{c}h_a'(x)h_b(y)f(y)dy + \frac{1}{c}h_a(x)h_b(x)f(x) \\
&\quad + \int_a^x \frac{1}{c}h_a(y)h_b'(x)f(y)dy - \frac{1}{c}h_a(x)h_b(x)f(x) \\
&= \int_a^x \frac{1}{c}h_a'(x)h_b(y)f(y)dy + \int_a^x \frac{1}{c}h_a(y)h_b'(x)f(y)dy.
\end{aligned}$$

$$\begin{aligned}
(\mathscr{G}f)''(x) &= \int_a^x \frac{1}{c}h_a''(x)h_b(y)f(y)dy + \frac{1}{c}h_a'(x)h_b(x)f(x) \\
&\quad + \int_a^x \frac{1}{c}h_a(y)h_b''(x)f(y)dy - \frac{1}{c}h_a(x)h_b'(x)f(x) \\
&= q(x)(\mathscr{G}f)(x) - f(x),
\end{aligned}$$

and (a) follows. (b) can be proved by a similar computation. We leave it as an exercise. □

The next proposition describes a regularizing effect of the Green's operator.

Proposition 6.9. *The operator* $\overline{\mathscr{G}} \in B(L^2(a,b), E)$*, that is, there exists a constant* C *such that*

$$\|\overline{\mathscr{G}}f\|_\infty + \|(\overline{\mathscr{G}}f)'\|_\infty \ \le C\|f\|_2.$$

By a density argument it suffices to establish this estimate for $\mathscr{G}$ and f from $C([a,b])$. But then it follows more or less directly from the above formulas for the first and second derivatives of $\mathscr{G}f$. For instance,

$$\begin{aligned} |\mathscr{G}f(x)| &\le \sup_{x,y}|G(x,y)|\int_a^b |f(y)|dy \\ &\le \sqrt{b-a}\sup_{x,y}|G(x,y)|\sqrt{\int_a^b |f|^2} \\ &\equiv C_1\|f\|_2. \end{aligned}$$

Similarly, we obtain $\|(\mathscr{G}f)'\|_\infty \le C_2\|f\|_2$.

Proposition 6.10. $\overline{\mathscr{G}}$ *is self-adjoint on* $L^2((a,b))$.

Proof. For $f,g \in C[a,b]$,

$$\begin{aligned} \langle \overline{\mathscr{G}}f, g\rangle = \langle \mathscr{G}f, g\rangle &= \int_a^b (\int_a^b G(x,y)f(y)dy)g(x)dx \\ &= \int_a^b (\int_a^b G(x,y)g(x)dx)f(y)dy \quad \text{(Fubini's theorem)} \\ &= \int_a^b (\int_a^b G(y,x)g(x)dx)f(y)dy \quad \text{(By } G(x,y)=G(y,x)) \\ &= \langle \mathscr{G}g, f\rangle = \langle f, \mathscr{G}g\rangle \\ &= \langle f, \overline{\mathscr{G}}g\rangle. \end{aligned}$$

Now, for $f \in L^2(a,b)$ and $g \in C[a,b]$, pick $f_n \to f$ in $L^2(a,b)$ where $f_n \in C[a,b]$. We have

$$\lim_{n\to\infty}\langle \overline{\mathscr{G}}f_n, g\rangle = \langle \overline{\mathscr{G}}f, g\rangle$$

and

$$\lim_{n\to\infty}\langle f_n, \overline{\mathscr{G}}g\rangle = \langle f, \overline{\mathscr{G}}g\rangle.$$

It follows that $\langle \overline{\mathscr{G}}f, g\rangle = \langle f, \overline{\mathscr{G}}g\rangle$, for all $f \in L^2(a,b)$ and $g \in C[a,b]$. Next, let $g_n \to g \in L^2(a,b)$,

$$\langle \overline{\mathscr{G}}f, g\rangle = \lim_{n\to\infty}\langle \overline{\mathscr{G}}f, g_n\rangle = \lim_{n\to\infty}\langle f, \overline{\mathscr{G}}g_n\rangle = \langle f, \overline{\mathscr{G}}g\rangle,$$

so $\overline{\mathscr{G}}$ is self-adjoint. □

Proposition 6.11. $N(\overline{\mathscr{G}}) = \{0\}$.

Proof. Suppose $\overline{\mathscr{G}}f = 0$ we want to show that $f = 0$. Pick $f_n \in C([a,b])$ so that $f_n \to f$ in L^2-norm. By Proposition 6.9,

$$\|\mathscr{G}f_n - \overline{\mathscr{G}}f\|_\infty \le C_1\|f_n - f\|_2 \to 0.$$

$\overline{\mathscr{G}}f = 0$ implies $\mathscr{G}f_n \to 0$ in sup-norm. For any smooth function φ vanishing at a and b,

$$\begin{aligned}\langle f_n, \varphi\rangle &= \int_a^b f_n(x)\varphi(x)dx \\ &= \int_a^b L\mathscr{G}f_n\varphi(x)dx \\ &= \int_a^b \left(-\frac{d^2}{dx^2} + q(x)\right)\mathscr{G}f_n(x)\varphi(x)dx \\ &= \int_a^b \mathscr{G}f_n(x)(L\varphi)(x)dx \quad \text{(after twice integration by parts)} \\ &\to 0\end{aligned}$$

as $\|\mathscr{G}f_n\|_\infty \to 0$. If follows that for all smooth φ vanishing at a and b,

$$\langle f, \varphi\rangle = 0.$$

As every continuous function can be approximated in L^2-norm by smooth functions vanishing at endpoints, and any L^2-function can be approximated by continuous functions in L^2-norm, this relation implies that

$$\langle f, g\rangle = 0, \quad \text{for all } g \in L^2(a, b).$$

In particular, taking $g = f$ we get $f = 0$. □

Now, we can appeal to Theorem 6.7 to conclude $\{\varphi_j, \varphi_j'\}_{1,1}^{N,N'}, N, N' \le \infty$, with

$$\overline{\mathscr{G}}\varphi_j = \lambda_j\varphi_j, \ \lambda_j > 0$$

$$\overline{\mathscr{G}}\varphi_j' = \lambda_j\varphi_j', \ \lambda_j' < 0$$

and $L^2(a, b) = \overline{\langle\varphi_j, \varphi_j'\rangle}$. By Proposition 6.11 these eigenfunctions form a complete orthonormal set of $L^2(a, b)$.

Theorem 6.12. *Let $q \in C[a, b]$ in (6.2). The eigenvalue problem (6.2) has infinitely many positive eigenvalues $\mu_j \to \infty$ and at most finitely many negative eigenvalues. Each eigenvalue has simple multiplicity. The normalized eigenfunctions φ_j, φ_j' belong to $C^2[a, b]$, satisfying $\varphi_j(a) = \varphi_j(b) = \varphi_j'(a) = \varphi_j'(b) = 0$. They form a complete orthonormal set in $L^2(a, b)$.*

Proof. Let us assume q is positive and so all eigenvalues are positive. First of all, we claim that (μ, ϕ) is an eigen-pair for L if and only if $(\lambda, \phi), \lambda = \mu^{-1}$ is an eigen-pair for $\overline{\mathscr{G}}$. In fact, from Proposition 6.8 (b) we know that $L\phi = \mu\phi$ implies $\mathscr{G}\phi = \mu^{-1}\phi$. On the other hand, let $\phi \in L^2((a, b))$ satisfy $\overline{\mathscr{G}}\phi = \lambda\phi$. By Proposition 6.11 $\lambda \ne 0$. We need to show $\phi \in C^2([a, b]) \cap E$. Using Proposition 6.9 and $\phi = \lambda^{-1}\overline{\mathscr{G}}\phi \in C^1([a, b]) \cap E$ we know in particular that ϕ is continuous, hence $\overline{\mathscr{G}}\phi = \mathscr{G}\phi \in C^2([a, b])$ and $L\phi = \lambda^{-1}\phi$ by Proposition 6.8 (a).

Next, if $\phi_j, \ j = 1, 2$, are eigenfunctions of (6.2) for the same eigenvalue μ. We can find a suitable constant C such that $\phi_1'(a) + C\phi_2'(a) = 0$. It follows that the function $\phi_3 = \phi_1 + C\phi_2$ satisfies (6.2) as well as $\phi_3(a) = \phi_3'(a) = 0$. From the uniqueness of the initial value problem for differential equations we conclude that $\phi_3 \equiv 0$. Hence ϕ_1 and ϕ_2 are linear dependent. Therefore each eigenvalue is simple.

The proof of the theorem is completed.

□

We point out that a large part of this theorem remains valid for compact operators on Banach spaces. We refer to chapter 21 in [Lax] for the so-called Riesz or Riesz-Schauder theory.

The eigenvalue problem (6.2) is a special case of the Sturm-Liouville eigenvalue problem, an important topic in the theory of ordinary differential equations. The interested reader may consult [Dunford-Schwarz], or “Theory of Ordinary Differential Equations” by Coddington and Levinson.

Exercise

1. Let $T \in B(X)$.

   (a) Show that $T_1 = \frac{1}{2}(T + T^*)$ and $T_2 = \frac{1}{2i}(T - T^*)$ are self-adjoint.

   (b) Show that if $T = S_1 + iS_2$ where S_1 and S_2 are self-adjoint, then $T = T_1 + iT_2$.

2. Let $T \in B(X)$. Show that $\mathrm{Re}\langle Tx, x\rangle = 0$ implies $T + T^* = 0$

3. Under the identification of the dual space of a Hilbert space with itself by the Fréchet-Riesz theorem, show that the transpose of $T \in B(Y', X')$ becomes the adjoint $T^* \in B(Y, X)$. Note: The identification is sesquilinear.

4. Let $T \in L(X)$ be a compact operator. Show that

   (a) T is bounded,

   (b) for any $S \in B(X)$, TS and ST are compact operators; and

   (c) all compact operators form a closed set in $B(X)$. Hint: Use a diagonal sequence.

5. Show that the following bounded linear operators are not compact:

   (a) $Sf(x) = f(x+1),\ f \in L^2(\mathbb{R})$, and

   (b) $Tf(x) = xf(x),\ f \in L^2(\mathbb{R})$, and

   (c) $Lx = \sum_j a_{jk}x_k,\ x = (x_1, x_2, \cdots) \in \ell^2$ (over reals) where $a_{jk} = a_{kj},\ \sum_j a_{jk}^2 = 1$ for each j and $\sum_j a_{jk}a_{jm} = 0$ for distinct k, m.

6. Let $T \in B(X)$ be of finite rank.

   (a) Show that T is compact.

   (b) Show that the adjoint of T is also of finite rank.

7. Prove that for any compact $T \in B(X)$, there exist $T_n \in B(X), n \geq 1$, of finite rank such that $\|T_n - T\| \to 0$. Hint: Use the fact that the closure of $R(T)$ is separable when T is compact. Then use Problems 3(c) and 6. It is not true that every compact operator on a Banach space can be approximated by operators of finite rank, though.

8. Show that the adjoint of a compact operator is again compact.

9. Let $S \in B(\ell^2)$ and write $(Sx)_i = \sum \alpha_{ij}x_j$.

   (a) Show that S^* is given by $(S^*x)_i = \sum \overline{\alpha_{ji}}x_j$.

   (b) Under $\sum_{i,j} |\alpha_{ij}|^2 < \infty$, show that S is a compact operator.

10. Let $T \in B(\ell^2)$ and write $(Tx)_i = \sum a_{ij}x_j$. Suppose that except $a_j \equiv a_{jj},\ b_j \equiv a_{j\ j+1},\ c_j \equiv a_{j+1\ j}$ all other entries vanish. Show that T is compact if and only if $a_j, b_j, c_j \to 0$ as $j \to \infty$.

11. Let T be a compact, self-adjoint linear operator on X and $\lambda \neq 0$. Establish the “Fredholm alternative”: The equation

$$(Tx - \lambda x) = y,$$

   is solvable if and only if y is orthogonal to all solutions z of

$$(T - \lambda)z = 0.$$

   Hint: Use the theorem on reduction to principal axes. Is it still true when $\lambda = 0$?

12. Let T be a compact, self-adjoint linear operator on X and let

$$R(x) = \frac{\langle Tx, x\rangle}{\|x\|^2},\ x \neq 0$$

be its Rayleigh quotient. Let $\lambda_n,\ n \geq 1$, be the positive eigenvalues of T ordered in decreasing order. Show that

$$\lambda_n = \max_{E_n} \min_{x \in E_n/\{0\}} R(x).$$

This is called the Fischer's principle for eigenvalues. Formulate a corresponding principle for negative eigenvalues.

我最憐君中宵舞，道男兒到死心似鐵。
看試手，補天裂。
辛棄疾《賀新郎》

Chapter 7

Weak Compactness

An essential difference between finite and infinite dimensional normed spaces is that the closed unit ball is compact in the former but not compact in the latter. To compensate the loss of compactness in an infinite dimensional space, one may impose additional conditions to sustain compactness. A complete answer is known for the space of continuous functions under the sup-norm, see the discussion on Ascoli-Arzela theorem in Chapter 2. Yet there is a more radical way of thinking, namely, we search for a weaker concept of compactness which an infinite dimensional closed unit ball satisfies. This leads us to the study of both weakly sequential and weak compactness. A weak topology contains less open sets than the topology induced by the norm (the strong topology), so the chance of obtaining compact sets is higher. In the first section of this chapter, we discuss weak sequential convergence and prove the widely used result that the closed unit ball is weakly sequentially compact in a reflexive space. To study the problem in a general normed space, new concepts of weak and weak* topologies are introduced. In Section 2 we discuss some basic properties of the topologies induced by a family of linear functionals on a vector space. By specifying these families to X' on X and X on X' (through the canonical identification), we obtain the weak and weak* topologies on the spaces X and X' respectively. In Section 3 we prove the central results in this chapter, namely, Alaoglu theorem and a theorem characterizing reflexive spaces by weak compactness. We conclude this chapter by a discussion on extreme points in a convex set and a proof of Krein-Milman theorem, a cornerstone in functional analysis.

For simplicity we will take the scalar field to be real. The reader should have no difficulty in extending the results to the complex field.

In Sections 2-4 knowledge of point set topology is assumed.

7.1 Weak Sequential Compactness

Let $(X, \|\cdot\|)$ be a normed space and X' its dual. A sequence $\{x_n\}$ in X is called **weakly convergent** to some $x \in X$ if for every $\Lambda \in X'$,

$$\Lambda x_n \to \Lambda x, \quad \text{as } n \to \infty.$$

Denote it by $x_n \rightharpoonup x$.

We will call the convergence of a sequence "strong convergence" in contrast to weak convergence. The following proposition clarifies the relationship between these two notions of convergence.

Proposition 7.1. *Let $(X, \|\cdot\|)$ be a normed space and $\{x_n\} \subset X$.*

(a) $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ implies $x = y$.

(b) $x_n \to x$ implies that $x_n \rightharpoonup x$.

(c) If $x_n \rightharpoonup x$, then $\|x_n\| \le C, \quad \forall n$ for some constant C.

(d) If $x_n \rightharpoonup x$, then $\|x\| \le \underline{\lim}_{n\to\infty} \|x_n\|$.

(e) If $x_n \rightharpoonup x$, then x belongs to the closure of the convex hull of $\{x_n\}$.

Note that (d) can be deduced from (e) in this proposition. However, a short, direct proof is preferred.

Proof. (a) From $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ we deduce that $\Lambda(x-y)=0$ for all $\Lambda \in X'$. From Corollary 3.11, $x=y$.

(b) $\{x_n\}$ converges to x strongly means that $\|x_n - x\| \to 0$ as $n \to \infty$. For $\Lambda \in X'$,

$$|\Lambda x_n - \Lambda x| \le \|\Lambda\| \|x_n - x\| \to 0,$$

so $\{x_n\}$ converges to x weakly.

(c) Since for each $\Lambda \in X'$, $\Lambda x_n \to \Lambda x$, so Λx_n is bounded. The conclusion follows immediately from the uniform boundedness principle. You should note that X' is a Banach space.

(d) Pick $\Lambda_1 \in X'$ satisfying $\Lambda_1 x = \|x\|$ and $\|\Lambda_1\| = 1$, that is, it is a dual point of x. For any convergent subsequence of $\{\|x_n\|\}$, $\{\|x_{n_j}\|\}$, we have

$$\begin{aligned} \|x\| &= |\Lambda_1 x| \\ &\le |\Lambda_1(x - x_{n_j})| + |\Lambda_1 x_{n_j}| \\ &\le |\Lambda_1(x - x_{n_j})| + \|x_{n_j}\| \\ &\to \lim_{j\to\infty} \|x_{n_j}\|, \end{aligned}$$

whence (d) follows.

(e) Let K be the closure of the convex hull of $\{x_n\}$. If, on the contrary, x does not belong to K, by the separation form of the Hahn-Banach theorem in Chapter 2, there exist some $\Lambda \in X'$ and α such that

$$\Lambda x < \alpha < \Lambda y, \ \forall y \in K.$$

In particular, taking $y = x_n$ and letting $n \to \infty$, we have

$$\Lambda x < \alpha \le \lim_{n\to\infty} \Lambda x_n = \Lambda x,$$

contradiction holds. □

Proposition 7.1(b) shows that strong convergence implies weak convergence. When X is of finite dimension, every element is of the form $x = \sum \alpha_j z_j$ after a basis $\{z_1, \cdots, z_n\}$ has been chosen. Consider the n many linear functionals given by $\Lambda_j(x) = \alpha_j, j = 1, \cdots, n$. When $x_k \rightharpoonup x$ where $x_k = \sum \alpha_j^k z_j$ and $x = \sum \alpha_j z_j$, we have $\Lambda_j(x_k) = \alpha_j^k \to \Lambda_j(x) = \alpha_j$. It shows that $x_k \to x$, that is, weak convergence also implies strong convergence. So they are equivalent when the space is finite dimensional. However, for infinite dimensional spaces this is rare. There are plenty weakly sequentially convergent sequences which are not strongly convergent. Let us look at two examples.

Example 7.1 Consider ℓ^p-space, $1 < p < \infty$ and $\{e_j\} \subset \ell^p$ where e_j's are the "canonical vectors". It is clear that $\{e_j\}$ does not have any convergent subsequence as $\|e_i - e_j\|_p = 2^{1/p}$ for distinct i and j. On the other hand, we claim that $e_j \rightharpoonup 0$. To see this, recall that any bounded linear functional Λ on ℓ^p can be identified with

$$\Lambda x = \sum_{j=1}^{\infty} y_j x_j, \ x = (x_1, x_2, \cdots, x_n, \cdots),$$

where $y = (y_1, y_2, \cdots, y_n, \cdots)$, $\sum |y_j|^q < \infty$, by ℓ^p-ℓ^q duality. We have $|\Lambda e_j| = |y_j|$. As $\sum |y_j|^q < \infty$, $|y_j| \to 0$, that is, $e_j \rightharpoonup 0$ as $j \to \infty$.

Example 7.2 Consider $\{f_n\}$, $f_n(x) = \sin nx$, in $L^2(0,1)$. By a direct calculation, we have

$$\int_0^1 |f_n - f_m|^2 = 1 + O(\frac{1}{n}) + O(\frac{1}{m}), \quad \text{as } n, m \to \infty,$$

which means that this sequence does not converge in $L^2(0,1)$. Nevertheless, let us show that it is weakly convergent to zero. First, we claim that

$$\int_0^1 x^m \sin nx dx \to 0, \quad \text{as } n \to \infty,$$

for every monomial x^m. Indeed, this follows easily from the formulas

$$\int_0^1 x^m \sin nx dx = -\frac{\cos n}{n} + \frac{m}{n}\int_0^1 x^{m-1}\cos nx dx,$$

and

$$\int_0^1 \sin nx dx = \frac{1-\cos n}{n}.$$

Consequently,

$$\int_0^1 p(x)\sin nx dx \to 0, \quad \text{as } n \to \infty,$$

for every polynomial p. As all polynomials form a dense set in $L^2(0,1)$, a density argument shows that the above formula also holds when p is replaced by an L^2-function. By self-duality, we conclude that $\{\sin nx\}$ converges to 0 weakly in $L^2(0,1)$.

Later, we will see that in an infinite dimensional reflexive (Banach) space, divergent sequences which are weakly convergent always exist. However, in some exceptional cases things may behave differently. A result of Schur asserts that any weakly convergent sequence in ℓ^1 is also strongly convergent, see exercise.

The most important and useful result concerning weak sequential compactness is the following theorem.

Theorem 7.2. *Every closed ball in a reflexive space is weakly sequentially compact.*

A set E in $(X, \|\cdot\|)$ is **weakly sequentially compact** if every sequence in it contains a weakly convergent subsequence in E.

Proof. Without loss of generality we assume the ball is given by $\{x \in X : \|x\| \le 1\}$. Let $\{x_n\}$ be a sequence contained in this ball. We would like to extract a weakly convergent subsequence from it.

Let $Y = \overline{\langle x_n \rangle}$ be the closed subspace of X spanned by $\{x_n\}$. It is clear that Y is separable. As any closed subspace of a reflexive space is reflexive, Y is also reflexive. Recalling that a normed space is separable when its dual is separable, we conclude from the relation $(Y')' = Y$ and the separability of Y that Y' is also separable. Let S be a countable dense set in Y'. By extracting a diagonal sequence, we find a subsequence of $\{x_n\}, \{y_n\}$, such that

$$\lim_{n\to\infty} \Lambda y_n \text{ exists for every } \Lambda \in S. \tag{7.1}$$

For any $\Lambda \in Y'$, we can pick a sequence $\{\Lambda_j\}$ from S such that $\|\Lambda - \Lambda_j\| \to 0$ as $j \to \infty$. We claim that $\{\Lambda y_n\}$ is a Cauchy sequence in $\mathbb{R}$. For, taking any $\varepsilon > 0$, we fix j_0 such that $\|\Lambda - \Lambda_{j_0}\| < \varepsilon$. Then

$$\begin{aligned} |\Lambda y_n - \Lambda y_m| &\le |(\Lambda - \Lambda_{j_0})y_n| + |\Lambda_{j_0} y_n - \Lambda_{j_0} y_m| + |(\Lambda_{j_0} - \Lambda)y_m| \\ &< 2\varepsilon + |\Lambda_{j_0} y_n - \Lambda_{j_0} y_m|. \end{aligned}$$

By (7.1), there exists n_0 such that $|\Lambda_{j_0} y_n - \Lambda_{j_0} y_m| < \varepsilon$, for all $n, m \ge n_0$, so

$$|\Lambda y_n - \Lambda y_m| < 3\varepsilon,$$

that is, $\{\Lambda y_n\}$ is a Cauchy sequence. Define a real-valued function ℓ on Y' by

$$\ell(\Lambda) = \lim_{n\to\infty} \Lambda y_n.$$

It is readily checked that $\ell(\Lambda)$ is linear. Moreover, we have

$$\begin{aligned} |\ell\Lambda| &= \lim_{n\to\infty} |\Lambda y_n| \\ &\leq \|\Lambda\| \overline{\lim_{n\to\infty}} \|y_n\| \\ &\leq \|\Lambda\|, \end{aligned}$$

which means that $\ell \in Y''$.

By the reflexivity of Y, there exists some $y \in Y$ such that $\Lambda y = \ell(\Lambda)$. We conclude that $\Lambda y_n \to \Lambda y$ for every $\Lambda \in Y'$. Since each $\Lambda \in X'$ is a bounded functional on Y by restriction, $y_n \rightharpoonup y$. By Proposition 7.1 (c), $\|y\| \leq 1$. The proof is completed.

□

Corollary 7.3. *Let C be a non-empty convex set in a reflexive space X. It is weakly sequentially compact if and only if it is closed and bounded.*

Proof. As C is bounded, it is contained in some closed ball B. By Theorem 7.2, any sequence $\{x_n\}$ in C has a subsequence $\{x_{n_i}\}$ weakly converging to some $x \in B$. As C is closed and convex, $x \in C$ according to Proposition 7.1(d), so C is weakly sequentially compact.

Conversely, let $\{x_n\}$ be a sequence in C which converges to some x in X. We would like to show that x belongs to C, so that C is closed. In fact, as C is weakly sequentially compact, there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ which converges weakly to some y in C. By the uniqueness of limit, we conclude that x is equal to y, so it belongs to C. On the other hand, if there is some $\{x_n\} \in C$, $\|x_n\| \to \infty$, by weak sequential compactness there exists a weakly convergent subsequence x_{n_j}. However, by Proposition 7.1 (c), this subsequence is bounded, contradiction holds. Hence C must be bounded. □

Recall that in a finite dimensional normed space a set is sequentially compact if and only if it is closed and bounded. We have generalized it to convex sets in a reflexive space simply by replacing sequential compactness by weak sequential compactness.

Parallel to weak sequential convergence, we call a sequence $\{\Lambda_k\}$ in the dual space X' **weakly* sequentially convergent** to some Λ if $\Lambda_k x \to \Lambda x$ for every $x \in X$. Weak* sequential compactness for a set in X' can be defined correspondingly. We refer to the exercises for some properties of this notion.

We conclude this section with an application of weak sequential compactness. More applications can be found in exercises. We examine again the problem of best approximation. In Theorem 5.8 we showed that this problem always admits a unique solution in a Hilbert space. Now, we have

Theorem 7.4. *Let X be a reflexive space and C a nonempty closed, convex subset. Then for any $x \in X$, there exists $z \in C$ such that*

$$\|x - z\| = \inf\{\|x - y\| : y \in C\}.$$

In other words, the best approximation problem always has a solution in a reflexive space.

Proof. Let $\{y_n\}$ be a minimizing sequence of the problem, that is,

$$\|x - y_n\| \to d \equiv \inf\{\|x - y\| : \quad y \in C\}.$$

From

$$\|y_n\| \leq \|x - y_n\| + \|x\| \to d + \|x\|,$$

we see that $\{y_n\}$ is a bounded sequence in X. By Theorem 7.2, it contains a weakly convergent subsequence $\{y_{n_j}\}, y_{n_j} \rightharpoonup z$ for some z. Proposition 7.1.(d) asserts that $z \in C$. Moreover, by the same proposition we have

$$\|x - z\| \leq \varliminf_{j\to\infty} \|x - y_{n_j}\| = d,$$

so z is a point in C realizing the distance. □

7.2 Topologies Induced by Functionals

In this section we will first give a quick review on some basic topological concepts, especially those concerning the topology induced by a family of functions on a set. Next, we examine more closely about the case where the set is a vector space and the functions are linear functionals on this vector space.

Recall that (X, τ) where X is a set and τ is a collection of subsets of X is called a topological space if τ satisfies

(a) The empty set $\varnothing$ and X belong to τ,

(b) unions of elements in τ belongs to τ, and

(c) intersections of finitely many elements in τ belongs to τ.

Any element in τ is called an open set. A set F is closed if its complement is open. Immediately we deduce from (a), (b), and (c) the following facts:

(d) X and $\varnothing$ are closed sets,

(e) intersections of closed sets are closed sets, and

(f) unions of finitely many closed sets are closed sets.

For any subset E of X, its closure is defined to be

$$\overline{E} \equiv \bigcap \{F : \ F \text{ is a closed set containing } E\}.$$

Note that X is closed and it contains E. By (e) $\overline{E}$ is a closed set. Clearly, it is the smallest closed set containing E. A subset K is compact if every open covering of K has a finite subcover.

With open sets at hand, we can talk about convergence and continuity. For instance, a sequence $\{x_n\}$ in X is convergent to some x in X if for each open set G containing x, there exists some n_0 such that $x_n \in G$ for all $n \geq n_0$. A mapping $f : (X, \tau) \mapsto (Y, \sigma)$ between two topological spaces is **continuous** at x if $f^{-1}(G)$ is open for any open set G containing $f(x)$. It is continuous in a subset E if it is continuous at every x in E.

In a metric space (X, d), G is an open set if for every $x \in G$, there exists some metric ball $B_\rho(x) \subset G$. One can verify that the collection of all these open sets makes X into a topological space. This is the topology induced by the metric d. The notions of open set, closed set, the closure of a set, convergence of a sequence and continuity of functions all coincide with those previously defined for a metric space in Chapter 2.

However, caution must be made as many facts valid in a metric space are no longer true in a general topological space. For instance, a set in a metric space is closed if and only if the limit of any convergent sequence belongs to the set. In a general topological space, the "only if" part holds but the "if" part does not. Further, a set in a metric space is compact if and only if it is sequentially compact. This is not always true for a general topological space. There are compact topological spaces admitting sequences which do not have convergent sequences. On the other hand, there are non-compact topological spaces in which all sequences have convergent sequences. When it comes to continuity, a function f is continuous at x in a metric space if and only if for any sequence $\{x_n\}$ converging to x, $f(x_n)$ converges to $f(x)$. In a topological space, convergence of $f(x_n)$ to $f(x)$ for any $\{x_n\} \to x$ does not ensure continuity, although it holds when f is continuous at x. In a word, topological properties cannot be fully described in terms of sequences in a general topological space.

Now, we turn to the topology induced by functions on a set.

Let X be a non-empty set. For a non-empty collection of functions $\mathcal{F}$ from X to $\mathbb{R}$, we introduce a topology $\tau(X, \mathcal{F})$ on X by the following way. First, Let $\mathcal{U}_1$ be the collection of subsets of X of the form $f^{-1}(a, b)$ where $a, b \in \mathbb{R}$ and f in $\mathcal{F}$. (Define $\varnothing = f^{-1}(\varnothing)$.) Next, let $\mathcal{U}_2$ be the collection of all finite intersections of unions of elements from $\mathcal{U}_1$. Finally, let $\tau = \tau(X, \mathcal{F})$ contain all unions of elements from $\mathcal{U}_2$. One can verify that τ forms a topology on X. By this construction, each f in $\mathcal{F}$ is a continuous function in (X, τ). In fact, for any (X, τ_1) in which every function in $\mathcal{F}$ is continuous, τ_1 must contain τ. In this sense τ is the weakest topology to make each element of $\mathcal{F}$ continuous. We call it the **induced topology** by $\mathcal{F}$.

Intuitively speaking, the induced topology is finer (containing more open sets) if there are more functions in $\mathcal{F}$ and coarser (containing less open sets) if there are less functions in $\mathcal{F}$. A topological space (X, τ) is a Hausdorff space if for any two distinct points in X, there exist two disjoint open sets containing these points respectively. In analysis Hausdorff space is preferred for many of its nice properties. For instance, a compact set is closed in a Hausdorff space. A metric space is always Hausdorff, as any distinct x_1 and x_2 are contained in the disjoint open sets $\{z \in X : d(z, x_1) < 1/2d(x_1, x_2)\}$ and $\{z \in X : d(z, x_2) < 1/2d(x_1, x_2)\}$ respectively. To make an induced topology a Hausdorff one, $\mathcal{F}$ cannot contain too few functions. It is called **separating** if for any two distinct points x and y in X, there exists a function $f \in \mathcal{F}$ such that $f(x) \neq f(y)$.

Proposition 7.5. *The space $(X, \tau(X, \mathcal{F}))$ is a Hausdorff space if $\mathcal{F}$ is separating.*

Proof. For distinct x_1 and x_2, let $f \in \mathcal{F}$ satisfy $f(x_1) < \alpha < f(x_2)$ for some α, say. Then $G_1 \equiv \{x : f(x) < \alpha\}$ and $G_2 \equiv \{x : f(x) > \alpha\}$ are two disjoint open sets containing x_1 and x_2 respectively. □

So far, X has been taken to be a non-empty set without any extra structure and $\mathcal{F}$ is a set of real functions on X. Now, let us assume that X is a vector space and $\mathcal{F}$ a subset of $L(X, \mathbb{R})$, that is, it is composed of linear functionals. We would like to know more about the induced topology in this setting.

Proposition 7.6. *Consider the induced topology $\tau(X, \mathcal{F})$ where X is a vector space and $\mathcal{F} \subset L(X, \mathbb{R})$. Let G be a non-empty set in X. We have*

(a) G is open if and only if for each $x_0 \in G$, there exists U of the form

$$U = \{x : |\Lambda_j x| < \alpha,\ j = 1, \cdots, N\} \tag{7.2}$$

for some $\Lambda_j \in \mathcal{F}$ and $\alpha > 0$ such that $U + x_0 \subset G$.

(b) G is open if and only if $G + x_0$ is open for every $x_0 \in X$.

(c) G is open if and only if λG is open for every $\lambda \neq 0$.

From (b) and (c) we see that translations and multiplications by non-zero scalars are homeomorphisms with respect to $\tau(X, \mathcal{F})$.

Proof. (a) Let $x_0 \in G$. By the definition of $\tau(X, \mathscr{F})$ there exists a set of the form

$$V = \{x : \Lambda_j x \in (\alpha_j, \beta_j), j = 1, \cdots, N\}$$

containing x_0 in G. It follows that $x - x_0 \in \Lambda^{-1}((\alpha_j - \Lambda_j x_0, \beta_j - \Lambda_j x_0))$ for $x \in V$. So, the set in (7.2) by taking $\alpha = \min_j\{|\alpha_j - \Lambda_j x_0|, |\beta_j - \Lambda_j x_0|\}$ is an open set containing x_0 in G. The converse is trivial from definition.

(b) For $x \in G$, there exists some U as in (7.2) such that $U + x \subset G$. But then $U + (x + x_0) \subset G + x_0$.

(c) Argue as in (b). □

It is convenient to call a set of the form (7.2) a "τ-ball centered at 0" or simply a "τ-ball". Unlike a metric ball, a τ-ball is not only specified by its "radius" α (now a vector), but also the functionals Λ_j's.

Under the induced topology $\tau(X, \mathcal{F})$, every element in $\mathcal{F}$ is continuous. In a normed space we know that a linear functional is continuous if and only if it is bounded, see Proposition 3.2. Here we have a similar result.

Proposition 7.7. *Let Λ be a linear functional on $(X, \tau(X, \mathcal{F}))$.*

(a) Λ is continuous if and only if it is continuous at one point.

(b) Λ is continuous if and only if it is bounded on a τ-ball.

Proof. (a) Assume that Λ is continuous at x_0. For any (α, β) containing Λx_0, $\Lambda^{-1}(\alpha, \beta)$ has an open subset U containing x_0. Let (α', β') be any open interval containing Λx. Then $(\alpha, \beta) = (\alpha', \beta') + \Lambda x_0 - \Lambda x$ is an open interval containing Λx_0, using $\Lambda^{-1}(\alpha, \beta) = \Lambda^{-1}(\alpha', \beta') + x_0 - x$ and Proposition 7.6 (b), we see that $U - x_0 + x$ is an open subset of $\Lambda^{-1}(\alpha', \beta')$ containing x, so Λ is continuous at x.

(b) $\Lambda^{-1}(-1, 1)$ is open for a continuous Λ. As $0 \in \Lambda^{-1}(-1, 1)$, there is an open set of U the form (7.2) contained in Λ^{-1} by Proposition 7.6 (a). So $|\Lambda(U)| \leq 1$, and Λ is bounded on U. Conversely, if $|\Lambda(U)| \leq M$ for some constant M where U is a τ-ball. By (a) it suffices to show that Λ is continuous at 0, that is, $\Lambda^{-1}(a, b)$ is open for every $a, b, a < 0 < b$. Pick any $x_0 \in \Lambda^{-1}(a, b)$, there is an $\varepsilon > 0$ such that $(\Lambda x_0 - \varepsilon, \Lambda x_0 + \varepsilon) \subset (a, b)$. Letting $V = \frac{\varepsilon}{2M} U$, it is easy to see that $V + x_0$ is an open set containing x_0 and $V + x_0 \in \Lambda^{-1}(a, b)$, so $\Lambda^{-1}(a, b)$ is open. □

Under the topology $\tau(X, \mathcal{F})$, every element in $\mathcal{F}$ is continuous by definition. Are there more? Consider the very special case where $\mathcal{F}$ consists of a single functional Λ. Clearly, any constant multiple of Λ is continuous. Furthermore, one can show that the sum of two linear functionals from $\mathcal{F}$ is continuous. The following proposition asserts that these are the only cases.

Proposition 7.8. *Consider $(X, \tau(X, \mathcal{F}))$ where X is vector space and $\mathcal{F} \subset L(X, \mathbb{R})$. The collection of all continuous linear functionals is given by $\mathcal{F}$ if and only if $\mathcal{F}$ is a subspace of $L(X, \mathbb{R})$.*

Proof. We will only prove the "if " part and leave the "only if " part as exercise.

Let Λ be continuous in $\tau(X, \mathcal{F})$. There exists an open set

$$U = \{x : \Lambda_j x \in (-\alpha, \alpha), j = 1, \cdots, N\} \subset \Lambda^{-1}(-1, 1).$$

We claim that Λ vanishes on $\bigcap_{j=1}^N N(\Lambda_j)$. For, if z satisfies $\Lambda_j z = 0$, $j = 1, \cdots, N$, then $\Lambda_j(\lambda z) = 0$ for all λ, so $\lambda z \in U$. From

$$|\lambda||\Lambda z| = |\Lambda(\lambda z)| \leq 1$$

that $\Lambda z = 0$ after letting $|\lambda\|$ go to infinity. By the lemma below, Λ is a linear combination of Λ_j, so $\Lambda \in \mathcal{F}$ by assumption. □

Lemma 7.9. *Let $\Lambda, \Lambda_1, \cdots, \Lambda_n$ be in $L(X, \mathbb{R})$. If $\Lambda x = 0$ whenever $x \in \bigcap_{j=1}^n N(\Lambda_j)$, then Λ is a linear combination of $\Lambda_1, \cdots,$ and Λ_n.*

Proof. Let $Z = \{(\Lambda x, \Lambda_1 x, \cdots, \Lambda_n x) : \ x \in X\}$. Clearly Z is a subspace of $\mathbb{R}^{n+1}$ and it is proper because the point $(1, 0, \cdots, 0)$ does not belong to it by assumption. We can find a hyperplane $az + a_1 z_1 + \cdots + a_n z_n = 0$ which contains Z but not $(1, 0, \cdots, 0)$. In other words,

$$a\Lambda x + a_1 \Lambda_1 x + \cdots + a_n \Lambda_n x = 0,$$

for all $x \in X$ and $a1 + a_1 0 + \cdots + a_n 0 \neq 0$. The second expression shows that a is non-zero, so the desired conclusion follows from the first expression. □

In Chapter 3 we discussed the separation theorem as a consequence of Hahn-Banach theorem. Now we establish a separation theorem in induced topology. It will be our main tool in later development. We start with a lemma.

Lemma 7.10. *Let C be an open, convex set in $(X, \tau(X, \mathcal{F}))$ containing 0 and p its gauge. Then*

$$C = \{x : \ p(x) < 1\}.$$

Recall that the gauge of a convex set is given by

$$p(x) = \inf\{\mu > 0 : \ \frac{1}{\mu}\, x \in C\},$$

and $p(x) = \infty$ if no such μ exists. It is a positive homogeneous, subadditive function. When C is open and contains 0, it contains some τ-ball. Therefore, for every $x \in X$, we can find some small $\varepsilon > 0$ so that εx belongs to this τ-ball and hence C, so $p(x)$ is always finite.

Proof. We claim that $\{p < 1\} \subset C$ for any convex set C (not necessarily open) containing 0. Indeed, if $p(x) < 1$ for some x, then there exists some $\mu \in (0, 1)$ such that $\mu^{-1}x \in C$. By convexity $x = (1-\mu)0 + \mu(\mu^{-1}x) \in C$.

To show the inclusion from the other direction, we observe for each x in the open C, we can find a τ-ball such that $U + x \subset C$. From the definition of U, there exists some small $\varepsilon > 0$ such that $\varepsilon x \in U$. Thus, $x + \varepsilon x \in C$ and it implies that $p(x) \le 1/(1+\varepsilon) < 1$, the desired conclusion follows. □

Theorem 7.11. *Let A and B be two disjoint, non-empty convex sets in $(X, \tau(X, \mathcal{F}))$ where X is a vector space and $\mathcal{F} \subset L(X, \mathbb{R})$.*

(a) When A is open, there exists a continuous linear functional Λ such that

$$\Lambda x < \Lambda y, \quad \textit{for all } x \in A, \ y \in B.$$

(b) When A is compact and B is closed, there exist a continuous linear functional Λ, α and β such that

$$\Lambda x < \alpha < \beta < \Lambda y, \quad \textit{for all } x \in A, \ y \in B.$$

Proof. (a) Consider the convex set $C = A - B + x_0$ where x_0 is a point in $B - A$. It is open because $C = \bigcup_{x \in B} A - x + x_0$ and A is open. Moreover, it contains the origin as x_0 is located outside C. Let p be the gauge of C. Define Λ_0 on the one-dimensional subspace $\langle x_0 \rangle$ by $\Lambda_0(\alpha x_0) = \alpha$. Then $\Lambda_0 \le p$ on this subspace. This is trivial when $\alpha \le 0$. When $\alpha > 0$, by Lemma 7.10 $p(\alpha x_0) = \alpha p(x_0) \ge \alpha$ as x_0 lies outside C. Appealing to the general Hahn-Banach theorem, we find an extension of Λ_0, $\Lambda \in L(X, \mathbb{R})$, satisfying $\Lambda \le p$ in X. For, $x \in A$ and $y \in B$,

$$\Lambda(x - y + x_0) \le p(x - y + x_0)$$

holds. It follows that $\Lambda x < \Lambda y$ after using $\Lambda x_0 = 1$ and Lemma 7.10.

We still have to show that Λ is continuous. We pick a τ-ball U in C. Noting that $x \in U$ implies $-x \in U$, we have $|\Lambda x| \le p(x) < 1$ in U, so Λ is continuous by Proposition 7.7(b).

(b) We use a compactness argument to show that there is an open set V such that $A + V$ is disjoint from B. For each $x \in A$, as $X \setminus B$ is open, there exists $V_x = \{y : |\Lambda_j y| \le 2\gamma_x, j = 1, \cdots, N\}, \gamma_x > 0$, so that $U_x \equiv V_x + x$ is disjoint from B. The collection of all open sets $\frac{1}{2}V_x + x$ forms an open cover of A. As A is compact, there is a finite subcover given by, say, finitely many $\frac{1}{2}V_{x_k} + x_k, \ k = 1, \cdots, m$. Taking

$$V = \{y : |\Lambda_j y| \le \gamma\}, \quad \text{some } \gamma > 0,$$

where the Λ_j's are taken from all those linear functionals appearing in the definition of V_{x_k}, one verifies that $A + V$ is an open, convex set disjoint from B.

By (a) there exists a continuous linear functional Λ satisfying $\Lambda x < \Lambda y$ for all $x \in A + V$ and $y \in B$. It is elementary to show that a non-zero linear functional is an open map, so $\Lambda(A + V)$ is an open set in $\mathbb{R}$. On the other hand, ΛA is compact as the image of a compact set by a continuous functional. So (b) holds for some α and β. □

7.3 Weak and Weak* Topologies

Let $(X, \|\cdot\|)$ be a normed space. The topology $\tau(X, X')$ is called the **weak topology** of X. This is the weakest topology to make every bounded linear functional continuous. As ensured by the Hahn-Banach theorem, there are sufficiently many elements in X' to separate points, the weak topology is Hausdorff. However, it contains much less open sets than the strong topology does when the space is infinite dimensional. In sharp contrast to norm topology, we have the following results.

Proposition 7.12. *Let X be an infinite dimensional normed space. Every weakly open set contains an infinite dimensional subspace of X.*

A set is weakly open means that it is open in $\tau(X, X')$. As a consequence, every non-empty weakly open set is unbounded in norm.

Proof. As every weakly open set contains a weak ball (that's, $\tau(X, X')$-ball) U, it suffices to prove the result for U. Consider the linear map from X to $\mathbb{R}^N$ given by $\Phi(x) = (\Lambda_1 x, \cdots, \Lambda_N x)$ where Λ_j's are the bounded linear functionals defining U. The kernel of Φ is of infinite dimension. For any $x \in N(\Phi)$, $\Lambda_j x = 0$ for all j, so $N(\Phi) \subset U$. □

A topological space is called **metrizable** if its topology is induced by some metric.

Proposition 7.13. *The weak topology is not metrizable when X is an infinite dimensional normed space.*

Proof. Assume the weak topology on X comes from a metric d. As the topology induced from a metric admits a countable local base given by $\{x : d(x, x_0) < 1/n\}$, $n \geq 1$, at every point x_0, in particular, there is a countable base at 0 consisting of weak balls $U_n = \{x : |\Lambda_j^n x| < \alpha_n\}$, where $j = 1, 2, \cdots, N(n)$, $n \geq 1$. All these Λ_j^n's form a countable set in X'. As X' is a Banach space and every Hamel basis of a Banach space is uncountable (Proposition 4.14), we can find some $T \in X'$ which is independent of all these Λ_j^n's. Consider the open set G given by $\{x : |Tx| < 1\}$. It must contain some U_{n_0}, so T vanishes on $\bigcap_j N(\Lambda_j^{n_0})$. However, by Lemma 7.9, T is a linear combination of $\Lambda_j^{n_0}$'s, contradiction holds. Hence the weak topology is not metrizable. □

Although the weak and norm topologies are very different as seen from the above propositions, they have something in common.

Proposition 7.14. *A convex set in a normed space X is weakly closed if and only if it is closed.*

Proof. Since the weak topology is weaker than the norm topology, any weakly open set is open in the norm topology, so any weakly closed set must be closed. Conversely, let C be closed and convex. For $x_0 \notin C$, by Theorem 7.11 there exist some $\Lambda \in X'$ and scalar α such that

$$\Lambda x_0 < \alpha < \Lambda y, \quad \forall y \in C.$$

Thus the open set $V = \{x : \Lambda x < \alpha\}$ is disjoint from C, so $X \setminus C$ is weakly open. □

Next, consider the dual space X' of a normed space X. We know that it is a Banach space under the operator norm. Furthermore, under the canonical identification X can be viewed as a subspace of X''.

The **weak* topology** on X' is given by $\tau(X', X)$. It is clearly Hausdorff. A local base at 0 consists of "weak* balls"

$$U = \{\Lambda : |\Lambda x_j| < \alpha,\ j = 1, \cdots, N\},$$

for some N and $\alpha > 0$.

The most important result in weak* topology is the following theorem.

Theorem 7.15 (**Alaoglu**). *The closed ball in X' is weakly* compact.*

Proof. Let P be the product space $\prod_{x\in X}[-\|x\|, \|x\|]$ endowed with the product topology. By Tychonoff theorem P is compact. We set up a mapping Φ from $\mathcal{B}$, the closed unit ball in X', to P by setting $\Phi(\Lambda) = p$ if and only if $\Lambda x = p_x$, where p_x is the projection of P to $[-\|x\|, \|x\|]$. By the definition of the product topology, its local base at p is given by sets of the form

$$\{q : \ |q_{x_j} - p_{x_j}| < \alpha,\ j = 1, \cdots, N\},$$

for some x_j's and $\alpha > 0$. By comparing the weak* balls in X' with this local base, we know that Φ is a homeomorphism from $(\mathcal{B}, \tau(X', X))$ to P. To establish the theorem it suffices to show that $\Phi(\mathcal{B})$ is closed in P, since a closed subset of a compact Hausdorff space is compact.

Let p be in the closure of $\Phi(\mathcal{B})$. We define $\Lambda x = p_x$. To show $p \in \Phi(\mathcal{B})$, we must prove that Λ is linear and $\|\Lambda\| \leq 1$.

First, we claim $\Lambda(x + y) = \Lambda x + \Lambda y$, that is, $p_{x+y} = p_x + p_y$. For, consider the open set V containing p given by $\{q : |q_x - p_x|, |q_y - p_y|, |q_{x+y} - p_{x+y}| < \alpha\}$. As p belongs to the closure of $\Phi(\mathcal{B})$, for each $\alpha > 0$, there exists some Λ_1 in $\mathcal{B}$ in V, that is, $|\Lambda_1 x - p_x|, |\Lambda_1 y - p_y|, |\Lambda_1(x + y) - p_{x+y}| < \alpha$. It follows that

$$\begin{aligned} |p_{x+y} - p_x - p_y| &\leq |p_{x+y} - \Lambda_1(x + y)| + |\Lambda_1 x + \Lambda_1 y - p_x - p_y| \\ &\leq |p_{x+y} - \Lambda_1(x + y)| + |\Lambda_1 x - p_x| + |\Lambda_1 y - p_y| \\ &< 3\alpha, \end{aligned}$$

which implies $p_{x+y} = p_x + p_y$. Similarly, one can show that $p_{\alpha x} = \alpha p_x$, so Λ is linear. Furthermore, from $|\Lambda x| = |p_x| \leq \|x\|$ we have $\|\Lambda\| \leq 1$, so $\Lambda \in \mathcal{B}$. The proof of this theorem is completed. □

Corollary 7.16. *A bounded set in X' is weakly* compact if and only if it is weakly* closed.*

Proof. Since the weak* topology is Hausdorff, any weakly* compact set must be weakly* closed. Conversely, every bounded set K is contained in some closed ball. By Alaoglu's theorem, any closed ball is weakly* compact, so K is also weakly* compact if it is weakly* closed. □

We have obtained a satisfactory compactness result for the dual space of a normed space. Yet the old question remains unanswered, namely, is the closed unit ball in a normed space weakly compact? Now we give an answer.

Theorem 7.17. *The closed unit ball in a normed space is weakly compact if and only if the space is reflexive.*

Proof. When X is reflexive, the closed unit ball of X, $\mathcal{B}$, is weakly compact by Alaoglu's theorem. Conversely, $\mathcal{B}$ is compact in $\tau(X, X')$ means $J(\mathcal{B})$ is compact and so is closed in $\tau(X'', X')$. Here J is the canonical identification of $\mathcal{B}$ in X''. By the lemma below, $J(\mathcal{B})$ coincides with $\mathcal{B}''$, the closed unit ball in X''. We conclude that the canonical identification is surjective and X is reflexive. The proof of this theorem is completed. □

Lemma 7.18. *The closed ball in a normed space X is dense in the closed unit ball in X'' under the $\tau(X'', X')$-topology.*

Proof. Suppose on the contrary there is some $p \in \mathcal{B}''$ disjoint from $\mathcal{C}$, the weakly* closure of $J(\mathcal{B})$. By Alaoglu theorem, $\mathcal{B}''$ is weakly* compact, so is weakly* closed. It implies $\mathcal{C} \subset \mathcal{B}''$. By Theorem 7.11(b) and Proposition 7.8, there exists some Λ_1 in X' such that, for some α and β,

$$q\Lambda_1 < \alpha < \beta < p\Lambda_1,$$

holds for all $q \in \mathcal{C}$. Taking q be $Jx,\ x \in \mathcal{B}$, we have $\Lambda_1 x = q\Lambda_1 < \alpha$ which implies $\|\Lambda_1\| \leq \alpha$ after taking supremum over all $x \in \mathcal{B}$. On the other hand, we have $\beta < |p\Lambda_1| \leq \|p\|\|\Lambda_1\| \leq \|\Lambda_1\|$, contradiction holds. □

We pointed out that weak and weak sequential compactness are two different concepts; neither one implies the other in a general topological space. However, a remarkable theorem of Eberlein-Šmulian asserts that every subset of a normed space is weakly compact if and only if it is weakly sequentially compact [DS]. In other words, these two concepts are equivalent in a normed space. Using this theorem, the converse of Theorem 7.2 holds. It is true that the closed ball is weakly sequentially compact if and only if the space is reflexive.

7.4 Extreme Points in Convex Sets

Let X be a vector space and $\mathcal{F}$ a subspace of $L(X, \mathbb{R})$. We will consider compact, convex sets in the topological space $(X, \tau(X, \mathcal{F}))$. A well-known result of Caratheodory states that any point in a compact, convex subset of $\mathbb{R}^n$ can be expressed as the linear combination of at most $n+1$ many extreme points. We would like to extend this theorem to infinite dimensional spaces.

Let E be a non-empty subset of X. A point x in E is called an **extreme point** if whenever it is expressed as $\lambda x_1 + (1-\lambda)x_2$ for some $x_1, x_2 \in E$ and $\lambda \in (0,1)$, x_1 and x_2 must equal to x itself. Vertices of a polygon and points on the boundary of a ball are examples of extreme points.

Let us examine some examples in infinite dimensional spaces.

Example 7.3 Let $\{e_j\}$ be the canonical vectors in ℓ^∞. It is easy to show that $\{\pm e_i\}_1^\infty$ form the set of all extreme points in the closed unit ball $\{x \in \ell^\infty : \ \|x\|_\infty \leq 1\}$. Furthermore, every x in this ball belongs to the closure of linear combinations of $\{e_j\}'$s.

Example 7.4 Let C_1 be $\{f \in C[0,1] : \ |f(x)| \leq 1, f(0) = f(1) = 0\}$. Clearly C_1 is a closed, convex set in $C[0,1]$. However, it has no extreme points. For, by continuity for every function f in C_1 there is some subinterval $[a,b]$ of $(0,1)$ on which $\alpha < f(x) < \beta,\ \alpha, \beta \in (-1,1)$. Fix a continuous function ϕ compactly supported in $[a,b]$. Then, for sufficiently small ε, $f \pm \varepsilon\phi$ belong to C_1. The relation $f = [(f+\varepsilon\phi) + (f - \varepsilon\phi)]/2$ shows that f is not an extreme point.

If instead we consider $C_2 = \{f \in C[0,1] : \ 0 \leq f(x) \leq 1, f(0) = f(1) = 0\}$, C_2 is again a closed, convex set. As in the previous case one can show that every non-zero function is not an extreme point. Thus, the only extreme point in C_2 is the zero function.

The following theorem provides the "right" generalization of Carathedory theorem to infinite dimensional spaces. It asserts that points in any compact, convex set can be approximated by linear combinations of its extreme points.

Theorem 7.19 (**Krein-Milman**). *Let X be either a normed space, or a vector space with topology induced by a separating subset $\mathcal{F}$ of $L(X, \mathbb{R})$. Then every non-empty compact, convex set K in X has extreme points. In fact, the closed convex hull of the extreme points is equal to K.*

Proof. Denote the extreme points of K by K_e and its closed convex hull by $\mathrm{cco}K_e$. Clearly $\mathrm{cco}K_e \subset K$. We need to show the inclusion from the other direction. Suppose this is not true. Letting $x_0 \in K/\mathrm{cco}K_e$, by strong separation theorem for normed spaces or by Theorem 7.11, there is a continuous linear functional Λ satisfying

$$\Lambda x < \alpha < \beta < \Lambda x_0, \quad \text{for all } x \in \mathrm{cco}K_e. \tag{7.3}$$

Consider the set $J = \{x : \Lambda x = M\}$, $M \equiv \max_{y\in K} \Lambda y$. From the above expression we know that J is disjoint from $\text{cco}K_e$. On the other hand, it is a non-empty, compact, convex subset in X. By the lemma below, it contains an extreme point z. We claim that z is also an extreme point in K. For, if $z = \lambda x_1 + (1-\lambda)x_2$ for some x_1 and x_2 in K, $M = \Lambda z = \lambda\Lambda x_1 + (1-\lambda)\Lambda x_2$ which forces $\Lambda x_1 = \Lambda x_2 = M$, hence x_1 and x_2 belong to J. As z is an extreme point in J, x_1 and x_2 must equal to z, so z belongs to K_e. By (7.3) $\Lambda z < M$, contradiction holds. We conclude that $K \subset \text{cco}K_e$. □

Lemma 7.20. *Setting as in Krein-Milman's theorem, there exists an extreme point in K.*

Proof. Let us call a subset E of K has "property X" if for every $x \in E$, $x = \lambda x_1 + (1-\lambda)x_2$, for some $x_1, x_2 \in K$, $\lambda \in (0,1)$, implies that $x_1, x_2 \in E$. Denote the collection of all closed, non-empty subsets of K satisfying property X by $\mathcal{E}$. It is non-empty as it contains K. $\mathcal{E}$ is endowed with the partial ordering, namely, $E_1 \leq E_2$ if and only if $E_2 \subset E_1$.

We would like to use Zorn's lemma to show that $\mathcal{E}$ contains a maximal element. Let $\mathcal{C}$ be a chain in $\mathcal{E}$. Consider the set

$$E^* = \bigcap_{E\in\mathcal{C}} E.$$

Being the intersection of closed sets, E^* is closed. As $\mathcal{F}$ is separating, $\tau(X,\mathcal{F})$ is Hausdorff. (Of course this is true for the norm topology.) Consequently, every closed subset of a compact set is again compact, so E^* is a non-empty compact subset of K. Furthermore, it is easy to verify that E^* has property X, so it is an upper bound of the chain. By Zorn's lemma, there exists a maximal element E_1 in $\mathcal{E}$. We claim that this maximal element consists of a single point, hence an extreme point in K. For, suppose on the contrary, there are two distinct points, x_1 and x_2, in E_1. We choose a continuous linear functional Λ satisfying, say, $\Lambda x_1 < \Lambda x_2$, and consider the set $E_2 = \{x \in E_1 : \Lambda x = \max_{y\in E_1} \Lambda y\}$. By the continuity of Λ and the compactness of E_1, this set is non-empty and closed. Moreover, it has property X by an argument used in the proof of Theorem 7.19 above. In this way we produce a proper subset of E_1 which belongs to $\mathcal{E}$, contradicting the maximality of E_1. We conclude that E_1 is a singleton. □

Corollary 7.21. *Any closed and bounded convex subset in a reflexive space is the closed convex hull of its extreme points.*

It is amazing that the extreme point in a convex set is an algebraic concept, yet its existence is established by exploring the topological structure of the space. The theorem of Carathedory stated in the beginning of this section gives more information about the relation between the point and the extreme points in $\mathbb{R}^n$. In the infinite dimensional setting one is led to Choquet theory, see [L] for details.

In these notes, we have encountered two types of topologies on a vector space, that is, those induced by a norm and those induced by a set of linear functionals. Yet there is one type we do not cover, namely, those with topologies induced by a metric (or by a sequence of norms). It includes the space of all infinitely differentiable functions as a special case. These three types of spaces can be unified under the concept of locally convex topological vector space. The reader is referred to [R2] for a systematic discussion.

Exercise 7

1. Let f_k be the piecewise function whose graph connects $(0,0),(1/n,1),(2/n,0)$ and $(1,0)$. Show that $\{f_k\}$ converges weakly but not strongly to 0 in $C[0,1]$.
2. Deduce Proposition 7.1 (d) from (e).
3. Let $\{x_k\}$ converge weakly to x in some normed space. Prove that there is a sequence $\{y_k\} \to x$ satisfying,
$$y_k = \sum_1^{N(k)} \alpha_j^k x_j, \quad \sum_j \alpha_j^k = 1, \quad \alpha_j^k \in [0,1],$$
and
$$j_0(k) = \min\{j : \ \alpha_j^k > 0\} \to \infty,$$
as $k \to \infty$. This is a sharpened version of Proposition 7.1 (e).
4. Show that a sequence $\{x^n\}$ in $\ell^p, 1 < p < \infty$, is weakly convergent to x if and only if for each k, $x_k^n \to x_k$ as $n \to \infty$.
5. Show that a weakly convergent sequence in ℓ^1 also converges strongly. This result is called Schur's theorem. Google for more about it.
6. Show that a weakly convergent sequence in $C[a,b]$ must converge pointwisely. Given an example to show that the converse may not hold.
7. Show that in a Hilbert space H, $\{x_n\} \to x$ if and only if (a) $x_n \rightharpoonup x$ and (b) $\|x_n\| \to \|x\|$.
8. Show that $\{f_n\} \rightharpoonup f$ in $C[a,b]$ if and only if (a) $\{f_n\} \to f$ pointwisely and (b) $\{\|f_n\|_\infty\}$ is uniformly bounded.
9. Show that $\{f_n\} \rightharpoonup f$ in $L^p(a,b), p \in (1,\infty)$, if and only if (a)
$$\int_a^x f_n(s)ds \to \int_a^x f(s)ds, \quad \forall x \in [a,b],$$
and (b) $\{\|f_n\|_p\}$ is uniformly bounded in p-norm.
10. Show that every bounded linear functional attains its minimum/maximum in a closed convex subset of a reflexive space. In fact, a theorem of James (1971) asserts that the converse is also true, namely, if every bounded linear functional attains its minimum/maximum over the closed unit ball of a Banach space, this space must be reflexive.
11. Let X be a normed space and $\{\Lambda_k\}$ a sequence in X' which is weakly* convergent to some Λ. Show that
(a) $\|\Lambda\| \le \underline{\lim}_{k\to\infty} \|\Lambda_k\|$,
(b) $\|\Lambda_k\| \le C$, for some constant C for all k, and
(c) $\Lambda_k x_k \to \Lambda x$ whenever $x_k \to x$ in X.
12. Recall that the dual space of $C[-1,1]$ is given by $V[-1,1]$ (see Section 3.5). Show that the sequence $\{\Lambda_k\}$ given by
$$\Lambda_k(f) \equiv \frac{k}{2}\int_{-1/k}^{1/k} f(x)dx,$$
converges weakly* but does not converge weakly to 0. Hint: Each Riemann integrable h function induces a bounded linear functional on $BV_0[-1,1]$ by
$$\int_{-1}^1 h(x)dg, \quad \forall g \in BV_0[-1,1].$$

13. Let X be a separable Banach space and C' the closed unit ball in X'. Prove the following theorem of Helly: Every sequence in C' contains a weakly* convergent subsequence. In other words, C' is weakly* sequentially compact.

14. Let $\mathcal{F}$ be a class of linear functionals defined on the vector space X. Show that for $f, g \in \mathcal{F}, \alpha, \beta \in \mathbb{F}$, $\alpha f + \beta g$ is continuous in $(X, \mathcal{F})$.

15. Show that in a finite dimensional normed space the weak topology always coincides with the strong one.

16. A sequence $\{x_n\}$ in a normed space X converges with respect to its weak topology if for every weakly open set G containing x, there exists some n_0 such that $x_n \in G$ for all $n \geq n_0$. Show that this holds if and only if $\{x_n\}$ converges weakly to x.

17. Show that the unit ball $B = \{x \in X : \ \|x\| < 1\}$ in a normed space X is not open with respect to its weak topology. In fact, its weak interior is empty.

18. Show that the weak closure of the sphere $S = \{x \in X : \ \|x\| = 1\}$ is the closed unit ball $\{x \in X : \|x\| \leq 1\}$.

19. Let T be a linear operator from normed space X to another normed space Y. Prove that it is bounded if and only if it is continuous with respect to the weak topologies. Suggestion: Apply the closed graph theorem in the "if" part.

20. Let C' be the closed unit ball in X' where X is a normed space. Show that C' is metrizable in weak* topology if and only if X is separable. Here "metrizable" means the weak* topology is induced by some metric. Suggestion: Let $\{z_n\}$ be a dense subset of X. Define

$$d(\Lambda_1, \Lambda_2) = \sum_{j=1} \frac{1}{2^j} \left|\Lambda_1 z_j - \Lambda_2 z_j\right|,$$

and show it is the desired metric.

21. Show that the closed ball under the weak topology in a separable Banach space is metrizable. The converse is also true, look up [DS] for a proof.

22. Let $\{f_n\}$ be a sequence of functions on $[a, b]$ satisfying $\|f_n\|_1 \leq M$ for some M. Show that there is a subsequence $\{f_{n_k}\}$ and $f \in L^1(a, b)$ such that

$$\int_a^b \varphi f_{n_k} \to \int_a^b \varphi f, \quad \forall \varphi \in C[a, b],$$

as $k \to \infty$. Hint: Every function in $L^1(a, b)$ may be viewed as a bounded linear functional on $C[a, b]$.

23. Show that for every $f, \ \|f\|_1 \leq 1$, in $L^1(a, b)$, there exist two distinct functions g and h, both with L^1-norms less than or equal to 1, such that $f = (g + h)/2$. This shows that the closed unit ball in $L^1(0, 1)$ does not have extreme point.

24. Find all extreme points of the closed unit ball in $C[a, b]$. There are very few.

25. Show that a Banach space cannot be the dual space of another Banach space if the extreme set of its closed unit ball is a finite set.

26. Let C the unit closed ball in $\mathcal{C}_0$, the space of sequences converging to 0. Show that it does not have any extreme point. Use this fact to deduce that $\mathcal{C}_0$ is not the dual space of any normed space.

27. Prove that $\mathcal{C}$ is linear isomorphic but not norm-preserving linear isomorphic to $\mathcal{C}_0$. Hint: For $x = (x_1, \cdots, x_n, \cdots) \in \mathcal{C}, \ x_n \to x^*$, define $Tx = y, \ y = (x^*, x_1 - x^*, \cdots, x_n - x^*, \cdots)$. Verify that this is a linear isomorphism. To establish the second assertion use the fact that $(1, 1, 1, \cdots)$ is an extreme point in the closed unit ball of $\mathcal{C}$.

蘭佩紫，菊簪黃，殷勤理舊狂。

晏幾道《阮郎歸》

Chapter 8

Nonlinear Operators

Finally we come to nonlinear functional analysis. In the past seven chapters we have been concerned with the properties of infinite dimensional spaces and linear operators including linear functionals between them. In some sense we were working on generalizations of linear algebra. In analysis one studies linear and nonlinear functions as well. Of course, there are far more nonlinear functions than linear functions. You just have to recall only degree one polynomials are linear and all polynomials of higher degree are nonlinear, let alone transcendental functions. In this chapter we give a very brief introduction to nonlinear functional analysis. The main theme is to extend results in calculus, especially in differentiation theory, to infinite dimensional settings. You will see that linearization plays a dominating role in the study.

There are three sections in this chapter. In the first section several fixed point theorems are discussed, starting from the contraction principle, Brouwer fixed-point theorem on finite dimensional space and ending on Schauder fixed point theorem. Their applications are illustrated by examples. In the next section we develop calculus on Banach space. Of crucial importance are the implicit and inverse function theorems which are proved by the contraction mapping principle. Finally, we discuss how to minimize a nonlinear functional over a subset in a Banach space. Recall that one valuable application of differentiation is to determine the critical points of a function. Similarly, in a function space a minimum of a certain functional is a critical point of this functional. This is a huge topic which has been split into different branches of mathematics such as the calculus of variations, optimization theory, control theory, etc. The reader may appreciate the use of convexity and weak topology in this context.

8.1 Fixed-Point Theorems

For a map from a set to itself, a fixed point of this map is an element in this set which is not moved by it. Many theoretical and practical problems can be formulated as problems of finding fixed points of certain maps. The general question of solving equation, symbolically written as $f(x) = 0$, is equivalent to solving $g(x) = x$, where $g(x) \equiv f(x) + x$. Consequently any root of f is a fixed point of g. In this section we will discuss three widely known fixed-point theorems, starting with the fixed-point theorem established by Banach in 1920. Since its discovery, this theorem remains as one of the most frequently used results in analysis.

The setting of Banach fixed-point theorem, or contraction mapping principle, is formulated on a complete metric space. Let (X, d) be a metric space. A map $T : (X, d) \to (X, d)$ is called a *contraction* if there exists some $\gamma \in (0, 1)$ such that

$$d\left(T(x), T(y)\right) \leqslant \gamma d(x, y)\ ,\ \forall x, y \in X\ .$$

It is clear that any contraction is necessarily continuous.

Theorem 8.1. *Every contraction on a complete metric space has a unique fixed point.*

Proof. Let (X, d) be a complete metric space. Pick any x_0 from X and define a sequence $\{x_n\}$ by iteration: $x_n = T^n(x_0)$, $n \geqslant 1$. We claim that $\{x_n\}$ is a Cauchy sequence. For, we have

$$\begin{aligned} d(x_n, x_m) &= d\left(T^n(x_0)\ ,\ T^{n-m}(x_0)\right) \\ &\leqslant \gamma d\left(T^{n-1}(x_0)\ , T^{m-1}(x_0)\right) \\ &\ \vdots \\ &\leqslant \gamma^m d\left(T^{n-m}(x_0)\ , x_0\right) \end{aligned}$$

for any $n, m,\ n > m$. On the other hand, for $l \geq 1$,

$$\begin{aligned} d\left(T^l(x_0), x_0\right) &\leqslant d\left(T^l(x_0), T^{l-1}(x_0)\right) + d\left(T^{l-1}(x_0), T^{l-2}(x_0)\right) + \cdots \\ &\quad + d\left(T(x_0), x_0\right) \\ &\leqslant \left(\gamma^{l-1} + \gamma^{l-2} + \cdots + 1\right) d\left(T(x_0), x_0\right) \\ &\leqslant \frac{d\left(T(x_0), x_0\right)}{1-\gamma}\ . \end{aligned}$$

Taking $l = n - m$, we have

$$d\left(x_n, x_m\right) \leqslant \frac{d\left(T(x_0), x_0\right)}{1-\gamma}\gamma^m \longrightarrow 0, \quad \text{as } m \to \infty,$$

hence $\{x_n\}$ is a Cauchy sequence. By completeness of the space, $z \equiv \lim_{n\to\infty} T^n(x_0)$ exists. By the continuity of T, $T(z) = T(\lim_{n\to\infty} T^n(x_0)) = \lim_{n\to\infty} T^{n+1}(x_0) = z$, in other words, z is a fixed point of T.

If w is another fixed point of T, then

$$0 \leqslant d(w, z) = d\left(T(w), T(z)\right) \leqslant \gamma d(w, z).$$

As $\gamma \in (0, 1)$, it forces $d(w, z) = 0$, i.e., $w = z$, so the fixed point is unique. □

It is worthwhile to note that the above proof provides a constructive way to find the fixed point. Starting from any initial point, the fixed point can be found as the limit of an iteration scheme. The contraction mapping principle has wide applications. You should have learned how it is used to establish the local solvability of the initial value problem of ordinary differential equations. Another standard application is the proof of the implicit function theorem. We shall, in the next section, shows that it can be used to prove the same theorem in the infinite dimensional setting.

Banach fixed-point theorem asserts the existence of fixed points for special maps (contractions) in a general space (a complete metric space). There are fixed-points theorems which hold for general maps in a special space. The Brouwer fixed-point theorem is the most famous one among them. It is concerned with continuous functions from the closed unit ball of the n-dimensional Euclidean space to itself. The complete statement was first proved by Brouwer in 1912 using homotopy, a newly invented topological concept. Over the years there are many different proofs and generalizations.

Let B be a closed ball in $\mathbb{R}^n$, $n \geqslant 1$.

Theorem 8.2. *Every continuous map from B to itself has a fixed point.*

This theorem is not valid when the closed ball is replaced by the open one. For instance, the map $T(x) = (1 + x)/2$ which maps $(0, 1)$ to itself is free of fixed points.

In the following we will take the ball to be the closed unit ball centered at the origin. We begin with a computational lemma.

Lemma 8.3. *Let f be twice continuously differentiable from B to B. Denote its Jacobian matrix by $J_f(x) = (\partial f^i/\partial x_j), i, j = 1, \cdots, n$. Let c_{ij} be its (i, j)-th cofactor. Then for each i,*

$$\sum_{j=1}^{n} \frac{\partial c_{ij}}{\partial x_j} = 0.$$

Proof. Without loss of generality take $i = n$. Let $\mathbf{g}^j$ be the $j-$th $(n-1)$-column vector

$$\mathbf{g}^j = \begin{bmatrix} \dfrac{\partial f^1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f^{n-1}}{\partial x_j} \end{bmatrix}.$$

We have, by the definition of the cofactor matrix,

$$c_{nj} = (-1)^{n+j} \det \left[\mathbf{g}^1, \cdots, \check{\mathbf{g}}^j, \cdots, \mathbf{g}^n\right],$$

where "v" means the j-th column $\mathbf{g}^j$ is removed. Note that $\left[\mathbf{g}^1, \cdots, \check{\mathbf{g}}^j, \cdots, \mathbf{g}^n\right]$ is an $(n-1) \times (n-1)$-matrix. By the rule of differentiation, we have

$$\begin{aligned} \frac{\partial c_{nj}}{\partial x_j} &= (-1)^{n+j} \sum_{k<j} \det \left[\mathbf{g}^1, \cdots, \frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^j, \cdots, \mathbf{g}^n\right] \\ &\quad + (-1)^{n+j} \sum_{k>j} \det \left[\mathbf{g}^1, \cdots, \check{\mathbf{g}}^j, \cdots, \frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \mathbf{g}^n\right]. \end{aligned}$$

Using the elementary properties of the determinant, we have

$$\begin{aligned} \frac{\partial c_{nj}}{\partial x_j} &= (-1)^{n+j} \sum_{k<j} (-1)^{k-1} \det \left[\frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^k, \cdots, \check{\mathbf{g}}^j, \cdots, \mathbf{g}^n\right] \\ &\quad + (-1)^{n+j} \sum_{k>j} (-1)^{k-2} \det \left[\frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^j, \cdots, \check{\mathbf{g}}^k, \cdots, \mathbf{g}^n\right]. \end{aligned}$$

Set σ_{kj} equal to 1 if $k < j$, to 0 if $k = j$ and to -1 if $k > j$. Then $\sigma_{jk} = -\sigma_{kj}$ and

$$\frac{\partial c_{nj}}{\partial x_j} = (-1)^n \sum_{k=1}^n (-1)^{j+k-1} \sigma_{kj} \ \det \left[\frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^k, \cdots, \check{\mathbf{g}}^j \cdots, \mathbf{g}^n\right].$$

So,

$$\begin{aligned} \sum_{j=1}^n \frac{\partial c_{nj}}{\partial x_j} &= (-1)^n \sum_{k,j} (-1)^{j+k-1} \sigma_{kj} \det \left[\frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^k, \cdots, \check{\mathbf{g}}^j, \cdots, \mathbf{g}^n\right]. \\ &= (-1)^n \sum_{k,j} (-1)^{k+j-1} \sigma_{jk} \det \left[\frac{\partial \mathbf{g}^j}{\partial x_k}, \cdots, \check{\mathbf{g}}^j, \cdots, \check{\mathbf{g}}^k, \cdots, \mathbf{g}^n\right] \\ &= (-1)^{n+1} \sum_{k,j} (-1)^{j+k-1} \sigma_{kj} \det \left[\frac{\partial \mathbf{g}^k}{\partial x_j}, \cdots, \check{\mathbf{g}}^k, \cdots, \mathbf{g}^j, \cdots, \mathbf{g}^n\right] \\ &= -\sum_{j=1}^n \frac{\partial c_{nj}}{\partial x_j}, \end{aligned}$$

after using $\partial \mathbf{g}^k / \partial x_j = \partial \mathbf{g}^j / \partial x_k$ in the last line, we are done. □

Proof of Theorem 8.2. Let us first prove the theorem assuming that $F : B \to B$ is twice continuously differentiable. Assume on the contrary F that does not have a fixed point, that's, $F(x) - x \neq 0$, $\forall x \in B$. For each x, consider the equation for λ,

$$\|x + \lambda(x - F(x))\| = 1 ,$$

where $\| \cdot \|$ is the Euclidean norm. This is a quadratic equation; indeed, by expanding it we have

$$\|x\|^2 + 2\langle x, x - F(x)\rangle\lambda + \|x - F(x)\|^2\lambda^2 = 1 ,$$

where $\langle \cdot, \cdot \rangle$ stands for the Euclidean inner product. There are two real roots given by

$$\lambda = \frac{\langle x, F(x) - x\rangle \pm \sqrt{\langle x, x - F(x)\rangle^2 - \|x - F(x)\|^2(\|x\|^2 - 1)}}{\|x - F(x)\|^2}$$

It is clear that the larger root $a(x)$, regarded as a function of x, is given by

$$a(x) = \frac{\langle x, F(x) - x\rangle + \sqrt{\langle x, x - F(x)\rangle^2 + (1 - \|x\|^2)\|x - F(x)\|^2}}{\|x - F(x)\|^2}.$$

It is readily checked that a is continuously differentiable in B and vanishes on ∂B, the boundary of B. (You should note that $\langle x, F(x) - x\rangle < 0$ by the characterization of equality sign in Cauchy-Schwarz inequality).

Now, consider the one-parameter maps on B to itself given by

$$F(x, \lambda) = x + \lambda a(x)\big(x - F(x)\big).$$

We have $F(x, 0) = x$ and $F(x, 1) \in \partial B$. Consider the integral

$$I_\lambda = \int_B \det J_F(x)dx.$$

It is helpful to keep in mind that this integral gives the volume of the set $F(B)$ in $\mathbb{R}^n$ in view of the formula of change of variables. We claim that

$$\frac{\partial I_\lambda}{\partial \lambda} = 0. \tag{8.1}$$

For,

$$\begin{aligned}
\frac{\partial I_\lambda}{\partial \lambda} &= \frac{\partial}{\partial \lambda}\int_B \det J_F dx \\
&= \int_B \sum_{i,j} c_{ij}\frac{\partial^2 F^i}{\partial\lambda\partial x_j}dx \\
&= \int_B \sum \frac{\partial}{\partial x_j}\left(c_{ij}\frac{\partial F^i}{\partial \lambda}\right)dx \qquad \text{(Lemma 8.3)} \\
&= \int_{\partial B} c_{ij}\frac{\partial F^i}{\partial \lambda}\nu_j dx. \qquad \text{(by the divergence theorem)}
\end{aligned}$$

Recall that the divergence theorem asserts that for any vector field $\mathbf{v} = (v_1, \cdots, v_n)$ in the domain Ω,

$$\int_\Omega \sum_i \frac{\partial v_j}{\partial x_j} = \int_{\partial\Omega} \sum_j v_i\nu_j ds,$$

where $\nu = (\nu_1, \cdots, \nu_n)$ is the unit outer normal at $\partial\Omega$, the boundary of Ω. Since $\partial F/\partial\lambda = a(x)(x - F(x))$ vanishes on ∂B, (8.1) follows.

From (8.1) we conclude that I_λ is a constant. In particular, $I_1 = I_0$. Since $F(x, 0) = x$, $I_0 = |B|$, the volume of B. However, on the other hand, as $F(\cdot, 1)$ maps B to ∂B, $\det J_F(x, 1) \equiv 0$ which implies that $I_1 = 0$, contradiction holds. (To see why $\det J_F(x, 1) \equiv 0$, we may reason as follows: If $\det J_F(x_0, 1) \neq 0$ at some x_0. By the continuity of $\det J_F(x, 1)$ we may assume x_0 is located in the interior of B. Non-vanishing of the determinant implies that the matrix $J_F(x_0, 1)$ is invertible. By the inverse function theorem, the image of $F(\cdot, 1)$ would contain an open set surrounding the point $F(x_0, 1)$, which would be in conflict with $F(B, 1) \subset \partial B$.)

From this contradiction we conclude that every twice continuously differentiable map from B to itself has a fixed point.

Now the general case. Let $F = (F^1, \cdots, F^n)$ be any continuous map from B to itself. For each F^j, we can find a sequence of polynomials $\{P_k^j\}$ which approximate it uniformly in B. Therefore, the map $F_k = (P_k^1, \cdots, P_k^n)$ is smooth from B to $\mathbb{R}^n$. It is not hard to see that we can find $\lambda_k \in (0, 1)$, $\lambda_k \to 1$ such that, $G_k = \lambda_k F_k : B \to B$. Let z_k by a fixed point of G_k. Then $G_k(z_k) = z_k$. By Bolzano-Weierstrass theorem we can extract from z_k a convergent subsequence, still denoted by $\{z_k\}$ which converges to some z Then

$$\begin{aligned} \|F(z) - z\| &\leqslant \|F(z) - G_k(z)\| + \|G_k(z) - G_k(z_k)\| \\ &\quad + \|G_k(z_k) - z_k\| + \|z_k - z\| \\ &\to 0. \end{aligned}$$

that's, z is a fixed point for F. The proof of Brouwer fixed-point theorem is completed.

Theorem 8.2 clearly is a topological result. It implies that every continuous map from a set homeomorphic to the closed unit ball to itself has a fixed point. In particular, this is true on compact convex sets in $\mathbb{R}^n$, see exercise.

An obvious difference between the contraction mapping principle and Brouwer fixed-point theorem is the lack of uniqueness in the latter. In fact, trivial examples show that the fixed point may not be unique.

It is a usual practise in mathematics that people try to approach an important theorem from various angles and obtain different proofs. There is no exception for Brouwer fixed-point theorem. After Brouwer's topological proof, many different proofs have emerged. Our analytic proof is adapted from Dunford-Schwartz. One may consult any book on algebraic topology topological proofs. The book by J. Franklin, Methods of Mathematical Economics, Linear and Nonlinear Programming, Fixed-Points Theorems, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1980, contains three more additional proofs as well as other fixed-point theorems.

Brouwer fixed-point theorem is a theorem on a finite dimensional space. In functional analysis the emphasis is on infinite dimensional spaces. Can this theorem be generalized to infinite dimension? We have learned that an essential difference between finite and infinite dimensions is the loss of compactness. It turns this phenomenon plays a role. Here is a counterexample. Consider the map Φ defined in the closed unit ball of ℓ^2, $\{x \in \ell^2 : \|x\|_2 \leq 1\}$, given by $\Phi(x) = ((1 - \|x\|_2^2)^{1/2}, x_1, x_2, \cdots)$. It is clear that this map is continuous into the ball itself (in fact, to its boundary). However, it does not have a fixed point. For, if $\Phi(z) = z$ for some z in this ball, by equalling the components of $\Phi(z)$ and z we have $z_1 = z_2 = z_3 = \cdots$ which implies $z = (0, 0, 0 \cdots)$. But this is impossible from the first component: $1 - \|z\|_2^2 = z_1^2$.

This example shows that continuity is not sufficient to ensure the existence of fixed points in infinite dimensional spaces. A most direct way is to restrict our attention to compact sets.

The following result is a fixed-point theorem established by Schauder in 1930.

Theorem 8.4. *Let C be a non-empty compact, convex set in the Banach space X. Every continuous map from C to itself has a fixed point.*

Proof. By compactness, for each $1/n$ we can cover C by finitely many balls $B_{1/n}(z_1), \cdots, B_{1/n}(z_N)$ where the centers $z_j,\ j = 1, \cdots, N$, belong to C. Let C_n be the convex hull of these centers, that is,

$$C_n = \{\sum_j \lambda_j z_j : \ \sum \lambda_j = 1,\ \lambda_j \in [0,1]\}.$$

Each C_n is a compact convex set in some finite dimensional space. We define a map P_n from C to C_n by

$$P_n(x) = \frac{\sum_j \operatorname{dist}(x, C \setminus B_{1/n}(z_j))z_j}{\sum_j \operatorname{dist}(x, C \setminus B_{1/n}(z_j))}.$$

It is straightforward to verify that P_N is continuous and satisfies

$$\|P_n(x) - x\| < \frac{1}{n},$$

in C. Now, consider the composite map $P_n \circ T$ and restrict it to C_n to obtain a continuous map from C_n to itself. Applying Brouwer fixed-point theorem to it, we obtain some x_n in C_n satisfying $P_n(T(x_n)) = x_n$. As $C_n \subset C$ and C is compact, by passing to a subsequence if necessary, we may assume $x_0 = \lim_{n\to\infty} x_n$ exists in C. Using the above estimate, we have

$$\|x_n - T(x_n)\| = \|P_n(T(x_n)) - T(x_n)\| < \frac{1}{n}.$$

Letting $n \to \infty$, we conclude that $\|x_0 - T(x_0)\| = 0$, that is, x_0 is a fixed point of T. The proof of Schauder fixed-point theorem is complete.

□

Schauder fixed-point theorem is a very common tool in the study of partial differential equations. Let's demonstrate its power through a simple case.

Consider the ordinary differential equation

$$\frac{d^2u}{dx^2} = f(x, u, \frac{du}{dx}) \quad , \quad x \in (0,1). \tag{8.2}$$

There are two kinds of problems associated to this equation. The first one is the initial value problem, namely, we look for a solution $u(x)$ satisfies (8.3) together with the initial conditions $u(0) = a, u'(0) = b$ where a and b are prescribed values. The fundamental existence theorem of ODE's asserts that this problem has a unique solution in some interval containing 0 when $f(x, z, p)$ is sufficiently regular, for instance, it is continuously differentiable in (x, z, p) near $(0, a, b)$. Alternatively, one may consider boundary value problems. For example, one may seek a solution of (8.3) which also satisfies the boundary conditions $u(0) = \alpha$ and $u(1) = \beta$. Boundary value problems arise from separation of variables in partial differential equations.

Here for simplicity assume the continuous function f is independent of p and satisfies the structural condition

$$|f(x, z)| \leqslant C_1 (1 + |z|^\gamma) \ , \ (x, z) \in [0,1] \times \mathbb{R}, \tag{8.3}$$

for some constants C_1 and $\gamma \in (0,1)$.

Proposition 8.5. *Under condition (8.3), (8.2) has a solution satisfying* $u(0) = u(1) = 0$.

Proof. From the discussion in Section 6.3 we know that (8.2) is equivalent to the integral equation

$$u(x) = \int G(x, y) f(y, u(y)) \ dy$$

where the integration is over $[0,1]$ and G is the Green function of the linear problem ($q \equiv 0$). It is known that $G, \partial G/\partial x, \partial G/\partial y$ are continuous on $[0,1]\times[0,1]$. We choose the space $X = \{C[0,1] : u(0) = u(1) = 0\}$ with the sup-norm and define

$$Tu(x) = \int G(x,y)f(y,u(y))dy.$$

It is clear that $T : X \to X$ is continuous. Consider the closed and convex subset $C = \{u \in X : \|u\|_\infty, \|u'\|_\infty \leqslant R\}$. As a direct consequence of Ascoli-Arzela theorem C is also compact. We claim that T maps C into C when R is sufficiently large. For, from (8.4),

$$\begin{aligned} |Tu(x)| &\leqslant \sup_x \int |G(x,y)f(y,u(y)dy| \\ &\leqslant MC_1 \int \left(1 + |u(y)|^\gamma\right) dy \\ &\leqslant MC_1 \left(1 + R^\gamma\right), \end{aligned}$$

where $M = \sup_{x,y} |G(x,y)|$. Similarly,

$$\begin{aligned} \Big|\frac{d}{dx}Tu(x)\Big| &= \Big|\int \frac{\partial G}{\partial x}(x,y)f(y,u(y))dy\Big| \\ &\leqslant M_1C_1(1 + R_0^\gamma), \end{aligned}$$

where $M_1 = \sup_{x,y} |\partial G/\partial x(x,y)|$. Since $\gamma < 1$, we can choose a large R_0 so that

$$MC(1 + R_0^\gamma), \quad MC_1(1 + R_0^\gamma) \leqslant R_0.$$

With this choice of R_0, T maps C into itself.

Now we can apply Schauder fixed-point theorem to conclude that T admits a fixed point $u \in C$. In other words,

$$u(x) = Tu(x) = \int G(x,y)f(y,u(y))dy),$$

so u solves (8.2).

□

What happens when the exponent γ in (8.4) is larger or equal to one? Things become more delicate. We just point out that then a solution may not exist. Consider the special case

$$\begin{cases} \dfrac{d^2u}{dx^2} = -4\pi^2 u + \varphi(x) \\ u(0) = u(1) = 0. \end{cases}$$

Multiplying the equation by $\sin 2\pi x$ and then integrating over $[0,1]$, we obtain a necessary condition for solvability, namely,

$$\int_0^1 \varphi(x)\sin 2\pi x dx = 0.$$

In particular, this problem does not admit a solution when $\varphi(x) = \sin 2\pi x$.

8.2 Calculus in Normed Spaces

The concept of differentiability has a natural generalization to infinite dimensional space. Let F be a map from a set E in the normed space X to another normed space Y and x_0 a point in E. The map F is said to be *differentiable* at x_0 if there exists a bounded linear operator $L \in B(X,Y)$ such that

$$\|f(x) - f(x_0) - L(x - x_0)\| = \circ(\|x - x_0\|), \quad \text{as } x \to x_0,$$

in other words,

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|}{\|x - x_0\|} = 0.$$

The linear operator is called the *Frèchet derivative* or simply the derivative of F at x_0. The map F is called differentiable on E if it is differentiable at every point of E. In this case, the derivative is a bounded linear operator depending on $x \in E$ and usually is denoted by $F'(x)$ or $DF(x)$. We call F a C^1*-map* if $x \mapsto F'(x)$ is continuous from E to $B(X, Y)$.

Let's consider two examples.

First, let $X = C^1[0,1]$ and $Y = C[0,1]$ under the C^1- and sup-norms respectively. The map

$$F(u)(t) = t \sin u(t) + (u'(t))^2$$

maps X to Y. To find its derivative we need to determine the linear operator $F'(u)$ such that

$$\lim_{w \to u} \frac{\|F(w) - F(u) - F'(u)(w - u)\|_\infty}{\|w - u\|_{C^1}} = 0.$$

Setting $w = u + \varepsilon\varphi$ in the above, we see that in case $F'(u)$ exists, we must have

$$F'(u)\varphi(t) = \lim_{\varepsilon \to 0} \frac{F(u(t) + \varepsilon\varphi(t)) - F(u(t))}{\varepsilon},$$

at each $u(t)$ and $\varphi(t)$. Applying the chain rule in the variable ε, we readily obtain

$$F'(u)\varphi = t\cos(u(t))\ \varphi(t) + 2u'(t)\varphi'(t).$$

It is now a direct check using the definition that this expression indeed gives the derivative of F. Observe that it is linear on φ but nonlinear in u.

Second, consider the nonlinear functional $S:\ C^1(\overline{\Omega}) \to \mathbb{R}$ given by

$$S(u) = \int_\Omega \sqrt{1 + |\nabla u(x)|^2} dx,$$

where Ω is a bounded domain in $\mathbb{R}^n$. This functional gives the surface area of the hypersurface $\{(x, u(x))\}$ over Ω. Proceeding as above, its derivative is given by

$$DS(u)\varphi = \int_\Omega \frac{\langle \nabla u, \nabla\varphi \rangle}{\sqrt{1 + |\nabla u|^2}},$$

a bounded linear functional on $C^1(\overline{\Omega})$ under the C^1-norm.

Here are some elementary properties of differentiability.

Proposition 8.6. *Let X, Y and Z be normed spaces and $E \subset X$ and $N \subset Y$.*

(i) Let $F, G : E \to Y$ be differentiable at x. Then $\alpha F + \beta G$ is differentiable at x and

$$D(\alpha F + \beta G)(x) = \alpha DF(x) + \beta DG(x).$$

(ii) Let $F : E \to N$ and $G : N \to Z$ be differentiable at x and $F(x)$ respectively. Then $G \circ F$ is differentiable at x and

$$D(G \circ F)(x) = DG(F(x))DF(x).$$

We leave the proof of this proposition as an exercise.

When F is differentiable (resp. C^1) in E, $f(E) \subset N$ and G is differentiable (resp. C^1) in N, this proposition shows that $G \circ F$ is differentiable (resp. C^1) in E.

Let φ be a continuous map from the interval $[a, b]$ to the Banach space X. When $X = \mathbb{R}$, we can define the Riemann integral of φ over $[a, b]$. Likewise the same thing can be done in a Banach space. The old definition works, namely, φ is *integrable* on $[a, b]$ if there exists an element z in X such that, for every $\varepsilon > 0$, there exists some $\delta > 0$, so that

$$\Big\| \sum_j \varphi(t_j)\Delta t_j - z \Big\|_X < \varepsilon,$$

for any partition P of $[a, b]$ whose length is less than δ. The number z is called the *integral* of φ and will be denoted by $\int_a^b \varphi(t)dt$. Be careful it is an element in X.

Same as in the one dimensional case, any continuous map on $[a, b]$ is integrable.

Proposition 8.7. *Let $\varphi : [a, b] \to X$ be continuous where X is a Banach space.*

(i) There holds

$$\Big\| \int_a^b \varphi(t)dt \Big\| \le \int_a^b \|\varphi(t)\|dt,$$

(ii) For every $\Lambda \in X'$,

$$\int_a^b (\Lambda\varphi)(t)dt = \Lambda\Big(\int_a^b \varphi(t)dt \Big).$$

(iii) If φ is a C^1-map, then

$$\varphi(b) - \varphi(a) = \int_a^b \varphi'(t)dt.$$

Proof. (i) and (ii) follow directly from definition. For (iii), we observe that if $\varphi(b) - \varphi(a)$ is not equal to $\displaystyle\int_a^b \varphi'(t)dt$, Hahn-Banach theorem tells us that there exists some $\Lambda_1 \in X'$ such that

$$\Lambda_1 (\varphi(b) - \varphi(a)) \neq \Lambda_1 \left(\int_0^1 \varphi'(t)dt \right).$$

However, by linearity and (ii),

$$\begin{aligned} \Lambda_1 (\varphi(b) - \varphi(a)) &= (\Lambda_1\varphi)(b) - (\Lambda_1\varphi)(a) \\ &= \int_a^b \big(\Lambda_1\varphi\big)'(t)dt \\ &= \int_a^b \Lambda_1\varphi'(t)dt \\ &= \Lambda_1 \int_a^b \varphi'(t)dt, \end{aligned}$$

contradiction holds. □

Now we come to the fundamental inverse function theorem. Roughly speaking, it tells us that a map is locally invertible at a particular point if its linearization at the same point is invertible.

Theorem 8.8. *Let $F : U \to Y$ be a C^1-map where X and Y are Banach spaces and U is open in X. Suppose that $F(x_0) = y_0$ and $F'(x_0)$ is invertible. There exist open sets V and W containing x_0 and y_0 respectively such that the restriction of F on V is a bijection onto W with a C^1-inverse.*

Recall that a bounded linear operator is invertible if its inverse exists and is bounded.

Proof. Without loss of generality take $x_0,\ y_0 = 0$. First we would like to show that there is a unique solution for the equation $F(x) = y$ for y near 0. We shall use the contraction mapping principle to achieve our goal. For a fixed y, define the map in U by

$$T(x) = L^{-1}\left(Lx - F(x) + y\right)$$

where $L = F'(0)$. It is clear that any fixed point of T is a solution to $F(x) = y$. We have

$$\begin{aligned} \|T(x)\| &\leqslant \|L^{-1}\|\ \|F(x) - Lx - y\| \\ &\leqslant \|L^{-1}\|\left(\|F(x) - Lx\| + \|y\|\right) \\ &\leqslant \|L^{-1}\|\left(\circ(\|x\|) + \|y\|\right). \end{aligned}$$

We can find a small ρ_0 such that

$$\|L^{-1}\| \circ (\|x\|) \leq \frac{1}{4}\|x\|, \quad \forall x, \quad \|x\| \leq \rho_0. \tag{8.4}$$

Then for for each y in $B_R(0)$, $\|L^{-1}\|R \leq \rho_0/2$, T maps $\overline{B_{\rho_0}(0)}$ to itself. Moreover, for x_1, x_2 in $B_{\rho_0}(0)$, we have

$$\begin{aligned} \|T(x_2) - T(x_1)\| &= \left\|L^{-1}\left(F(x_2) - Lx_2 - y\right) - L^{-1}\left(F(x_1) - Lx_1 - y\right)\right\| \\ &\leqslant \|L^{-1}\|\ \|F(x_2) - F(x_1) - F'(0)(x_2 - x_1)\| \\ &\leqslant \|L^{-1}\| \left\| \int_0^1 F'\left(x_1 + t(x_2 - x_1)\right)(x_2 - x_1)dt - F'(0)(x_2 - x_1) \right\|, \end{aligned}$$

where we have applied Proposition 8.6 (ii) to $\varphi(t) = F(x_1 + t(x_2 - x_1))$. Since F' is continuous in U, by further restricting ρ_0 we may assume

$$\|F'(x) - F'(0)\| < \frac{1}{2\left(\|L^{-1}\| + 1\right)}\ ,\ \forall x \in B_{\rho_0}(0)\ .$$

Consequently,

$$\begin{aligned} \|T(x_2) - T(x_1)\| &\leqslant \|L^{-1}\| \frac{1}{2(1 + \|L^{-1}\|)} \|x_2 - x_1\| \\ &< \frac{1}{2}\|x_2 - x_1\|\ . \end{aligned}$$

We have shown that $T : \overline{B_{\rho_0}(0)} \to \overline{B_{\rho_0}(0)}$ is a contraction. By the contraction mapping principle, there is a unique fixed point for T, in other words, for each y in the ball $B_R(0)$ there is a unique point x in $\overline{B_{\rho_0}(0)}$ solving $F(x) = y$. Defining $G:\ B_R(0) \to B_{\rho_0}(0) \subset X$ by setting $G(y) = x$, G is inverse to F.

Next, we claim that G is continuous. In fact, for $G(y_i) = x_i,\ i = 1, 2,$

$$\begin{aligned}
\|G(y_2) - G(y_1)\| &= \|x_2 - x_1\| \\
&= \|T(x_2) - T(x_1)\| \\
&\leqslant \|L^{-1}\| \, \| (\circ\|x_2 - x_1\| + \|y_2 - y_1\|) \\
&\leq \|L^{-1}\| \, \| (\circ\|x_2\| + \circ\|x_1\| + \|y_2 - y_1\|) \\
&\leq \frac{1}{2}\|x_2 - x_1\| + \|L^{-1}\|\|y_2 - y_1\| \\
&= \frac{1}{2}\|G(y_2) - G(y_1)\| + \|L^{-1}\|\|y_2 - y_1\|,
\end{aligned}$$

which, by (8.4), implies

$$\|G(y_2) - G(y_1)\| \leqslant 2\|L^{-1}\|\|y_2 - y_1\| \ , \tag{8.5}$$

that's, G is continuous on $B_R(0)$.

Finally, let's show that G is a C^1-map in $B_R(0)$. In fact, for $y_1, y_1 + y$ in $B_R(0)$, using

$$\begin{aligned}
y &= F(G(y_1 + y)) - F(G(y_1)) \\
&= \int_0^1 F'(G(y_1) + t(G(y_1 + y) - G(y_1))dt \ (G(y_1 + y) - G(y_1)),
\end{aligned}$$

we have

$$G(y_1 + y) - G(y_1) = F'^{-1}(G(y_1))y + R,$$

where R is given by

$$F'^{-1}(G(y_1)) \int_0^1 \Big(F'(G(y_1)) - F'(G(y_1) + t(G(y_1 + y) - G(y_1))\Big)(G(y_1 + y) - G(y_1))dt.$$

As G is continuous and F is C^1, we have

$$G(y_1 + y) - G(y_1) - F'^{-1}(G(y_1))y = \circ(1)(G(y_1 + y) - G(y_1))$$

for small y. Using (8.5), we see that

$$G(y_0 + y) - G(y_0) - F'^{-1}(G(y_0))y = \circ(\|y\|) \ ,$$

as $\|y\| \to 0$. We conclude that G is differentiable with derivative equal to $F'^{-1}(G(y_0))$. The proof of the inverse function theorem is completed by taking $W = B_R(0)$ and $V = F^{-1}(W)$. □

Remark. Under the setting of Theorem 8.8, what happens if the invertibility of $F'(x_0)$ is replaced by surjectivity? Well, assume that X is the direct sum of X_1 and $X_2 \equiv \ker F'(x_0)$. Then the following conclusion holds: There exist V_1, V_2 and W open subsets of X_1, X_2 and Y respectively and C^1-map $G:\ W \to V_1 \times V_2$ such that for each x_2 in X_2, $G(\cdot, x_2)$ is the inverse to $F(\cdot, x_2)$.

Next we deduce the implicit function theorem from the inverse function theorem. In fact, these two theorems are equivalent; in the exercise you are asked to give a self-contained proof of the implicit function theorem and deduce the inverse function theorem from the implicit function theorem.

Theorem 8.9. *Consider C^1-map $F : U \to Z$ where U is an open set in the Banach spaces $X \times Y$. Suppose $(x_0, y_0) \in U$ and $F(x_0, y_0) = 0$. If $F_y(x_0, y_0)$ is invertible from Y to Z, then there exist an open subset $U_1 \times V_1$ of U containing (x_0, y_0) and a C^1-map $\varphi : U_1 \to V_1$, $\varphi(x_0) = y_0$, such that*

$$F(x, \varphi(x)) = 0 \ , \qquad \forall x \in U_1 \ .$$

Moreover, if ψ is a C^1-map from U_2, an open set containing x_0, to Y satisfying $F(x, \psi(x)) = 0$ and $\psi(x_0) = y_0$, then ψ equals to φ in $U_1 \cap U_2$.

The notation $F_y(x_0, y_0)$ stands for the "partial derivative" of F in y, that is, the derivative of F at y_0 while x_0 is fixed as a constant.

Proof. Consider $\Phi : U \to X \times N$ given by

$$\Phi(x, y) = (x, F(x, y)).$$

By assumption

$$\Phi'(x_0, y_0)(x, y) = (x, F_y(x_0, y_0)(x, y))$$

is invertible from $X \times Y$ to $X \times Z$. By the inverse function theorem, there exists some $\Psi = (\Psi_1, \Psi_2) : U_1 \times N_1 \to U$ which is inverse to Φ. For every $(x, z) \in U_1 \times N_1$, we have

$$\Phi(\Psi_1(x, z), \Psi_2(x, z)) = (x, z),$$

which immediately implies

$$\Psi_1(x, z) = x, \text{ and } F((\Psi_1(x, z), \Psi_2(x, z)) = z.$$

In particular, taking $z = 0$ gives

$$(x, 0) = \Phi(\Psi(x, 0)) = (x, F(x, \Psi_2(x, 0)), \quad \forall x \in U_1,$$

so the function $\varphi(x) \equiv \Psi_2(x, 0)$ satisfies our requirement. The uniqueness assertion can be easily established and is left to the reader. □

The implicit function theorem is indispensable in analysis. You will encounter many of its applications as you go along. Here we give a very simple one about the multiplicity of solutions to differential equations. Consider the boundary value problem

$$\begin{cases} \dfrac{d^2u}{dx^2} = -\lambda u + g(u), \\ u(0) = u(1) = 0, \end{cases}$$

where λ is a given number and $g(y)$ is a function satisfying $g(0) \equiv 0$. Clearly the zero function is a trivial solution of this problem. An interesting question is, could it admit another solution? Taking the special case $g \equiv 0$ where the equation can be solved explicitly, we see that it has a nonzero solution if and only if $\lambda = n^2\pi^2$, $n \in \mathbb{N}$. Indeed, the solutions are given by $u(x) = c \sin n\pi x$, where c is an arbitrary nonzero constant. A value at which nontrivial solutions exist arbitrarily near the trivial solution is called a bifurcation point. In this problem, every number $n^2\pi^2$ is a bifurcation point. In the general case, the zero function is still a trivial solution. We would like to know which λ is a bifurcation point. To this end, we take X, Y, and Z respectively to be $\mathbb{R}, \{u \in C^2[0, 1] : \ u(0) = u(1) = 0\}$, and $C[0, 1]$ and $F(\lambda, u) = u'' + \lambda u + g(u)$. We have $F(\lambda, 0) = 0$ and

$$F_y(\lambda, 0)v = v'' + \lambda v + g_y(0)v.$$

Clearly $F_y(\lambda, 0)$ is invertible if and only if λ is not equal to $n^2\pi^2 - g_y(0)$, $n \in \mathbb{N}$. By the implicit function theorem, there exists an open set containing the zero function which does not contain any additional solution to the problem. Hence values not equal to $n^2\pi^2 - g_y(0)$ cannot be bifurcation points. What happens when λ is equal to $n^2\pi^2 - g_y(0)$? This is bifurcation theory. More information is required from g to obtain a conclusion.

The technique in the proof of the inverse function theorem can be used to establish a nonlinear version of the open mapping theorem.

Theorem 8.10. *Let F be a C^1-map from U to Y where U is open in X and X, Y are Banach spaces. Suppose that $F'(x)$ maps X onto Y for every x in U. Then $F(U)$ is open in Y.*

Lemma 8.11. *Let $T \in B(X, Y)$ be surjective where X and Y are Banach spaces. There exists a constant C such that*

$$\inf\{\|x - z\| : \ z \in \ker T\} \le C\|Tx\|, \quad \forall x \in X.$$

Proof. Consider the quotient Banach space $\tilde{X} = X/\ker T$ under the norm

$$\|\tilde{x}\| = \inf\{\|x - z\| : z \in \ker T\}.$$

The induced map $\tilde{T} : \tilde{X} \to Y$ given by $\tilde{T}\tilde{x} = Tx, \ x \in \tilde{x}$, is a bounded linear operator onto Y. By Banach inverse mapping theorem $\tilde{T}$ is invertible, that is,

$$\|\tilde{x}\| \leq C\|\tilde{T}\tilde{x}\|,$$

and the lemma follows. □

Proof of Theorem 8.10. It suffices to show that if $F'(x_0)(X) = Y$ where $x_0 \in U$, there exist balls $B_\rho(x_0)$ and $B_R(y_0)$, $y_0 = F(x_0)$, such that $\subset B_R(y_0)F(B_\rho(x_0))$.

With ρ and R both small to be specified, for any fixed y in $B_R(y_0)$ we define a sequence $\{x_n\}$ in $B_{\rho/2}(x_0)$ as follows. First, find $x'_{n+1} \in X$ such that

$$Tx'_{n+1} = Tx_n - (F(x_n) - y), \quad T = F'(x_0).$$

Of course such point exists as T is onto. As

$$\inf \|x_n - (x'_{n+1} + z)\| \leq C\|Tx_n - Tx_{n+1}\|,$$

for all $z \in \ker T$ by the above lemma, we could modify x'_{n+1} by some element in $\ker T$ to get x_{n+1} satisfying

$$\|x_{n+1} - x_n\| \leq (C+1)\|Tx_n - Tx_{n+1}\|.$$

Starting from $n = 0$, we have

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq (C+1)\|F(x_n) - y\| \\ &= (C+1)\|F(x_n) - F(x_{n-1}) - Tx_n + Tx_{n-1}\| \\ &\leq (C+1) \circ (\|x_n - x_{n-1}\|). \end{aligned}$$

We can choose a small ρ such that

$$\|x_{n+1} - x_n\| \leq \frac{1}{2}\|x_n - x_{n-1}\|,$$

as long as $\{x_n\}$ stays in $B_{\rho/2}(x_0)$. From the above estimate we have

$$\|x_{n+1} - x_0\| \leq \sum_{j=0}^{n} \frac{1}{2^j}\|x_1 - x_0\| \leq 2\|y_0 - y\|.$$

By choosing $R < \rho/4$, for every $y \in B_R(y_0)$, $\{x_n\} \subset B_{\rho/2}(x_0)$. Moreover,

$$\|x_n - x_m\| \leq \frac{1}{2^m}\|x_{n-m} - x_0\| \leq \frac{\rho}{2^{m+1}} \to 0,$$

as $n > m, m \to \infty$. By completeness there is some x in $B_\rho(x_0)$ such that $x = \lim_{n\to\infty} x_n$. From $Tx = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Tx_{n+1}$ we deduce that x solves $F(x) = y$.

Let $F : U \to Y$ be differentiable where U is open in X. Its derivative $F'(x)$ belongs to $B(X, Y)$. It is *twice differentiable* at x if $F' : U \to B(X, Y)$ is differentiable at x. The second (Frèchet) derivative at x, denoted by $F''(x)$ or $D^2F(x)$, belongs to $B(X, B(X, Y))$. F is a C^2-map if $x \mapsto F''(x)$ is continuous from U to $B(X, B(x, Y))$.

There is a natural way to identify the space $B(X, B(X, Y))$ with the multi-linear space $M_2(X, Y)$ where X and Y are normed spaces. A map $T : X \times X \to Y$ is a *bilinear form* from X to Y if $T(x_1, x_2)$

is linear in x_1 (resp. x_2) while x_2 (resp. x_1) is fixed. All continuous, bilinear maps from $X \times X$ to Y form a vector space $M_2(X,Y)$. For any such map T, define

$$\|T\| = \sup_{x,y} \{\|T(x_1,x_2)\|_Y : \|x_1\|_X, \|x_2\|_X \leqslant 1\} \ .$$

It is readily checked that $(M_2(X,Y), \|\cdot\|)$ forms a normed space, and it is complete when Y is complete.

Given $T \in B(X, B(X,Y))$, define

$$\widehat{T}(x_1,x_2) = T(x_1)x_2 \ , \qquad x_1,x_2 \in X \ .$$

It is routine to verify that $T \longmapsto \widehat{T}$ established a norm-preserving isomorphism from $B(X,B(X,Y))$ to $M_2(X,Y)$. Under this isomorphism we may identify $B(X,B(X,Y))$ with $M_2(X,Y)$. It follows that the second derivative $F''(x)$ may be regarded as a bilinear form with value in Y. In fact, the following proposition shows that it is symmetric.

Proposition 8.12. *Let F be a C^2-map from U to Y where U is open in X and X, Y are normed spaces. Then*

$$F''(x)(x_1,x_2) = F''(x)(x_2,x_1) \ , \qquad \forall x \in U \ , \ x_1,x_2 \in X \ .$$

Proof. For $x_1, x_2 \in U$ and $\varepsilon_1, \varepsilon_2$ small, $x + \varepsilon_1 x_1 + \varepsilon_2 x_2 \in U$. Consider the C^2-function φ given by

$$\varphi(\varepsilon_1,\varepsilon_2) = \Lambda\left(F(x+\varepsilon_1 x_1 + \varepsilon_2 x_2)\right)$$

where Λ is in Y'. By the chain rule,

$$\frac{\partial \varphi}{\partial \varepsilon_1} = \Lambda F'(x+\varepsilon_1 x_1+\varepsilon_2 x_2)x_1 \ ,$$

and

$$\frac{\partial \varphi}{\partial \varepsilon_2} = \Lambda F'(x+\varepsilon_1 x_1+\varepsilon_2 x_2)x_2 \ ,$$

So, at $(\varepsilon_1,\varepsilon_2) = (0,0)$,

$$\begin{aligned}
\frac{\partial^2 \varphi}{\partial \varepsilon_2 \partial \varepsilon_1} &= \Lambda F''(x+\varepsilon_1 x_1+\varepsilon_2 x_2)(x_1,x_2) \\
&= \Lambda F''(x)(x_1,x_2) \\
\frac{\partial^2 \varphi}{\partial \varepsilon_1 \partial \varepsilon_2} &= \Lambda F''(x)(x_2,x_1) \ .
\end{aligned}$$

The desired result follows from the relation $\partial^2\varphi/\partial\varepsilon_2\partial\varepsilon_1 = \partial^2\varphi/\partial\varepsilon_1\partial\varepsilon_2$. □

Similarly one can define the m-th derivative of F and identify it with an m-linear function. Same as in this proposition, $F^{(m)}(x)(x_1,\cdots,x_m)$ is symmetric in $(x_1,\cdots,x_m)$ when F is a C^m-map. With this preparation the reader may formulate and prove a version of Taylor's expansion theorem in the infinite dimensional setting.

8.3 Minimization Problems

One remarkable application of calculus is to the determination of the extremum of a function. For a differentiable function f defined on (a,b), its extremal point must be a critical point and all critical points can be found by solving the equation $f'(x) = 0$. Thus the problem of finding extremal points reduces to a much simpler problem. Of course, a critical point may not be an extremal point, as seen in the example $f(x) = x^3$ at $x = 0$. However, in many cases critical points of a function are very few and one can determine which one is minimum or maximum by simply comparing their values.

In the infinite dimensional setting the situation is similar and in fact even more powerful. To be specified consider the functional $J : U \to \mathbb{R}$ where U is an open set in the normed space X. A point x_0 in U is called a *critical point* of J if J' exists at x_0 and $J'(x_0) = 0$. Just as in the finite dimensional case, a point x_0 is called a *local minimum* (resp. *local maximum*) of J if $J(x_0) \leqslant J(x)$ (resp. $J(x_0) \geqslant J(x)$) for all x in a neighborhood of x_0, and a *strict local minimum* or (resp. *strict local maximum*) if the inequality is strict when x is not equal to x_0.

Proposition 8.13. *Let U be an open set in the normed space X and $J : U \to \mathbb{R}$. (a) Let x_0 be a local minimum or maximum of J in U. Then x_0 is a critical point of J if J is differentiable at x_0.*

(b) When J is twice differentiable near the local minimum/maximum x_0, $J''((x_0)x, x) \geq 0$ (resp. ≤ 0) for all x in X. Moreover, it is a strict local minimum or maximum if the inequality is strict when x is not equal to 0.

Proof. For (a), let x_0 be a local minimum of J. By the definition of differentiability,

$$J(x) = J(x_0) + J'(x_0)(x - x_0) + o(\|x - x_0\|) \ , \ \text{as } x \to x_0 \ .$$

Setting $x - x_0 = \varepsilon z$, $\|z\| = 1$, and letting $\varepsilon \to 0$ in the above expression,

$$0 \leqslant \lim_{\varepsilon \to 0} \frac{J(x_0 + \varepsilon z) - J(x_0)}{\varepsilon} = J'(x_0)z.$$

Replacing z by $-z$ we obtain the reversed inequality, hence $J'(x_0) = 0$.

On the other hand, (b) follows immediately from Taylor's expansion,

$$J(x) = J(x_0) + \frac{1}{2}(J''(x_0)(x - x_0), x - x_0) + o(\|x - x_0\|^2).$$

□

Here let's consider an example. Let u be a continuous function defined on the unit interval $[0, 1]$. The graph $\{(x, u(x)) : \ x \in [0, 1]\}$ is a curve in the plane. From calculus we know that its length is given by the formula

$$L(u) = \int_0^1 \sqrt{1 + u_x^2} \ dx.$$

We consider the problem of finding the shortest curve which is in the form of a graph and with ends resting at $(0, 0)$ and $(1, 1)$. To set up the problem we let $X = \{u \in C^1[0, 1] : \ u(0) = 0, u(1) = 1\}$, $Y = \{u \in C^1[0, 1] : \ u(0) = u(1) = 0\}$ and define $M(u) = L(u + x)$ on the Banach space Y (under the C^1-norm). It is easy to verify that M is differentiable in Y. Indeed, its derivative is given by

$$M'(u)\varphi = \int_0^1 \frac{(u_x + 1)\varphi_x}{\sqrt{1 + (u_x + 1)^2}} dx,$$

for every φ in Y. According to the above proposition, if there is a function w in X minimizing the length, it must satisfy the equation

$$\int_0^1 \frac{w_x \varphi_x}{\sqrt{1 + w_x^2}} dx = 0,$$

for all φ in Y. Using the lemma below, we conclude that $w_x/\sqrt{1 + w_x^2}$ must be a constant. Solving this differential equation with the boundary conditions $w(0) = 0$ and $w(1) = 1$, we see that it is given by $w(x) = x$. We have shown that if there is a function in X minimizing the length, then it must be the linear function connecting $(0, 0)$ and $(1, 1)$. We caution the reader that we did not solve the problem; the existence of a minimum has yet to be established. But we know what it looks like if it exists.

Lemma 8.14. *Let f be a continuous function which satisfies*

$$\int_0^1 f(x)\varphi_x(x)dx = 0, \qquad \forall \varphi \in Y.$$

Then f is a constant.

The reader may prove this lemma as an exercise.

Very often we encounter extremal problems with constraints where the Lagrange multipliers come up. The following result justifies their presence by the implicit function theorem.

Proposition 8.15. *Let $J, L : U \to \mathbb{R}$ be C^1 where U is open in the Banach space X. Suppose that*

$$J(x_0) \leqslant J(x),$$

for all $x \in U$ satisfying

$$L(x) = 0.$$

Assume that $L'(x_0)$ does not vanish identically. Then there exists some $\lambda \in \mathbb{R}$ such that

$$J'(x_0) + \lambda L'(x_0) = 0 \quad in \quad X' \ .$$

Proof. By assumption $L'(x_0)z_0 \neq 0$ for some z_0 in X. Consider the function $\Phi(s,t) = L(x_0 + sx + tz_0)$ where x is a given point in X. It is well-defined for sufficiently small s and t and satisfies $\Phi(0,0) = 0$. As

$$\frac{\partial \Phi}{\partial t} = L'(x_0)z_0 \neq 0 \quad \text{at} \quad (0,0) \ ,$$

by the implicit function theorem there exists $\eta(s)$, $s \in (-\varepsilon_0, \varepsilon_0)$, ε_0 small, such that

$$\Phi\big(s, \eta(s)\big) = 0.$$

Differentiating this relation gives

$$L'(x_0)x + L'(x_0)\eta'(0)z_0 = 0,$$

that is,

$$\eta'(0) = -\frac{L'(x_0)x}{L'(x_0)z_0}.$$

As $x_0 + sx + \eta(s)z_0$ lies on the constraint set,

$s \longmapsto J\big(x_0 + sx + \eta(s)z_0\big)$ is minimized at $s = 0$, so

$$0 = J'(x_0)x + J'(x_0)\eta'(0)z_0 \ ,$$

or

$$J'(x_0)x + \lambda L'(x_0)x = 0 \ ,$$

where

$$\lambda = -\frac{J'(x_0)z_0}{L'(x_0)z_0} \ .$$

□

Constrained optimization problems are very common. Let's consider an isoperimetric problem. Adopting the notation in the previous example, we look at the following problem:

$$\inf\{L(u) : \ A(u) = \alpha > 0, \ u \in Y\},$$

where the area functional A is given by

$$A(u) = \int_0^1 |u(x)|dx.$$

Assume that w is a minimum of this problem which is positive in $(0,1)$. Then L and A are differentiable at w. Since $A'(w)w > 0$, by Proposition 8.15 there is a constant λ such that $L'(w) + \lambda A'(w) = 0$. So w satisfies

$$\int_0^1 \Big(\frac{w_x \varphi_x}{\sqrt{1+w_x^2}} + \lambda\varphi\Big) = 0.$$

To proceed further, let us assume that w is in $C^2[0,1]$. By integration by parts and using the abundance of φ, we have the following second order differential equation for the minimum,

$$\frac{d}{dx}\frac{w_x}{(1+w_x^2)^{\frac{1}{2}}} = \frac{w_{xx}}{(1+w_x^2)^{\frac{3}{2}}} = \lambda,$$

where the Lagrange multiplier λ is nonzero due to the area constraint. Observing that

$$\frac{d}{dx}\frac{1}{\sqrt{1+w_x^2}} = \frac{w_{xx}}{(1+w_x^2)^{3/2}},$$

we integrate this equation to obtain

$$w(x) = \frac{1}{\lambda}\Big[c - \sqrt{1-(\lambda x + d)^2}\Big],$$

for some constants c and d. This formula can be put into the form

$$\Big(w - \frac{c}{\lambda}\Big)^2 + \Big(x + \frac{d}{\lambda}\Big)^2 = \frac{1}{\lambda},$$

and from the boundary conditions we further infer $d/\lambda = -1/2$. Therefore, the graph of the minimum w is a circular arc connecting $(0,0)$ and $(1,0)$ with its center lying on the lower vertical line $x = 1/2, w \le 0$. From the area constraint it exists if and only if $A \in (0, \pi/8)$.

Any discussion on minimization problems cannot be called complete without touching the issue of existence of the minimum, although the consideration in the previous paragraphs has already provided efficient and practical ways to determine them. No matter have you known it or not, the following statement from calculus is easy to understand. Let J be a functional defined in E, a closed set in some Euclidean space. Assume that it is continuous and coercive. Then the problem $\inf\{J(x) : x \in E\}$ has a solution. Here a functional J is *coercive* in a subset of a normed space if $J(x)$ tends to ∞ whenever $\|x\|$ tends to ∞. A continuous functional is always coercive when the subset is bounded. The proof of this assertion is a simple application of Bolzano-Weierstrass theorem.

In infinite dimensional settings, compactness is lost and from past experience we know that it is necessary to turn to weak topology. We call a functional J defined in a subset E of some normed space *weakly sequentially continuous* if $J(x_n) \to J(x)$ whenever $\{x_n\}$ tends to x weakly in E. It is *weakly sequentially lower semicontinuous* if

$$\liminf_{n\to\infty} J(x_n) \ge J(x),$$

whenever $\{x_n\}$ tends to x weakly in E.

Theorem 8.16. *Let E be a weakly sequentially closed set in the reflexive space X and J a weakly sequentially lower semicontinuous function defined in E. Suppose that it is also coercive. Then there is some x^* in E satisfying*

$$J(x^*) \le J(x), \quad \forall x \in E.$$

Proof. Fix some x_0 in E and consider the functional J in the closed set $E_1 = \{x \in E : J(x) \le J(x_0)\}$. Coercivity implies that any minimizing sequence of $\inf\{J(x) : x \in E_1\} = \inf\{J(x) : x \in E\}$ is bounded in norm. By sequentially weak compactness, we can extract a subsequence $\{x_{n_j}\}$ which converges weakly to some x^* in E_1, noting that E_1 is weakly sequentially closed. By weakly sequential lower semicontinuity,

$$J(x^*) \le \liminf_{j\to\infty} J(x_{n_j}) = \inf_E J,$$

so x^* achieves the minimal value of J over E. □

The whole space is of course weakly sequentially closed. The following proposition, which is contained in Corollary 7.3, provides many easily verified examples.

Proposition 8.17. *Any bounded, closed, convex subset of a reflexive Banach space is weakly sequentially compact.*

Concerning weakly sequentially lower semicontinuous functions we have

Proposition 8.18. *Let C be a convex subset of the normed space X. Suppose that a function J defined in C can be written as $J_1 + J_2$ where J_1 is convex and J_2 is weakly sequentially continuous. Then J is weakly sequentially lower semicontinuous.*

Proof. It suffices to prove the proposition assuming that F is convex. For any two points x and y in C, the function $\varphi(s) = J((1-s)x + sy)$ is convex on $[0, 1]$. From

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(s)ds \ge \varphi'(0),$$

we obtain

$$J(y) - J(x) \ge J'(x)(y - x).$$

Plugging $y = x_n$ where $\{x_n\}$ tends to x weakly in this inequality, we immediately have

$$\lim_{n\to\infty} J(x_n) - J(x) \ge \lim_{n\to\infty} J'(x)(x_n - x) = 0.$$

□

We conclude by applying a variational approach to the Dirichlet problem (8.2). In Section 1 we solved it by Schauder fixed point theorem. We would like to find the solution as a minimum of a functional on a suitable space.

To this end, define

$$E(u) = \frac{1}{2}\int_0^1 u_x^2(x)dx + \int_0^1 F(x, u(x))\, dx,$$

where F is given by

$$F(x, z) = \int_0^z f(x, s)ds.$$

Formally its derivative is given by

$$E'(u)\varphi = \int_0^1 u_x(x)\varphi_x(x)dx + \int_0^1 f(x, u(x))\varphi(x)\, dx,$$

for all φ vanishing at endpoints. When u is critical, $E'(u) = 0$, we have

$$0 = \int_0^1 \big(u_x\varphi_x + f(x, u)\varphi\big)dx = 0.$$

Assuming that u is C^2, an integration by parts shows that u solves our problem.

Consider the space $X = \{u \in C^1[a,b] : \ u(a) = u(b) = 0\}$. We have seen that this space is complete under the C^1-norm. But now, in order to apply Theorem 8.16, we consider it under the weaker norm

$$\|u\|_a = \|u_x\|_2 + \|u\|_2.$$

It is straightforward to verify that $\|\cdot\|_a$ forms a norm.

Proposition 8.19. *Let $u \in X$. We have*

(a)

$$\|u\|_\infty \le \Big(\int u_x^2\Big)^{\frac{1}{2}},$$

and

(b)

$$\int u^2 \le \int u_x^2;$$

We leave the proof of this proposition to you. Note that (a) shows that every element in $\mathcal{X}$ can be identified with a continuous function and (b) shows that the norm $\|\cdot\|_a$ is equivalent to

$$\|u\|_b = \sqrt{\int u_x^2}.$$

Let $\mathcal{X}$ be the completion of X under $\|\cdot\|_a$. Write $E = E_1 + E_2$ where

$$E_1(u) = \frac{1}{2}\int u_x^2,$$

and

$$E_2(u) = \int F(x,u).$$

Clearly E_1, E_2 and consequently E, extend to $\mathcal{X}$ and we will use the same notations to denote their extensions. In the following we will verify that they fulfill the hypotheses in Proposition 8.18. We will do this assuming the points are in X. A straightforward approximation argument will show that they hold in $\mathcal{X}$ as well. Claim: E_1 is convex. For, E_1 is convex if and only if the function $h(s) = E_1((1-s)u + sv)$ is convex for any u, v in X. And we have

$$h''(s) = \int (u_x^2 - 2u_x v_x + v_x^2) \ge 0$$

by Cauchy-Schwarz inequality. Next, E_2 is weakly sequentially continuous. Suppose on the contrary that while $\{u_n\}$ weakly converges to some w in X, $E_2(u_n)$ does not converge to $E_2(w)$. From the estimate

$$|u_n(y) - u_n(x)| \le \Big|\int_x^y u_{nx}(z)dz\Big| \le |y-x|^{\frac{1}{2}}\|u_n\|_b,$$

and the fact that weakly sequential convergence implies boundedness in norm, we see that $\{u_n\}$ is equicontinuous. It is also uniformly bounded by Proposition 8.19 (a). By Ascoli-Arzela theorem, $\{u_n\}$ sub-converges to w uniformly. Since F is uniformly continuous in (x,z), $E_2(u_n)$ tends to $E_2(w)$, contradiction holds.

According to Proposition 8.18, E is weakly sequentially lower semicontinuous. In order to apply

Theorem 8.16, it remains to verify that E is coercive. Using (8.3) we have

$$\begin{aligned} E(u) &= \frac{1}{2}\int u_x^2 - \int F(x,u) \\ &\geq \frac{1}{2}\int u_x^2 - C_1\int (|u| + \frac{|u|^{\gamma+1}}{\gamma+1}) \\ &\geq \frac{1}{2}\int u_x^2 - C_1\Big(\int u^2\Big)^{\frac{1}{2}} - \frac{C_1}{\gamma+1}\Big(\int u^2\Big)^{\frac{\gamma+1}{2}} \\ &\geq \frac{1}{2}\int u_x^2 - C_1\Big(\int u_x^2\Big)^{\frac{1}{2}} - \frac{C_1}{\gamma+1}\Big(\int u_x^2\Big)^{\frac{\gamma+1}{2}} \end{aligned}$$

As $\gamma \in (0,1)$, it is clear from this estimate that $E(u)$ tends to ∞ as $\|u\|_b$ becomes unbounded. Hence E is coercive.

Applying Theorem 8.16 we have the following result.

Proposition 8.20. *Under (8.3), the minimization problem*

$$\inf\{E(u) :\ u \in \mathcal{X}\}$$

has a minimum. It is a solution to (8.2) provided it is in $C^2[a,b]$.

Although it is possible to show that the minimum is really C^2, we will not do it here. Since Hilbert posed the regularity of solutions to variational problems as one of his famous problems in mathematics in 1900, there have been great advances on this topic. We will not step into this highly developed area in these notes but refer the interested reader to Folland's book for a taste. An advantage of the variational approach is that more information concerning the solution can be obtained. For instance, since the solution is a minimum, we have the inequality $E''(u)(\varphi,\varphi) \geq 0$, that is,

$$\int \big(\varphi_x^2 - f_z(x,u)\varphi^2\big) \geq 0,$$

for every φ in X, of course, provided f is differentiable in its second component.