Functional Analysis

A FIRST COURSE

M. Thamban Nair

FUNCTIONAL ANALYSIS
A FIRST COURSE

FUNCTIONAL ANALYSIS
A FIRST COURSE

M. THAMBAN NAIR
Department of Mathematics
Indian Institute of Technology Madras
Chennai

PHI Learning Private Limited
Delhi-110092
2014

₹ **350.00**

FUNCTIONAL ANALYSIS—A FIRST COURSE
by M. Thamban Nair

ISBN-978-81-203-1735-2

The export rights of this book are vested solely with the publisher.

Fourth Printing ... ... January, 2014

Published by Asoke K. Ghosh, PHI Learning Private Limited, Rimjhim House, 111, Patparganj Industrial Estate, Delhi-110092 and Printed by Glorious Printer, Delhi-110092.

Contents

Preface

This book is an outgrowth of my teaching of Functional Analysis for postgraduate students of mathematics for the last 15 years – from February 1987 to December 1995 at Goa University, and the remaining period at Indian Institute of Technology Madras.

Although there are several books on Functional Analysis by authors in India and abroad, students often find it difficult to choose a proper text which is student-friendly and helps them understand the basics. Somehow, the available books fail to motivate through examples or are not exhaustive enough to prepare them for further study. It is often the case with university education that a teacher who teaches a course need not necessarily be an expert in that field and, therefore, sometimes he finds himself in the same situation as the students. Therefore, this is an earnest attempt on my part to help both the students and the teachers to get acquainted with minimum amount of topics in Functional Analysis with maximum understanding of its concepts.

This book is intended to be an introductory text on Functional Analysis. The material in this text can be used for two courses of about 45 lectures each – a course on *Basic Functional Analysis* covering Chapters 1 - 7, and a course on *Basic Operator Theory* covering the remaining Chapters, 8 - 14.

Chapter 1 is meant only as a review, as its title indicates. However, the reader is encouraged to go through it, as some of the basic notions and notations that we use throughout this course are introduced here. The reader will be acquainted with many examples of infinite dimensional spaces which he will encounter on later occasions.

Some of the important, but simple to follow, theorems in the setting of Hilbert spaces, such as *best approximation theorems*, *projection theorem*, and *Riesz representation theorem*, are proved in Chapters 2 and 3. The concept of a closed operator, and the need to prove *closed group theorem* are discussed in Chapter 3, while the proof and more results related to closed graph theorem are given in Chapter 7. Also, a discussion on sequence of operators considered in Chapter 3 enables to give a sufficient condition for the convergence of a sequence of quadrature formulas, without having to wait up to Chapter 6 for *uniform boundedness principle* to have its converse as well.

Some results, which are special to inner product spaces and Hilbert spaces, are considered in Chapter 4, whereas the three important theorems of Functional Analysis, namely, the *Hahn-Banach theorem, the uniform boundedness principle, and closed graph theorem,* are proved in Chapters 5 - 7.

Some notions such as separability, reflexivity, and weak convergence are discussed in Chapter 8. These lead to the *Enberlein-Shmulyan* theorem, and a best approximation theorem in a reflexive space.

The results proved on compact operators in Chapter 9 are used effectively in later chapters on spectral results and operator equation. In Chapter 10, the spectral theory is initiated with the most elementary notions of eigenvalues and approximate eigenvalues. Many useful results are proved at this level, specially for compact operators.

In Chapter 11, the notion of adjoint of an operator between Hilbert spaces is introduced, and some basic properties of normal, self-adjoint, and unitary operators are analyzed.

Spectral results with the underlying space as Hilbert space are considered in Chapter 12. It is shown that some of the results available for compact operators in the setting of Banach spaces have simpler proofs if they are either normal or self-adjoint on a Hilbert space. Also, the notions of integration of operator-valued functions and spectral projection introduced in Chapter 10 are found useful in proving that an isolated spectral value of a normal operator is an eigenvalue. This result is usually derived as a consequence of spectral theorem for a self-adjoint operator.

Chapter 13 starts with spectral representation of a compact self-adjoint operator. As an application of this, the singular value representation of any compact operator on a Hilbert space is given. Although spectral theorem for a self-adjoint operator is stated after introducing all the necessary background material, its detailed proof is omitted. However, the main steps involved in the proof are indicated.

In Chapter 14, certain issues pertaining to the solution of an operator equation have been discussed. Results in this chapter show how some of the important results of functional analysis can be used for representation and approximation of solutions of operator equations.

The source for the materials in the text is my understanding of the subject covered in such text books as those by Bachman and Naricci, Goffman and Pedrick, Kato, Limaye, Simmons, and Taylor and Lay.

Since Functional Analysis is considered as a bridge between abstract mathematics and applied mathematics, I have endeavoured to provide examples and illustrations which are useful to numerical analysis and solution of operator equations. Most of the examples are designed and discussed in such a way that they not only illustrate a particular result or a sequence of results, but also point towards the limitations of the previous results so that subsequent stronger results follow. Specific remarks after theorems and examples often satisfy many of the questions which the readers may ask themselves, and instill a hope that such questions would be addressed in future sections. Also, I believe that insertions such as *Why, How, Verify, Show* etc. in the text would encourage the reader to be inquisitive and to think independently on the topics under discussion.

Exercises are interspersed throughout the book. They are meant for students to work out themselves, perhaps with some hints from the teacher, to enhance their understanding of the concepts. Additional exercises are given as problems at the end of each chapter which the students can take as home assignments or present them during tutorial hours.

I would be very grateful for critical comments, corrections, and suggestions from the users of this book.

M. Thamban Nair

Acknowledgements

I am indebted to Professor B.V. Limaye who was not only the supervisor of my Ph.D. thesis and a research collaborator, but also a friend, philosopher and guide. My meeting with him in the summer of 1981 in the lounge of Tata Institute of Fundamental Research, Mumbai, was a turning point in my life, both mathematically and in many other ways. My mathematical writings have been greatly influenced by him.

Other people who have influenced me indirectly in the exposition of topics in this book are my research collaborators Professor R. S. Anderssen of Australian National University and Professor E. Schock of University of Kaiserslautern (Germany).

Many of my colleagues and friends were always forthcoming in their many useful suggestions. In this regard, I specially thank my colleague Dr. Arindama Singh who carefully read through most of the chapters and offered valuable comments.

The help rendered by P.S. Srinivasan (Research Scholar), by introducing me to some of the subtle nuances in LaTeX in the book-style and by providing me with all the necessary style-files, is gratefully acknowledged. Also, I am grateful to the Publishers, Prentice Hall of India, specially their editors, for their valuable suggestions in the editorial and production aspects of the book.

Finally, I thank my wife Sunita for her constant support and encouragement throughout the course of preparation of this book, and my mother Meloth Parvathi Amma for her blessings.

M. Thamban Nair

Note to the Reader

A few words about some of the usual set-theoretic notations that we use throughout the book, and about the way mathematical results, examples, and exercises and remarks have been numbered:

Set-theoretic notations. We shall use standard set-theoretic notations such as

$$\cup, \ \cap, \ \subseteq, \ \subset, \ \in$$

to denote 'union', 'intersection', 'subset of', 'proper subset of', 'belong(s) to', respectively. For sets S_1 and S_2, the set $\{x \in S_1 : x \notin S_2\}$ is denoted by $S_1 \backslash S_2$. If f is a function with domain S_1 and codomain S_2, then we use the notation $f : S_1 \to S_2$, and it is also called a 'map' from S_1 to S_2.

Also, we use the following standard notations and symbols without defining them explicitly:

$\mathbb{N}$	:	set of all positive integers
$\mathbb{Z}$	:	set of all integers
$\mathbb{R}$	:	set of all real numbers
$\mathbb{C}$	:	set of all complex numbers
$:=$	:	is defined by
$\forall$	:	for all
$\exists$	:	there exists or there exist
$\Longrightarrow$	:	implies or imply
$\Longleftrightarrow$	:	if and only if
$\mapsto$	:	maps to

For 'if and only if', sometimes we use the Halmos' symbol 'iff'.

To mark the end of a proof (of a Lemma, Proposition, Theorem, or Corollary), we use the symbol ▮, while □ is used for the end of an exercise. **Bold face** is used when a new terminology is defined, and *italics* are used to emphasize a terminology or statement.

About numbering. Four sequences of numbers have been used for specifying mathematical results, examples, exercises, and remarks, respectively. To be more precise, mathematical results are numbered consecutively within each chapter without distinguishing their status (such as lemmas, propositions, theorems, and corollaries), whereas examples, exercises, and remarks are numbered separately. For instance, if a corollary, a lemma, and a proposition come after Theorem 2.3, in that order, then they are numbered as Corollary 2.4, Lemma 2.5, Proposition 2.6, respectively. Example 2.5 comes after Example 2.4, Exercise 2.5 comes after Exercise 2.4, and Remark 2.5 comes after Remark 2.4. In all the above illustrations, number 2 refers to Chapter 2.

1

Review of Linear Algebra

We begin our course with some basic definitions, results and examples from linear algebra. Although we assume that the reader has already undergone a course in linear algebra, he is advised to have a quick look at all the portions in this chapter, as some of the basic notions and notations that we use throughout this course are introduced here.

1.1 Linear Spaces

We shall denote by $\mathbb{K}$ the field of real numbers or the field of complex numbers. If special emphasis is required, then the fields of real numbers and complex numbers will be denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively.

A **linear space** over $\mathbb{K}$ is a set X together with two operations

$$(x, y) \mapsto x + y \in X, \qquad (x, y) \in X \times X,$$
$$(\alpha, x) \mapsto \alpha x \in X, \qquad (\alpha, x) \in \mathbb{K} \times X,$$

called *addition* and *scalar multiplication*, respectively, satisfying the following axioms:

(a) $x + y = y + x \quad \forall x, y \in X.$

(b) $(x + y) + z = x + (y + z) \quad \forall x, y, z \in X.$

(c) $\exists\, \theta \in X$ such that $x + \theta = x \quad \forall x \in X.$

(d) $\forall x \in X, \exists\, \tilde{x} \in X$ such that $x + \tilde{x} = \theta.$

(e) $\alpha(x + y) = \alpha x + \alpha y \quad \forall \alpha \in \mathbb{K}, \forall x, y \in X.$

(f) $(\alpha + \beta)x = \alpha x + \beta x \quad \forall \alpha, \beta \in \mathbb{K}, \forall x \in X.$

(g) $(\alpha\beta)x = \alpha(\beta x) \quad \forall \alpha, \beta \in \mathbb{K}, \forall x \in X.$

(h) $1x = x \quad \forall x \in X.$

Linear spaces are also known as **vector spaces**, and elements of a vector space are sometimes called **vectors**. Elements of the field $\mathbb{K}$ (over which the vector space is defined) are often called **scalars**.

The element θ in axiom (c) in the definition of a linear space X is called the *zero* in X. Assuming that there will not be any ambiguity, we denote the zero element in X and the zero in $\mathbb{K}$ by the same symbol 0. Also, for $x \in X$, the element $\tilde{x}$ in the axiom (d) is called the *additive inverse* of x and is denoted by $-x$. It is an easy exercise to show that the zero element and the additive inverse of any element in a linear space are unique.

We observe that a linear space X, by definition, cannot be an empty set. It contains at least one element, viz., the zero element. If X contains at least one nonzero element, then it contains infinitely many elements: if x is a nonzero element in X, and if α, β are scalars such that $\alpha \neq \beta$, then $\alpha x \neq \beta x$. This is a consequence of axiom (h) (*Verify*).

Unless otherwise specified, we always assume that the linear space under discussion is *nontrivial*, i.e., it contains at least one nonzero element.

EXAMPLE 1.1 (i) The set $\mathbb{K}^n$ of all ordered n-tuples of scalars, is a linear space over $\mathbb{K}$ with addition and scalar multiplication defined coordinatewise, i.e., if

$$x = (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n, \qquad y = (\beta_1, \ldots, \beta_n) \in \mathbb{K}^n, \qquad \alpha \in \mathbb{K},$$

then

$$x + y := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n), \qquad \alpha x := (\alpha\alpha_1, \ldots, \alpha\alpha_n).$$

In this space,

$$-x = (-\alpha_1, \ldots, -\alpha_n), \qquad \theta = (0, \ldots, 0).$$

We may recall that an ordered n-tuple $(\alpha_1, \ldots, \alpha_n)$ is characterized by the property

$$(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n) \iff \alpha_i = \beta_i \quad \forall i = 1, \ldots, n.$$

(ii) For $n \in \{0, 1, 2, \ldots\}$, let $\mathcal{P}_n$ be the set of all polynomials of degree at most n, with coefficients in $\mathbb{K}$, i.e., $x \in \mathcal{P}_n$ if and only if x is of the form

$$x = a_0 + a_1 t + \cdots + a_n t^n$$

for some scalars $a_0, a_1 \ldots, a_n$. Then $\mathcal{P}_n$ is a linear space with addition and scalar multiplication defined as follows:

For $x = a_0 + a_1 t + \cdots + a_n t^n$, $y = b_0 + b_1 t + \cdots + b_n t^n$ in $\mathcal{P}_n$, and $\alpha \in \mathbb{K}$,

$$x + y := (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n,$$
$$\alpha x := \alpha a_0 + \alpha a_1 t + \cdots + \alpha a_n t^n.$$

The zero polynomial, i.e., the polynomial with all its coefficients zero, is the zero element of the space, and

$$-x = -a_0 - a_1 t - \cdots - a_n t^n.$$

(iii) Let $\mathcal{P}$ be the set of all polynomials with coefficients in $\mathbb{K}$, i.e., $x \in \mathcal{P}$ if and only if $x \in \mathcal{P}_n$ for some $n \in \{0, 1, 2, \ldots\}$. For $x, y \in \mathcal{P}$, let n, m be such that $x \in \mathcal{P}_n$ and $y \in \mathcal{P}_m$. Then we have $x, y \in \mathcal{P}_k$, where $k = \max\{n, m\}$. Hence, we can define $x + y$ and αx for $\alpha \in \mathbb{K}$ as in $\mathcal{P}_k$. With this addition and scalar multiplication, it follows that $\mathcal{P}$ is a linear space.

(iv) Let $X = \mathbb{K}^{m \times n}$ be the set of all $m \times n$ matrices with entries in $\mathbb{K}$. Then X is a linear space with respect to the addition and scalar multiplication defined as follows: For $A = (a_{ij})$, $B = (b_{ij})$ in X, and $\alpha \in \mathbb{K}$,

$$A + B := (a_{ij} + b_{ij}), \qquad \alpha A := (\alpha a_{ij}).$$

In this space, $-A = (-a_{ij})$, and the matrix with all its entries are zeroes is the zero element.

(v) Let X be the set of all scalar sequences. For (α_n) and (β_n) in X and $\alpha \in \mathbb{K}$, we define

$$(\alpha_n) + (\beta_n) = (\alpha_n + \beta_n), \qquad \alpha(\alpha_n) = (\alpha\alpha_n).$$

With this addition and scalar multiplication, X is a linear space with its zero element as the sequence of zeroes and $-(\alpha_n) = (-\alpha_n)$.

(vi) Let Ω be a metric space and $C(\Omega)$ be the set of all $\mathbb{K}$-valued bounded continuous functions defined on Ω. For $x, y \in C(\Omega)$ and $\alpha \in \mathbb{K}$, we define $x + y$ and αx *pointwise*, i.e.,

$$(x + y)(t) = x(t) + y(t), \qquad (\alpha x)(t) = \alpha x(t) \quad \forall t \in \Omega.$$

Then it can be shown (*Verify*) that $x + y$, $\alpha x \in C(\Omega)$, and $C(\Omega)$ is a linear space with zero element as the zero function, and additive inverse of $x \in C(\Omega)$ as $-x$ defined by $(-x)(t) = -x(t)$, $t \in \Omega$.

NOTATION: If $\Omega = [a, b]$, we shall denote the space $C(\Omega)$ by $C[a, b]$. In case we want to emphasis that the scalar field is $\mathbb{R}$, then we shall write $C([a, b], \mathbb{R})$ in place of $C[a, b]$.

(vii) For $k \in \mathbb{N}$, let $C^k[a, b]$ be the set of all $\mathbb{K}$-valued functions defined on $[a, b]$ such that for each $j \in \{1, \ldots, k\}$, its j-th derivative $x^{(j)}$ of x exists and $x^{(j)} \in C[a, b]$. It can be seen that $C^k[a, b]$ is a linear space with respect to the addition and scalar multiplication defined as in $C[a, b]$.

(viii) Let $\mathcal{R}[a, b]$ be the set of all $\mathbb{K}$-valued Riemann integrable functions on $[a, b]$. Recall that a function $x : [a, b] \to \mathbb{C}$ is Riemann integrable if and only if its real part and imaginary parts are Riemann integrable. From the theory of Riemann integration, it follows that if $x, y \in \mathcal{R}[a, b]$ and $\alpha \in \mathbb{K}$, then $x + y$ and αx defined pointwise belongs to $\mathcal{R}[a, b]$. Thus, we see that $\mathcal{R}[a, b]$ is a linear space.

(ix) Now we give a general example. For a nonempty set S, let $\mathcal{F}(S, \mathbb{K})$ be the set of all functions from S into $\mathbb{K}$. For $x, y \in \mathcal{F}(S, \mathbb{K})$ and $\alpha \in \mathbb{K}$, let $x + y$ and αx be defined pointwise, i.e.,

$$(x + y)(s) = x(s) + y(s), \qquad (\alpha x)(s) = \alpha x(s) \quad \forall s \in S.$$

Let $-x$ and θ be defined by

$$(-x)(s) = -x(s), \qquad \theta(s) = 0 \quad \forall s \in S.$$

Then it is easy to see that $\mathcal{F}(S, \mathbb{K})$ is a linear space over $\mathbb{K}$.

For showing that $\mathcal{F}(S, \mathbb{K})$ is a linear space, what one essentially requires is the linear structure on $\mathbb{K}$. Thus, in a similar fashion, we can show that $\mathcal{F}(S, Y)$, the set of all functions from S into a linear space Y, is a linear space.

Consider the following particular cases:

(a) Let $S = \{1, \ldots, n\}$. Then, it is easily seen that the function $J : \mathcal{F}(S, \mathbb{K}) \to \mathbb{K}^n$ defined by

$$J(x) = (x(1), \ldots, x(n)), \quad x \in X,$$

is bijective. Moreover, it is seen that for every $x, y \in \mathcal{F}(S, \mathbb{K})$ and $\alpha \in \mathbb{K}$,

$$J(x + y) = J(x) + J(y), \qquad J(\alpha x) = \alpha J(x).$$

Such a map is called a *linear operator.* Linear operators will be considered in more detail in the next section. Here we give only its definition.

Let X and Y be linear spaces. A function $A : X \to Y$ is called a **linear operator** or a **linear transformation** if

$$A(x+y) = A(x) + A(y), \qquad A(\alpha x) = \alpha A(x)$$

for all $x, y \in X$, $\alpha \in \mathbb{K}$.

A bijective linear transformation is sometimes called a **linear isomorphism.** Thus, if there is a linear isomorphism $A : X \to Y$ between linear spaces X and Y, then as far as their linear structures are concerned, they are indistinguishable. Hence, we may regard them as same, upto a linear isomorphism. Thus, if $S = \{1, \ldots, n\}$, then the linear spaces $\mathcal{F}(S, \mathbb{K})$ and $\mathbb{K}^n$ can be considered as same. We may say that these two spaces are same upto isomorphism. With this identification in mind, we shall denote the j-th entry α_j of an element $x = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{K}^n$ by $x(j)$ for $j = 1, \ldots, n$.

Similarly, if S is a set with n elements, say $S = \{s_1, \ldots, s_n\}$, then also $\mathcal{F}(S, \mathbb{K})$ can be identified with $\mathbb{K}^n$ by the map

$$x \mapsto (x(s_1), \ldots, x(s_n)), \qquad x \in \mathcal{F}(S, \mathbb{K}).$$

(b) Next, suppose $S = \mathbb{N}$, the set of all positive integers. Then $\mathcal{F}(S, \mathbb{K})$ can be identified with the set of all scalar sequences. The identification is given by

$$x \mapsto (x(1), x(2), \ldots), \qquad x \in \mathcal{F}(\mathbb{N}, \mathbb{K}).$$

With this identification, the n-th entry α_n of a sequence $x = (\alpha_n)$ is also denoted by $x(n)$.

Similarly, if S is a denumerable set, say $S = \{s_1, s_2, \ldots\}$, then $\mathcal{F}(S, \mathbb{K})$ can be identified with the set of all scalar sequences by the map

$$x \mapsto (x(s_1), x(s_2), \ldots), \qquad x \in \mathcal{F}(S, \mathbb{K}).$$

(c) Consider $S = \{1, \ldots, m\} \times \{1, \ldots, n\}$. Then the resulting linear space $\mathcal{F}(S, \mathbb{K})$ is in one-one correspondence with the set $\mathbb{K}^{m \times n}$ of all $m \times n$ matrices with entries in $\mathbb{K}$. The bijective map, in this case, is

$$x \mapsto A_x,$$

where A_x is the $m \times n$ matrix whose ij-th entry is $x(i, j)$.

Exercise 1.1 Show that the maps considered in (a), (b), (c) above are linear transformations. $\square$

EXAMPLE 1.1 (*cont.*) (x) Let J be an interval and $\mathcal{P}_n(J)$ be the set of all polynomials of degree at most n considered as $\mathbb{K}$-valued functions on J. Thus, $x \in \mathcal{P}_n(J)$ if and only if $x \in \mathcal{F}(J, \mathbb{K})$ and there exist scalars $a_0, a_1, \ldots, a_n$ such that

$$x(t) = a_0 + a_1 t + \cdots + a_n t^n, \qquad t \in J.$$

Then as in example (ii) above, $\mathcal{P}_n(J)$ is a linear space.

(xi) Let J be an interval and $\mathcal{P}(J) = \cup_{n=0}^{\infty} \mathcal{P}_n(J)$. Then as in example (iii) above, $\mathcal{P}(J)$ is a linear space.

NOTATION: If $J = [a, b]$, then we may write $\mathcal{P}_n(J)$ and $\mathcal{P}(J)$ as $\mathcal{P}_n[a, b]$ and $\mathcal{P}[a, b]$, respectively.

Next, we give two examples of linear spaces which are constructed out of some given linear spaces.

Product space

(xii) Let $X_1, \ldots, X_n$ be linear spaces. Then the *cartesian product*

$$X = X_1 \times \cdots \times X_n,$$

the set of all of n-tuples $(x_1, \ldots, x_n)$ with $x_j \in X_j$, $j \in \{1, \ldots, n\}$, is a linear space with respect to the addition and scalar multiplication defined by

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n),$$

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n)$$

with zero element $(0, \ldots, 0)$ and additive inverse of $x = (x_1, \ldots, x_n)$ defined by $-x = (-x_1, \ldots, -x_n)$.

The above linear space is called the **product space** of $X_1, \ldots X_n$.

As a particular example, the space $\mathbb{K}^n$ can be considered as the product space $X_1 \times \cdots \times X_n$ with $X_j = \mathbb{K}$ for $j = 1, \ldots, n$.

1.1.1 Span and Subspace

Let X be a linear space. If $u_1, \ldots, u_n$ are some elements of X, then, by a **linear combination** of $u_1, \ldots, u_n$, we mean an element in X of the form $\alpha_1 u_1 + \cdots + \alpha_n u_n$ with $\alpha_j \in \mathbb{K}$, $j = 1, \ldots, n$.

Let S be a subset of X. Then the set of all linear combinations of elements of S is called the **span** of S, and is denoted by $\operatorname{span} S$.

Thus, for $S \subseteq X$, $x \in \operatorname{span} S$ if and only if there exists $x_1, \ldots, x_n$ in S and scalars $\alpha_1, \ldots, \alpha_n$ such that $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$.

As a convention, span of the empty set is taken to be the singleton set $\{0\}$.

Remember! By a linear combination, we always mean a linear combination of a *finite number* of elements in the space. An expression of the form $\alpha_1 x_1 + \alpha_2 x_2 + \cdots$ with $x_1, x_2, \ldots$ in X and $\alpha_1, \alpha_2, \ldots$ in $\mathbb{K}$ has no meaning in a linear space, unless there is some additional structure which allows such expression.

NOTATION: If $x_1, \ldots, x_n$ are in a linear space X and $E \subseteq X$, then the span of $\{x_1, \ldots, x_n\} \cup E$ is denoted by span $\{x_1, \ldots, x_n; E\}$.

It is easy to see that, for any subset S of X, $\operatorname{span} S$ is a vector space with respect to the vector space operations on X. This motivates us to have the following definition:

A subset X_0 of a linear space X is called a **subspace** of X if X_0 is a vector space with respect to the vector space operations on X. A subspace which is a proper subset of the whole space is called a **proper subspace**.

Observe that for every subset S of X, we have $S \subseteq \operatorname{span} S$. But, the reverse inclusion need not hold.

Let X be a linear space. The following results can be easily verified:

(1) Let S be a subset of X. Then $\operatorname{span} S$ is the smallest subspace containing S, i.e., if X_0 is a subspace containing S, then $\operatorname{span} S \subseteq X_0$. Equivalently,

$$\operatorname{span} S = \bigcap \{Y : Y \text{ is a subspace of } X \text{ containing } S\}.$$

(2) A subset X_0 of X is a subspace if and only if $X_0 = \operatorname{span} X_0$.

(3) A subset X_0 of X is a subspace of X if and only if for every $x, y \in X_0$ and for every $\alpha \in \mathbb{K}$, $x + y \in X_0$ and $\alpha x \in X_0$.

(4) Suppose X_0 is a subspace of X and $x_0 \in X \setminus X_0$. Then for every $x \in \operatorname{span}\{x_0; X_0\}$, there exist a unique $\alpha \in \mathbb{K}$, $y \in X_0$ such that $x = \alpha x_0 + y$.

(5) Suppose Y is a subspace of a linear space X and Z is a subspace of Y. Then Z is a subspace of X.

EXAMPLE 1.2 (i) Let X be $C[a,b]$ or $C^k[a,b]$ for some $k \in \mathbb{N}$ and let $u_j(t) = t^{j-1}$ for $t \in [a,b]$, $j \in \mathbb{N}$. Then $\mathcal{P}_n[a,b]$ is the span of $\{u_1, \ldots, u_{n+1}\}$, and $\mathcal{P}[a,b]$ is the span of $\{u_1, u_2, \ldots\}$, and hence, they are subspaces of X.

(ii) Let $X = \mathbb{K}^n$ and for each $j \in \{1, \ldots, n\}$, let $e_j \in \mathbb{K}^n$ be the element with its j-th coordinate 1 and all other coordinates 0's. Then $\mathbb{K}^n$ is the span of $\{e_1, \ldots, e_n\}$.

NOTATION: For $(i,j) \in \mathbb{N} \times \mathbb{N}$, let

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

Thus, (ii) above, the i-th coordinate of e_j is δ_{ij} for $i, j = 1, \ldots, n$.

(iii) For $1 \leq k < n$, $Y = \{(\alpha_1, \ldots, \alpha_n) : \alpha_j = 0, j = k+1, \ldots, n\}$ is a subspace of $\mathbb{K}^n$ and it is the span of $\{e_1, \ldots, e_k\}$, where $e_j(i) = \delta_{ij}$ with $j = 1, \ldots, k$; $i = 1, \ldots, n$.

(iv) Let

$$c_{00} := \bigcup_{k=1}^{\infty} \{(\alpha_1, \alpha_2, \ldots) : \alpha_j = 0 \ \forall j \geq k\}.$$

Clearly, c_{00} is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{K})$, and it is the span of $\{e_1, e_2, \ldots\}$, where $e_j(i) = \delta_{ij}$ with $i, j \in \mathbb{N}$. Note that an element $x \in \mathcal{F}(\mathbb{N}, \mathbb{K})$ belongs to c_{00} if and only if there exists $N \in \mathbb{N}$ such that $x(n) = 0$ for all $n \geq N$.

(v) Let

$$\ell^1(\mathbb{N}) := \left\{ x \in \mathcal{F}(\mathbb{N}, \mathbb{K}) : \sum_{j=1}^{\infty} |x(j)| < \infty \right\},$$

the set of all absolutely summable ($\mathbb{K}$-valued) functions on $\mathbb{N}$. We show that $\ell^1(\mathbb{N})$ is a subspace of $\mathcal{F}(\mathbb{N}, \mathbb{K})$: For $x, y \in \ell^1(\mathbb{N})$ and $\alpha, \beta \in \mathbb{K}$, and $n \in \mathbb{N}$, we have

$$\begin{aligned} \sum_{j=1}^{n} |\alpha x(j) + \beta y(j)| &\leq |\alpha| \sum_{j=1}^{n} |x(j)| + |\beta| \sum_{j=1}^{n} |y(j)| \\ &\leq |\alpha| \sum_{j=1}^{\infty} |x(j)| + |\beta| \sum_{j=1}^{\infty} |y(j)| \\ &< \infty. \end{aligned}$$

Hence, by letting $n \to \infty$, we have $\sum_{j=1}^{\infty} |\alpha x(j) + \beta y(j)| < \infty$ so that $\alpha x + \beta y \in \ell^1(\mathbb{N})$.

Remark 1.1 Suppose S is a countable set. Then, as in the case of $\ell^1(\mathbb{N})$, we can still see that the set

$$\ell^1(S) := \Big\{x \in \mathcal{F}(S, \mathbb{K}) : \sum_{s \in S} |x(s)| < \infty\Big\},$$

the set of all absolutely summable ($\mathbb{K}$-valued) functions on S is a subspace of $\mathcal{F}(S, \mathbb{K})$. If S is a denumerable set, then by the sum $\sum_{s \in S} |x(s)|$, we mean $\sum_{j=1}^{\infty} |x(s_j)|$, for any enumeration $\{s_1, s_2, \ldots\}$ of S. This is justifiable, because we know from the theory of convergence of series that if (a_n) is a sequence of real numbers, and if (b_n) is a sequence obtained from (a_n) by some re-arrangements of its terms, then

$$\sum_{n=1}^{\infty} |a_n| < \infty \iff \sum_{n=1}^{\infty} |b_n| < \infty.$$

EXAMPLE 1.2 (*cont.*) (vi) For a nonempty set S, let

$$\ell^\infty(S) := \big\{x \in \mathcal{F}(S, \mathbb{K}) : \sup_{s \in S} |x(s)| < \infty\big\}.$$

Note that $\ell^\infty(S)$ is the set of all bounded functions on S. Thus, $x \in \ell^\infty$ if and only there exists $M_x > 0$ such that $|x(s)| \le M_x$ for all $s \in S$. We show that $\ell^\infty(S)$ is a subspace of $\mathcal{F}(S, \mathbb{K})$: To see this, let $x, y \in \ell^\infty(S)$ and $\alpha, \beta \in \mathbb{K}$. Suppose $M_x > 0$, $M_y > 0$ such that $|x(s)| \le M_x$ and $|y(s)| \le M_y$ for all $s \in S$. Then,

$$|\alpha x(s) + \beta y(s)| \le |\alpha| M_x + |\beta| M_y \quad \forall\, s \in S.$$

Thus, $\sup_{s \in S} |\alpha x(s) + \beta y(s)| < \infty$, and hence $\alpha x + \beta y \in \ell^\infty$.

In this example, if S is a finite set, then $\ell^\infty(S) = \mathcal{F}(S, \mathbb{K})$. But, if S is an infinite set, then $\ell^\infty(S)$ is a proper subspace of $\mathcal{F}(S, \mathbb{K})$. To see this, let $\{s_1, s_2, \ldots\}$ be a denumerable subset of S, and let $x \in \mathcal{F}(S, \mathbb{K})$ be defined by $x(s_j) = j$ for all $j \in \mathbb{N}$. Then we see that x does not belong to $\ell^\infty(S)$.

NOTATION: We shall use the notation ℓ^1 and ℓ^∞ for $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$, respectively. For a nonempty set S, it is also customary to use the notation $B(S)$ for $\ell^\infty(S)$, as it is the set of all bounded functions on S.

(vii) The set c_{00} introduced in example (iv) above is a subspace of $\ell^1(\mathbb{N})$, and the sets

$$\begin{aligned} c_0 &:= \{x \in \mathcal{F}(\mathbb{N},\mathbb{K}) : x(n) \to 0 \text{ as } n \to \infty\},\\ c &:= \{x \in \mathcal{F}(\mathbb{N},\mathbb{K}) : (x(n)) \text{ converges }\} \end{aligned}$$

are subspaces of $\ell^\infty(\mathbb{N})$. We observe that

$$c_{00} \subseteq \ell^1(\mathbb{N}) \subseteq c_0 \subseteq c \subseteq \ell^\infty(\mathbb{N}).$$

The above inclusions are, in fact, proper. To see this, let x, y, u, v in $\mathcal{F}(\mathbb{N},\mathbb{K})$ be defined

$$x(j) = (-1)^j, \qquad y(j) = \frac{j}{j+1}, \qquad u(j) = \frac{1}{j}, \qquad v(j) = \frac{1}{j^2},$$

for $j \in \mathbb{N}$. Then we see that

$$x \in \ell^\infty(\mathbb{N}) \setminus c, \qquad y \in c \setminus c_0,$$

$$u \in c_0 \setminus \ell^1(\mathbb{N}), \qquad v \in \ell^1(\mathbb{N}) \setminus c_{00}.$$

(viii) If Ω is a metric space, then $C(\Omega)$ is a subspace of $\ell^\infty(\Omega)$.

(ix) The linear space $C^k[a,b]$, $k \in \mathbb{N}$, is a subspace of $C[a,b]$.

(x) The linear space $C[a,b]$ is a subspace of $\mathcal{R}[a,b]$.

Sum of sets and subspaces

Let X be a linear space, $x \in X$, and E, E_1, E_2 be subsets of X. Then we define the following:

$$x + E := \{x + u : u \in E\},$$

$$E_1 + E_2 := \{x_1 + x_2 : x_1 \in E_1,\ x_2 \in E_2\}.$$

It can be seen that if X_1 and X_2 are subspaces of X, then their sum $X_1 + X_2$ is a subspace of X and it is, in fact, the span of $X_1 \cup X_2$. It is also easy to see (*Verify*) that if $X_1 \cap X_2 = \{0\}$, then every $x \in X_1 + X_2$ can be written uniquely as

$$x = x_1 + x_2 \quad \text{with} \quad x_1 \in X_1,\ x_2 \in X_2.$$

1.1.2 Quotient Space

Let us recall from set theory the definition of an *equivalence relation*.

A relation $\mathcal{R}$ on a set S is a subset of the cartesian product $S \times S$, and it is called an **equivalence relation** if it is

(a) reflexive, i.e., $(x, x) \in \mathcal{R}\ \forall x \in S$,

(b) symmetric, i.e., $(x, y) \in \mathcal{R}$ implies $(y, x) \in \mathcal{R}$, and

(c) transitive, i.e., $(x, y) \in \mathcal{R}$ and $(y, z) \in \mathcal{R}$ imply $(x, z) \in \mathcal{R}$.

If $\mathcal{R}$ is an equivalence relation on a set S and if $x \in S$, then the set

$$[x] := \{u \in S : (x, u) \in \mathcal{R}\}$$

is called the **equivalence class** of x with respect to $\mathcal{R}$, and the set of all equivalence classes is called the *quotient set*, and it is denoted by $S/\mathcal{R}$.

Usually, an equivalence relation is denoted by $\sim$, and we shall write $x \sim y$ whenever (x, y) belongs to the relation $\sim$.

Now, let X be a linear space and X_0 be a subspace of X. For $x, y \in X$, define

$$x \sim y \iff x - y \in X_0.$$

Then it is easily seen that $\sim$ is an equivalence relation on X, and the equivalence class $[x]$ of $x \in X$ is the set

$$x + X_0 = \{x + u : u \in X_0\}.$$

Moreover, the quotient set $X/\sim$ is a linear space with respect to the addition and scalar multiplication defined by

$$[x] + [y] = [x + y], \qquad \alpha[x] = [\alpha x]$$

for all $x, y \in X$ and $\alpha \in \mathbb{K}$ (*Verify*). The class $[0]$ is the zero element of the space and $-[x] := [-x]$ for every $x \in X$. This linear space $X/\sim$ is called the **quotient space** of X corresponding to the subspace X_0, and it is usually denoted by X/X_0.

1.1.3 Linear Independence, Basis and Dimension

Let X be a linear space.

A subset E of X is said to be **linearly dependent** if there exists $u \in E$ such that $u \in \operatorname{span}(E \setminus \{u\})$.

A subset of X is said to be **linearly independent** in X if it is not linearly dependent.

Thus, a subset E of X is linearly dependent if and only if there exists $\{u_1, \ldots, u_n\} \subseteq E$ and scalars $\alpha_1, \ldots, \alpha_n$, with at least one of them nonzero, such that

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0;$$

and a subset E of X is linearly independent if and only if for every finite subset $\{u_1, \ldots, u_n\}$ of E,

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \implies \alpha_i = 0 \quad \forall i = 1, \ldots, n.$$

If $\{u_1, \ldots, u_n\}$ is a linearly independent (respectively, dependent) subset of a linear space X, then we may also say that $u_1, \ldots, u_n$ are linearly independent (respectively, dependent) in X.

Note that a linearly dependent set cannot be empty. In other words, the empty set is linearly independent!

Exercise 1.2 Let X be a linear space.

(i) Show that a subset $\{u_1, \ldots, u_n\}$ of X is linearly dependent if and only if there exists a nonzero $(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{K}^n$ such that $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$.

(ii) Show that a subset $\{u_1, \ldots, u_n\}$ of X is linearly independent if and only if the function $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_1 u_1 + \cdots + \alpha_n u_n$ from $\mathbb{K}^n$ into X is injective.

(iii) Show that if $E \subseteq X$ is linearly independent in X, then $0 \notin E$.

(iv) Show that if $E \subseteq X$ is linearly dependent in X, then every superset of E is also linearly dependent.

(v) Show that if $E \subseteq X$ is linearly independent in X, then every subset of E is also linearly independent.

(vi) Show that if $\{u_1, \ldots, u_n\}$ is a linearly independent subset of X, and if Y is a subspace of X such that span $\{u_1, \ldots, u_n\} \cap Y = \varnothing$, then every x in the span of $\{u_1, \ldots, u_n, Y\}$ can be written uniquely as $x = \alpha_1 u_1 + \cdots + \alpha_n u_n + y$ with $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, $y \in Y$.

(vii) Show that if E_1 and E_2 are linearly independent subsets of X such that $(\operatorname{span} E_1) \cap (\operatorname{span} E_2) = \{0\}$, then $E_1 \cup E_2$ is linearly independent. $\square$

A subset E of X is said to be a **Hamel basis** of X if it is linearly independent and $\operatorname{span} E = X$.

When there are more structures on a linear space, we shall come across some other types of bases also, such as *Schauder basis* and *orthonormal basis*.

Throughout this book, a Hamel basis is simply referred to as a **basis**.

EXAMPLE 1.3 (i) For each $j \in \{1, \ldots, n\}$, let $e_j \in \mathbb{K}^n$ be such that $e_j(i) = \delta_{ij}$, $i, j = 1, \ldots, n$. Then $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{K}^n$.

The above basis of $\mathbb{K}^n$ is often called the **standard basis** of $\mathbb{K}^n$.

(ii) Let $u_j(t) = t^{j-1}$ for $t \in [a, b]$, $j \in \mathbb{N}$. Then $\{u_1, \ldots, u_{n+1}\}$ is a basis of $\mathcal{P}_n[a, b]$, and $\{u_1, u_2, \ldots\}$ is a basis of $\mathcal{P}[a, b]$.

(iii) For $i = 1, \ldots, m$; $j = 1, \ldots, n$, let E_{ij} be the $m \times n$ matrix with its (i, j)-th entry as 1 and all other entries 0. Then

$$\{E_{ij} : i = 1, \ldots, m;\ j = 1, \ldots, n\}$$

is a basis of $\mathbb{K}^{m \times n}$.

(iv) For $\lambda \in [a, b]$, let $u_\lambda(t) = \exp(\lambda t)$, $t \in [a, b]$. Then it is seen that $\{u_\lambda : \alpha \in [a, b]\}$ is an uncountable linearly independent subset of $C[a, b]$.

Clearly, a linearly independent subset of a subspace remains linearly independent in the whole space. Thus, the set $\{u_1, u_2, \ldots\}$ in Example 1.3(ii) is linearly independent in $C[a, b]$ and $\mathcal{F}([a, b], \mathbb{K})$.

Exercise 1.3 If $\{u_1, \ldots, u_n\}$ is a basis of a linear space X, then show that for every $x \in X$, there exists a unique n-tuple $(\alpha_1, \ldots, \alpha_n)$ of scalars such that $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$. □

A linear space X is said to be a **finite dimensional space** if there is a finite basis for X. A linear space which is not a finite dimensional space is called an **infinite dimensional space**.

Theorem 1.1 *If a linear space has a finite spanning set, then it has a finite basis.*

Proof. Let X be a linear space and S be a finite subset of X such that $\operatorname{span} S = X$. If S itself is linearly independent, then we are through. Suppose S is not linearly independent. Then there exists $u_1 \in S$ such that $u_1 \in \operatorname{span}(S \setminus \{u_1\})$. Let $S_1 = S \setminus \{u_1\}$. Clearly,

$$\operatorname{span} S_1 = \operatorname{span} S = X.$$

If S_1 is linearly independent, then we are through. Otherwise, there exists $u_2 \in S_1$ such that $u_2 \in \operatorname{span}(S_1 \setminus \{u_2\})$. Let $S_2 = S \setminus \{u_1, u_2\}$. Then, we have

$$\operatorname{span} S_2 = \operatorname{span} S_1 = X.$$

If S_2 is linearly independent, then we are through. Otherwise, continue the above procedure. This procedure will stop after a finite number of steps, as the original set S is a finite set, and we end up with a subset S_k of S which is linearly independent and $\operatorname{span} S_k = X$. ∎

By definition, an infinite dimensional space cannot have a finite basis. Is it possible for a finite dimensional space to have an infinite basis, or an infinite linearly independent subset? The answer is, as expected, negative. In fact, we have the following result.

Theorem 1.2 *Let X be a finite dimensional linear space with a basis consisting of n elements. Then every subset of X with more than n elements is linearly dependent.*

Proof. Let $\{u_1, \ldots, u_n\}$ be a basis of X, and $\{x_1, \ldots, x_{n+1}\} \subset X$. We show that $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent.

If $\{x_1, \ldots, x_n\}$ is linearly dependent, then $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent. So, let us assume that $\{x_1, \ldots, x_n\}$ is linearly independent. Now, since $\{u_1, \ldots u_n\}$ is a basis of X, there exist scalars $\alpha_1, \ldots, \alpha_n$ such that

$$x_1 = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Since $x_1 \neq 0$, one of $\alpha_1, \ldots, \alpha_n$ is nonzero. Without loss of generality, assume that $\alpha_1 \neq 0$. Then we have

$$u_1 \in \operatorname{span}\{x_1, u_2, \ldots, u_n\}$$

so that

$$X = \operatorname{span}\{u_1, u_2, \ldots, u_n\} = \operatorname{span}\{x_1, u_2, \ldots, u_n\}.$$

Let $\alpha_1^{(2)}, \ldots, \alpha_n^{(2)}$ be scalars such that

$$x_2 = \alpha_1^{(2)} x_1 + \alpha_2^{(2)} u_2 + \cdots + \alpha_n^{(2)} u_n.$$

Since $\{x_1, x_2\}$ is linearly independent, at least one of $\alpha_2^{(2)}, \ldots, \alpha_n^{(2)}$ is nonzero. Without loss of generality, assume that $\alpha_2^{(2)} \neq 0$. Then we have

$$u_2 \in \operatorname{span}\{x_1, x_2, u_3, \ldots, u_n\}$$

so that

$$X = \operatorname{span}\{x_1, u_2, \ldots, u_n\} = \operatorname{span}\{x_1, x_2, u_3, \ldots, u_n\}.$$

Now, let $1 \leq k \leq n-1$ be such that

$$X = \operatorname{span}\{x_1, x_2, \ldots, x_k, u_{k+1}, \ldots, u_n\}.$$

Suppose $k < n-1$. Then there exist scalars $\alpha_1^{(k+1)}, \ldots, \alpha_n^{(k+1)}$ such that

$$x_{k+1} = \alpha_1^{(k+1)} x_1 + \cdots + \alpha_k^{(k+1)} x_k + \alpha_{k+1}^{(k+1)} u_{k+1} + \cdots + \alpha_n^{(k+1)} u_n.$$

Since $\{x_1, \ldots, x_{k+1}\}$ is linearly independent, at least one of the scalars $\alpha_{k+1}^{(k+1)}, \ldots, \alpha_n^{(k+1)}$ is nonzero. Without loss of generality, assume that $\alpha_{k+1}^{(k+1)} \neq 0$. Then we have

$$u_{k+1} \in \operatorname{span}\{x_1, \ldots, x_{k+1}, u_{k+2}, \ldots, u_n\}$$

so that

$$\begin{aligned} X &= \operatorname{span}\{x_1, \ldots, x_k, u_{k+1}, \ldots, u_n\} \\ &= \operatorname{span}\{x_1, \ldots, x_{k+1}, u_{k+2}, \ldots, u_n\} \end{aligned}$$

Thus, the above procedure leads to $X = \operatorname{span}\{x_1, \ldots, x_{n-1}, u_n\}$ so that there exist scalars $\alpha_1^{(n)}, \ldots, \alpha_n^{(n)}$ such that

$$x_n = \alpha_1^{(n)} x_1 + \cdots + \alpha_{n-1}^{(n)} x_{n-1} + \alpha_n^{(n)} u_n.$$

Since $\{x_1, \ldots, x_n\}$ is linearly independent, it follows that $\alpha_n^{(n)} \neq 0$. Hence,

$$u_n \in \operatorname{span}\{x_1, \ldots, x_n\}.$$

Consequently,

$$X = \operatorname{span}\{x_1, x_2, \ldots, x_{n-1}, u_n\} = \operatorname{span}\{x_1, x_2, \ldots, x_{n-1}, x_n\}.$$

Thus, $x_{n+1} \in \operatorname{span}\{x_1, \ldots, x_n\}$, showing that $\{x_1, \ldots, x_{n+1}\}$ is linearly dependent. ∎

The following corollaries are easy consequences of Theorem 1.2.

Corollary 1.3 *If X is a finite dimensional linear space, then any two bases of X have the same number of elements.*

Corollary 1.4 *If a linear space contains an infinite linearly independent subset, then it is an infinite dimensional space.*

Corollary 1.5 *If (a_{ij}) is an $m \times n$ matrix with $a_{ij} \in \mathbb{K}$ and $n > m$, then there exists a nonzero $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ such that*

$$a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n = 0, \qquad i = 1, \ldots, m.$$

Exercise 1.4 (i) Write detailed proofs of Corollaries 1.3 – 1.5.

(ii) Assuming Corollary 1.5, give an alternate proof for Theorem 1.2. □

By Corollary 1.5, we see that if $\mathbf{A}$ is an $m \times n$ matrix with entries from $\mathbb{K}$ and $n > m$, then there exists an $n \times 1$ nonzero matrix $\mathbf{x}$ such that

$$\mathbf{Ax} = \mathbf{0},$$

where $\mathbf{0}$ is the $m \times 1$ zero matrix.

An $n \times 1$ matrix is also called an n-**vector**.

In view of Corollaries 1.3 and 1.4, we define the **dimension** of a finite dimensional linear space X to be the number of elements in a basis of X, and this number is denoted by $\dim X$. If X is infinite dimensional, then its dimension is defined to be infinity and we write $\dim X = \infty$.

EXAMPLE 1.4 (i) Suppose S is a finite set consisting of n elements. Then $\mathcal{F}(S, \mathbb{K})$ is of dimension n. To see this, let $S = \{s_1, \ldots, s_n\}$, and for each $j \in \{1, \ldots, n\}$, define $f_j \in \mathcal{F}(S, \mathbb{K})$ by

$$f_j(s_i) = \delta_{ij}, \qquad i \in \{1, \ldots, n\}.$$

Then the set $\{f_1, \ldots, f_n\}$ is a basis of $\mathcal{F}(S, \mathbb{K})$: Clearly,

$$\sum_{j=1}^{n} \alpha_j f_j = 0 \implies \alpha_i = \sum_{j=1}^{n} \alpha_j f_j(s_i) = 0 \quad \forall i.$$

Thus, $\{f_1, \ldots, f_n\}$ is linearly independent. To see that it spans the space, it is enough to note that

$$f = \sum_{j=1}^{n} f(s_j) f_j \quad \forall f \in \mathcal{F}(S, \mathbb{K}).$$

(ii) It is seen that the set $\{e_1, e_2, \ldots, \} \subseteq \mathcal{F}(\mathbb{N}, \mathbb{K})$ with $e_j(i) = \delta_{ij}$ is a linearly independent subset of the spaces $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$. Hence, it follows that $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$ are infinite dimensional spaces.

(iii) We have already observed in Example 1.3(ii) that $\{u_1, u_2, \ldots, \}$ with $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$, is a basis of $\mathcal{P}[a, b]$. In particular, $\mathcal{P}[a, b]$ is an infinite dimensional space. Since $\mathcal{P}[a, b]$ is a subspace of the spaces $C^k[a, b]$ for every $k \in \mathbb{N}$, the space $C^k[a, b]$ for each $k \in \mathbb{N}$ is infinite dimensional.

By Theorem 1.1 we know that every space having a finite spanning set has a finite basis. Does every linear space have a basis? Of course, if the space is the trivial space, that is the zero space, then it has no basis. What about for a nontrivial space? The answer is affirmative. For its proof we need the statement of *Zorn's lemma*. Before stating it let us recall some definitions from set theory.

Partial order and Zorn's lemma

A (binary) relation $\mathcal{R}$ on a set S is called a **partial order** if it is

(a) reflexive, i.e., $(x, x) \in \mathcal{R} \quad \forall x \in S$,

(b) antisymmetric, i.e., $(x, y) \in \mathcal{R}$, $(y, x) \in \mathcal{R}$ imply $x = y$ $\forall x, y \in S$, and

(c) transitive, i.e., $(x, y) \in \mathcal{R}$, $(y, z) \in \mathcal{R}$ imply $(x, z) \in \mathcal{R}$ $\forall x, y \in S$.

A set, together with a partial order, is called a **partially ordered set**. A partial order is usually denoted by $\preccurlyeq$.

A subset T of a partially ordered set with partial order $\preccurlyeq$ is called a **totally ordered set** if for every $x, y \in T$, either $x \preccurlyeq y$ or $y \preccurlyeq x$.

Suppose S is a partially ordered set with partial order $\preccurlyeq$. An element $x \in S$ is called a **maximal element** of S if for $y \in S$, $x \preccurlyeq y$ implies $y = x$. An element $x \in S$ is called an **upper bound** of $T \subseteq S$ if $t \preccurlyeq x$ for every $t \in T$.

EXAMPLE 1.5 (i) Let $X = \{re^{i\theta} : 0 \le r \le 1,\ 0 \le \theta < 2\pi\}$. Here $re^{i\theta}$ represents a point $(r\cos\theta, r\sin\theta)$ in the plane $\mathbb{R}^2$. On X define the relation $\preccurlyeq$ by

$$r_1 e^{i\theta_1} \preccurlyeq r_2 e^{i\theta_2} \iff \theta_1 = \theta_2, \quad r_1 \le r_2.$$

For each $\theta \in [0, 2\pi)$, $e^{i\theta}$ is a maximal element of X. Also, for each $\theta \in [0, 2\pi)$, the set $S_\theta = \{re^{i\theta} : 0 \le r \le 1\}$ is a totally ordered subset of X and the element $e^{i\theta}$ is an upper bound of S_θ.

(ii) Let $Y = \{re^{i\theta} : 0 \leq r < 1, \ 0 \leq \theta < 2\pi\}$. Then Y is a partially ordered set with partial order as in (i). This partially ordered set has no maximal element. Also, for each $\theta \in [0, 2\pi)$, the set $\{re^{i\theta} : 0 \leq r < 1\}$ is a totally ordered subset of X which has no upper bound in Y.

(iii) Let $Z = Y \cup \{1\}$, where Y is as in (ii). Then Z is a partially ordered set with partial order as in (i). The element 1, a maximal element of Z, is the only maximal element. Also, for each $\theta \in [0, 2\pi)$, if S_θ is as in (ii), then S_θ has an upper bound if and only if $\theta = 0$, and in that case, 1 is the upper bound of S_0.

Zorn's lemma: If X is a nonempty partially ordered set with partial order $\preccurlyeq$ such that its every totally ordered subset has an upper bound, then X contains a maximal element.

Theorem 1.6 *Every (nontrivial) linear space has a basis. In fact, if E_0 is a linearly independent subset of linear space X, then there exists a basis E for X such that $E_0 \subseteq E$.*

Proof. Let $X \neq \{0\}$ be a linear space and $E_0 \subseteq X$ be a linearly independent subset of X. Consider the family $\mathcal{X}$ of all linearly independent subsets of X which contains E_0. Then $\mathcal{X}$ is a partially ordered set under the relation $\preccurlyeq$ of inclusion, i.e., for S_1, S_2 in $\mathcal{X}$,

$$S_1 \preccurlyeq S_2 \iff S_1 \subseteq S_2.$$

If $\mathcal{T}$ is a totally ordered subset of $\mathcal{X}$, then it is seen that $\cup\{S : S \in \mathcal{T}\}$ is an upper bound for $\mathcal{T}$. Hence, by Zorn's lemma, $\mathcal{X}$ has a maximal element, say E.

Clearly, since $E \in \mathcal{X}$, it is linearly independent. Hence, it is enough to show that $\operatorname{span} E = X$. Suppose there exists $x \in X$ such that $x \notin \operatorname{span} E$. Then we see that the set $\tilde{E} = E \cup \{x\}$ belongs to $\mathcal{X}$, and

$$E \subseteq \tilde{E}, \qquad E \neq \tilde{E},$$

contradicting the maximality of E. Thus, $\operatorname{span} E = X$. ∎

We have already observed in Corollary 1.3 that any two bases of a finite dimensional space have the same number of elements. Is there an analogous result for infinite dimensional spaces? The answer is given by the following theorem. For its proof, one may refer Goffman and Pedrick [15], Section 2.4.

Theorem 1.7 *Let X be a linear space. Then any two bases of X have the same cardinality.*

Using Theorems 1.6 and 1.7, we can infer that $C[a, b]$ cannot have a denumerable basis. To see this, recall from Example 1.3(iv) that $C[a, b]$ has an uncountable linearly independent subset. This set can be extended to a basis E of $C[a, b]$, which has to be uncountable.

1.2 Linear Operators

Let X and Y be linear spaces. Recall that a function $A : X \to Y$ is said to be a **linear operator** from X to Y if

$$A(x + y) = A(x) + A(y),$$

$$A(\alpha x) = \alpha A(x)$$

for every $x, y \in X$ and for every $\alpha \in \mathbb{K}$.

For $x \in X$, we write the element $A(x)$ by Ax also.

We shall denote the set of all linear operators from a linear space X to a linear space Y by $\mathcal{L}(X, Y)$. If $X = Y$, then we write $\mathcal{L}(X, Y)$ by $\mathcal{L}(X)$.

On the set $\mathcal{L}(X, Y)$, define addition and scalar multiplication *pointwise*, i.e., for A, B in $\mathcal{L}(X, Y)$ and $\alpha \in \mathbb{K}$, linear operators $A + B$ and αA are defined by

$$(A + B)(x) = Ax + Bx,$$

$$(\alpha A)(x) = \alpha Ax$$

for all $x \in X$. Then it is seen that $\mathcal{L}(X, Y)$ is a linear space with its zero element as the zero operator $O : X \to Y$ defined by

$$Ox = 0 \quad \forall x \in X$$

and the additive inverse of $A \in \mathcal{L}(X, Y)$ is the operator $-A : X \to Y$ defined by

$$(-A)(x) = -Ax \quad \forall x \in X.$$

A linear operator with its codomain as the scalar field $\mathbb{K}$ is called a **linear functional** on X. The space $\mathcal{L}(X, \mathbb{K})$ of all linear functionals on X is called the **dual** of the space X.

Usually, linear functionals are denoted by small letters such as f, g, etc., whereas linear operators between general linear spaces are denoted by capital letters A, B, T, etc.

EXAMPLE 1.6 (i) Let $(a_{ij}) \in \mathbb{K}^{m\times n}$. For $x = (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, let

$$Ax = (\beta_1, \ldots, \beta_m), \qquad \beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j, \quad i = 1, \ldots, m.$$

Then $A : \mathbb{K}^n \to \mathbb{K}^m$ is a linear operator.

More generally, let X be an n-dimensional space and Y be an m-dimensional space. Let $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_m\}$ be a bases of X and Y, respectively. For $x = \sum_{j=1}^{n} \alpha_j u_j \in X$, define $A : X \to Y$ by

$$Ax = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}\alpha_j \Big) v_i.$$

Then A is a linear operator.

(ii) Let X be a finite dimensional space and let $\{u_1, \ldots, u_n\}$ be a basis of X. For $x = \alpha_1 u_1 + \cdots + \alpha_n u_n \in X$ let $f_j : X \to \mathbb{K}$ be defined by

$$f_j(x) = \alpha_j, \qquad j = 1, \ldots, n.$$

Then f_j is a linear functional for each $j \in \{1, \ldots, n\}$.

The linear functionals $f_1, \ldots, f_n$ defined above are called the **coordinate functionals** on X with respect to the basis U. We observe that

$$f_i(u_j) = \delta_{ij} \quad \forall i, j = 1, \ldots, n.$$

It is to be remarked that $f_1, \ldots, f_n$ defined above depend not only on the basis $U = \{u_1, \ldots, u_n\}$, but also on the order in which $u_1, \ldots, u_n$ appear in the representation of any $x \in X$. When we have to take this order into account, we say that U is an **ordered basis** of X.

CONVENTION: Whenever we deal with a basis $\{u_1, \ldots, u_n\}$ of a finite dimensional space explicitly, we assume it to be an ordered basis.

Theorem 1.8 *Let X be a finite dimensional linear space, and let $U = \{u_1, \ldots, u_n\}$ be a basis of X. If $f_1, \ldots, f_n$ are the coordinate functionals on X with respect to U, then we have the following:*

(i) *Every $x \in X$ can be written as $x = \sum_{j=1}^{n} f_j(x) u_j$.*

(ii) *$\{f_1, \ldots, f_n\}$ is a basis of $\mathcal{L}(X, \mathbb{K})$.*

Proof. Since $U = \{u_1, \ldots, u_n\}$ is a basis of X, for every $x \in X$, there exist unique scalars $\alpha_1, \ldots \alpha_n$ such that $x = \sum_{j=1}^n \alpha_j u_j$. Now, using the relation $f_i(u_j) = \delta_{ij}$, it follows that

$$f_i(x) = \sum_{j=1}^n \alpha_j f_i(u_j) = \alpha_i, \qquad i = 1, \ldots, n.$$

Therefore, the result in (i) follows.

To see (ii), first we observe that if $\sum_{i=1}^n \alpha_i f_i = 0$, then

$$\alpha_j = \sum_{i=1}^n \alpha_i f_i(u_j) = 0 \quad \forall j = 1, \ldots, n.$$

Hence, $\{f_1, \cdots, f_n\}$ is linearly independent in $\mathcal{L}(X, \mathbb{K})$. It remains to show that the span $\{f_1, \ldots, f_n\} = \mathcal{L}(X, \mathbb{K})$. For this, let $f \in \mathcal{L}(X, \mathbb{K})$ and $x \in X$. Then using the representation of x in (i), we have

$$f(x) = \sum_{j=1}^n f_j(x) f(u_j) = \left(\sum_{j=1}^n f(u_j) f_j\right)(x)$$

for all $x \in X$. Thus, $f = \sum_{j=1}^n f(u_j) f_j$ so that $f \in \text{span}\{f_1, \cdots, f_n\}$. This completes the proof. ∎

Let X be a finite dimensional space and let $U = \{u_1, \ldots, u_n\}$ be a basis of X, and $f_1, \ldots, f_n$ be the associated coordinate functionals. In view of the above theorem, we say that $F = \{f_1, \ldots, f_n\}$ is the **dual basis of** $\mathcal{L}(X, \mathbb{K})$ with respect to the (ordered) basis U of X.

Exercise 1.5 Let X be finite dimensional linear space, and $U = \{u_1, \ldots, u_n\}$ be a basis of X. If $F = \{f_1, \ldots, f_n\}$ is the dual basis of $\mathcal{L}(X, \mathbb{K})$ with respect to U, then show that the map

$$x \mapsto (f_1(x), \ldots, f_n(x)), \qquad x \in X,$$

is a bijective linear operator from X to $\mathbb{K}^n$. □

EXAMPLE 1.6 (*cont.*) (iii) For $x = a_0 + a_1 t + \cdots + a_n t^n$ in $\mathcal{P}_n$, let

$$(Ax)(t) = a_1 + 2a_2 t + \cdots + n a_n t^{n-1},$$

$$(Bx)(t) = a_0 t + \frac{a_1}{2} t^2 + \cdots + \frac{a_n}{n+1} t^{n+1}.$$

Then A is a linear operator from $\mathcal{P}_n$ to $\mathcal{P}_m$ for $m \geq n-1$ and B is linear operator from $\mathcal{P}_n$ to $\mathcal{P}_\ell$ for $\ell \geq n+1$.

(iv) Let $X = C^1[a,b]$ and $Y = C[a,b]$. For $x \in X$, define

$$\begin{aligned}(Ax)(t) &= \frac{d}{dt}x(t), \qquad x \in X,\ t \in [a,b],\\ (Bx)(t) &= \int_a^t x(s)\,ds, \qquad x \in Y,\ t \in [a,b],\\ f(x) &= \int_a^b x(s)\,ds, \qquad x \in Y.\end{aligned}$$

Then A is a linear operator from X to Y, B is a linear operator from Y to X and f is a linear functional on Y.

(v) Let $t_1, \ldots, t_n$ be distinct points in $[a,b]$ and $w_1, \ldots, w_n$ be elements from $\mathbb{K}$. For $x \in C[a,b]$, define

$$f(x) = \sum_{j=1}^{n} x(t_j)w_j, \qquad x \in C[a,b].$$

Then f is a linear functional on $C[a,b]$.

(vi) Let X be a linear space and X_0 be a subspace of X. From the definition of the *quotient space* X/X_0, it follows that $\eta : X \to X/X_0$ defined by $\eta(x) = [x]$, $x \in X$, is a linear operator.

The map η defined above is called the **quotient map** associated with the quotient space X/X_0.

(vii) Let X be a linear space and X_0 be a subspace of X. Then the map $A : X_0 \to X$ defined by $Ax = x$ for all $x \in X_0$ is a linear operator.

The operator A in (vii) above is called an **inclusion operator**. If $X_0 = X$, then the above inclusion operator is called the **identity operator** on X. The identity operator on a linear space X is usually denoted by I_X, or simply by I, if the space under consideration is understood.

Exercise 1.6 (i) Let X and Y be finite dimensional linear spaces, and $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_m\}$ be bases of X and Y, respectively. Let $F = \{f_1, \ldots, f_n\}$ be the dual basis of $\mathcal{L}(X, \mathbb{K})$ with respect to U and $G = \{g_1, \ldots, g_m\}$ be the dual basis of $\mathcal{L}(Y, \mathbb{K})$ with respect to V. For $i = 1, \ldots, m$; $j = 1, \ldots, n$, let $A_{ij} : X \to Y$ defined by

$$A_{ij}(x) = f_j(x)v_i, \qquad x \in X.$$

Show that $\{A_{ij} : i = 1, \ldots, m; j = 1, \ldots, n\}$ is a basis of $\mathcal{L}(X, Y)$.

(ii) Let X and Y be linear spaces, and X_0 be a subspace of X. Let $A_0 : X_0 \to Y$ be a linear operator. Show that there exists a linear operator $A : X \to Y$ such that $A_{|X_0} = A_0$. $\square$

1.2.1 Matrix Representations

Let X and Y be finite dimensional linear spaces, and $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_m\}$ be bases of X and Y, respectively. Let $A : X \to Y$ be a linear operator. Since $Au_j \in Y$ for each $j = 1, \ldots, n$ and $\{v_1, \ldots, v_m\}$ is a basis of Y, there exist scalars a_{ij} such that

$$Au_j = \sum_{i=1}^{m} a_{ij} v_i.$$

We may observe that the $m \times n$ matrix (a_{ij}) depends not only on the bases U and V but also on the order in which the elements of U and V are arranged.

The $m \times n$ matrix (a_{ij}) is called the **matrix representation** of A with respect to the *ordered bases* U and V. We may denote this matrix by $[A]_{U,V}$.

Next, let $(a_{ij}) \in \mathbb{K}^{m \times n}$. Then we know that $A : X \to Y$ defined by

$$Ax = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\alpha_j\right)v_i, \qquad x = \sum_{j=1}^{n} \alpha_j u_j \in X,$$

is a linear operator. Note that

$$Au_j = \sum_{i=1}^{m} a_{ij} v_i, \qquad j = 1, \ldots, n.$$

This shows that $(a_{ij}) = [A]_{U,V}$. Thus, we have proved that, given the ordered bases U and V of finite dimensional spaces X and Y, respectively, there is a one-one correspondence between the space $\mathcal{L}(X, Y)$ and the space of all $m \times n$ matrices.

Exercise 1.7 Let X and Y be finite dimensional linear spaces, and $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_m\}$ be bases of X and Y, respectively. Show the following:

(i) If $\{g_1, \ldots, g_m\}$ is the ordered dual basis of $\mathcal{L}(Y, \mathbb{K})$ with respect to the basis V of Y, then $[A]_{U,V} = (g_i(Au_j))$.

(ii) If $A, B \in \mathcal{L}(X, Y)$ and $\alpha \in \mathbb{K}$, then

$$[A + B]_{U,V} = [A]_{U,V} + [B]_{U,V}, \qquad [\alpha A]_{U,V} = \alpha[A]_{U,V}.$$

(iii) Suppose $\{E_{ij} : i = 1 \ldots, m;\ j = 1, \ldots, n\}$ is a basis of $\mathbb{K}^{m \times n}$. If $T_{ij} \in \mathcal{L}(X, Y)$ is the linear transformation such that $[T_{ij}]_{U,V} = E_{ij}$, then $\{T_{ij} : i = 1 \ldots, m;\ j = 1, \ldots, n\}$ is a basis of $\mathcal{L}(X, Y)$. $\square$

1.2.2 Range and Null Space

Let X and Y be linear spaces and $A : X \to Y$ be a linear operator. Then it is easily seen that the sets

$$R(A) = \{Ax : x \in X\}, \qquad N(A) = \{x \in X : Ax = 0\}$$

are subspaces of Y and X, respectively. These subspaces are, respectively, called the **range** and **null space** of A.

The dimension of $R(A)$ is called the **rank** of A, denoted by $\operatorname{rank} A$, and the dimension of $N(A)$ is called the **nullity** of A, denoted by $\operatorname{null} A$.

A linear operator $A : X \to Y$ is said to be of **finite rank** if $\operatorname{rank} A < \infty$.

Exercise 1.8 Let $A : X \to Y$ be a linear operator between linear spaces X and Y. Show that A is of finite rank if and only if there exists $n \in \mathbb{N}$, $\{v_1, \ldots, v_n\} \subset Y$ and $\{f_1, \ldots, f_n\} \subset \mathcal{L}(X, \mathbb{K})$ such that $Ax = \sum_{j=1}^{n} f_j(x) v_j$ for all $x \in X$. $\square$

Let $A : X \to Y$ be a linear operator. Clearly, A is onto or surjective if and only if $R(A) = Y$. Using linearity of A, it is easily seen (*Verify*) that A is one-one or injective if and only if $N(A) = \{0\}$.

Moreover, we observe the following:

(1) If E is a linearly independent subset of X and if A is injective, then the set $A(E) := \{Ax : x \in E\}$ is a linearly independent subset of Y.

(2) If $\{u_1, \ldots, u_n\} \subseteq X$ is such that $\{Au_1, \ldots, Au_n\}$ is a linearly independent subset of Y, then $\{u_1, \ldots u_n\}$ is a linearly independent subset of X.

From these results, we obtain the following theorem.

Theorem 1.9 *Let X and Y be finite dimensional linear spaces and $A : X \to Y$ be a linear operator. Then A is injective if and only if* $\operatorname{rank} A = \dim X$. *In particular, if* $\dim X = \dim Y$, *then A is injective if and only if A is surjective.*

In fact, the above result is a particular case of the following theorem as well.

Theorem 1.10 *Let X and Y be linear spaces and $A : X \to Y$ be a linear operator. Then*

$$\operatorname{rank} A + \operatorname{null} A = \dim X.$$

Proof. First we observe that, if either $\operatorname{null} A = \infty$ or $\operatorname{rank} A = \infty$, then $\dim X = \infty$(*Why?*). Therefore, assume that both

$$r := \operatorname{rank} A < \infty, \qquad k := \operatorname{null} A < \infty.$$

Suppose $U = \{u_1, \ldots, u_k\}$ is a basis of $N(A)$ and $V = \{v_1, \ldots, v_r\}$ is a basis of $R(A)$. Let $W = \{w_1, \ldots, w_r\} \subseteq X$ such that $Aw_j = v_j$, $j = 1, \ldots, r$. We show that $U \cup W$ is a basis of X.

Let $x \in X$. Since V is a basis of $R(A)$, there exist scalars $\alpha_1, \ldots, \alpha_r$ such that

$$Ax = \sum_{i=1}^{r} \alpha_j v_j = \sum_{i=1}^{r} \alpha_j Aw_j$$

Hence, $A\left(x - \sum_{i=1}^{r} \alpha_j w_j\right) = 0$ so that $x - \sum_{i=1}^{r} \alpha_j w_j \in N(A)$. Since U is a basis of $N(A)$, there exist scalars $\beta_1, \ldots, \beta_k$ such that

$$x - \sum_{i=1}^{r} \alpha_j w_j = \sum_{i=1}^{k} \beta_j u_j.$$

Thus, $x \in \operatorname{span}(U \cup W)$. It remains to show that $U \cup W$ is linearly independent. For this, suppose $a_1, \ldots, a_k$ and $b_1, \ldots, b_r$ are scalars such that

$$\sum_{i=1}^{k} a_j u_j + \sum_{i=1}^{r} b_j w_j = 0.$$

Applying A to the above equation, it follows that $\sum_{i=1}^{r} b_j v_j = 0$ so that, by the linear independence of V, $b_j = 0$ for all $j = 1, \ldots, r$. Therefore, we have $\sum_{i=1}^{k} a_j u_j = 0$. Now, by the linear independence of U, $a_j = 0$, for all $j = 1, \ldots, k$. This completes the proof. ∎

The following result is, in fact, hidden in the proof of the above theorem.

Theorem 1.11 *Let $A : X \to Y$ be a finite rank operator, and let $w_1, \ldots, w_r$ in X be such that $\{Aw_1, \ldots, Aw_r\}$ is a basis of $R(A)$. Then for every $x \in X$, there exist unique $(\alpha_1, \ldots, \alpha_r) \in \mathbb{K}^r$ and $u \in N(A)$ such that*

$$x = u + \sum_{j=1}^{r} \alpha_j w_j.$$

In particular, if $f : X \to \mathbb{K}$ is a nonzero linear functional and $x_0 \in X$ is such that $f(x_0) \neq 0$, then for every $x \in X$, there exists a unique $u \in N(A)$ such that

$$x = u + \frac{f(x)}{f(x_0)} x_0.$$

Proof. By the hypothesis, for every $x \in X$, there exist scalars $\alpha_1, \ldots, \alpha_r$ such that $Ax = \sum_{j=1}^{r} \alpha_j A w_j$. Hence, $u := x - \sum_{j=1}^{r} \alpha_j w_j$ belongs to $N(A)$ so that

$$x = u + \sum_{j=1}^{r} \alpha_j w_j \quad \text{with} \quad u \in N(A).$$

The uniqueness of the above representation is easy to verify using the fact that $\{Aw_1, \ldots, Aw_r\}$ is linearly independent.

The proof of the particular case easily follows from the first part of the theorem. ∎

1.2.3 Linear Functionals and Hyperspaces

Theorem 1.11 shows that if f is a nonzero linear functional on a linear space X and if $x_0 \in X \setminus N(f)$, then

$$X = \operatorname{span}\{x_0; N(f)\}.$$

This motivates the following definition.

A subspace H of a linear space X is called a **hyperspace** of X if there exists $x_0 \in X \setminus H$ such that

$$\text{span}\{x_0; H\} = X.$$

As an immediate example, null spaces of nonzero linear functionals are hyperspaces.

Is every hyperspace a null space of a nonzero linear functional? The answer is in the affirmative.

Theorem 1.12 *A subspace of a linear space is a hyperspace if and only if it is the null space of a nonzero linear functional.*

Proof. Let X be a linear space. We have already seen that null spaces of nonzero linear functionals are hyperspaces.

Conversely, suppose H is a hyperspace of X and $x_0 \in X \setminus H$ such that $X = \text{span}\{x_0; H\}$. Then, for every $x \in X$, there exists unique pair (α, u) in $\mathbb{K} \times H$ such that $x = \alpha x_0 + u$. Define

$$f(\alpha x_0 + u) = \alpha, \qquad \alpha \in \mathbb{K}, \ u \in H.$$

It is seen that f is a linear functional with $f(x_0) = 1$, $N(f) = H$. ∎

Let us observe that if X is a linear space, then the family S of all proper subspaces of X is a partially ordered set with respect to the set inclusion. The following theorem shows that maximal elements of S are precisely the hyperspaces of X.

Theorem 1.13 *Let X be a linear space and S be the partially ordered set of all proper subspaces of X. Let H be a subspace of X. Then H is a hyperspace of X if and only if it is a maximal element of S.*

Proof. Let H be a hyperspace of X and let $x_0 \in X \setminus H$ such that $\text{span}\{x_0; H\} = X$. Suppose Y is a proper subspace of X such that $H \subseteq Y$. We show that $Y = H$. Suppose $Y \neq H$, and let $y_0 \in Y \setminus H$. Since $\text{span}\{x_0; H\} = X$, there exists $\alpha_0 \neq 0$ in $\mathbb{K}$ and $u \in H$ such that $y_0 = \alpha_0 x_0 + u$. Hence,

$$x_0 = \frac{1}{\alpha_0}(y_0 - u) \in \text{span}\{y_0; H\} \subseteq Y.$$

Consequently,

$$X = \text{span}\{x_0; H\} \subseteq \text{span}\{y_0; H\} \subseteq Y.$$

This contradicts the fact that $Y \in \mathcal{S}$. Thus, every hyperspace is a maximal element of $\mathcal{S}$.

Conversely, suppose that H is a maximal element in $\mathcal{S}$. Let $u_0 \in X \setminus H$. Then H is a proper subspace of $\operatorname{span}\{u_0; H\}$ so that by the maximality of H we have

$$\operatorname{span}\{u_0; H\} = X,$$

showing that H is a hyperspace of X. ∎

Exercise 1.9 Show that a linear functional on a linear space is completely determined by its null space and an element not in the null space; i.e., if f and g are linear functionals on a linear space X such that $N(f) = N(g)$, and $f(x_0) = g(x_0)$ for some $x_0 \in X \setminus N(f)$, then $f = g$. □

1.2.4 Product of Operators and Inverse of an Operator

Let X, Y, Z be linear spaces. For $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(Y, Z)$ the **product of operators** $BA : X \to Z$ is defined by

$$(BA)(x) = B(Ax) \quad \forall x \in X.$$

Clearly, $BA \in \mathcal{L}(X, Z)$. If $X = Y = Z$, then both AB, BA belong to $\mathcal{L}(X)$.

It can be seen easily that for $A \in \mathcal{L}(X, Y)$, if there exists a linear operator $B : R(A) \to X$ such that

$$B(Ax) = x \quad \forall x \in X, \qquad A(By) = y \quad \forall y \in R(A),$$

then A is injective. Conversely, if $A : X \to Y$ is an injective linear operator, then for every $y \in R(A)$, there exists a unique $x \in X$ such that $Ax = y$ so that a function from $R(A)$ to X can be defined in a natural way by associating each $y = Ax \in R(A)$ the element $x \in X$. Denoting this function by $B : R(A) \to X$, it is easily seen that $B : R(A) \to X$ is a linear operator, and it satisfies

$$B(Ax) = x \quad \forall x \in X, \qquad A(By) = y \quad \forall y \in R(A).$$

This operator B, which is uniquely defined by the above properties, is called the **inverse** of A, and is denoted by $A^{-1} : R(A) \to X$.

Clearly, if $A \in \mathcal{L}(X,Y)$ is bijective, then its inverse is defined on all of Y. It can be easily seen (*Verify*) that $A \in \mathcal{L}(X,Y)$ is bijective if and only if there exists a unique $B \in \mathcal{L}(Y,X)$ such that

$$AB = I_Y, \qquad BA = I_X,$$

and in that case B is also bijective, and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Having defined product of operators, we can define *powers* of operators.

For $A \in \mathcal{L}(X)$ and $n \in \mathbb{N}$, $A^n : X \to X$ is defined inductively by

$$A^n(x) = A\left(A^{n-1}(x)\right), \qquad x \in X,$$

where $A^0(x) := x$ for every $x \in X$.

Using this, we can define **polynomials of an operator**.

For $A \in \mathcal{L}(X)$ and $p \in \mathcal{P}_n$, say $p(t) = a_0 + a_1 t + \cdots + a_n t^n$, we define $p(A) : X \to X$ by

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n.$$

Exercise 1.10 (i) If $A : X \to Y$ and $B : Y \to X$ are linear operators such that AB is bijective, then show that B injective and A is surjective.

(ii) Let X, Y, Z be finite dimensional linear spaces and U, V, W be basis of X, Y, Z, respectively. If $A \in \mathcal{L}(X,Y)$ and $B \in \mathcal{L}(Y,Z)$, then show that $[BA]_{U,W} = [B]_{V,W}[A]_{U,V}$. $\square$

In the following, we consider a special class of linear operators called *projections.*

1.2.5 Projections

Let X be a linear space, and X_1 and X_2 be subspaces of X such that

$$X = X_1 + X_2 \quad \text{with} \quad X_1 \cap X_2 = \{0\}.$$

Then we know that every $x \in X$ can be written uniquely as

$$x = x_1 + x_2 \quad \text{with} \quad x_1 \in X_1, \, x_2 \in X_2.$$

Now define $P : X \to X$ by

$$P(x_1 + x_2) = x_1, \qquad (x_1, x_2) \in X_1 \times X_2.$$

It can be seen (*Verify*) that P is a linear operator and it satisfies

$$Pu = u \quad \forall u \in R(P), \quad Pv = 0 \quad \forall v \in X_2.$$

From this it follows that

$$R(P) = X_1, \qquad N(P) = X_2, \qquad P^2 = P.$$

It can be seen that a linear operator P satisfying these properties is the unique.

A linear operator $P : X \to X$ is called a **projection operator** or simply a **projection** if

$$Pu = u \quad \forall u \in R(P).$$

Exercise 1.11 Show that a linear operator $P : X \to X$ is a projection if and only if $P^2 = P$. □

If $P : X \to X$ is a projection with $R(P) = M$ and $N(P) = N$, then we say that P is a **projection onto** M **along** N.

Thus, we know that if $X = X_1 \oplus X_2$, then there is a unique projection $P : X \to X$ onto X_1 along X_2, and in that case $I - P$ is the unique projection onto X_2 along X_1.

EXAMPLE 1.7 (i) The simplest examples of projections on a linear space X are the zero operator and the identity operator. The identity operator is usually denoted by I and it is defined by

$$Ix = x \quad \forall x \in X.$$

(ii) For $x = (\alpha_1, \alpha_2) \in \mathbb{K}^2$, define $Px = \alpha_1$. Then, $P : \mathbb{K}^2 \to \mathbb{K}^2$ is a projection onto $M = \{(\alpha_1, \alpha_2) : \alpha_2 = 0\}$ along $N = \{(\alpha_1, \alpha_2) : \alpha_1 = 0\}$.

(iii) For $\tau \in [a, b]$ let $u \in C[a, b]$ be such that $u(\tau) = 1$. For $x \in C[a, b]$, define

$$Px = x(\tau)u.$$

Then $P : C[a, b] \to C[a, b]$ is a projection onto $M := \operatorname{span}\{u\}$ along $N := \{x \in C[a, b] : x(\tau) = 0\}$.

(iv) Associated with every subspace X_0 of X, there is a projection with range X_0. To see this, let X_0 be a subspace of X and let E_0 be a basis of X_0. Then we know that there is a basis E of X such that $E_0 \subseteq E$. Taking

$$Y_0 = \text{span}\,(E \setminus E_0),$$

it is seen that $X = X_0 \oplus Y_0$ so that there is a unique projection $P_0 : X \to X$ such that

$$R(P_0) = X_0, \qquad N(P_0) = Y_0.$$

1.2.6 Eigenvalues and Eigenvectors

Let $A : X \to X$ be a linear operator on a linear space X.

A scalar λ is called an **eigenvalue** of A if there exists a nonzero vector $x \in X$ such that

$$Ax = \lambda x,$$

and in that case, x is called an **eigenvector** of A corresponding to the eigenvalue λ.

The set of all eigenvectors of A corresponding to an eigenvalue, together with the zero vector, is called an **eigenspace** of A, and the set of all eigenvalues of A is called the **eigenspectrum** of A.

We denote the eigenspectrum of A by $\sigma_{\text{eig}}(A)$.

Thus, $\lambda \in \mathbb{K}$ is an eigenvalue of A if and only if $A - \lambda I$ is not injective, and in that case, $N(A - \lambda I)$ is the corresponding eigenspace of A.

EXAMPLE 1.8 The conclusions in (i)-(vi) below can be verified easily:

(i) Let $A : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$A : (\alpha_1, \alpha_2, \alpha_3) \mapsto (\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3)$$

Then $\sigma_{\text{eig}}(A) = \{1\}$ and $N(A - I) = \text{span}\,\{(0, 0, 1)\}$.

(ii) Let $A : \mathbb{K}^2 \to \mathbb{K}^2$ be defined by $A : (\alpha_1, \alpha_2) \mapsto (\alpha_2, -\alpha_1)$. If $\mathbb{K} = \mathbb{R}$, then A has no eigenvalues, i.e., $\sigma_{\text{eig}}(A) = \varnothing$.

(iii) Let A be as in (ii) above. If $\mathbb{K} = \mathbb{C}$, then $\sigma_{\text{eig}}(A) = \{i, -i\}$, $N(A - iI) = \text{span}\,\{(1, i)\}$ and $N(A + iI) = \text{span}\,\{(1, -i)\}$.

(iv) Let $X = c_{00}$, and let (λ_n) be a sequence of scalars. Let $A : X \to X$ be defined by

$$(Ax)(i) = \lambda_i x(i) \quad \forall x \in X,\ i \in \mathbb{N}.$$

Then $\sigma_{\text{eig}}(A) = \{\lambda_1, \lambda_2, \ldots\}$, and for each $j \in \mathbb{N}$, e_j is an eigenvector corresponding to the eigenvalue λ_j. In case $\lambda_1, \lambda_2, \ldots$ are distinct, then $N(A - \lambda_j I) = \text{span}\{e_j\}$ for all $j \in \mathbb{N}$. Here, $e_j \in c_{00}$ is such that $e_j(i) = \delta_{ij}$, for $i, j \in \mathbb{N}$.

(v) Let $A : \mathcal{P} \to \mathcal{P}$ be defined by

$$(Ax)(t) = tx(t), \qquad x \in \mathcal{P}.$$

Then $\sigma_{\text{eig}}(A) = \varnothing$.

(vi) Let X be $\mathcal{P}[a,b]$ and $A : X \to X$ be defined by

$$(Ax)(t) = \frac{d}{dt}x(t), \qquad x \in \mathcal{P}.$$

Then $\sigma_{\text{eig}}(A) = \{0\}$ and $N(A) = \text{span}\{x_0\}$, where $x_0(t) = 1$ for all $t \in [a,b]$.

Existence of an eigenvalue

From the above examples we observe that in those cases in which the eigenspectrum is empty, either the scalar field is $\mathbb{R}$ or the linear space is infinite dimensional. The next result shows that if the space is finite dimensional and if the scalar field is the set of all complex numbers, then the eigenspectrum is nonempty. The proof given below is due to Axler [5].

Theorem 1.14 *Let X be a finite dimensional linear space over $\mathbb{C}$. Then every linear operator on X has at least one eigenvalue.*

Proof. Let X be an n-dimensional linear space over $\mathbb{C}$, and $A : X \to X$ be a linear operator. Let x be a nonzero element in X. Since $\dim X = n$, the set $\{x, Ax, A^2x, \ldots, A^n x\}$ is linearly dependent. Let $a_0, a_1 \ldots, a_n$ be scalars with at least one of them being nonzero such that

$$a_0 x + a_1 Ax + \cdots + a_n A^n x = 0.$$

Let $k = \max\{j : a_j \neq 0,\ j = 1, \ldots, n\}$. Then writing

$$p(t) = a_0 + a_1 t + \cdots + a_k t^k, \quad p(A) = a_0 I + a_1 A + \cdots + a_k A^k,$$

we have

$$p(A)(x) = 0.$$

By fundamental theorem of algebra, there exist $\lambda_1, \ldots, \lambda_k$ in $\mathbb{C}$ such that

$$p(t) = a_k(t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_k).$$

Thus, we have

$$a_k(A - \lambda_1 I)(A - \lambda_2 I) \ldots (A - \lambda_k I)(x) = p(A)(x) = 0.$$

The above relation shows that at least one of $A - \lambda_1 I, \ldots, A - \lambda_k I$ is not injective so that at least one of $\lambda_1, \ldots, \lambda_k$ is an eigenvalue of A. ∎

Theorem 1.15 *Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of a linear operator $A : X \to X$ with corresponding eigenvectors $u_1, \ldots, u_n$, respectively. Then the set $\{u_1, \ldots, u_n\}$ is linearly independent.*

Proof. We prove this result by induction. The result is obvious if $n = 1$. Hence, we consider the case of $n > 1$. Let $k \in \mathbb{N}$ be such that $k < n$, and assume that $\{u_1, \ldots, u_k\}$ is linearly independent. We have to show that $\{u_1, \ldots, u_k, u_{k+1}\}$ is linearly independent. For scalars $c_1, \ldots, c_k, c_{k+1}$, let

$$x = c_1 u_1 + \cdots + c_k u_k + c_{k+1} u_{k+1}.$$

We have to show that, if $x = 0$, then $c_j = 0$ for $j = 1, \ldots, k + 1$.

We note that

$$Ax = c_1 \lambda_1 u_1 + \cdots + c_k \lambda_k u_k + c_{k+1} \lambda_{k+1} u_{k+1}$$

so that

$$Ax - \lambda_{k+1} x = c_1(\lambda_1 - \lambda_{k+1}) u_1 + \cdots + c_k(\lambda_k - \lambda_{k+1}) u_k.$$

Now suppose that $x = 0$. Then we have $Ax - \lambda_{k+1} x = 0$, i.e.,

$$c_1(\lambda_1 - \lambda_{k+1}) u_1 + \cdots + c_k(\lambda_k - \lambda_{k+1}) u_k = 0.$$

From this, using the fact that $\{u_1, \ldots, u_k\}$ is linearly independent in X, and $\lambda_1, \ldots, \lambda_k, \lambda_{k+1}$ are distinct, it follows that $c_j = 0$ for $j = 1, \ldots, k$. Therefore, $0 = x = c_{k+1} u_{k+1}$ so that $c_{k+1} = 0$. This completes the proof. ∎

By the above theorem we can immediately infer that if X is finite dimensional, then the eigenspectrum of every linear operator on X is a finite set.

With this we complete our review of basic definitions and results from linear algebra. The remaining concepts we need from linear algebra will be introduced as and when we require them.

PROBLEMS

1. Let $t_0, t_1, \ldots, t_n$ be in $[a, b]$ such that $a = t_0 < t_1 < \ldots < t_n = b$. For $k \in \mathbb{N}$, let $X_{k,n}$ be the set of all those functions $x \in C([a, b], \mathbb{R})$ such that the restriction of x to each interval $[t_{j-1}, t_j]$ is a polynomial of degree atmost k. Then show that $X_{k,n}$ is a linear space over $\mathbb{R}$. What is the dimension of $X_{k,n}$?

2. Given real numbers $a_0, a_1, \ldots, a_k$, let X be the set of all solutions $x \in C^k[a, b]$ of the differential equation

$$a_0 \frac{d^k x}{dt^k} + a_1 \frac{d^{k-1} x}{dt^{k-1}} + \cdots + a_k x = 0.$$

Show that X is a linear space over $\mathbb{R}$. What is the dimension of X?

3. Show that the span of the set $\{1,\ 1+t,\ 1+t+t^2,\ \ldots,\ 1+t+\cdots+t^n\}$ is the space of all polynomials of degree atmost n.

4. Let $t_0, t_1, \ldots, t_n$ be in $[a, b]$ such that $a = t_0 < t_1 < \ldots < t_n = b$. For each $j \in \{1, \ldots, n\}$, let u_j be in $C([a, b], \mathbb{R})$ such that

$$u_j(t_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

and the restriction of u_j to each interval $[t_{j-1}, t_j]$ is a polynomial of degree atmost 1. Show that the span of $\{u_1, \ldots, u_n\}$ is the space $X_{1,n}$ in Problem 1.

5. Show that a subset $E = \{u_1, \ldots, u_n\}$ of a linear space X is linearly independent if and only if for every $x \in \operatorname{span} E$, there exists a unique $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ such that $x = \alpha_1 u_1 + \cdots + \alpha_n u_n$.

6. For each λ in the interval $(0, 1)$, let $u_\lambda = (1, \lambda, \lambda^2, \ldots)$. Show that $u_\lambda \in \ell^1$ for each $\lambda \in (0, 1)$, and the set $\{u_\lambda : 0 < \lambda < 1\}$ is linearly independent in ℓ^1. Infer that every basis of the spaces c_0, c, ℓ^∞ is an uncountable set.

7. Let X and Y be finite dimensional linear spaces of same dimension. Find a linear isomorphism between X and Y.

8. Let $t_1, \ldots, t_n$ be distinct points in $[a, b]$ and $u_1, \ldots, u_n$ are in $C[a, b]$ such that $u_i(t_j) = \delta_{ij}$ for $i, j = 1, \ldots, n$. Show that

$$Px = \sum_{j=1}^{n} x(t_j)u_j, \quad x \in C[a, b],$$

defines a projection operator on $C[a, b]$ with $R(P) = \operatorname{span}\{u_1, \ldots, u_n\}$.

The above projection P is called an **interpolatory projection** onto the span of $\{u_1, \ldots, u_n\}$. We shall discuss these types of projections in Chapter 3 as well.

9. Let X be a linear space and $A \in \mathcal{L}(X)$. If λ is an eigenvalue of A, then show that, for every polynomial p, $p(\lambda)$ is an eigenvalue of $p(A)$.

10. Let X be a finite dimensional linear space, and $A \in \mathcal{L}(X)$. Show that, if X has a basis consisting of eigenvectors of A, then A can be represented as a diagonal matrix.

11. Let $t_1, \ldots, t_n$ be distinct points in $[a, b]$ and for $j = 1, \ldots, n$, let

$$u_j(t) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{t - t_i}{t_j - t_i}.$$

Show that these $u_1, \ldots, u_n$ satisfy the requirements of the Problem 8 and the range of the corresponding projection P is the set of all polynomials of degree atmost $n - 1$.

12. Let $t_1, \ldots, t_n$ be distinct points in $[a, b]$. Show that for every $x \in C[a, b]$, there exists a unique polynomial $p(t)$ of degree atmost $n - 1$ such that $p(t_j) = x(t_j)$ for $j = 1, \ldots, n$.

2

Normed Linear Spaces

Linear spaces and linear operators arise naturally in applications, wherein a physical problem is often modelled as a mathematical equation. In such situations not only does one want to find a solution of the problem, but also one wants to know whether the solution is stable under perturbations in the data. That is, if the data is perturbed *slightly* (possibly by inaccurate measurements, or by the process of numerical approximations), then one would like to know whether the resulting solution is *close to* the actual solution. To address such issues, one has to be clear about the meaning of words *slightly*, *close to*, etc., in the context of the underlying linear space. In most of the linear spaces that one encounters in applications, there are certain *natural norms* by which one can decide whether two elements x and y are close to each other or not; in other words, whether the element $x - y$ is close to zero or not.

For example, if x and y are continuous functions on a closed interval $[0, 1]$, then one may say that x is close to y within an error level $\varepsilon > 0$, if

$$\sup_{0 \le t \le 1} |x(t) - y(t)| < \varepsilon.$$

If $x = (\alpha_1, \ldots, \alpha_n)$ and $y = (\beta_1, \ldots, \beta_n)$ are two vectors in $\mathbb{R}^n$, then we may make such inference if

$$\sqrt{\sum_{i=1}^{n} |\alpha_i - \beta_i|^2} < \varepsilon.$$

There are certain natural and simple properties, such as *positive definiteness, homogeneity, triangle inequality,* that the maps

$$x \mapsto \sup_{0 \le t \le 1} |x(t)|, \quad x \in C[0, 1], \qquad x \mapsto \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}, \quad x \in \mathbb{R}^n$$

have, which motivate us to introduce the notion of a *norm* on an arbitrary linear space.

2.1 Norm on a Linear Space

A **norm** $\|\cdot\|$ on a linear space X (over the field $\mathbb{K}$ of real or complex numbers) is a function

$$x \mapsto \|x\|, \qquad x \in X,$$

from X to the set $\mathbb{R}$ of all real numbers such that for every $x, y \in X$ and $\alpha \in \mathbb{K}$,

(a) $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$,

(b) $\|\alpha x\| = |\alpha| \, \|x\|$.

(c) $\|x + y\| \leq \|x\| + \|y\|$.

The properties (a), (b), (c) above of a norm are called *positive definiteness*, *homogeneity*, and *triangle inequality*, respectively.

A linear space X, together with a norm, is called a **normed linear space.**

Instead of saying that X is a linear space with a norm $\|\cdot\|$, we may also say that the pair $(X, \|\cdot\|)$ is a normed linear space. This is sometimes useful when we consider different norms on the same linear space.

Assuming that there will not be any confusion, a norm on a linear space will be often denoted by $\|\cdot\|$ without specifying the space concerned. In case a need arises to distinguish one norm from another, we shall use certain subscripts for the norms.

It is to be observed that if X is a normed linear space with norm $\|\cdot\|$ and X_0 is a subspace of the linear space X, then the norm $\|\cdot\|$ restricted to X_0 is a norm on X_0.

Remark 2.1 In the definition of a norm, the assumptions $\|x\| \geq 0$ for all $x \in X$, and $\|0\| = 0$ in (a) are, in fact, redundant. This is seen as follows: For any $x \in X$, taking $\alpha = 0$ in (b),

$$\|0\| = \|0x\| = |0| \, \|x\| = 0,$$

and hence, using (c),

$$0 = \|0\| = \|x - x\| \leq \|x\| + \| - x\| = 2\|x\|.$$

From the definition of a norm $\|\cdot\|$, it can be seen easily that the function

$$(x, y) \mapsto \|x - y\|, \qquad x, y \in X,$$

is a *metric* on X, called the metric induced by or associated with the norm.

Recall that a **metric** d on a set Ω is a function $d : \Omega \times \Omega \mapsto \mathbb{R}$ such that for every $x, y, z \in \Omega$,

(a) $d(x, y) \geq 0$, and $d(x, y) = 0$ iff $x = y$,

(b) $d(x, y) = d(y, x)$,

(c) $d(x, y) \leq d(x, z) + d(z, y)$.

A set, together with a metric, is called a **metric space.** If d is a metric on Ω, then we may also say that the pair (Ω, d) is a metric space.

One may ask whether every metric on a linear space is induced by a norm. The answer is, in general, negative. For example, a bounded metric on nontrivial linear space can never be induced by a norm. To see this, let d be a bounded metric on a (nontrivial) linear space X, i.e., there exists $c > 0$ such that

$$d(x, y) \leq c \quad \forall x, y \in X.$$

Let x_0 be a nonzero element in X. If d is induced by a norm $\|\cdot\|$, then we get

$$|\alpha| \, \|x_0\| = \|\alpha x_0\| = d(\alpha x_0, 0) \leq c \quad \forall \alpha \in \mathbb{K},$$

which is impossible.

An example of a bounded metric is the discrete metric, i.e., the metric d on Ω defined by

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

for $x, y \in \Omega$. Also, given any metric d on Ω,

$$d_1(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \qquad d_2(x, y) = \min\{d(x, y), 1\}$$

for all $x, y \in \Omega$ define bounded metrics on X. In fact,

$$d_1(x, y) \leq 1, \qquad d_2(x, y) \leq 1 \quad \forall x, y \in \Omega.$$

A subset S of a normed linear space X is said to be a **bounded set** if it is bounded with respect to the metric induced by the norm.

Exercise 2.1 Let X be a normed linear space with norm $\|\cdot\|$. Show the following:

(i) A subset S of X is bounded if and only if there exists $c > 0$ such that $\|x\| \le c$ for every $x \in S$.

(ii) A subspace Y of X is bounded if and only if $Y = \{0\}$. □

Some basic notions from the theory of metric spaces

We shall use freely most of the elementary concepts from the theory of metric spaces with the above metric in mind. However, we define a few terminologies which we shall use very frequently in this course.

In the following; Ω, Ω_1, Ω_2 are metric spaces with metrics d, d_1, d_2, respectively.

A sequence (x_n) in metric space Ω is said to be **convergent** in Ω if there exists $x \in \Omega$ such that for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x, x_n) < \varepsilon \quad \forall n \ge N,$$

and in that case, we say (x_n) **converges** to x and write this fact by $\lim_{n\to\infty} x_n = x$ or $x_n \to x$ as $n \to \infty$ or simply $x_n \to x$.

It can be easily seen that, if (x_n) is a convergent sequence in Ω, then for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_n, x_m) < \varepsilon \quad \forall n, m \ge N.$$

A sequence (x_n) having the above conclusion is called a **Cauchy sequence.**

One of the important properties of a Cauchy sequence is that it converges whenever it has a convergent subsequence (*Verify*).

A metric space or the associated metric is said to be **complete** if every Cauchy sequence in the space is convergent.

We shall show that every finite dimensional normed linear space is complete, whereas an infinite dimensional space need not be complete.

A normed linear space which is complete with respect to the induced metric is called a **Banach space**, named after the Polish mathematician Banach who made significant contributions to the theory of complete normed linear spaces and operators on them. In the next section, we shall consider Banach spaces in detail.

A function $f : \Omega_1 \to \Omega_2$ is said to be **continuous** at $u \in \Omega_1$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$x \in \Omega_1, \quad d_1(x, u) < \delta \implies d_2(f(x), f(u)) < \varepsilon.$$

We say that $f : \Omega_1 \to \Omega_2$ is continuous if it is continuous at every $u \in \Omega_1$. If the δ in the above definition does not depend on the element u, then f is said to be a **uniformly continuous** function.

A function $f : \Omega_1 \to \Omega_2$ is a **homeomorphism** if f is bijective, continuous, and its inverse $f^{-1} : \Omega_2 \to \Omega_1$ is also continuous; and it is an **isometry** if

$$d_2(f(x), f(y)) = d_1(x, y) \quad \forall x, y \in \Omega_1.$$

Another concept closely related to continuity is that of openness: For $x \in \Omega$ and $r > 0$, the set

$$B(x, r) := \{y \in \Omega : d(x, y) < r\}$$

is called an **open ball** in Ω with centre x and radius r. A subset E of Ω is said to be an **open set** in Ω if for every $x \in E$, there exists $r > 0$ such that $B(x, r) \subseteq E$.

A subset $E \subseteq \Omega$ is said to be a **closed set** in Ω if its complement $E^c := \Omega \setminus E$ is open in Ω.

Let $E \subseteq \Omega$, and let $x_0 \in \Omega$. Then $x_0 \in \Omega$ is called an **interior point** of E if there exists an open ball with centre x_0 and contained in E. The set of all interior points of E is called the **interior** of E, and is denoted by E°.

A point $x_0 \in \Omega$ is called an **accumulation point** or a **limit point** of $E \subseteq \Omega$ if every open ball with centre x_0 contains at least one point of E other than possibly x_0.

The set of all points of E together with all its accumulation points is called the **closure** of E, and it is denoted by $\overline{E}$ or $\mathrm{cl}\,(E)$. Elements of $\overline{E}$ are also called **closure points** of E.

A point $x_0 \in \Omega$ is called a **boundary point** of $E \subseteq \Omega$ if every open ball with centre x_0 contains some point of E as well as some point of E^c. The set of all boundary points of E is called the **boundary** of E, and we denote it by ∂E.

A subset E of Ω is said to be **dense** in Ω if $\overline{E} = \Omega$.

Let $E \subseteq \Omega$. It can be easily verified that

(1) E is open if and only if $E = E^\circ$,

(2) E is closed if and only if $E = \overline{E}$,

(3) $x \in \overline{E}$ if and only if there is a sequence (x_n) in E such that $x_n \to x$, and

(4) $\overline{E} = E^\circ \cup \partial E$.

For $E \subseteq \Omega$ and $x \in \Omega$, the **distance** from x to E is defined by

$$\text{dist}\,(x, E) := \inf\,\{d(x, u) : u \in E\}.$$

It can be seen (*Verify*) that

$$\text{dist}\,(x, E) = 0 \quad\Longleftrightarrow\quad x \in \overline{E},$$

and the map

$$x \mapsto \text{dist}\,(x, E), \qquad x \in \Omega,$$

is a continuous function on Ω.

A function $f : \Omega_1 \to \Omega$ is an **open map** if f maps every open subset of Ω_1 onto an open subset of Ω_2, i.e., for every open set U in Ω_1, the set $f(U)$ is open in Ω_2.

Exercise 2.2 Prove the following:

(i) A function $f : \Omega_1 \to \Omega_2$ is continuous at $x_0 \in \Omega_1$ if and only if for every sequence (x_n) in Ω_1 with $x_n \to x_0$, we have $f(x_n) \to f(x_0)$ in Ω_2.

(ii) A function $f : \Omega_1 \to \Omega_2$ is continuous at $x_0 \in \Omega_1$ if and only if for every open set V in Ω_2 with $f(x_0) \in V$, there exists an open set U in Ω_1 with $x_0 \in U$ such that $f(U) \subseteq V$.

(iii) A bijective function $f : \Omega_1 \to \Omega_2$ is open if and only if its inverse $f^{-1} : \Omega_2 \to \Omega_1$ continuous. $\square$

Let X be a normed linear space. From the triangle inequality of the norm (Axiom (iii)), we have

$$\|x - y\| \geq |\,\|x\| - \|y\|\,| \quad \forall x, y \in X.$$

Hence, it follows that the norm $x \mapsto \|x\|$ is a uniformly continuous function.

Let X and Y be normed linear spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and let $A : X \to Y$ be a linear operator. Then it can be seen, by using the linearity of A, that A is an isometry if and only if

$$\|Ax\|_Y = \|x\|_X \quad \forall x \in X.$$

As we shall see in the examples that follow, there can be many norms on a linear space. In applications, the choice of certain norm is decided by the kind of requirement in a given situation.

2.1.1 Examples of Normed Linear Spaces

In this subsection we shall consider many examples of normed linear spaces. While discussing these examples we shall introduce many useful inequalities such as Cauchy-Schwarz inequalities, Hölder's inequalities and Minkowski's inequalities in different spaces. Also, we discuss relations connecting the spaces in these examples. Examples of normed linear spaces which depend on concepts from measure theory, and some of the spaces derived from other spaces are considered in separate sections. An important class of examples, namely, inner product spaces is also dealt with in a separate section.

EXAMPLE 2.1 (i) Let $X = \mathbb{K}^n$, the linear space of all n-tuples of scalars. For $x = (\alpha_1, \ldots, \alpha_n) \in X$, let

$$\|x\|_1 = |\alpha_1| + \cdots + |\alpha_n|, \qquad \|x\|_\infty = \max\{|\alpha_1|, \ldots, |\alpha_n|\}.$$

Then $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms on X.

More generally, let X be a linear space of dimension $n < \infty$ and let $E = \{u_1, \ldots, u_n\}$ be a basis of X. For $x \in X$, let $\alpha_1, \ldots, \alpha_n$ be the unique scalars such that $x = \sum_{j=1}^n \alpha_j u_j$. We define

$$\|x\|_1 = |\alpha_1| + \cdots + |\alpha_n|, \qquad \|x\|_\infty = \max\{|\alpha_1|, \ldots, |\alpha_n|\}.$$

Then $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms on X. In fact,

$$\|x\|_1 = |f_1(x)| + \cdots + |f_n(x)|, \qquad \|x\|_\infty = \max\{|f_1(x)|, \ldots, |f_n(x)|\},$$

where $f_1, \ldots, f_n$ are the coordinate functionals associated with the ordered basis $E = \{u_1, \ldots, u_n\}$ (see Example 1.6(ii)). In fact, if $\|\cdot\|$ is any norm on $\mathbb{K}^n$, then

$$\|x\|_E := \|(f_1(x), \ldots, f_n(x))\|, \qquad x \in X,$$

defines a norm on X (*Verify*).

(ii) Let $X = \mathbb{K}^n$ and for $x = (\alpha_1, \ldots, \alpha_n) \in X$, let

$$\|x\|_2 = \left(|\alpha_1|^2 + \cdots + |\alpha_n|^2\right)^{1/2}.$$

Then $x \mapsto \|x\|_2$ is a norm on $\mathbb{K}^n$. To verify the triangle inequality,

$$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2 \quad \forall\, x, y \in \mathbb{K}^n,$$

we make use of *Cauchy-Schwarz inequality.*

Cauchy-Schwarz inequality in $\mathbb{K}^n$

Proposition 2.1 *For $x = (\alpha_1, \ldots, \alpha_n)$ and $y = (\beta_1, \ldots, \beta_n)$ in $\mathbb{K}^n$,*

$$\sum_{j=1}^{n} |\alpha_j \beta_j| \leq \|x\|_2 \|y\|_2.$$

Proof. Note that the result holds if one of x and y is zero. Next, assume that $x \neq 0$ and $y \neq 0$. Let us observe that for every pair of non-negative real numbers a, b,

$$ab \leq \frac{1}{2}\left(a^2 + b^2\right). \tag{2.1}$$

Now taking

$$a = \frac{|\alpha_j|}{\|x\|_2}, \qquad b = \frac{|\beta_j|}{\|y\|_2},$$

we have

$$\frac{|\alpha_j \beta_j|}{\|x\|_2 \|y\|_2} \leq \frac{1}{2}\left(\frac{|\alpha_j|^2}{\|x\|_2^2} + \frac{|\beta_j|^2}{\|y\|_2^2}\right).$$

Taking the sum over $j = 1, \ldots, n$, we get

$$\frac{1}{\|x\|_2 \|y\|_2} \sum_{j=1}^{n} |\alpha_j \beta_j| \leq \frac{1}{2}\left(\frac{1}{\|x\|_2^2} \sum_{j=1}^{n} |\alpha_j|^2 + \frac{1}{\|y\|_2^2} \sum_{j=1}^{n} |\beta_j|^2\right) = 1.$$

Thus, the inequality is established. ∎

Next, observe that for each $j \in \{1, \ldots, n\}$,

$$|\alpha_j + \beta_j|^2 \leq (|\alpha_j| + |\beta_j|)^2 = |\alpha_j|^2 + |\beta_j|^2 + 2|\alpha_j||\beta_j|.$$

Taking the sum over $j = 1, \ldots, n$ and using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|x+y\|_2^2 &\leq \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2 \|y\|_2 \\ &= (\|x\|_2 + \|y\|_2)^2. \end{aligned}$$

From this we have $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$.

Let us recall that for a nonempty set S, the set $\mathcal{F}(S,\mathbb{K})$ of all functions from S into $\mathbb{K}$ is a linear space with respect to addition and scalar multiplication defined pointwise, i.e., for every $x, y \in \mathcal{F}(S,\mathbb{K})$ and $\alpha \in \mathbb{K}$,

$$(x+y)(s) = x(s) + y(s), \qquad (\alpha x)(s) = \alpha x(s) \quad \forall\, s \in S.$$

Also, recall that if $S = \{1, \ldots, n\}$, then $\mathcal{F}(S,\mathbb{K}) = \mathbb{K}^n$, and if $S = \mathbb{N}$, then $\mathcal{F}(S,\mathbb{K})$ is the space of all scalar sequences, up to a linear isomorphism.

EXAMPLE 2.1 (*cont.*) (iii) We recall from Example 1.2(v) that

$$\ell^1(\mathbb{N}) = \Big\{x \in \mathcal{F}(\mathbb{N},\mathbb{K}) : \sum_{j=1}^{\infty} |x(j)| < \infty\Big\}$$

is a linear space. It is easily seen that the map

$$x \mapsto \|x\|_1 := \sum_{j=1}^{\infty} |x(j)|, \qquad x \in \ell^1(\mathbb{N}),$$

is a norm on $\ell^1(\mathbb{N})$.

(iv) Recall from Example 1.2(vi) that for a nonempty set S,

$$\ell^\infty(S) := \{x \in \mathcal{F}(S,\mathbb{K}) : \sup_{s\in S} |x(s)| < \infty\}$$

is a linear space. It is easily seen that the map

$$x \mapsto \|x\|_\infty := \sup_{s\in S} |x(s)|$$

is a norm on $\ell^\infty(S)$. Also, for a metric space Ω,

$$C(\Omega) := \{x \in \ell^\infty(\Omega) : x \text{ continuous on } \Omega\}$$

is a subspace of the linear space $\ell^\infty(\Omega)$, and hence

$$\|x\|_\infty = \sup\{|x(t)| : t \in \Omega\}, \qquad x \in C(\Omega)$$

is a norm on $C(\Omega)$ as well.

We Observe that if $\Omega = \{1, \ldots, n\}$, then $C(\Omega) = \mathbb{K}^n$, and if $\Omega = \mathbb{N}$, then $C(\Omega) = \ell^\infty(\mathbb{N})$. Here, equalities in the spaces are to be understood in the sense of canonical isomorphisms. We have already mentioned in Chapter 1 that it is also customary to use the symbol $B(\Omega)$ in place of $\ell^\infty(\Omega)$.

(v) Let

$$\ell^2(\mathbb{N}) = \left\{x \in \mathcal{F}(\mathbb{N}, \mathbb{K}) : \sum_{j=1}^{\infty} |x(j)|^2 < \infty\right\}.$$

We show that $\ell^2(\mathbb{N})$ is a normed linear space with the norm given by

$$\|x\|_2 = \left(\sum_{j=1}^{\infty} |x(j)|^2\right)^{1/2}, \qquad x \in \ell^2(\mathbb{N}).$$

In order to show that $\ell^2(\mathbb{N})$ is a linear space, first we have to see whether $x + y \in \ell^2(\mathbb{N})$ and $\alpha x \in \ell^2(\mathbb{N})$ for all $x, y \in \ell^2(\mathbb{N})$ and for all $\alpha \in \mathbb{K}$. The second requirement is easy to see. To see the first requirement, we make use of the *Cauchy-Schwarz inequality in* $\mathcal{F}(\mathbb{N}, \mathbb{K})$.

Cauchy-Schwarz inequality in $\mathcal{F}(\mathbb{N}, \mathbb{K})$

For $x \in \mathcal{F}(\mathbb{N}, \mathbb{K})$, let

$$\|x\|_2 = \left(\sum_{j=1}^{\infty} |x(j)|^2\right)^{1/2}$$

We may observe that for $x \in \ell^2(\mathbb{N})$, $\|x\|_2$ is a finite quantity, whereas for an arbitrary element $x \in \mathcal{F}(\mathbb{N}, \mathbb{K})$, $\|x\|_2$ can be infinity.

Proposition 2.2 *For $x, y \in \mathcal{F}(\mathbb{N}, \mathbb{K})$,*

$$\sum_{j=1}^{\infty} |x(j)y(j)| \leq \|x\|_2 \|y\|_2.$$

Proof. Note that the above inequality holds if one of $\|x\|_2$ and $\|y\|_2$ is either zero or infinity. The remaining case is

$$0 < \|x\|_2 < \infty, \qquad 0 < \|y\|_2 < \infty.$$

In this case, the result follows as in (ii), by taking

$$a = \frac{|x(j)|}{\|x\|_2}, \qquad b = \frac{|y(j)|}{\|y\|_2},$$

in (2.1), and taking the sum with $j \in \mathbb{N}$. ∎

Next, we show the *triangle inequality*

$$\|x+y\|_2 \le \|x\|_2 + \|y\|_2 \quad \forall x, y \in \mathcal{F}(\mathbb{N}, \mathbb{K}),$$

from which it also follows that

$$x + y \in \ell^2(\mathbb{N}) \quad \text{whenever} \quad x, y \in \ell^2(\mathbb{N}).$$

Clearly, the inequality holds if one of $\|x\|_2$ and $\|y\|_2$ is infinity. Hence, assume that

$$\|x\|_2 < \infty, \qquad \|y\|_2 < \infty.$$

Then, from the inequality

$$|x(j) + y(j)|^2 \le (|x(j)| + |y(j)|)^2 = |x(j)|^2 + |y(j)|^2 + 2|x(j)|\,|y(j)|,$$

we obtain, by taking the sum over $j \in \mathbb{N}$ and applying the Cauchy-Schwarz inequality, the relation

$$\|x+y\|_2^2 \le \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2\|y\|_2 = (\|x\|_2 + \|y\|_2)^2.$$

Hence, it follows that $\|x+y\|_2 < \infty$ and $\|x+y\|_2 \le \|x\|_2 + \|y\|_2$.

(vi) The maps

$$x \mapsto \|x\|_1 = \int_a^b |x(t)|\, dt, \qquad x \mapsto \|x\|_\infty = \max\{|x(t)| : t \in [a,b]\}$$

are norms on $C[a,b]$. Observe that both $\|x\|_1$ and $\|x\|_\infty$ are finite quantities for all $x \in C[a,b]$.

(vii) For $x \in C[a,b]$, let

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

The map $x \mapsto \|x\|_2$ is a norm on $C[a,b]$. All the axioms except the triangle inequality

$$\|x+y\|_2 \le \|x\|_2 + \|y\|_2 \quad \forall x, y \in C[a,b],$$

are easy to verify. To verify the triangle inequality we make use of the *Cauchy-Schwarz inequality* in $C[a,b]$

Cauchy-Schwarz inequality in $C[a,b]$

Proposition 2.3 *For $x, y \in C[a,b]$,*

$$\int_a^b |x(t)y(t)|\,dt \le \|x\|_2\|y\|_2.$$

Proof. Note that the above inequality holds if either $\|x\|_2$ or $\|y\|_2$ is zero. Hence, assume that $\|x\|_2 \neq 0$ and $\|y\|_2 \neq 0$. Now the result follows as in (ii) by taking

$$a = \frac{|x(t)|}{\|x\|_2}, \qquad b = \frac{|y(t)|}{\|y\|_2},$$

in (2.1), and taking the integral. ∎

Now the triangle inequality follows as in (ii) and (v) by making use of the Cauchy-Schwarz inequality.

Remark 2.2 In order to see the positive definiteness of $\|\cdot\|_1$ and $\|\cdot\|_2$ on $C[a,b]$, one has to recall the fact that for $u \in C[a,b]$,

$$\int_a^b |u(t)|\,dt = 0 \iff u(t) = 0 \quad \forall t \in [a,b].$$

EXAMPLE 2.1 (*cont.*)(viii) For $1 < p < \infty$, let

$$\ell^p(\mathbb{N}) := \left\{x \in \mathcal{F}(\mathbb{N},\mathbb{K}) : \sum_{j=1}^{\infty} |x(j)|^p < \infty\right\}.$$

For $x \in \mathcal{F}(\mathbb{N},\mathbb{K})$, let

$$\|x\|_p := \left(\sum_{j=1}^{\infty} |x(j)|^p\right)^{1/p}$$

Note that $\|x\|_p$ can be infinity. It is seen that, if we have the triangle inequality

$$\|x+y\|_p \le \|x\|_p + \|y\|_p, \qquad x, y \in \ell^p(\mathbb{N}), \tag{2.2}$$

then $\ell^p(\mathbb{N})$ is a linear space and the map

$$x \mapsto \|x\|_p, \qquad x \in \ell^p(\mathbb{N}),$$

is a norm on $\ell^p(\mathbb{N})$. For proving the inequality (2.2), referred to as *Minkowski's inequality*, we make use of another inequality, called *Hölder's inequality*:

$$\sum_{j=1}^{\infty} |x(j)y(j)| \le \|x\|_p \|y\|_q, \qquad x, y \in \mathcal{F}(\mathbb{N}, \mathbb{K}), \tag{2.3}$$

where $q \in (1, \infty)$ is such that

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{2.4}$$

Using the convention $1/\infty = 0$, we can say that (2.4) holds for $p = 1$ and $p = \infty$ as well, with the corresponding q as ∞ and 1, respectively.

If $p, q \in [1, \infty]$ satisfy (2.4), we say that p and q are **conjugate exponents**. If p and q are conjugate exponents, then we shall also say that q (respectively, p) is the conjugate exponent of p (respectively, q).

For establishing (2.3), we require the relation

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \tag{2.5}$$

in place of (2.1) for $a \ge 0$, $b \ge 0$, where p and q are conjugate exponents with $1 < p < \infty$. First let us prove (2.5).

For $\alpha \ge 0$ and $\beta \ge 0$, we prove

$$\alpha^{1/p}\beta^{1/q} \le \frac{\alpha}{p} + \frac{\beta}{q}. \tag{2.6}$$

Once this is proved, (2.5) follows by taking $\alpha = a^p$ and $\beta = b^q$.

Clearly, (2.6) holds if $\alpha = \beta$ or if one of α and β is zero. Hence, assume that $0 < \alpha < \beta$. Consider the continuously differentiable function $f(t) = t^{1/q}$, $t > 0$. Then, by the mean value theorem, we have

$$f(\beta) - f(\alpha) = f'(\xi)(\beta - \alpha)$$

for some $\xi \in (\alpha, \beta)$, i.e.,

$$\beta^{1/q} - \alpha^{1/q} = \frac{1}{q}\xi^{-1/p}(\beta - \alpha).$$

Since $\xi^{-1/p} < \alpha^{-1/p}$, it follows that

$$\beta^{1/q} - \alpha^{1/q} < \frac{1}{q}\,\alpha^{-1/p}(\beta - \alpha).$$

From this we have

$$\alpha^{1/p}\beta^{1/q} - \alpha < \frac{1}{q}\,(\beta - \alpha),$$

i.e.,

$$\alpha^{1/p}\beta^{1/q} \le \frac{\alpha}{p} + \frac{\beta}{q}.$$

Thus, the proof of (2.6) is completed.

Now we prove (2.3).

Hölder's inequality in $\mathcal{F}(\mathbb{N}, \mathbb{K})$

Proposition 2.4 *Suppose* $1 < p < \infty$, *and* q *is a conjugate exponent of* p. *Then*

$$\sum_{j=1}^{\infty} |x(j)y(j)| \le \|x\|_p \|y\|_q \quad \forall\, x, y \in \mathcal{F}(\mathbb{N}, \mathbb{K}). \tag{2.7}$$

Proof. It is seen that the inequality holds if $\|x\|_p\|y\|_q$ is zero or infinity. Therefore, we assume that

$$0 < \|x\|_p < \infty, \qquad 0 < \|y\|_q < \infty.$$

In this case, using relation (2.5), we obtain

$$\frac{|x(j)y(j)|}{\|x\|_p\|y\|_q} \le \frac{|x(j)|^p}{p\|x\|_p^p} + \frac{|y(j)|^q}{q\|y\|_q^q}.$$

Taking the sum by varying j over $\mathbb{N}$, we get the inequality in (2.7). ∎

Minkowski's inequality in $\mathcal{F}(\mathbb{N}, \mathbb{K})$

Proposition 2.5 *Suppose* $1 < p < \infty$. *Then*

$$\|x + y\|_p \le \|x\|_p + \|y\|_p \quad \forall\, x, y \in \mathcal{F}(\mathbb{N}, \mathbb{K}).$$

Proof. Clearly, the inequality follows if at least one of $\|x\|_p$ and $\|y\|_p$ is infinity. Hence, assume that

$$\|x\|_p < \infty, \ \|y\|_p < \infty.$$

We observe that for every $\alpha, \beta \in \mathbb{K}$,

$$|\alpha + \beta|^p \le (|\alpha| + |\beta|)^p \le 2^p \left(\max\{|\alpha|, |\beta|\}\right)^p \le 2^p\{|\alpha|^p + |\beta|^p\}.$$

Thus we have

$$|x(j) + y(j)|^p \le 2^p \left(|x(j)|^p + |y(j)|^p\right)$$

so that, taking the sum over $j \in \mathbb{N}$,

$$\|x + y\|_p < \infty.$$

Next, we observe that

$$\begin{aligned}
\sum_{j=1}^{\infty} |x(j) + y(j)|^p &= \sum_{j=1}^{\infty} (|x(j)| + |y(j)|)|x(j) + y(j)|^{p-1} \\
&\le \sum_{j=1}^{\infty} |x(j)|\,|x(j) + y(j)|^{p-1} \\
&\quad + \sum_{j=1}^{\infty} |y(j)|\,|x(j) + y(j)|^{p-1}
\end{aligned}$$

so that, applying Hölder's inequality (2.7) on each term on the right hand side of the above relation, we have

$$\|x + y\|_p^p \le (\|x\|_p + \|y\|_p)\, \|x + y\|_p^{p/q}.$$

Since $\|x + y\|_p < \infty$, we get the required inequality. ∎

Thus, we have shown that

$$\|x\|_p := \left(\sum_{j=1}^{\infty} |x(j)|^p\right)^{1/p}, \qquad x \in \ell^p(\mathbb{N}).$$

defines the norm on $\ell^p(\mathbb{N})$.

In the above example, if we replace infinite sum by finite sum, say with j from 1 to n, then we see that

$$x = (\alpha_1, \ldots, \alpha_n) \mapsto \begin{cases} \left(\sum_{i=1}^{n} |\alpha_i|^p\right)^{1/p} & \text{if } 1 \le p < \infty \\ \max\{|\alpha_i| : i = 1, \ldots, n\} & \text{if } p = \infty, \end{cases}$$

defines a norm on $\mathbb{K}^n$ for each $p \in [1, \infty]$.

Remark 2.3 From the remark at the end of example (i) above, it follows that, if X is any linear space of dimension n, $E = \{u_1, \ldots, u_n\}$ is a basis of X, and if $f_1, \ldots, f_n$ are the coordinate functionals associated with E, then

$$\|x\|_{E,p} = \begin{cases} \left(\sum_{i=1}^{n} |f_i(x)|^p\right)^{1/p} & \text{if } 1 \le p < \infty \\ \max\{|f_i(x)| : i = 1, \ldots, n\} & \text{if } p = \infty, \end{cases}$$

defines a norm on X for each $p \in [1, \infty]$. This norm $\|\cdot\|_{E,p}$ will be called the p-norm on X with respect to the basis E.

NOTATION: Let $1 \le p \le \infty$. Then the space $\mathbb{K}^n$ with $\|\cdot\|_p$ will be denoted by $\ell^p(n)$, and the space $\ell^p(\mathbb{N})$ will be denoted simply by ℓ^p.

Remark 2.4 Suppose S is any countable set. Then, as in the case of $\ell^p(\mathbb{N})$, we can show that, for $1 \le p < \infty$,

$$\ell^p(S) := \left\{ x \in \mathcal{F}(S, \mathbb{K}) : \sum_{s \in S} |x(s)|^p < \infty \right\},$$

is a subspace of $\mathcal{F}(S, \mathbb{K})$, and

$$x \mapsto \|x\|_p := \left(\sum_{s \in S} |x(s)|^p\right)^{1/p}, \qquad x \in \ell^p(S),$$

is a norm on $\ell^p(S)$.

EXAMPLE 2.1 (*cont.*) (ix) For $1 < p < \infty$, let

$$\|x\|_p = \left(\int_a^b |x(t)|^p \, dt\right)^{1/p}, \qquad x \in C[a, b].$$

We show that $x \mapsto \|x\|_p$ is a norm on $C[a,b]$. Remark 2.2 holds good for this case as well. All the axioms of a norm except the triangle inequality

$$\|x+y\|_p \le \|x\|_p + \|y\|_p, \quad x, y \in C[a,b],$$

which is also known as the *Minkowski's inequality* on $C[a,b]$, are easy to verify. To prove Minkowski's inequality, we require *Hölder's inequality* in $C[a,b]$.

Hölder's inequality in $C[a,b]$

Proposition 2.6 *Suppose* $1 < p < \infty$, *and* q *is a conjugate exponent of* p. *Then*

$$\int_a^b |x(t)y(t)|\,dt \le \|x\|_p \|y\|_q \quad \forall\, x, y \in C[a,b]. \tag{2.8}$$

Proof. Clearly, the inequality holds if either $\|x\|_p$ or $\|y\|_q$ is zero. Hence, assume that both $\|x\|_p$ and $\|y\|_q$ are nonzero. By taking

$$a = \frac{|x(t)|}{\|x\|_p}, \qquad b = \frac{|y(t)|}{\|y\|_q}$$

in (2.5), we obtain

$$\frac{|x(t)y(t)|}{\|x\|_p\|y\|_q} \le \frac{|x(t)|^p}{p\|x\|_p^p} + \frac{|y(t)|^q}{q\|y\|_q^q}.$$

Now the result follows by taking integrals on both the sides. ∎

Minkowski's inequality in $C[a,b]$

Proposition 2.7 *Suppose* $1 < p < \infty$. *Then*

$$\|x+y\|_p \le \|x\|_p + \|y\|_p \quad \forall x, y \in C[a,b].$$

Proof. We observe that

$$\begin{aligned}
\|x+y\|_p^p &= \int_a^b |x(t)+y(t)|^p\,dt \\
&\le \int_a^b (|x(t)| + |y(t)|)\,|x(t)+y(t)|^{p-1}\,dt \\
&= \int_a^b |x(t)||x(t)+y(t)|^{p-1}\,dt \\
&\quad + \int_a^b |y(t)||x(t)+y(t)|^{p-1}\,dt.
\end{aligned}$$

Hence, using Hölder's inequality (2.8), we have

$$\|x+y\|_p^p \le (\|x\|_p + \|y\|_p)\,\|x+y\|_p^{p/q}.$$

From this we obtain the required inequality. ∎

EXAMPLE 2.1 (*cont.*) (x) For $k \in \mathbb{N}$ let $X = C^k[a,b]$, and let $1 \le p \le \infty$. Then the maps

$$x \mapsto \sum_{j=0}^{k} \|x^{(j)}\|_p, \qquad x \mapsto \max\{\|x^{(j)}\|_p : j = 0, 1, \ldots, k\}$$

are norms on X. Here $x^{(0)} = x$, and $x^{(j)}$ denotes the j-th derivative of x for $j = 1, \ldots, k$.

(xi) For $1 \le p \le \infty$, the map

$$x \mapsto |x(a)| + \|x^{(1)}\|_p, \qquad x \in C^1[a,b],$$

is a norm on $C^1[a,b]$. But the map

$$x \mapsto \|x^{(1)}\|_p, \qquad x \in C^1[a,b],$$

is *not* a norm on $C^1[a,b]$. To see the last statement, it is enough to note that $\|x^{(1)}\|_p = 0$ implies $x^{(1)} = 0$ so that x is a constant function, not necessarily the zero function.

Remark 2.5 (a) We have seen in Example 1.2(vii) that

$$c_{00} \subset \ell^1 \subset c_0 \subset c \subset \ell^\infty,$$

and the above inclusions are strict. In fact, we can replace ℓ^1 by ℓ^p for any p with $1 \le p < \infty$, and the same sequences in Example 1.2(vii) will serve to show that the inclusions

$$c_{00} \subset \ell^p \subset c_0 \subset c \subset \ell^\infty,$$

are strict.

(b) One may look at the *unit circle* in the space $\mathbb{R}^2$ with respect to the norm $\|\cdot\|_p$ for $1 \le p \le \infty$, i.e., the set

$$\Gamma_p = \{x \in \mathbb{R}^2 : \|x\|_p = 1\}.$$

It is seen that Γ_2 is the *usual circle* in $\mathbb{R}^2$, Γ_1 is the square with vertices $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$, and Γ_∞ is the square with vertices $(1,1)$, $(-1,1)$, $(-1,-1)$, $(1,-1)$. Thus, Γ_2 lies outside Γ_1 and inside Γ_∞. It can be verified that Γ_p lies outside Γ_1 and inside Γ_∞ for every p with $1 < p < \infty$. If we denote the closed unit ball in $\mathbb{R}^2$ with respect to $\|\cdot\|_p$ by Δ_p, then it is seen (*Verify*) that for $1 \leq r \leq 2 \leq s \leq \infty$,

$$\Delta_1 \subseteq \Delta_r \subseteq \Delta_2 \subseteq \Delta_s \subseteq \Delta_\infty.$$

Another property of the unit circle Γ_p is that if $1 < p < \infty$, then (*Verify*) for every $x, y \in \Gamma_p$,

$$x \neq y \implies \left\|\frac{1}{2}(x+y)\right\|_p < 1.$$

This property of a norm is called *strict convexity*.

A normed linear space X is said to be **strictly convex** if for $x, y \in X$,

$$\|x\| = 1 = \|y\|, \quad x \neq y \implies \left\|\frac{1}{2}(x+y)\right\| < 1.$$

It can be seen that the spaces $\ell^p(n)$ and $\ell^p(\mathbb{N})$ for $1 < p < \infty$ are strictly convex.

In sections 2.1.2 - 2.1.4, we consider a few special classes of normed linear spaces, namely, quotient spaces, product spaces and inner product spaces. We shall also show that *inner product spaces* are strictly convex.

2.1.2 Seminorms and Quotient Spaces

After seeing examples of norms on $C[a,b]$, one may wonder whether

$$\nu_p(x) = \left(\int_a^b |x(t)|^p \, dt\right)^{1/p}, \qquad x \in \mathcal{R}[a,b],$$

for $1 \leq p < \infty$, defines a norm on $\mathcal{R}[a,b]$, the linear space of all Riemann integrable functions on $[a,b]$. As in the case of $C[a,b]$, it can be seen (*Verify*) that

$$\nu_p(x+y) \leq \nu_p(x) + \nu_p(y), \qquad \nu_p(\alpha x) = |\alpha| \nu_p(x)$$

for all $x, y \in \mathcal{R}[a,b]$, $\alpha \in \mathbb{K}$. The only property of the norm that ν_p does not share is positive definiteness. There are nonzero functions $x \in \mathcal{R}[a,b]$ with $\nu_p(x) = 0$. The above function ν_p is the motivation for us to define the notion of a *seminorm.*

Let X be a linear space. A function $\nu : X \to \mathbb{R}$ is called a **seminorm** on X if

$$\nu(x+y) \leq \nu(x) + \nu(y), \qquad \nu(\alpha x) = |\alpha|\nu(x)$$

for all $x, y \in X$, $\alpha \in \mathbb{K}$.

As we have remarked after the definition of a norm, it can be seen that if $\nu(\cdot)$ is a seminorm on X, then it also satisfies

$$\nu(0) = 0, \qquad \nu(x) \geq 0 \quad \forall x \in X.$$

Clearly, every norm is a seminorm, whereas a seminorm $\nu(\cdot)$ on X is a norm if and only if $\nu(x) \neq 0$ for every nonzero $x \in X$.

Exercise 2.3 If X is a normed linear space with norm $\|\cdot\|$ and $\nu : X \to \mathbb{R}$ is a seminorm on X, then show that $\|x\|_* := \|x\| + \nu(x)$, $x \in X$, defines another norm on X. $\square$

A seminorm induces a norm on certain quotient space in a natural way.

Let X be a linear space and $\nu : X \to \mathbb{R}$ be a seminorm on X. Then it is seen (*Verify*) that

$$Z_\nu = \{x \in X : \nu(x) = 0\}$$

is a subspace of X, and the map

$$[x] \mapsto \nu(x), \quad [x] \in X/Z_\nu,$$

is a norm on the **quotient space** X/Z_ν. The above norm on the quotient space is called the **quotient norm** associated with the seminorm ν.

Recall from Section 1.1.2 that the quotient space X/Z_ν is the linear space consisting of all equivalence classes $[x]$, i.e.,

$$[x] = \{u \in X : x - u \in Z_\nu\}, \qquad x \in X,$$

with addition and scalar multiplication defined by

$$[x] + [y] := [x+y], \qquad \alpha\,[x] := [\alpha x]$$

for every $x, y \in X$ and for every $\alpha \in \mathbb{K}$.

EXAMPLE 2.2 (i) The map $x \mapsto \nu(x) := \int_a^b |x(t)|\,dt$ is a seminorm on $\mathcal{R}[a,b]$.

(ii) Let X be a linear space and Y be a normed linear space with norm $\|\cdot\|$, and $A : X \to Y$ be a linear operator. Then the map

$$x \mapsto \nu_A(x) := \|Ax\|, \qquad x \in X,$$

is a seminorm on X. This seminorm ν_A is a norm if and only if A is an injective operator (*Verify*).

As a particular case of the above, for any linear functional f on a linear space X,

$$x \mapsto |f(x)|, \qquad x \in X,$$

defines a seminorm on X. As concrete examples of the above situations, we see that

$$x \mapsto \|x'\|_\infty, \qquad x \in C^1[a,b],$$

is a seminorm on $C^1[a,b]$, and for each $t \in [a,b]$, the map

$$x \mapsto |x(t)|, \qquad x \in C[a,b],$$

is a seminorm on $C[a,b]$.

(iii) Let X be a normed linear space and X_0 be subspace of X. Let $\nu : X \to \mathbb{R}$ be defined by

$$\nu(x) = \operatorname{dist}(x, X_0), \qquad x \in X.$$

Then it is easily seen (*Verify*) that ν is a seminorm on X. Since

$$\operatorname{dist}(x, X_0) = 0 \iff x \in \overline{X_0},$$

we have

$$Z_\nu := \{x \in X : \nu(x) = 0\} = \overline{X_0},$$

and the norm on the quotient space X/Z_ν is given by

$$\|[x]\|_* := \nu(x) = \operatorname{dist}(x, X_0), \qquad [x] \in X/Z_\nu.$$

Clearly, $Z_\nu = X_0$ if and only if X_0 is a closed subspace of X.

Thus, given any closed subspace X_0 of a linear space X, the quotient space X/X_0 is a normed linear space and the norm is given by

$$\|[x]\| = \operatorname{dist}(x, X_0), \qquad [x] \in X/X_0.$$

We note that

$$\nu(x) = \|x\| \quad \forall x \in X \iff X_0 = \{0\},$$

$$X/Z_\nu = \{0\} \iff X_0 \text{ dense in } X.$$

2.1.3 Measurable Functions and L^p Spaces

To consider the next example, we recall some basic definitions and results from the theory of Lebesgue measure and integration. For more details, and for proofs of the results stated in this section, one may refer to the book by Royden [26] or de Barra [8].

For an open interval I, we denote by $\ell(I)$ the length of I, and for $E \subseteq \mathbb{R}$, the (Lebesgue) **outer measure** of E is defined as the quantity

$$\mu^*(E) = \inf_{\mathcal{I}} \left\{ \sum_{n=1}^{\infty} \ell(I_n) : E \subseteq \bigcup_{n=1}^{\infty} I_n \right\},$$

where the infimum is taken over the set $\mathcal{I}$ of all sequences (I_n) of open intervals such that $E \subseteq \cup_{n=1}^{\infty} I_n$.

A set $E \subseteq \mathbb{R}$ is said to be (Lebesgue) **measurable** if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c) \quad \forall A \subseteq \mathbb{R}.$$

Here E^c denotes the complement of E. The set $\mathcal{M}$ of all Lebesgue measurable sets has the following properties:

(a) $\varnothing,\ \mathbb{R} \in \mathcal{M}$,

(b) $E \in \mathcal{M} \implies E^c \in \mathcal{M}$,

(c) $\{E_n : n \in \mathbb{N}\} \subseteq \mathcal{M} \implies \cup_{n=1}^{\infty} E_n \in \mathcal{M}$.

It can be shown that open sets belong to $\mathcal{M}$ so that countable unions and countable intersections of open sets as well as closed sets also belong to $\mathcal{M}$. It can also be shown (using *axiom of choice*) that not every subset of $\mathbb{R}$ is in $\mathcal{M}$.

For $E \in \mathcal{M}$, $\mu^*(E)$ is called the *Lebesgue measure of E*, and the extended real-valued function $E \mapsto \mu^*(E)$, $E \in \mathcal{M}$, i.e., the restriction of μ^* to $\mathcal{M}$, is called the **Lebesgue measure** on $\mathbb{R}$. We shall denote the Lebesgue measure by μ, i.e., $\mu(E) = \mu^*(E)$ for all $E \in \mathcal{M}$. It can be established that $\mu(\varnothing) = 0$, and if (E_n) is a sequence of pairwise disjoint members of $\mathcal{M}$, i.e., $E_n \in \mathcal{M}$ for every $n \in \mathbb{N}$ and $E_n \cap E_m = \varnothing$ whenever $n \neq m$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) = \sum_{n=1}^{\infty} \mu(E_n).$$

An important property of Lebesgue measure is its *regularity*, i.e., for every $E \in \mathcal{M}$ and $\varepsilon > 0$, there exist an open set $G \supseteq E$ and a closed set $F \subseteq E$ such that

$$\mu(G \setminus E) < \varepsilon, \qquad \mu(E \setminus F) < \varepsilon.$$

Thus, every measurable set can be *approximated* measure theoretically by open sets *from outside* and closed sets *from inside*.

An extended real-valued function f defined on $E \in \mathcal{M}$ is said to be (Lebesgue) **measurable**, if for every $a \in \mathbb{R}$,

$$\{t \in E : f(t) > a\} \in \mathcal{M}.$$

A complex-valued function f defined on $E \in \mathcal{M}$ is said to be *measurable* if its real and imaginary parts are measurable.

It can be shown that the set of all $\mathbb{K}$-valued measurable functions on $E \in \mathcal{M}$ is a linear space over $\mathbb{K}$.

Suppose ϕ is a *simple* measurable function, i.e., ϕ is of the form $\phi = \sum_{j=1}^{k} c_j \chi_{E_j}$ for some $E_j \in \mathcal{M}$ and $c_j \in \mathbb{K}$. Then the *integral* of ϕ over $E \in \mathcal{M}$ is defined by

$$\int_E \phi \, d\mu = \sum_{j=1}^{k} c_j \mu(E_j \cap E).$$

Here, χ_E denotes the characteristic function of the set E.

If f is a non-negative extended real-valued measurable function on $E \in \mathcal{M}$, then its integral over E is defined by

$$\int_E f \, d\mu = \sup_S \int_E \phi \, d\mu,$$

where S is the set of all simple non-negative measurable functions ϕ such that $0 \leq \phi \leq f$. In fact, this definition is motivated by the following result.

Proposition 2.8 *Let f be a non-negative extended real-valued measurable function on $E \in \mathcal{M}$, and for each $n \in \mathbb{N}$, let*

$$E_{j,n} = \left\{ t \in E : \frac{j-1}{2^n} \leq f(t) < \frac{j}{2^n} \right\}, \qquad j = 1, \ldots, n\,2^n,$$

$$F_n = \{t \in E : f(t) \geq n\}, \qquad \phi_n = \sum_{j=1}^{n\,2^n} \frac{j-1}{2^n} \chi_{j,n} + n\,\chi_n,$$

where $\chi_{j,n}$ and χ_n are the characteristic functions of $E_{j,n}$ and F_n, respectively. Then $0 \leq \phi_n \leq \phi_{n+1} \leq f$ for every $n \in \mathbb{N}$, and

$$\phi_n(t) \to f(t) \quad \textit{as} \quad n \to \infty \quad \forall t \in E.$$

Suppose $E \in \mathcal{M}$ and f is an extended real-valued measurable function defined on E. If one of the integrals $\int_E f^+ \, d\mu$, $\int_E f^- \, d\mu$ is finite, then the integral of f over E is defined by

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

where f^+ and f^- are, respectively, the positive and negative parts of f, i.e.,

$$f^+(t) = \max\{f(t), 0\}, \qquad f^-(t) = \max\{-f(t), 0\} \quad \forall t \in E.$$

Next, suppose that f is a complex-valued measurable function defined on $E \in \mathcal{M}$. We say that f is **integrable** over E if

$$\int_E |f| \, d\mu < \infty.$$

Here, $|f|$ denotes the function defined by $|f|(t) = |f(t)|$, $t \in E$.

It can be seen that, if f is integrable, then both $\mathrm{Re}\,(f)$ and $\mathrm{Im}\,(f)$ are integrable over E and the integrals $\int_E \mathrm{Re}\,(f) \, d\mu$ and $\int_E \mathrm{Im}\,(f) \, d\mu$ are finite. In that case, we define the integral of f over E by

$$\int_E f \, d\mu = \int_E \mathrm{Re}(f) \, d\mu + i \int_E \mathrm{Im}\,(f) \, d\mu.$$

Then we have the relation

$$\left| \int_E f \, d\mu \right| \le \int_E |f| \, d\mu.$$

We say that a property P concerning elements of $E \in \mathcal{M}$ holds **almost everywhere** on E if the set of points in E at which P does not hold is of measure zero. That is, if $E_P = \{t \in E : P \text{ holds at } t\}$, then $\mu(E \setminus E_P) = 0$. In such a case, we say that "P holds a.e. on E". For example, P could be the property that two given functions f and g being equal so that

$$f = g \text{ a.e. on } E \iff \mu(\{t \in E : f(t) \ne g(t)\}) = 0,$$

or P could be the property that a given sequence (f_n) of measurable functions converges so that (f_n) converges a.e. on E if and only if $\mu(\{t \in E : (f_n(t)) \text{ does not converge}\}) = 0$.

Suppose f is a non-negative measurable function defined on $E \in \mathcal{M}$, and suppose $\mu(E) > 0$. Then the following results hold:

$$\int_E f\,d\mu < \infty \implies f \text{ is finite a.e. on } E,$$

$$\int_E f\,d\mu = 0 \implies f = 0 \text{ a.e. on } E.$$

We shall have occasions to refer to the following results.

Theorem 2.9 (Fatou's lemma) *If (f_n) is a sequence of non-negative measurable functions on $E \in \mathcal{M}$, then*

$$\int_E (\liminf_n f_n)\,d\mu \le \liminf_n \int_E f_n\,d\mu.$$

Theorem 2.10 (Monotone convergence theorem) *If (f_n) is a sequence of non-negative measurable functions on $E \in \mathcal{M}$ such that $f_n \le f_{n+1}$ for all $n \in \mathbb{N}$ and $(f_n(t))$ converges to $f(t)$ for each $t \in E$, then*

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

Theorem 2.11 (Dominated convergence theorem) *Suppose (f_n) is a sequence of real or complex-valued measurable functions defined on $E \in \mathcal{M}$ such that $|f_n(t)| \le g(t)$ for every $n \in \mathbb{N}$ and for every $t \in E$, where g is an integrable function on E. If $(f_n(t))$ converges for each $t \in E$, then the function f defined by*

$$f(t) = \lim_{n\to\infty} f_n(t), \qquad t \in E,$$

is integrable, and

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

The space $L^p[a, b]$

Let $E \in \mathcal{M}$ with $\mu(E) > 0$, and $1 \le p \le \infty$. For a measurable function $f : E \to \mathbb{K}$ and for $1 \le p < \infty$, let

$$\nu_p(f) := \left(\int_E |f|^p\,d\mu\right)^{1/p},$$

and let

$$\nu_\infty(f) := \inf\{\alpha > 0 : |f| \le \alpha \text{ a.e. on } E\}.$$

Also, for $1 \le p \le \infty$, let

$$\mathcal{L}^p(E) = \{f \in \mathcal{F}(E, \mathbb{K}) : f \text{ measurable and } \nu_p(f) < \infty\}.$$

For $f \in \mathcal{L}^\infty(E)$, the quantity

$$\nu_\infty(f) := \inf\{\alpha > 0 : |f| \le \alpha \text{ a.e. on } E\}$$

is called the **essential supremum** of f.

Clearly,

$$\nu_p(\alpha f) = |\alpha| \nu_p(f) \quad \forall f \in \mathcal{L}^p(E),\ \alpha \in \mathbb{K}.$$

Thus, it is seen that once the inequality

$$\nu_p(f+g) \le \nu_p(f) + \nu_p(g), \qquad f, g \in \mathcal{L}^p(E)$$

is satisfied, then $\mathcal{L}^p(E)$ is a linear space and the map $f \mapsto \nu_p(f)$ is a seminorm on $\mathcal{L}^p(E)$. The above triangle inequality, also referred to as Minkowski's inequality, will be proved after proving the *Hölder's inequality* for measurable functions.

Hölder's inequality for measurable functions

Proposition 2.12 *Let f and g be measurable functions on E. Then*

$$\int_E |fg|\, d\mu \le \nu_p(f)\, \nu_q(g), \qquad f, g \in \mathcal{L}^p(E),$$

where p and q are conjugate exponents.

Proof. It is easily seen that the inequality holds if $p = 1$ or if $p = \infty$. Hence, assume that $1 < p < \infty$ so that $1 < q < \infty$. Also, it is seen that the inequality holds if one of $\nu_p(f)$, $\nu_p(g)$ is zero or infinity. Therefore, we assume that

$$0 < \nu_p(f) < \infty, \qquad 0 < \nu_p(g) < \infty.$$

In this case we obtain the inequality as in Example 2.1(ix). ∎

Minkowski's inequality for measurable functions

Proposition 2.13 *Let f and g be measurable functions on E, and p and q be conjugate exponents. Then*

$$\nu_p(f+g) \le \nu_p(f) + \nu_p(g), \qquad f, g \in \mathcal{L}^p(E),$$

Proof. We note that the inequality holds if one of $\nu_p(f)$, $\nu_p(g)$ is infinity. Hence, assume that

$$\nu_p(f) < \infty, \qquad \nu_p(g) < \infty.$$

It is easily seen that the inequalities hold if $p = 1$ or $p = \infty$. Next, suppose $1 < p < \infty$. Then from the relation

$$|f(t) + g(t)|^p \leq (|f(t)| + |g(t)|)^p \leq 2^p\,(|f(t)|^p + |g(t)|^p)$$

it follows, by taking integrals on both sides, that

$$\nu_p(f + g) < \infty.$$

Now, the required inequality follows as in Example 2.1(ix) above. ∎

Thus, we can conclude that $\mathcal{L}^p(E)$ is a linear space and ν_p is a seminorm on $\mathcal{L}^p(E)$. Hence, it follows that

$$[f] \mapsto \nu_p(f)$$

is a norm on the quotient space $\mathcal{L}^p(E)/\sim$, where $\sim$ is the equivalence relation on $\mathcal{L}^p(E)$ induced by ν_p, and $[f]$ is the equivalence class of $f \in \mathcal{L}^p(E)$. We denote this quotient space $\mathcal{L}^p(E)/\sim$ by $L^p(E)$.

In the sequel, for the sake of simplicity of notation, an $f \in \mathcal{L}^p(E)$ and the corresponding $[f] \in L^p(E)$ will be denoted by the same symbol f.

2.1.4 Product Space and Graph Norm

Let $X_1, \ldots, X_n$ be normed linear spaces with norms $\|\cdot\|_1, \ldots, \|\cdot\|_n$, respectively, and let $X = X_1 \times \cdots \times X_n$ be the product linear space. For $x = (x_1, \ldots, x_n) \in X$ and $j \in \{1, \ldots, n\}$, let $\pi_j(x)$ be the element in X with its j-th component as x_j and all other components as zeroes, i.e.,

$$\pi_j(x) = (y_1, \ldots, y_n), \qquad y_i = \begin{cases} x_j & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

A norm $\|\cdot\|_X$ on X is called a **product norm** if

$$\|\pi_j(x)\|_X = \|x_j\|_j \quad \forall j \in \{1, \ldots, n\}, \forall x = (x_1, \ldots, x_n) \in X.$$

It is easily seen that

$$\|(x_1, \ldots, x_n)\|_p = \begin{cases} \left(\sum_{j=1}^{n} \|x_j\|_j^p\right)^{1/p} & \text{if } 1 \leq p < \infty \\ \max\{\|x_1\|_1, \ldots, \|x_n\|_n\} & \text{if } p = \infty \end{cases}$$

defines a product norm on X. The proof is exactly the same as in the case of $\|\cdot\|_p$ on $\mathbb{K}^n$ with $\|\cdot\|_j$ in place of $|\cdot|$.

Let X and Y be normed linear spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $A : X \to Y$ be a linear operator. Then the subset

$$G(A) = \{(x, Ax) : x \in X\}$$

of the product space $X \times Y$ is called the **graph** of A. It can be easily seen that $G(A)$ is a subspace of $X \times Y$. Let $\|\cdot\|_{X \times Y}$ be a product norm on $X \times Y$. Then it is seen (*Verify*) that

$$x \mapsto \|x\|_A := \|(x, Ax)\|_{X \times Y}$$

is a norm on X, called the **graph norm** on X induced by the operator A.

For example, taking $\|(x, y)\| := \|x\| + \|y\|$ on $X \times Y$ as the product norm,

$$\|x\|_A := \|x\| + \|Ax\|, \qquad x \in X,$$

is the corresponding graph norm. As a concrete example, consider $X = C^1[a, b]$ and $Y = C[a, b]$, both with the norm $\|\cdot\|_\infty$. Then the graph norm on $C^1[a, b]$ induced by the linear operator $x \mapsto Ax := x'$ is

$$x \mapsto \|x\|_* := \|x\|_\infty + \|x'\|_\infty, \quad x \in C^1[a, b].$$

2.1.5 Inner Product Spaces

We recall from elementary vector algebra that the *dot product* of two nonzero vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ in $\mathbb{R}^3$ is defined by

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

and this dot product is related to the angle θ between $\mathbf{a}$ and $\mathbf{b}$ by the relation

$$\cos\theta = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{|\mathbf{a}|\,|\mathbf{b}|},$$

where $|\mathbf{a}| = \left(a_1^2 + a_2^2 + a_3^2\right)^{1/2}$ is the *absolute value* of $\mathbf{a} = (a_1, a_2, a_3)$.

We may observe that the map which associates each pair of vectors $\mathbf{a}$, $\mathbf{b}$ to the number $\mathbf{a} \cdot \mathbf{b}$ is *symmetric* and *linear* in each variable. Motivated by these properties, we introduce the notion of an *inner product* on an arbitrary linear space.

An **inner product** on a linear space X is a map

$$(x, y) \mapsto \langle x, y \rangle \in \mathbb{K}, \quad (x, y) \in X \times X,$$

which satisfies the following axioms:

(a) $\langle x, x \rangle \geq 0 \quad \forall x \in X$, and $\langle x, x \rangle = 0 \iff x = 0$,

(b) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall x, y, z \in X$,

(c) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \quad \forall \alpha \in \mathbb{K}$ and $\forall x, y \in X$, and

(d) $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x, y \in X$.

Here, and throughout this book, if α is a complex number, then the *complex conjugate* of α is denoted by $\bar{\alpha}$.

A linear space X together with an inner product $\langle \cdot, \cdot \rangle$ is called an **inner product space**.

As in the case of the definition of a norm, we see that the property $\langle 0, 0 \rangle = 0$ of an inner product follows from axioms (c) and (d). Indeed, for any $x \in X$, we have

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

Schwarz inequality

We may observe that the *dot product*

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3, \qquad \mathbf{a}, \mathbf{b} \in \mathbb{R}^3,$$

is an inner product on $\mathbb{R}^3$, and the *absolute value*

$$|\mathbf{a}| = (\mathbf{a} \cdot \mathbf{a})^{1/2}, \qquad \mathbf{a} \in \mathbb{R}^3,$$

defines a norm on $\mathbb{R}^3$. A question that naturally arises is: given an inner product $\langle \cdot, \cdot \rangle$ on a linear space X, is the map

$$x \mapsto \|x\| := \langle x, x \rangle^{1/2}, \qquad x \in X,$$

a norm on X? The answer is in the affirmative. To prove this we make use of the following theorem.

Theorem 2.14 (Schwarz inequality) *Let X be an inner product space, and $x, y \in X$. Then*

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

Equality holds in the above inequality if and only if x and y are linearly dependent.

Proof. The result is obvious if $y = 0$. Hence, assume that $y \neq 0$. Then for every $\alpha \in \mathbb{K}$ we have

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

By taking $\alpha = \langle x, y\rangle / \|y\|^2$, we have

$$|\alpha|^2\|y\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2}, \qquad \langle x, \alpha y\rangle = \bar{\alpha}\langle x, y\rangle = \frac{|\langle x, y\rangle|^2}{\|y\|^2}$$

so that

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

Thus, we get the inequality $|\langle x, y\rangle| \leq \|x\| \, \|y\|$. It also follows that

$$|\langle x, y\rangle| = \|x\| \, \|y\| \iff \|x - \alpha y\| = 0$$

and, in that case, x and y are linearly dependent. Also, if x and y are linearly dependent, then it is seen that $|\langle x, y\rangle| = \|x\| \, \|y\|$. ∎

Theorem 2.15 *If $\langle \cdot, \cdot \rangle$ is an inner product on a linear space X, then*

$$x \mapsto \|x\| := \langle x, x\rangle^{1/2}, \qquad x \in X,$$

is a norm on X.

Proof. Let $\langle \cdot, \cdot \rangle$ be an inner product on a linear space X and $x, y \in X$. Then using the Schwarz inequality, we obtain

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

Thus, $\|\cdot\|$ satisfies the inequality $\|x+y\| \leq \|x\| + \|y\|$ for every $x, y \in X$. The other conditions for $\|\cdot\|$ to be a norm are easily seen to be satisfied. ∎

Exercise 2.4 Let X be an inner product space and $x, y \in X$. Show the following:

(i) $\|x+y\| = \|x\| + \|y\|$ if and only if $y = 0$ or $x = \alpha y$ for some positive scalar α.

(ii) $\|x+\alpha y\| = \|x - \alpha y\|$ $\forall \alpha \in \mathbb{K}$ if and only if $\langle x, y\rangle = 0$. □

Remark 2.6 For nonzero vectors x and y in an inner product space X, by Schwarz inequality, we have

$$\frac{|\langle x, y\rangle|}{\|x\|\|y\|} \leq 1.$$

This relation motivates us to define the *angle* between two nonzero vectors x and y as

$$\theta_{x,y} := \cos^{-1}\left(\frac{|\langle x, y\rangle|}{\|x\|\,\|y\|}\right).$$

Note that if $x = cy$ for some nonzero scalar c, then $\theta_{x,y} = 0$, and if $\langle x, y\rangle = 0$, then $\theta_{x,y} = \pi/2$.

It is clear the if X_0 is a subspace of an inner product space X, then X_0 is also an inner product space with inner product as in X. Now we give a few more examples of inner product spaces.

EXAMPLE 2.3 (i) For $x = (\alpha_1, \ldots, \alpha_n)$ and $y = (\beta_1, \ldots, \beta_n)$ in $\mathbb{K}^n$, define

$$\langle x, y\rangle = \sum_{j=1}^{n} \alpha_j \bar{\beta}_j.$$

It is seen that $\langle \cdot, \cdot\rangle$ is an inner product on $\mathbb{K}^n$.

The above inner product is called the **standard inner product** on $\mathbb{K}^n$.

In general, if X is a finite dimensional linear space with basis $\{u_1, \ldots, u_n\}$ and the corresponding coordinate functionals $f_1, \ldots f_n$, then the function

$$(x, y) \mapsto \langle x, y\rangle = \sum_{j=1}^{n} f_j(x)\overline{f_j(y)}, \qquad (x, y) \in X \times X,$$

is an inner product on X.

(ii) The function

$$(x,y) \mapsto \langle x,y\rangle = \sum_{j=1}^{\infty} x(j)\overline{y(j)}, \qquad x,y \in \ell^2$$

is an inner product on ℓ^2: First one has to observe that $\sum_{j=1}^{\infty} x(j)\overline{y(j)}$ converges for all $x, y \in \ell^2$. This follows from the Cauchy-Schwarz inequality on ℓ^2. All the axioms for $\langle\cdot,\cdot\rangle$ to be an inner product are easy to verify.

(iii) The function

$$(f,g) \mapsto \langle f,g\rangle := \int_a^b f(t)\overline{g(t)}\, dt, \qquad f,g \in C[a,b]$$

is an inner product on $C[a,b]$.

(iv) For a measurable subset E of $\mathbb{R}$ with $\mu(E) > 0$, the function

$$(f,g) \mapsto \langle f,g\rangle := \int_E f\bar{g}\, d\mu, \qquad f,g \in L^2(E)$$

is an inner product on $L^2(E)$. Note that, in view of Cauchy-Schwarz inequality on $\mathcal{L}^p(E)$, $\int_E f\bar{g}\, d\mu \in \mathbb{K}$ for every $f, g \in L^2(E)$.

Parallelogram law

An important property of the norm induced by an inner product is that it satisfies the *parallelogram law*:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall\, x,y \in X.$$

The above equality follows by writing the norms in terms of inner products.

As a consequence of the above relation we have the following result.

Theorem 2.16 *Every inner product space is strictly convex.*

Proof. Let X be an inner product space. By parallelogram law, it follows that for every $x, y \in X$,

$$\|x\| = 1 = \|y\|, \quad x \neq y \implies \|x+y\|^2 = 4 - \|x-y\|^2 < 4.$$

Thus, X is strictly convex. ∎

The parallelogram law also helps us to show that certain norms cannot be induced by inner products. We have already seen that the norm $\|\cdot\|_2$ on each of the spaces $\ell^2(n)$, $\ell^2(\mathbb{N})$, $C[a,b]$ and $L^2[a,b]$ is induced by an inner product. Applying parallelogram law, we see below that if $p \neq 2$, then the norm $\|\cdot\|_p$ on $\ell^p(n)$, $\ell^p(\mathbb{N})$, $C[a,b]$ and $L^p[a,b]$ are not induced by inner products.

EXAMPLE 2.4 (i) For $1 \leq p \leq \infty$ let $X = \ell^p(S)$, where S is either $\{1,\dots,n\}$ or $\mathbb{N}$, and let $e_j \in X$ be such that $e_j(i) = \delta_{ij}$. Let

$$x = e_1 + e_2, \qquad y = e_1 - e_2.$$

Then we have

$$\|x+y\|_p = \|2e_1\|_p = 2, \qquad \|x-y\|_p = \|2e_2\|_p = 2,$$

$$\|x\|_p = \|y\|_p = \begin{cases} 2^{1/p} & \text{if } 1 \leq p < \infty \\ 1 & \text{if } p = \infty. \end{cases}$$

From these relations it follows that

$$\|x+y\|_p^2 + \|x-y\|_p^2 = 2\left(\|x\|_p^2 + \|y\|_p^2\right) \iff p = 2.$$

Thus, $\|\cdot\|_p$ is not induced by an inner product if $p \neq 2$.

(ii) For $1 \leq p \leq \infty$, let $Y = C[0,1]$ or $L^p[0,1]$ with $\|\cdot\|_p$. Taking

$$f(t) = t, \qquad g(t) = 1 - t \quad \forall t \in [0,1],$$

we have $\|f+g\|_p = 1$ and

$$\|f\|_p = \|g\|_p = \|f-g\|_p = \begin{cases} 1/(1+p)^{1/p} & \text{if } 1 \leq p < \infty \\ 1 & \text{if } p = \infty. \end{cases}$$

From these relations it follows that

$$\|f+g\|_p^2 + \|f-g\|_p^2 = 2\left(\|f\|_p^2 + \|g\|_p^2\right) \iff p = 2.$$

Thus, $\|\cdot\|_p$ is not induced by an inner product if $p \neq 2$.

Orthogonality

A notion special to inner product spaces as against a general normed linear space is *orthogonality*.

Let X be an inner product space, and $x, y \in X$. Then x and y in X are said to be **orthogonal** if $\langle x, y\rangle = 0$.

The following identity, called *Pythagoras theorem*, due to its geometric implication, can be easily deduced using the definition of an inner product.

Theorem 2.17 (Pythagoras theorem) *Let X be an inner product space, and x and y be orthogonal in X. Then*

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

It is easily seen that, if the scalar field is $\mathbb{R}$, then the converse of the Pythagoras theorem also holds. However, the converse need not be true if the scalar field is $\mathbb{C}$. A simple example shows this: Let $X = \mathbb{C}$ with standard inner product, and for nonzero real numbers $\alpha, \beta \in \mathbb{R}$, let $x = \alpha$, $y = i\,\beta$. Then we have

$$\|x+y\|^2 = \|\alpha + i\,\beta\|^2 = |\alpha|^2 + |\beta|^2 = \|x\|^2 + \|y\|^2,$$

but $\langle x, y\rangle = -i\,\alpha\beta \neq 0$.

A subset S of X is said to be an **orthogonal set** if every pair of distinct elements in S are orthogonal.

An orthogonal set S is said to be an **orthonormal set** if $\|x\| = 1$ for every $x \in S$. Thus, if E is an orthogonal set which does not contain the zero element, then the set

$$\widetilde{E} = \{x/\|x\| : x \in E\}$$

is an orthonormal set.

For a subset S of an inner product space X, we write

$$S^\perp = \{x \in X : \langle x, u\rangle = 0 \quad \forall u \in S\}.$$

The set $S^\perp$ is called the **annihilator** of S or the **orthogonal complement** of S.

We may note that

$$X^\perp = \{0\}, \qquad \{0\}^\perp = X.$$

The proof of the following proposition is very easy and is left to the reader.

Proposition 2.18 *Let X be an inner product space and $S \subseteq X$. Then*

(i) *$S^\perp$ is a closed subspace of X,*

(ii) *$S^\perp = \overline{S}^\perp$,*

(iii) *If S is dense in X, then $S^\perp = \{0\}$,*

(iv) *If S is an orthogonal set and $0 \notin S$, then S linearly independent.*

Gramm-Schmidt orthogonalization

Theorem 2.19 *Let X be an inner product space and $\{x_1, \ldots, x_n\}$ be linearly independent in X. Let $u_1 = x_1$, and for $j = 2, \ldots, n$, let*

$$u_j = x_j - \sum_{i=1}^{j-1} \frac{\langle x_j, u_i \rangle}{\langle u_i, u_i \rangle} u_i.$$

Then $\{u_1, \ldots, u_n\}$ is an orthogonal set, and

$$\operatorname{span}\{x_1, \ldots, x_j\} = \operatorname{span}\{u_1, \ldots, u_j\}, \qquad j = 1, \ldots, n.$$

Proof. Define $u_1 = x_1$ and

$$u_2 = x_2 - \frac{\langle x_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1.$$

Clearly, $u_2 \in \operatorname{span}\{x_1, x_2\}$ and $\langle u_2, u_1 \rangle = 0$. Since $\{x_1, x_2\}$ is linearly independent, $u_2 \neq 0$. Therefore, by Proposition 2.18(iv), $\{u_1, u_2\}$ is linearly independent, and hence, it follows that

$$\operatorname{span}\{u_1, u_2\} = \operatorname{span}\{x_1, x_2\}.$$

Having defined an orthogonal set $\{u_1, \ldots, u_{j-1}\}$ such that

$$\operatorname{span}\{u_1, \ldots, u_{j-1}\} = \operatorname{span}\{x_1, \ldots, x_{j-1}\},$$

let

$$u_j = x_j - \sum_{i=1}^{j-1} \frac{\langle x_j, u_i \rangle}{\langle u_i, u_i \rangle} u_i.$$

Then we have $u_j \in \operatorname{span}\{x_1, \ldots, x_j\}$ and $\langle u_j, u_k \rangle = 0$, $k < j$. Again, since $\{x_1, \ldots x_j\}$ is linearly independent, $u_j \neq 0$. Thus, $\{u_1, \ldots, u_n\}$ is the required orthogonal set. ∎

Remark 2.7 (a) From the above theorem, we can deduce that if $\{x_1, x_2, \ldots\}$ is a denumerable linearly independent set in X, then $\{u_1, u_2, \ldots\}$ defined by

$$u_1 = x_1, \qquad u_j = x_j - \sum_{i=1}^{j-1} \frac{\langle x_j, u_i\rangle}{\langle u_i, u_i\rangle} u_i, \qquad j = 2, 3, \ldots,$$

is an orthogonal set, and

$$\operatorname{span}\{x_1, \ldots, x_j\} = \operatorname{span}\{u_1, \ldots, u_j\} \quad \forall j \in \mathbb{N}.$$

(b) Theorem 2.19 also shows that every finite dimensional inner product space has a basis which is an orthonormal set. Let X be an inner product space of dimension n, and let $E = \{u_1, \ldots, u_n\}$ be a basis of X which is also an orthonormal set. Then, for every $x \in X$, we have

$$x = \sum_{j=1}^{n} \langle x, u_j\rangle u_j, \qquad \|x\|^2 = \sum_{j=1}^{n} |\langle x, u_j\rangle|^2.$$

Indeed, if $x = \alpha_1 u_1 + \cdots + \alpha_n u_n \in X$, then

$$\langle x, u_i\rangle = \alpha_1\langle u_1, u_i\rangle + \cdots + \alpha_n\langle u_n, u_i\rangle = \alpha_i \quad \forall i = 1, \ldots, n.$$

From this, the required representations of x and $\|x\|$ follow.

Do we have analogous representations in infinite dimensional spaces? We shall discuss this question in Chapter 4.

EXAMPLE 2.5 (i) It is seen that if $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{K}^n$, i.e., $e_j(i) = \delta_{ij}$ for $i, j = 1, \ldots, n$, then for each k with $1 \le k \le n$, $\{e_1, \ldots, e_k\}$ is an orthonormal subset of $\mathbb{K}^n$.

(ii) Let $e_j : \mathbb{N} \to \mathbb{K}$ be defined by $e_j(i) = \delta_{ij}$ for $i, j \in \mathbb{N}$. Then it is seen that $\{e_1, e_2, \ldots\}$ is an orthonormal subset of ℓ^2.

(iii) For $n \in \mathbb{Z}$, let

$$u_n(t) = \exp\left[\frac{2\pi i n(t-a)}{b-a}\right], \qquad t \in [a, b].$$

Then it is seen that $\{u_n : n \in \mathbb{Z}\}$ is an orthogonal subset of $C[a, b]$ and $L^2[a, b]$. Since $\|u_n\|_2 = \sqrt{b-a}$, the set $\{v_n : n \in \mathbb{Z}\}$ with $v_n = u_n/\sqrt{b-a}$, $n \in \mathbb{Z}$, is an orthonormal set.

(iv) Let $x_n(t) = t^{n-1}$ for $n \in \mathbb{N}$, $t \in [a,b]$. Since $\{x_1, x_2, \ldots\}$ is a linearly independent subset of $C[a,b]$, we obtain a set of orthogonal polynomials $\{u_1, u_2, \ldots\}$ using the Gramm-Schmidt orthogonalization procedure described in Theorem 2.19. If $a = -1$ and $b = 1$, then these orthogonal polynomials are called *Legendre polynomials.* These polynomials have many interesting and useful properties (see, for example, Wendroff [34] and Weinberger [33]).

2.1.6 Semi-inner Product and Sesquilinear Form

In the definition of an inner product, if the requirement "$\langle x, x\rangle = 0$ implies $x = 0$" is dropped, then such a map is called a *semi-inner product.* Thus, a map

$$\psi : X \times X \to \mathbb{K}$$

on a linear space X is called a **semi-inner product** if for every $x, y, u \in X$ and $\alpha \in \mathbb{K}$,

$$\begin{aligned}
\psi(x,x) &\geq 0,\\
\psi(x+u,y) &= \psi(x,y) + \psi(u,y),\\
\psi(\alpha x, y) &= \alpha\psi(x,y),\\
\psi(x,y) &= \overline{\psi(y,x)}.
\end{aligned}$$

EXAMPLE 2.6 (i) Suppose X is an inner product space with inner product $\langle\cdot,\cdot\rangle$ and $A : X \to X$ is a linear operator on X such that

$$\langle Ax, x\rangle \geq 0, \quad \langle Ax, y\rangle = \langle x, Ay\rangle \quad \forall x, y \in X.$$

Then it is easily seen that the map $(x,y) \mapsto \langle Ax, y\rangle$ is a semi-inner product.

(ii) The map

$$(f,g) \mapsto \int_a^b f\,\bar{g}\,d\mu, \qquad f, g \in \mathcal{L}^2[a,b],$$

is a semi-inner product on $\mathcal{L}^2[a,b]$.

(iii) Let X be an inner product space with inner product $\langle\cdot,\cdot\rangle$, and let $\{u_1, \ldots, u_n\} \subset X$. Then the map

$$(x,y) \mapsto \sum_{j=1}^{n} \langle x, u_j\rangle\langle u_j, y\rangle, \qquad x, y \in X,$$

defines a semi-inner product on X.

We may observe that this example is a particular case of the example (i) above, obtained by taking

$$Ax = \sum_{j=1}^{n} \langle x, u_j \rangle u_j, \qquad x \in X.$$

(iv) Let $t_1, \ldots, t_n$ be points in $[a, b]$. Then the map

$$(x, y) \mapsto \sum_{j=1}^{n} x(t_j)\overline{y(t_j)}, \qquad x, y \in C[a, b],$$

defines a semi-inner product on $C[a, b]$.

An important property that semi-inner products share with inner products is the Schwarz inequality. For its proof we make use of *polarization identities* which can be proved in a more general context.

Note that if $\psi(\cdot, \cdot)$ is a semi-inner product on a linear space X, then it is linear in first variable and conjugate linear in second variable, i.e., for every $x, y, u \in X$, and $\alpha, \beta \in \mathbb{K}$,

$$\begin{aligned} \psi(\alpha x + \beta y, u) &= \alpha\psi(x, u) + \beta\psi(y, u), \\ \psi(u, \alpha x + \beta y) &= \bar{\alpha}\psi(u, x) + \bar{\beta}\psi(u, y). \end{aligned}$$

A function $\psi : X \times X \to \mathbb{K}$ having the above properties is called a **sesquilinear form**. If a sesquilinear form $\psi(\cdot, \cdot)$ also has the property that $\psi(x, y) = \overline{\psi(y, x)}$ for all $x, y \in X$, then it is said to be a **symmetric sesquilinear form**.

Suppose $\psi(\cdot, \cdot)$ is a symmetric sesquilinear form on a linear space X. Then we see that $\psi(x, x) \in \mathbb{R}$ for all $x \in X$. The map $q : X \to \mathbb{R}$ defined by

$$q(x) = \psi(x, x), \qquad x \in X,$$

is called the **quadratic form** associated with $\psi(\cdot, \cdot)$.

A simple and important example of a sesquilinear form is the following.

Let X be an inner product space and $A : X \to X$ be a linear operator. Then $\psi(x, y) := \langle Ax, y \rangle$ for $x, y \in X$, defines a sesquilinear form. If A has the additional property that $\langle Ax, y \rangle = \langle x, Ay \rangle$, then we see that the above $\psi(\cdot, \cdot)$ is symmetric as well.

Exercise 2.5 Let X be a linear space and $\psi(\cdot,\cdot) : X \times X \to \mathbb{K}$. Show the following:

(i) If $\psi(\cdot,\cdot)$ is a symmetric sesquiliner form, then $\psi(0,0) = 0$.

(ii) If $\psi(\cdot,\cdot)$ is a semi-inner product, then the associated quadratic form $x \mapsto q(x) := \psi(x,x)$ is a seminorm on X. □

Theorem 2.20 (Polarization identities) *Let $\psi(\cdot,\cdot)$ be a symmetric sesquilinear form on a linear space X and $q(\cdot)$ be the associated quadratic form. Then we have the following:*

(i) *If $\mathbb{K} = \mathbb{R}$, then*

$$\psi(x,y) = \frac{1}{4}\left[q(x+y) - q(x-y)\right]. \tag{2.9}$$

(ii) *If $\mathbb{K} = \mathbb{C}$, then*

$$\psi(x,y) = \frac{1}{4}\left[q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)\right] \tag{2.10}$$

Proof. It is easily seen that

$$q(x+y) - q(x-y) = 4\,\mathrm{Re}\,\psi(x,y)$$

for every $x, y \in X$. If the scalar field is the set of complex numbers, then we have

$$q(x+iy) - q(x-iy) = 4\,\mathrm{Im}\,\psi(x,y)$$

for every $x, y \in X$. Thus, if $\mathbb{K} = \mathbb{R}$, then we have

$$\psi(x,y) = \frac{1}{4}\left[q(x+y) - q(x-y)\right],$$

and if $\mathbb{K} = \mathbb{C}$, then

$$\psi(x,y) = \frac{1}{4}\left[q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)\right]$$

This completes the proof. ∎

Relations (2.9) and (2.10) are called **polarization identities** associated with the semi-inner product $\psi(\cdot,\cdot)$.

We shall make use of the polarization identities to obtain a generalized form of the Schwarz inequality.

Theorem 2.21 (Generalized Schwarz inequality) *Let $\psi(\cdot,\cdot)$ be a semi-inner product on a linear space X. Then*

$$|\psi(x,y)|^2 \leq \psi(x,x)\psi(y,y) \quad \forall x,y \in X.$$

Proof. Let $x, y \in X$. If one of $\psi(x,x)$, $\psi(y,y)$ is nonzero, then the inequality follows as in the proof of Schwarz inequality. Hence, assume that $\psi(x,x) = 0$ and $\psi(y,y) = 0$. We have to show that $\psi(x,y) = 0$. For this, first we observe that

$$\psi(x+y, x+y) + \psi(x-y, x-y) = 0.$$

In case $\mathbb{K} = \mathbb{C}$, then we also have

$$\psi(x+iy, x+iy) + \psi(x-iy, x-iy) = 0.$$

Since $\psi(u,u) \geq 0$ for every $u \in X$, the above relations, together with the polarization identities (2.9) and (2.10), imply that $\psi(x,y) = 0$. ∎

2.2 Banach Spaces

Normed linear spaces which are complete with respect to their induced metrics are of special importance. Many of the normed linear spaces which are useful in applications are in fact complete.

Recall that a complete normed linear space is known as a **Banach space.**

A Banach space is called a **Hilbert space** if its norm is induced by an inner product.

A Banach space over $\mathbb{C}$ is called a **complex Banach space**, and a Banach space over $\mathbb{R}$ is called a **real Banach space**. Similarly, we can define a **complex Hilbert space** and a **real Hilbert space.**

It is easily seen that if $\|\cdot\|$ and $\|\cdot\|_*$ are two norms on a linear space X such that

$$a\|x\| \leq \|x\|_* \leq b\|x\| \quad \forall x \in X,$$

where a and b are positive real numbers, and if X is a Banach space with respect to $\|\cdot\|$, then X is a Banach space with respect to $\|\cdot\|_*$ as well.

Two norms $\|\cdot\|$ and $\|\cdot\|_*$ on a linear space X are said to be **equivalent** if there exist positive real numbers a and b such that

$$a\|x\| \leq \|x\|_* \leq b\|x\| \quad \forall x \in X.$$

Exercise 2.6 (i) Let X be a linear space and $\mathcal{N}$ be the set of all norms on X. Show that the relation of equivalence of two norms is an equivalence relation on $\mathcal{N}$.

(ii) Let $X_1, \ldots, X_n$ be linear spaces with norms $\|\cdot\|_1, \ldots, \|\cdot\|_n$, respectively. For $1 \le p \le \infty$, consider the product norm $\|\cdot\|_p$ on $X_1 \times \cdots \times X_n$ defined by

$$\|(x_1, \ldots, x_n)\|_p = \begin{cases} \left(\sum_{j=1}^{n} \|x_j\|_j^p\right)^{1/p} & \text{if } 1 \le p < \infty, \\ \max\{\|x_j\|_j : j = 1, \ldots, n\} & \text{if } p = \infty. \end{cases}$$

Then show that, for every $r, s \in [1, \infty]$, the norms $\|\cdot\|_r$ and $\|\cdot\|_s$ are equivalent. □

As a first example of a Banach space, we have the following.

Theorem 2.22 *Let X be a finite dimensional normed linear space of dimension $k < \infty$, and let $E = \{u_1, \ldots, u_k\}$ be a basis of X. Consider the norm*

$$\|x\|_E := \max\{|f_1(x)|, \ldots, |f_k(x)|\}, \qquad x \in X,$$

on X, where $f_1, \ldots, f_k$ are the coördinate functionals on X associated with E. Then X with $\|\cdot\|_E$ is a Banach space.

Proof. Let (x_n) be a Cauchy sequence in X with respect to $\|\cdot\|_E$. Then it follows that

$$|f_j(x_n) - f_j(x_m)| \to 0 \quad \text{as} \quad n, m \to \infty$$

for each $j \in \{1, \ldots, k\}$. Hence, by the completeness of $\mathbb{K}$, there exists $\alpha_j \in \mathbb{K}$ such that

$$|f_j(x_n) - \alpha_j| \to 0 \text{ as } n \to \infty.$$

Hence,

$$x_n = \sum_{j=1}^{k} f_j(x_n)u_j \to x = \sum_{j=1}^{k} \alpha_j u_j \quad \text{as} \quad n \to \infty. \quad ■$$

It is an easy exercise to see that the norm $\|\cdot\|_E$ in the above theorem is equivalent to the norm $\|\cdot\|_{E,p}$, $1 \le p < \infty$, defined by

$$\|x\|_{E,p} = \left(\sum_{j=1}^{n} |f_j(x)|^p\right)^{1/p}, \qquad 1 \le p < \infty.$$

Therefore, the space X with the basis $E = \{u_1, \ldots, u_k\}$ is complete with respect to the norms $\|\cdot\|_{E,p}$ for any p with $1 \le p < \infty$ as well. In fact, we show soon that the completeness of a finite dimensional space is not restricted to the above norms only. Our procedure is to show that any two norms on a finite dimensional space are equivalent, and use Theorem 2.22 to conclude that every finite dimensional space is complete with respect to any norm. For this, we shall make use of the following result.

Lemma 2.23 *Let Y be a closed subspace and Z be a finite dimensional subspace of a normed linear space X. Then $Y + Z$ is a closed subspace of X.*

In particular, every finite dimensional subspace of a normed linear space is closed.

Proof. We prove the result by induction on the dimension of Z. Suppose that Z is of dimension 1, say $Z = \operatorname{span}\{v_1\}$. If $v_1 \in Y$, then $Y + Z = Y$ and it is closed. Hence, assume that $v_1 \notin Y$ so that $\operatorname{dist}(v_1, Y) > 0$.

Let (x_n) be a sequence in $Y + Z$ such that $x_n \to x$ for some $x \in X$. Since $(x_n) \in Y + \operatorname{span}\{v_1\}$, for every $n \in \mathbb{N}$, there exist $y_n \in Y$ and $\alpha_n \in \mathbb{K}$ such that $x_n = y_n + \alpha_n v_1$. Also, the convergence of (x_n) implies that it is a Cauchy sequence. Hence,

$$\operatorname{dist}(x_n - x_m, Y) \le \|x_n - x_m\| \to 0 \quad \text{as} \quad n, m \to \infty.$$

But

$$\operatorname{dist}(x_n - x_m, Y) = \operatorname{dist}(\alpha_n v_1 - \alpha_m v_1, Y) = |\alpha_n - \alpha_m| \operatorname{dist}(v_1, Y).$$

Thus, (α_n) is a Cauchy sequence in $\mathbb{K}$. Let $\alpha \in \mathbb{K}$ be such that $\alpha_n \to \alpha$ as $n \to \infty$. Hence,

$$y_n = x_n - \alpha_n v_1 \to x - \alpha v_1 \quad \text{as} \quad n \to \infty.$$

Since Y is closed, $y := x - \alpha v_1 \in Y$ so that

$$x_n = y_n + \alpha_n v_1 \to y + \alpha v_1 \in Y + Z.$$

Thus, we have shown that $Y + Z$ is closed if $\dim Z = 1$.

Next, suppose that Z is of dimension k, $k > 1$, and let $\{v_1, \ldots, v_k\}$ be a basis of Z. Let

$$X_j = Y + Z_j, \qquad Z_j = \operatorname{span}\{v_1, \ldots, v_j\}, \qquad j = 1, \ldots, k.$$

We have already proved that X_1 is a closed subspace of X. Now, assume that X_{j-1} is a closed subspace of X. Then, since $X_j = X_{j-1} + \operatorname{span}\{v_j\}$, it follows, using the argument as in the case of $j = 1$, that X_j is a closed subspace of X. Thus, by induction, $X_k = Y + Z_k = X_{k-1} + \operatorname{span}\{v_k\}$ is a closed subspace of X.

The particular case follows by taking $Y = \{0\}$. ∎

Theorem 2.24 *Any two norms on a finite dimensional linear space are equivalent.*

Proof. Let X be a finite dimensional normed linear space with norm $\|\cdot\|$ and let $E = \{u_1, \ldots, u_k\}$ be a basis of X. Consider the norm $\|\cdot\|_E$ defined in Theorem 2.22, i.e.,

$$\|x\|_E = \max\{|f_1(x)|, \ldots, |f_k(x)|\}, \qquad x \in X,$$

where $f_1, \ldots, f_k$ are the coordinate functionals corresponding to E. It is enough to show that there exist positive real numbers $a > 0$ and $b > 0$ such that

$$a\|x\|_E \le \|x\| \le b\|x\|_E \quad \forall x \in X.$$

Since every $x \in X$ can be written as $x = f_1(x)u_1 + \cdots + f_k(x)u_k$, we have

$$\begin{aligned}
\|x\| &\le |f_1(x)|\|u_1\| + \cdots + |f_k(x)|\|u_k\| \\
&\le \|x\|_E(\|u_1\| + \cdots + \|u_k\|) \\
&= b\|x\|_E,
\end{aligned}$$

where $b = \|u_1\| + \cdots + \|u_k\|$. Also, writing $X_j = \operatorname{span}\{u_i : i \ne j\}$,

$$\begin{aligned}
\|x\| &= \|f_1(x)u_1 + \cdots + f_k(x)u_k\| \\
&\ge \operatorname{dist}(f_j(x)u_j, X_j) \\
&= |f_j(x)| \operatorname{dist}(u_j, X_j) \\
&\ge a|f_j(x)|,
\end{aligned}$$

where $a = \min\{\operatorname{dist}(u_j, X_j) : j = 1, \ldots, k\}$. Hence,

$$\|x\| \geq a\|x\|_E.$$

Clearly, $b > 0$. Since, by Lemma 2.23, each X_j is closed and since $u_j \notin X_j$, we have $a > 0$ as well. ∎

Now we state the theorem that we promised. Its proof follows from Theorems 2.24 and 2.22.

Theorem 2.25 *Every finite dimensional space is complete with respect to any norm on it.*

Exercise 2.7 Show that every finite dimensional subspace of a normed linear space is a Banach space. □

Now we consider a few specific examples.

EXAMPLE 2.7 (i) For a nonempty set S, the space $\ell^\infty(S)$ of all bounded $\mathbb{K}$-valued functions on S is a Banach space with respect to $\|\cdot\|_\infty$:

Let (x_n) be a Cauchy sequence in $\ell^\infty(S)$, i.e., $\|x_n - x_m\|_\infty \to 0$ as $n, m \to \infty$. Let $\varepsilon > 0$, and let $N \in \mathbb{N}$ such that

$$\|x_n - x_m\| < \varepsilon \quad \forall n, m \geq N.$$

Then we have

$$|x_n(s) - x_m(s)| \leq \|x_n - x_m\|_\infty < \varepsilon \quad \forall n, m \geq N \ \forall s \in S.$$

Thus, since $\mathbb{K}$ is complete, the sequence $(x_n(s))$ converges for each $s \in S$. Let

$$x(s) = \lim_{n\to\infty} x_n(s), \qquad s \in S.$$

Therefore, for all $s \in S$ and for all $n \geq N$, we have

$$|x_n(s) - x(s)| = \lim_{m\to\infty} |x_n(s) - x_m(s)| \leq \lim_{m\to\infty} \|x_n - x_m\|_\infty \leq \varepsilon.$$

Hence,

$$\sup_{s\in S} |x_n(s) - x(s)| \leq \varepsilon \quad \forall n \geq N.$$

This, in particular, implies that $x \in \ell^\infty(S)$ and $\|x_n - x\|_\infty \to 0$ as $n \to \infty$.

Thus, we have proved that, for any nonempty set S, $\ell^\infty(S)$ is a Banach space. In particular, taking $S = \mathbb{N}$, we see that $\ell^\infty := \ell^\infty(\mathbb{N})$ is a Banach space.

(ii) Let Ω be a metric space. Then $X = C(\Omega)$, the space of all $\mathbb{K}$-valued bounded continuous functions on Ω, is a Banach space with respect to the norm $\|\cdot\|_\infty$:

Since $C(\Omega)$ is a subspace of the Banach space $\ell^\infty(\Omega)$, it is enough to show that $C(\Omega)$ is a closed subspace of $\ell^\infty(\Omega)$. For this, let (x_n) be a sequence in $C(\Omega)$ such that $\|x_n - x\|_\infty \to 0$ as $n \to \infty$ for some $x \in \ell^\infty(\Omega)$. For $\varepsilon > 0$, let $N \in \mathbb{N}$ be such that

$$\|x_n - x\|_\infty < \varepsilon \quad \forall n \geq N.$$

It is enough to show that $x \in C(\Omega)$. For this, let $t_0 \in \Omega$ and $\varepsilon > 0$. By the above relation, we have

$$\begin{aligned}
|x(t) - x(t_0)| &\leq |x(t) - x_N(t)| + |x_N(t) - x_N(t_0)| \\
&\quad + |x_N(t_0) - x(t_0)| \\
&\leq \|x_N - x\|_\infty + |x_N(t) - x_N(t_0)| + \|x_N - x\|_\infty \\
&\leq 2\varepsilon + |x_N(t) - x_N(t_0)|,
\end{aligned}$$

for all $t \in \Omega$. Since $x_N \in C(\Omega)$, there exists $\delta > 0$ such that

$$|x_N(t) - x_N(t_0)| < \varepsilon \quad \text{whenever} \quad d(t, t_0) < \delta,$$

where $d(\cdot,\cdot)$ is the metric on Ω. Thus, we see that

$$|x(t) - x(t_0)| < 3\,\varepsilon \quad \text{whenever} \quad d(t, t_0) < \delta.$$

Since $t_0 \in \Omega$ is arbitrary, it follows that $x \in C(\Omega)$.

Exercise 2.8 Suppose Ω is a metric space. Show that every Cauchy sequence in $C(\Omega)$ converges. □

EXAMPLE 2.7 (*cont.*) (iii) Following the arguments similar to that used in the last example, it can be seen that the spaces

$$\begin{aligned}
c_0 &:= \{x \in \ell^\infty : (x(n)) \text{ converges to } 0\}, \\
c &:= \{x \in \ell^\infty : (x(n)) \text{ converges}\}
\end{aligned}$$

are closed subspaces of ℓ^∞. Hence, they are Banach spaces with respect to the norm $\|\cdot\|_\infty$. In fact, the completeness of c can also be seen as follows: We note that if $x \in c$ and $\alpha := \lim_{n\to\infty} x(n)$, then $x = y + \alpha e_0$, where

$$y(j) = x(j) - \alpha, \qquad e_0(j) = 1 \quad \forall j \in \mathbb{N}.$$

Note that $y \in c_0$. Thus,

$$c = c_0 + \operatorname{span}\{e_0\}.$$

Since c_0 is a closed subspace of ℓ^∞, it follows by Lemma 2.23, that c is also a closed subspace of the Banach space ℓ^∞. Hence, c is also Banach space.

CONVENTION. Hereafter, we consider the norms on the spaces c_0 and c to be the ones that they inherit from ℓ^∞. Also, unless otherwise specified, the norm on the space $C[a,b]$ is taken to be $\|\cdot\|_\infty$.

In the next example, we use the notion of *compactness*:

Let Ω be a metric space. A set $E \subseteq \Omega$ is said to be **totally bounded** if for every $\varepsilon > 0$, there exist finite number of points $x_1, \ldots, x_n$ (n depends on ε) in E such that $E \subseteq \cup_{j=1}^n B(x_j, \varepsilon)$. A set $K \subseteq \Omega$ is said to be **compact** if it is complete and totally bounded.

It can be seen that compact sets are closed and bounded, whereas a closed and bounded set need not be compact. Also, we know from analysis that compact subsets of $\mathbb{K}$ are precisely the closed and bounded sets. In Section 2.4, we shall prove that finite dimensional normed linear spaces are characterized by compactness of its closed unit ball.

(iv) For a metric space Ω, let $C_0(\Omega)$ be the set of all those functions $x \in C(\Omega)$ with the property that for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq \Omega$ such that $|x(t)| < \varepsilon$ for all $t \notin K_\varepsilon$. We show that $C_0(\Omega)$ is a closed subspace of the Banach space $C(\Omega)$ so that it is a Banach space. To see this, suppose (x_n) in $C_0(\Omega)$ is such that $\|x_n - x\|_\infty \to 0$ as $n \to \infty$ for some $x \in C(\Omega)$. For $\varepsilon > 0$, let $N \in \mathbb{N}$ be such that

$$\|x_n - x\|_\infty < \varepsilon \quad \forall n \geq N.$$

Then, for every $t \in \Omega$, we have

$$|x(t)| \leq |x(t) - x_N(t)| + |x_N(t)| \leq \varepsilon + |x_N(t)|.$$

Let K be a compact subset of Ω such that $|x_N(t)| < \varepsilon$ for all $t \notin K$. Then we have

$$|x(t)| \leq \varepsilon + |x_N(t)| < 2\,\varepsilon \quad \forall t \notin K,$$

showing that $x \in C_0(\Omega)$.

We observe that if Ω is a compact metric space, then $C_0(\Omega) = C(\Omega)$, and if $\Omega = \mathbb{N}$ with discrete metric, then $C_0(\mathbb{N}) = c_0$, the space of all sequences which converge to zero. These cases have already been discussed independently in examples (ii) and (iii) above.

Note that $C_0(\mathbb{R})$ is the set of all $x \in C(\mathbb{R})$ such that $|x(t)| \to 0$ as $|t| \to \infty$. For this reason, the space $C_0(\mathbb{R})$ is also called the space of continuous functions which vanish at infinity.

(v) For $1 \le p \le \infty$, $\ell^p := \ell^p(\mathbb{N})$ is a Banach space:

We have already seen in (i) that $\ell^\infty(\mathbb{N})$ is a Banach space. So let $1 \le p < \infty$. Let (x_n) be a Cauchy sequence in ℓ^p, and let $\varepsilon > 0$. Let $N \in \mathbb{N}$ be such that

$$\|x_n - x_m\|_p < \varepsilon \quad \forall n, m \ge N.$$

This implies that for each $k \in \mathbb{N}$,

$$\sum_{j=1}^{k} |x_n(j) - x_m(j)|^p < \varepsilon^p \quad \forall n, m \ge N.$$

In particular, $(x_n(j))$ is a Cauchy sequence in $\mathbb{K}$ for each $j \in \mathbb{N}$. Since $\mathbb{K}$ is complete, there exist scalars $\alpha_1, \alpha_2, \ldots$ such that $x_n(j) \to \alpha_j$ as $n \to \infty$ for each $j \in \mathbb{N}$. Therefore, from the above inequality, we have, by letting $m \to \infty$,

$$\sum_{j=1}^{k} |x_n(j) - \alpha_j|^p \le \varepsilon^p \quad \forall n \ge N.$$

Letting k tend to infinity, we get

$$\sum_{j=1}^{\infty} |x_n(j) - \alpha_j|^p \le \varepsilon^p \quad \forall n \ge N.$$

Thus, it is enough to show that $x = (\alpha_1, \alpha_2, \ldots) \in \ell^p$. Note that, by Minkowski's inequality

$$\begin{aligned}\left(\sum_{j=1}^{k} |\alpha_j|^p\right)^{1/p} &\le \left(\sum_{j=1}^{k} |\alpha_j - x_N(j)|^p\right)^{1/p} + \left(\sum_{j=1}^{k} |x_N(j)|^p\right)^{1/p} \\ &\le \varepsilon + \|x_N\|_p,\end{aligned}$$

for all $k \in \mathbb{N}$. Therefore, $x \in \ell^p$.

(vi) For $1 \leq p \leq \infty$, and for a measurable subset E of $\mathbb{R}$ with $\mu(E) > 0$, the space $L^p(E)$ is a Banach space:

Recall that $\mathcal{L}^p(E)$ is the linear space of all measurable functions $f : E \to \mathbb{K}$ such that $\nu_p(f) < \infty$, where

$$\nu_p(f) := \begin{cases} \left(\int_E |f|^p \, d\mu\right)^{1/p} & \text{if } 1 \leq p < \infty, \\ \inf\{\alpha > 0 : |f| \leq \alpha \text{ a.e. on } E\} & \text{if } p = \infty \end{cases}$$

and $L^p(E)$ is the quotient normed linear space corresponding to the equivalence relation on $\mathcal{L}^p(E)$ induced by the seminorm ν_p (Section 1.1.3). Recall also that the norm on $L^p(E)$ is $[f] \mapsto \|[f]\|_p := \nu_p(f)$. As we have mentioned at the end of Section 2.1.3, the elements of $L^p(E)$ are also denoted by f with the understanding that elements f and g are identified if they are equal almost everywhere.

Now we show the completeness of the space $L^p(E)$. First let $1 \leq p < \infty$. Let (f_n) be a Cauchy sequence in $L^p(E)$. Our idea is to obtain a convergent subsequence of (f_n), so that the sequence (f_n) itself would converge. For this, let (δ_n) be a sequence of positive real numbers, such that $M := \sum_{j=1}^{\infty} \delta_j < \infty$. For each $j \in \mathbb{N}$, let $n_j \in \mathbb{N}$ be such that

$$\|f_n - f_m\|_p < \delta_j \quad \forall n, m \geq n_j.$$

We can assume, without loss of generality that $n_j < n_{j+1}$ for all $j \in \mathbb{N}$. Hence,

$$\left\|f_{n_{j+1}} - f_{n_j}\right\|_p < \delta_j, \qquad j \in \mathbb{N}.$$

Let

$$g = \sum_{j=1}^{\infty} \left|f_{n_{j+1}} - f_{n_j}\right|, \qquad g_k = \sum_{j=1}^{k} \left|f_{n_{j+1}} - f_{n_j}\right| \quad \forall k \in \mathbb{N}.$$

Then we see that $g_k(t) \to g(t)$ as $k \to \infty$ for every $t \in E$, and by Minkowski's inequality,

$$\|g_k\|_p \leq \sum_{j=1}^{k} \left\|f_{n_{j+1}} - f_{n_j}\right\|_p \leq \sum_{j=1}^{k} \delta_j.$$

Since $g_k^p(t) \to g^p(t)$ as $k \to \infty$ for every $t \in E$, it follows by Fatou's lemma (Theorem 2.9) that

$$\int_E g^p \, d\mu \leq \liminf_k \int_E g_k^p \, d\mu \leq M^p.$$

Thus, $g \in L^p(E)$. Hence, $g(t)$ is finite a.e., that is, there exists $F \subseteq E$ such that the series $\sum_{j=1}^{\infty} \left(f_{n_{j+1}} - f_{n_j}\right)(t)$ converges absolutely for every $t \in F$ and $\mu(E \setminus F) = 0$. Let

$$f(t) = f_{n_1}(t) + \sum_{j=1}^{\infty} \left(f_{n_{j+1}} - f_{n_j}\right)(t) \quad \forall t \in F$$

and $f(t) = 0$ for $t \in E \setminus F$. Since

$$f_{n_{k+1}} = f_{n_1} + \sum_{j=1}^{k} \left(f_{n_{j+1}} - f_{n_j}\right), \qquad k \in \mathbb{N},$$

we have the convergence $f_{n_k}(t) \to f(t)$ as $k \to \infty$ for every $t \in F$. Again, by Fatou's lemma (Theorem 2.9), for every $m \in \mathbb{N}$, we have

$$\int_E |f - f_m|^p \, d\mu \leq \liminf_j \int_E |f_{n_j} - f_m|^p \, d\mu = \liminf_j \|f_{n_j} - f_m\|_p^p.$$

Since (f_n) is a Cauchy sequence in $L^p(E)$, from the above relation it follows that

$$f - f_m \in L^p(E), \qquad f \in L^p(E),$$

and $\|f - f_m\|_p \to 0$ as $m \to \infty$. Thus, we have shown that the Cauchy sequence (f_n) converges in $L^p(E)$.

Next, we consider the case $p = \infty$. Let (f_n) be a Cauchy sequence in $L^\infty(E)$. For $k, m, n \in \mathbb{N}$, we know that

$$|f_k(t)| \leq \|f_k\|_\infty, \qquad |f_n(t) - f_m(t)| \leq \|f_n - f_m\|_\infty$$

for all $k, m, n \in \mathbb{N}$ almost everywhere on E. Let

$$\begin{aligned} A_k &= \{t \in E : |f_k(t)| > \|f_k\|_\infty\}, \\ B_{m,n} &= \{t \in E : |f_n(t) - f_m(t)| > \|f_n - f_m\|_\infty\}. \end{aligned}$$

Then we have $\mu(A_k) = 0 = \mu(B_{m,n})$ for every $k, m, n \in \mathbb{N}$, so that by taking $\Omega = (\cup_k A_k) \cup (\cup_{m,n} B_{m,n})$, we have $\mu(\Omega) = 0$. Thus, for $t \in E \setminus \Omega$, and for all $k, m, n \in \mathbb{N}$,

$$|f_k(t)| \leq \|f_k\|_\infty, \qquad |f_n(t) - f_m(t)| \leq \|f_n - f_m\|_\infty.$$

For every $n \in \mathbb{N}$, let $\tilde{f}_n$ be the restriction of f_n to the set $E \setminus \Omega$. Then it follows that $\tilde{f}_n \in \ell^\infty(E \setminus \Omega)$ and $(\tilde{f}_n)$ is a Cauchy sequence in $\ell^\infty(E \setminus \Omega)$. Since $\ell^\infty(E \setminus \Omega)$ is a Banach space, the sequence $(\tilde{f}_n)$ converges to some g in $\ell^\infty(E \setminus \Omega)$. Now, let $f(t) = g(t)$ for $t \in E \setminus \Omega$ and $f(t) = 0$ for $t \in \Omega$. Then we have the convergence of (f_n) to f in $L^\infty(E)$.

Now we give a general example.

(vii) The Quotient space of a Banach space is a Banach space:

Theorem 2.26 *Let X_0 be a closed subspace of a normed linear space X. If X is a Banach space, then X/X_0 is a Banach space.*

Proof. Suppose X_0 is a closed subspace of a Banach space X. Recall that elements of the quotient space X/X_0, that is, the equivalence classes, are subsets of X of the form

$$[x] := \{x + u : u \in X_0\}, \qquad x \in X.$$

Let us denote $[x]$ by $x + X_0$ (for the obvious reason) and the *quotient norm* by $\|\cdot\|_*$. Thus, for $x \in X$,

$$\|x + X_0\|_* = \operatorname{dist}(x, X_0), \qquad x \in X.$$

Let $(x_n + X_0)$ be a Cauchy sequence in the quotient space X/X_0. We show that this sequence has a subsequence which converges in X/X_0 so that the sequence itself converges. Suppose (δ_n) is a sequence of positive real numbers, and for each $j \in \mathbb{N}$, let $n_j \in \mathbb{N}$ be such that

$$\|(x_n + X_0) - (x_m + X_0)\|_* < \delta_j \quad \forall n, m \geq n_j.$$

Without loss of generality, we can assume that $n_j \leq n_{j+1}$ for all $j \in \mathbb{N}$. Hence,

$$\|(x_{n_j} + X_0) - (x_{n_{j+1}} + X_0)\|_* < \delta_j \quad \forall j \in \mathbb{N}.$$

Then by the definition of the quotient norm, there exists $v_j \in X_0$ such that

$$\|(x_{n_j} + v_j) - (x_{n_{j+1}} + v_{j+1})\| < \delta_j.$$

Let us denote x_{n_j} by u_j for $j \in \mathbb{N}$. Then, for every $n, m \in \mathbb{N}$ with $n > m$, we have

$$\|(u_n + v_n) - (u_m + v_m)\| < \sum_{j=m}^{n-1} \delta_j.$$

Assume that the sequence (δ_n) satisfies the condition $\sum_{j=1}^{\infty} \delta_j < \infty$. Then it follows that $(u_n + v_n)$ is a Cauchy sequence in X. Since X is a Banach space, there exists $x \in X$ such that $u_n + v_n \to x$ as $n \to \infty$. Therefore, we have

$$\begin{aligned}\|(u_n + X_0) - (x + X_0)\|_* &= \|(u_n + v_n + X_0) - (x + X_0)\|_* \\ &\leq \|u_n + v_n - x\| \to 0 \quad \text{as} \quad n \to \infty.\end{aligned}$$

Thus, the subsequence $(u_n + X_0)$ of $(x_n + X_0)$ converges in X/X_0. ∎

The converse of the above theorem also holds.

Theorem 2.27 *If X_0 is complete subspace of a normed linear space X and X/X_0 is a Banach space, then X is a Banach space.*

Proof. Suppose X_0 is a complete subspace of normed linear space X and X/X_0 is a Banach space. We show that X is a Banach space. For this let (x_n) be a Cauchy sequence in X. Then we have

$$\|(x_n + X_0) - (x_m + X_0)\|_* \leq \|x_n - x_m\| \to 0 \quad \text{as} \quad n, m \to \infty.$$

Therefore, $(x_n + X_0)$ is a Cauchy sequence in X/X_0. Since X/X_0 is complete, it converges, say, to $x + X_0 \in X/X_0$. Now, let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that

$$\|x_n - x + X_0\|_* = \|(x_n + X_0) - (x + X_0)\|_* < \varepsilon \quad \forall n \geq N.$$

By the definition of the quotient norm, there exists $v_n \in X_0$ such that

$$\|(x_n - x) + v_n\| < \varepsilon \quad \forall n \geq \mathbb{N}.$$

This implies that $x_n + v_n \to x$ as $n \to \infty$. In particular, $(x_n + v_n)$ is a Cauchy sequence in X. Hence, (v_n) is also a Cauchy sequence in X_0. Since X_0 is complete, there exists $v \in X_0$ to which (v_n) converges. Hence,

$$x_n = (x_n + v_n) - v_n \to x - v \quad \text{as} \quad n \to \infty.$$

Thus, the proof is complete. ∎

2.2.1 Incomplete Normed Linear Spaces

In this section we give some examples of normed linear spaces which are not Banach spaces. We may observe that a normed linear space is not a Banach space if it is a proper dense subspace of a Banach space.

EXAMPLE 2.8 (i) Let $X = c_{00}$ with $\|\cdot\|_p$, $1 \le p \le \infty$. Then X is not a Banach space: We show that c_{00} is a dense subspace of ℓ^p for $1 \le p < \infty$, and it is dense in c_0 with respect to $\|\cdot\|_\infty$. Hence, the result follows since ℓ^p and c_0 are Banach spaces.

Suppose $x \in c_0$, and for $n \in \mathbb{N}$, let

$$x_n := (x(1), \ldots, x(n), 0, 0, \ldots).$$

Then $x_n \in c_{00}$, and we have

$$\|x - x_n\|_\infty = \sup\{|x(j)| : j > n\} \to 0 \quad \text{as} \quad n \to \infty.$$

Hence, c_{00} is dense in c_0 with respect to the norm $\|\cdot\|_\infty$.

We may also observe that if $1 \le p < \infty$ and $x \in \ell^p$, then with x_n as above,

$$\|x - x_n\|_p^p = \sum_{j=n+1}^{\infty} |x(j)|^p \to 0 \quad \text{as} \quad n \to \infty,$$

showing that c_{00} is dense in ℓ^p for $1 \le p < \infty$.

We shall soon see that c_{00} is not a Banach space with respect to any norm on it.

Exercise 2.9 Let $X = c_{00}$ with $\|\cdot\|_p$. Give examples of Cauchy sequences in X that do not converge in X. □

EXAMPLE 2.8 (*cont.*) (ii) For $1 \le p < \infty$, the space $C[a, b]$ with $\|\cdot\|_p$ defined by

$$\|x\|_p = \left(\int_a^b |x(t)|^p \, dt\right)^{1/p}, \qquad x \in C[a, b],$$

is not a Banach space. To see this, let $c \in (a, b)$ and $N \in \mathbb{N}$ be such that $c + 1/N \le b$. Consider the sequence (x_n) in $C[a, b]$ defined by

$$x_j(t) = 0, \qquad j = 1, \ldots, N-1; \; t \in [a, b],$$

and for $n \ge N$,

$$x_n(t) = \begin{cases} 0 & \text{if } a \le t < c \\ n(t-c) & \text{if } c \le t < c + 1/n \\ 1 & \text{if } c + 1/n \le t \le b. \end{cases}$$

It can be easily seen (*Verify*) that (x_n) is a Cauchy sequence with respect to the norm $\|\cdot\|_p$, $1 \le p < \infty$. But it does not converge to any $x \in C[a,b]$. To see this, suppose there exists $x \in C[a,b]$ such that $\|x_n - x\|_p \to 0$ as $n \to \infty$. Taking

$$y(t) = \begin{cases} 0 & \text{if } a \le t \le c \\ 1 & \text{if } c < t \le b, \end{cases}$$

we see that

$$\int_a^b |x_n(t) - y(t)|^p \, dt \to 0 \quad \text{as} \quad n \to \infty.$$

Writing $\nu_p(\phi) := \left(\int_a^b |\phi(t)|^p \, dt\right)^{1/p}$ for any Riemann integrable function ϕ in $[a,b]$, it follows, as in Minkowski's inequality on $C[a,b]$, that

$$\nu_p(x-y) \le \nu_p(x-x_n) + \nu_p(x_n-y).$$

Since $\nu_p(x-x_n) \to 0$ and $\nu_p(x_n-y) \to 0$ as $n \to \infty$, we have

$$\int_a^b |x(t) - y(t)|^p \, dt = 0.$$

In particular,

$$\int_a^c |x(t) - y(t)|^p \, dt = 0, \qquad \int_c^b |x(t) - y(t)|^p \, dt = 0.$$

Since both x and y are continuous in $[a,c) \cap (c,b]$, it follows that

$$x(t) = y(t) \quad \forall t \in [a,c) \cap (c,b], \ \forall n \in \mathbb{N},$$

This is a contradiction to the fact that x is a continuous function.

In fact, $C[a,b]$ is dense in the Banach space $L^p[a,b]$. More generally, we have the following result.

Theorem 2.28 *Let $\Omega \subseteq \mathbb{R}$ such that $0 < \mu(\Omega) < \infty$. Then $C(\Omega)$ is dense in $L^p(\Omega)$ for $1 \le p < \infty$.*

Proof. We have to prove that for every $f \in L^p(\Omega)$ and $\varepsilon > 0$, there exists $g \in C(\Omega)$ such that $\|f - g\|_p < \alpha\varepsilon$ for some $\alpha > 0$.

Let $f \in L^p(\Omega)$ and $\varepsilon > 0$. First assume that f is a non-negative function. Then by Proposition 2.8, there exists an increasing sequence (ϕ_n) of non-negative simple measurable functions on Ω such that $\phi_n(t) \to f(t)$ as $n \to \infty$ for every $t \in \Omega$. Since $(f - \phi_n)^p \le f^p$, by dominated convergence theorem (Theorem 2.11), $\|f - \phi_n\|_p \to 0$ as $n \to \infty$. This result implies that for every $\mathbb{K}$-valued function $f \in L^p(\Omega)$ and $\varepsilon > 0$, there exists a simple measurable function ϕ such that $\|f - \phi\|_p < \varepsilon$. Since ϕ is of the form $\sum_{j=1}^k \alpha_j \chi_{E_j}$ for some measurable subsets E_j of Ω and scalars α_j, $j = 1, \ldots, k$, it is enough to show the following:

(1) For every measurable subset E of Ω and $\varepsilon > 0$, there exists $h \in C(\Omega)$ such that $\|\chi_E - h\|_p < \varepsilon$.

By regularity of Lebesgue measure μ, for every measurable set E of Ω and $\varepsilon > 0$, there exists a closed set $F \subseteq E$ such that $\mu(E \backslash F) < \varepsilon$ so that

$$\|\chi_E - \chi_F\|_p^p = \int_\Omega |\chi_{E\backslash F}|^p \, d\mu = \mu(E \setminus F) < \varepsilon.$$

Hence, it is enough to show the following:

(2) For every closed set $F \subseteq \Omega$ and $\varepsilon > 0$, there exists $\tilde{h} \in C(\Omega)$ such that $\|\chi_F - \tilde{h}\|_p < \varepsilon$.

Let $F \subseteq \Omega$ be a closed set. Define

$$h_n(t) = \frac{1}{1 + n \operatorname{dist}(t, F)}, \qquad t \in \Omega.$$

Then we see that $h_n(t) \to \chi_F(t)$ for every $t \in \Omega$. Since we also have $|h_n(t) - \chi_F(t)| \le 2$ for all $t \in \Omega$ and for all $n \in \mathbb{N}$, by dominated convergence theorem, it follows that

$$\|h_n - \chi_F\|_p^p = \int_\Omega |h_n(t) - \chi_F(t)|^p \, d\mu \to 0 \quad \text{as} \quad n \to \infty.$$

This completes the proof. ∎

EXAMPLE 2.8 (*cont.*) (iii) The linear space

$$C_c(\mathbb{R}) := \{x \in C(\mathbb{R}) : \exists \text{ compact } K \subseteq \mathbb{R} \text{ with } x(t) = 0 \ \forall t \notin K\}$$

is not a closed subset of the Banach space $C(\mathbb{R})$ with respect to the norm $\|\cdot\|_\infty$, and consequently, $C_c(\mathbb{R})$ is not a Banach space:

Let (x_n) be a sequence of functions defined by

$$x_n(t) = \begin{cases} \exp(-t^2) & \text{if } -n \le t \le n \\ \exp(-n^2)[1 - n(t-n)] & \text{if } n \le t \le n + 1/n \\ \exp(-n^2)[1 + n(t+n)] & \text{if } -n - 1/n \le t \le -n \\ 0 & \text{if } |t| \ge n + 1/n. \end{cases}$$

Then, taking $x(t) = \exp(-t^2)$, $t \in \mathbb{R}$, it can be seen (*Verify*) that (x_n) is a sequence in $C_c(\mathbb{R})$ which converges to x in $C(\mathbb{R})$, but $x \notin C_c(\mathbb{R})$. Hence, $C_c(\mathbb{R})$ is not a closed subspace of $C(\mathbb{R})$.

Exercise 2.10 Show that $C_c(\mathbb{R})$ is a proper dense subspace of the Banach space $C_0(\mathbb{R})$ with respect to the norm $\|\cdot\|_\infty$. □

EXAMPLE 2.8 (*cont.*) (iv) The space $X = \mathcal{P}[a,b]$ of all polynomial functions defined on $[a,b]$ is not a Banach space with respect to the norm $\|\cdot\|_\infty$: By Weierstrass approximation theorem (cf. Rudin [27]), we know that for every $x \in C[a,b]$, there exists a sequence (x_n) of polynomials such that $\|x - x_n\|_\infty \to 0$ as $n \to \infty$. Thus, $\mathcal{P}[a,b]$ is dense in the Banach space $C[a,b]$ with $\|\cdot\|_\infty$.

What about $\mathcal{P}[a,b]$ with respect to $\|\cdot\|_p$ for $1 \le p < \infty$? Again, not a Banach space. To see this, let $x \in C[a,b] \setminus \mathcal{P}[a,b]$ and let (x_n) be in $\mathcal{P}[a,b]$ such that $\|x - x_n\|_\infty \to 0$ as $n \to \infty$. Then we have

$$\|x - x_n\|_p^p = \int_a^b |x(t) - x_n(t)|^p\,dt \le \|x - x_n\|_\infty^p (b - a) \to 0$$

as $n \to \infty$. Thus, for $1 \le p < \infty$, $\mathcal{P}[a,b]$ is dense in the space $C[a,b]$ with respect to $\|\cdot\|_p$, and $C[a,b]$ is dense in $L^p[a,b]$. Hence, $\mathcal{P}[a,b]$ is a dense proper subspace of the Banach space $L^p[a,b]$.

We shall soon show that $\mathcal{P}[a,b]$ is not a Banach space with respect to any norm.

Remark 2.8 We shall have occasions to invoke the following denseness properties of certain spaces which we have discussed in the above examples:

(1) c_{00} is dense in ℓ^p for $1 \le p < \infty$.

(2) c_{00} is dense in c_0.

(3) $\mathcal{P}[a,b]$ is dense in $C[a,b]$ with $\|\cdot\|_\infty$.

(4) $\mathcal{P}[a,b]$ is dense in $C[a,b]$ with $\|\cdot\|_p$ for $1 \le p \le \infty$.

(5) $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$.

(6) $C(\Omega)$ with $\Omega \subseteq \mathbb{R}$ and $0 < \mu(\Omega) < \infty$ is dense in $L^p(\Omega)$ for $1 \le p < \infty$.

Since for every $k \in \mathbb{N}$, $\mathcal{P}[a,b] \subseteq C^k[a,b] \subseteq C[a,b] \subseteq L^p[a,b]$, it also follows that

(7) $C^k[a,b]$ with $k \in \mathbb{N}$ is dense in $C[a,b]$ with $\|\cdot\|_p$ for $1 \le p \le \infty$,

(8) $C^k[a,b]$ with $k \in \mathbb{N}$ is dense in $L^p[a,b]$ for $1 \le p < \infty$.

2.2.2 Completion of Normed Linear Spaces

It can happen that a normed linear space X is not a Banach space, but it is a subspace of a Banach space. For example, we know that $C[a,b]$ with respect to the norm $\|\cdot\|_p$, $1 \le p < \infty$ is not a Banach space, but it is a dense subspace of the Banach space $L^p[a,b]$. One may ask: for every normed linear space X, does there exist a Banach space Y such that X is a dense subspace of Y? The answer is somewhat in the affirmative.

Let X be a normed linear space. Then a Banach space Y is said to be a **completion** of X if there exists a linear isometry $T: X \to Y$ such that $R(T)$ is a dense subspace of Y.

Now we indicate one of the ways of finding a completion of a normed linear space X, which is akin to the process of construction of the set of real numbers from the set of rational numbers .

Let X be a normed linear space and let $\mathcal{X}$ be the set of all Cauchy sequences in X. Using triangle inequality, it follows that if (x_n) and (y_n) are in $\mathcal{X}$, then $(\|x_n - y_n\|)$ is a Cauchy sequence in $\mathbb{K}$, and consequently, $\lim_{n\to\infty} \|x_n - y_n\|$ exists. In particular, for every $(x_n) \in \mathcal{X}$, $\lim_{n\to\infty} \|x_n\|$ exists. Now, for (x_n) and (y_n) in $\mathcal{X}$, define

$$(x_n) \sim (y_n) \iff \lim_{n\to\infty} \|x_n - y_n\| = 0.$$

Then it is seen that $\sim$ is an equivalence relation on $\mathcal{X}$. Let $\widetilde{\mathcal{X}}$ be the set of all equivalence classes.

If $(x_n) \in \mathcal{X}$, then let us denote the equivalence class of it by $[(x_n)]$. For $[(x_n)], [(y_n)]$ in $\widetilde{\mathcal{X}}$ and $\alpha \in \mathbb{K}$, define

$$[(x_n)] + [(y_n)] = [(x_n + y_n)], \qquad \alpha[(x_n)] = [(\alpha x_n)].$$

Note that these operations are well defined. With respect to these operations of addition and scalar multiplication, the set $\widetilde{\mathcal{X}}$ is a linear space with its zero element $[(\theta_n)]$, where $\theta_n = 0$ for every $n \in \mathbb{N}$. We define

$$\|[(x_n)]\|_* = \lim_{n\to\infty} \|x_n\|, \quad [(x_n)] \in \widetilde{\mathcal{X}}.$$

It can be seen (*Verify*) that $\|\cdot\|_*$ is a norm on $\widetilde{\mathcal{X}}$.

Theorem 2.29 *Let X be a normed linear space and $\widetilde{\mathcal{X}}$ be the normed linear space as defined above. Then $\widetilde{\mathcal{X}}$ is a completion of X.*

Proof. First we show that $\widetilde{\mathcal{X}}$ is a Banach space with respect to the norm $\|\cdot\|_*$. For this, let $(\tilde{x}_k)$ be a Cauchy sequence in $\widetilde{\mathcal{X}}$, and for each $k \in \mathbb{N}$, let $(x_{k,n}) \in \mathcal{X}$ be such that $\tilde{x}_k = [(x_{k,n})]$. For each $k \in \mathbb{N}$, let $\ell_k \in \mathbb{N}$ be such that $\|x_{k,n} - x_{k,m}\| < 1/k$ for all $n, m \geq \ell_k$. In particular,

$$\|x_{k,m} - x_{k,\ell_k}\| < 1/k \quad \forall m \geq \ell_k.$$

Note that, for every $n \geq \max\{\ell_k, \ell_m\}$,

$$\begin{aligned} \|x_{k,\ell_k} - x_{m,\ell_m}\| &\leq \|x_{k,\ell_k} - x_{k,n}\| + \|x_{k,n} - x_{m,n}\| \\ &\quad + \|x_{m,n} - x_{m,\ell_m}\| \\ &\leq \frac{1}{k} + \|x_{k,n} - x_{m,n}\| + \frac{1}{m}. \end{aligned}$$

Since $\|\tilde{x}_k - \tilde{x}_m\|_* = \lim_{n\to\infty} \|x_{k,n} - x_{m,n}\|$, it follows, by letting n tend to infinity, that

$$\|x_{k,\ell_k} - x_{m,\ell_m}\| \leq \frac{1}{k} + \|\tilde{x}_k - \tilde{x}_m\|_* + \frac{1}{m}.$$

Hence, (x_{n,ℓ_n}) is a Cauchy sequence in X. We may also note that, for all $n \geq \ell_k$,

$$\begin{aligned} \|x_{k,n} - x_{n,\ell_n}\| &\leq \|x_{k,n} - x_{k,\ell_k}\| + \|x_{k,\ell_k} - x_{n,\ell_n}\| \\ &\leq \frac{1}{k} + \|x_{k,\ell_k} - x_{n,\ell_n}\|. \end{aligned}$$

From this, using the fact that (x_{n,ℓ_n}) is a Cauchy sequence, it follows that $(\tilde{x}_k)$ converges to $[(x_{n,\ell_n})]$. Thus, we have shown that $\widetilde{\mathcal{X}}$ is a Banach space.

Now, we show that $\widetilde{\mathcal{X}}$ is a completion of X. For this, define $T : X \to \widetilde{\mathcal{X}}$ as follows: For $x \in X$, let $\tilde{y}_x$ be the equivalence class of the constant sequence $(x, x, \ldots)$. Define $T : X \to \widetilde{\mathcal{X}}$ by $Tx = \tilde{y}_x$. It is easily seen (*Verify*) that $T : X \to \widetilde{\mathcal{X}}$ is a linear isometry. It remains to show that $R(T)$ is a dense subspace of $\widetilde{\mathcal{X}}$. To see this, let $\tilde{x} := [(x_n)] \in \widetilde{\mathcal{X}}$, and let $\varepsilon > 0$. Let $N \in \mathbb{N}$ be such that $\|x_n - x_m\| < \varepsilon$ for all $n, m \geq N$. Taking $u_n = x_N$ for all $n \in \mathbb{N}$, it follows that $\tilde{u} := [(u_n)] \in R(T)$. Note that

$$\|\tilde{x} - \tilde{u}\|_* = \lim_{n\to\infty} \|x_n - u_n\| = \lim_{n\to\infty} \|x_n - x_N\| < \varepsilon.$$

Thus, we have shown that $R(T)$ is dense in $\widetilde{\mathcal{X}}$. This completes the proof. ∎

There can be many Banach spaces which are completions of a given normed linear space. But, as far as their linear structure and norm structure are concerned, they are all the same. That is, if Y_1 and Y_2 are completions of a normed linear space X, then there exists a bijective linear isometry between Y_1 and Y_2. To see this, suppose Y_1 and Y_2 are completions of a normed linear space X. Let

$$T_1 : X \to Y_1, \qquad T_2 : X \to Y_2$$

be linear isometries such that $R(T_1)$ dense in Y_1 and $R(T_2)$ dense in Y_2. We define a map $T : Y_1 \to Y_2$ as follows. Let $y \in Y_1$ and (y_n) be in $R(T_1)$ such that $y_n \to y$. Let (x_n) be in X such that $y_n = T_1 x_n$, $n \in \mathbb{N}$. Since T_1 is an isometry, it is seen that (x_n) is a Cauchy sequence in X. Let $z_n = T_2 x_n$, $n \in \mathbb{N}$. Again, since T_2 is an isometry, it follows that (z_n) is a Cauchy sequence in Y_2. Let $z = \lim_{n\to\infty} T_2 x_n$. Now, we define $Ty = z$, i.e.,

$$Ty = \lim_{n\to\infty} T_2 x_n, \qquad y \in Y_1,$$

where (x_n) is the sequence in X such that $y = \lim_{n\to\infty} T_1 x_n$. It is easy to show that $T : Y_1 \to Y_2$ is well defined. Moreover, T is a surjective linear isometry. Thus, Y_1 and Y_2 are linearly isometric.

In view of the above discussion, we may say that completion of a normed linear space is unique upto linear isometry.

For certain concrete cases, we can find their completions more specifically. We have already noticed the following:

(1) c_{00} is dense in ℓ^p for $1 \le p < \infty$.

(2) c_{00} is dense in c_0 with respect to $\|\cdot\|_\infty$.

(3) For every measurable set $\Omega \subseteq \mathbb{R}$ with nonzero finite measure, $C(\Omega)$ is dense in $L^p(\Omega)$ for $1 \le p < \infty$.

(4) $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$ with respect to $\|\cdot\|_\infty$.

Thus, for $1 \le p < \infty$, the Banach spaces ℓ^p and $L^p(\Omega)$ are completions of c_{00} and $C(\Omega)$, respectively, and the Banach spaces $(c_0, \|\cdot\|_\infty)$ and $(C_0(\mathbb{R}), \|\cdot\|_\infty)$ are completions of $(c_{00}, \|\cdot\|_\infty)$ and $(C_c(\mathbb{R}), \|\cdot\|_\infty)$, respectively.

Another class of Banach spaces which are useful in applications and obtained by the completion process of certain normed linear spaces is the class of 'Sobolev spaces'. For $k \in \mathbb{N}$ and $1 \le p < \infty$, it can be seen (*Verify*) that the space $C^k[a,b]$ is not a Banach space with respect to the norm

$$x \mapsto \sum_{j=0}^{k} \|x^{(j)}\|_p, \qquad x \in C^k[a,b],$$

where $x^{(0)} = x$, and for $j \geq 1$, $x^{(j)}$ is the j-th derivative of x. Its completion, denoted by $W^{k,p}[a,b]$, is an example of a Sobolev space. Note that, using the above procedure of completion, the space $L^p[a,b]$ can be thought of as $W^{0,p}[a,b]$, i.e., the completion of $C[a,b]$ with respect to the norm $\|\cdot\|_p$.

For finding the alternative ways of defining Sobolev spaces, and also to see their applications to partial differential equations, one may refer Kesavan [18].

2.2.3 Some Properties of Banach Spaces

A linear space may not be a Banach space with respect to certain norm, but can be a Banach space with respect to some other norm. For example, we know that $C[a,b]$ is not a Banach space with respect to the norm $\|\cdot\|_p$ for $1 \leq p < \infty$, but it a Banach space with respect to $\|\cdot\|_\infty$. A question that naturally arises is whether this is the case with all linear spaces. The answer is in the negative. In fact, the following theorem could supply lots of examples of linear spaces, each of which can never be a Banach space with respect to any norm. The spaces $\mathcal{P}[a,b]$ and c_{00} are certain specific examples.

Theorem 2.30 *A Banach space cannot have a denumerable basis.*

Before proving this theorem, let us observe that a basis of a normed linear space depends only on the linear structure of the space, not on the norm. Hence, from Theorem 2.30, we can infer that if a linear space is a Banach space with respect to some norm on it, then it cannot have a denumerable basis. To illustrate this point, consider the example of the space $X = C[a,b]$ with $\|\cdot\|_1$. We know that X is not a Banach space. But since $C[a,b]$ with $\|\cdot\|_\infty$ is a Banach space, we can conclude that $C[a,b]$ cannot have a denumerable basis, and consequently, the space X also cannot have a denumerable basis.

Now for proving the above theorem, we make use of two results. The first one is known as the *Baire category theorem*. It will be one of the crucial results in proving many other results.

Theorem 2.31 (Baire category theorem) *Let Ω be a complete metric space. If (Ω_n) is a sequence of subsets of Ω such that $\Omega = \cup_{n=1}^{\infty}\Omega_n$, then there is some $j \in \mathbb{N}$ such that interior of $\overline{\Omega_j}$ is nonempty.*

Proof. Let Ω be a complete metric space and (Ω_n) be a sequence of subsets of Ω such that $\Omega = \cup_{n=1}^{\infty}\Omega_n$. Since

$$\Omega = \cup_{n=1}^{\infty}\Omega_n \subseteq \cup_{n=1}^{\infty}\overline{\Omega_n} \subseteq \Omega,$$

we assume, without loss of generality, that each Ω_n is a closed subset of Ω, and $\Omega_n \neq \varnothing$ for every $n \in \mathbb{N}$.

Suppose the result is not true, i.e., for every $n \in \mathbb{N}$, Ω_n has no interior point. Let $x_1 \in \Omega_1$ and $r_1 = 1$. Since the ball $B(x_1, r_1)$ is not contained in Ω_1, and since Ω_1 is closed, the set $B(x_1, r_1) \cap \Omega_1^c$ is a nonempty open set. Let $x_2 \in B(x_1, r_1) \cap \Omega_1^c$ and $r_2 > 0$ be such that

$$\overline{B(x_2, r_2)} \subseteq B(x_1, r_1) \cap \Omega_1^c.$$

Having obtained $x_1, \ldots, x_n$ and $r_1, \ldots, r_n$ such that

$$x_j \in B(x_{j-1}, r_{j-1}) \cap \Omega_{j-1}^c, \qquad \overline{B(x_j, r_j)} \subseteq B(x_{j-1}, r_{j-1}) \cap \Omega_{j-1}^c,$$

for $j = 1, \ldots, n$, let $x_{n+1} \in B(x_n, r_n) \cap \Omega_n^c$ and $r_{n+1} > 0$ be such that

$$\overline{B(x_{n+1}, r_{n+1})} \subseteq B(x_n, r_n) \cap \Omega_n^c.$$

We may choose $r_1, r_2, \ldots,$ such that $r_n \to 0$ as $n \to \infty$. Clearly, $x_n \in B(x_m, r_m)$ for $n > m$ so that (x_n) is a Cauchy sequence in Ω. Since Ω is complete, it converges to some $x \in \Omega$. Then we have

$$d(x, x_m) = \lim_{n\to\infty} d(x_n, x_m) \leq r_m \quad \forall m \in \mathbb{N}$$

so that

$$x \in \overline{B(x_m, r_m)} \quad \forall m \in \mathbb{N}.$$

But

$$\overline{B(x_m, r_m)} \subseteq B(x_{m-1}, r_{m-1}) \cap \Omega_{m-1}^c \quad \forall m \in \mathbb{N}.$$

Thus, $x \notin \Omega_n$ for every $n \in \mathbb{N}$. This is a contradiction to the fact that $\Omega = \cup_{n=1}^{\infty}\Omega_n$. ∎

A subset of a metric space is said to be a **nowhere dense** set if its closure has empty interior.

Thus, the Baire category theorem states that a complete metric space cannot be written as a countable union of nowhere dense sets. For instance, by this theorem, the metric space of rational numbers is not complete, as it can be written as countable union of singleton sets, each of which is nowhere dense.

Another result which we require to prove Theorem 2.30 is the following.

Lemma 2.32 *Interior of a proper subspace of a normed linear space is empty.*

Proof. Let X_0 be a subspace of a normed linear space X such that its interior is nonempty. Let $x_0 \in X_0$ and $r > 0$ be such that $B(x_0, r) \subseteq X_0$. Then for every $x \in X$ with $x \neq 0$, we have

$$u = x_0 + \frac{r}{2\|x\|}x \in B(x_0, r) \subseteq X_0$$

so that

$$x = \frac{2\|x\|}{r}(u - x_0) \in X_0.$$

Thus, we have shown that $X_0 = X$. ∎

Exercise 2.11 Let S be a subset of a normed linear space X. Show that interior of S is empty if and only if its complement is dense in X. □

Proof of Theorem 2.30. Suppose X is a normed linear space with a denumerable basis, say, $\{u_1, u_2, \ldots\}$. Then, $X = \cup_n X_n$, where, for each $n \in \mathbb{N}$,

$$X_n = \text{span}\{u_1, \ldots, u_n\}.$$

Since each X_n is a closed proper subspace of X, by Lemma 2.32, interior of X_n is empty for each n. Therefore, by Baire category theorem (Theorem 2.31), X is not a Banach space. ∎

As application of Theorem 2.30, we consider the following.

EXAMPLE 2.9 (i) Let $X = \mathcal{P}$, the linear space of all polynomials with coefficients in $\mathbb{K}$. Since $\{u_1, u_2, \ldots\}$ with $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$, is a denumerable basis of $\mathcal{P}$, X is not a Banach space with respect to any norm.

(ii) The linear space c_{00} is not a Banach space with respect to any norm, since $\{e_1, e_2, \ldots\}$ with $e_i(j) = \delta_{ij}$, $i, j \in \mathbb{N}$, is a denumerable basis of c_{00}.

Exercise 2.12 Prove that a Banach space is finite dimensional if and only if every subspace of it is closed. □

Using the completeness of $\mathbb{K}$, we can immediately deduce that every absolutely convergent series of scalars converges. Similar proof works for a Banach space as well. More importantly, this property is a characterization of completeness of a normed linear space. First we have to define the notions of *convergence* and *absolute convergence* of a series, which we do as in the case of series of scalars.

Let X be a normed linear space. By a **series** in X we mean an expression of the form $\sum_{j=1}^{\infty} x_j$, where (x_n) is a sequence in X. We say that the series $\sum_{j=1}^{\infty} x_j$ is **convergent** if there exists $x \in X$ such that the sequence of its partial sums, namely $\left(\sum_{j=1}^{n} x_j\right)$, converges to x. In that case we say that the series $\sum_{j=1}^{\infty} x_j$ **converges** to x.

The series $\sum_{j=1}^{\infty} x_j$ is said to be **absolutely convergent** if the series $\sum_{j=1}^{\infty} \|x_j\|$ converges.

Theorem 2.33 *Let X be a normed linear space. Then X is a Banach space if and only if every absolutely convergent series of elements of X is convergent.*

Proof. Suppose X is a Banach space and $\sum_{j=1}^{\infty} x_j$ is an absolutely convergent series. Then writing $s_n = \sum_{j=1}^{n} x_j$ for $n \in \mathbb{N}$, we have, for $n > m$,

$$\|s_n - s_m\| \le \sum_{j=m+1}^{n} \|x_j\|.$$

Since $\sum_{j=1}^{\infty} \|x_j\| < \infty$, it follows that $\|s_n - s_m\| \to 0$ as $n, m \to \infty$, showing that (s_n) is a Cauchy sequence in X. Since X is a Banach space, there exists $x \in X$ such that $\|s_n - x\| \to 0$ as $n \to \infty$.

Conversely, suppose X is a normed linear space such that every absolutely convergent series of elements of X is convergent. We show that X is a Banach space. For this let (x_n) be a Cauchy sequence in X. It is enough to show that (x_n) has a convergent subsequence. Since (x_n) is a Cauchy sequence, we know that for each $j \in \mathbb{N}$, there exists $n_j \in \mathbb{N}$ such that $n_j \le n_{j+1}$ for all $j \in \mathbb{N}$, and

$$\|x_{n_j} - x_{n_{j+1}}\| < \frac{1}{2^j}.$$

Let us denote x_{n_j} by u_j for $j \in \mathbb{N}$. Then we observe that for every $n = 2, 3, \ldots$,

$$u_n = u_1 + \sum_{j=1}^{n-1} (u_{j+1} - u_j), \qquad j = 1, \ldots, n-1.$$

Since

$$\sum_{j=1}^{\infty} \|u_{j+1} - u_j\| \leq \sum_{j=1}^{\infty} \frac{1}{2^j} < \infty,$$

the series $\sum_{j=1}^{\infty} (u_{j+1} - u_j)$ is absolutely convergent. Hence, by hypothesis, it is convergent, say to u. Thus,

$$u_n = u_1 + \sum_{j=1}^{n-1} (u_{j+1} - u_j) \to u_1 + u$$

as $n \to \infty$. Note that (u_n) is a subsequence of the sequence (x_n). ∎

By the above theorem we can infer that in an arbitrary normed linear space the property of absolute convergence of a series is not stronger than convergence.

Exercise 2.13 Find an absolutely convergent series in c_{00} with $\|\cdot\|_p$, $1 \leq p \leq \infty$, which is not convergent. □

It is to be emphasized that the term *basis* used in Theorem 2.30 is for a *Hamel basis*. As we have pointed out in Chapter 1, there are other concepts of basis as well. One such concept is the *Schauder basis*.

2.3 Schauder Basis and Separability

A sequence (u_n) in an infinite dimensional normed linear space X is said to be a **Schauder basis** of X if for every $x \in X$, there exists a unique sequence (α_n) of scalars such that

$$x = \sum_{j=1}^{\infty} \alpha_j u_j.$$

It is to be observed that a Schauder basis of a normed linear space, if exists, is linearly independent. To see this, suppose $\{u_1, u_2, \ldots\}$

is a Schauder basis of a normed linear space X, and if $\alpha_1, \dots, \alpha_n$ are scalars such that $\sum_{j=1}^{n} \alpha_j u_j = 0$. Then, taking $\alpha_j = 0$ for all $j > n$, we have the representation of zero element in X as $\sum_{j=1}^{\infty} \alpha_j u_j$. Hence, by the uniqueness of the representation, it follows that $\alpha_j = 0$ for $j = 1, \dots, n$.

We note that if $\{u_1, u_2, \dots\}$ is a Schauder basis of a normed linear space X, then span$\{u_1, u_2, \dots\}$ is dense in X. Indeed, for every $x \in X$, we have

$$x = \sum_{j=1}^{\infty} \alpha_j u_j = \lim_{n\to\infty} \sum_{j=1}^{n} \alpha_j u_j,$$

and this belongs to the closure of span$(\{u_1, u_2, \dots\})$. From this it also follows (*How?*) that, if a normed linear space X has a denumerable (Hamel) basis E, then E is a Schauder basis as well. Thus, $\{e_1, e_2, \dots\}$ with $e_j(i) = \delta_{ij}$ is a Schauder basis of c_{00}, and $\{u_1, u_2, \dots\}$ with $u_j(t) = t^{j-1}$, $j \in \mathbb{N}$, is a Schauder basis of $\mathcal{P}[a, b]$ with respect to any norm on them. Since for every $x \in \ell^p$, $1 \le p < \infty$, we have

$$x = \lim_{n\to\infty} \sum_{j=1}^{n} x(j) e_j,$$

$\{e_1, e_2, \dots\}$ is a Schauder basis of ℓ^p, $1 \le p < \infty$.

We shall see in Chapter 4 that every *separable* Hilbert space has a Schauder basis, the so-called *orthonormal basis*. In this context, let us recall the definition of *separability*.

A metric space Ω is said to be **separable** if it has a countable dense subset.

Exercise 2.14 A normed linear space is separable if and only if it has a denumerable dense subset. □

A simple example of a separable normed linear space is the scalar field $\mathbb{K}$. Indeed, if $\mathbb{K} = \mathbb{R}$, then the set of rational numbers is a dense subset of $\mathbb{R}$, and if $\mathbb{K} = \mathbb{C}$, then the set of all those complex numbers with real and imaginary parts as rational numbers is a dense subset of $\mathbb{C}$.

What about the set of irrational numbers? It should be intuitively clear to the reader that it is also a separable metric space. To see this, let $\{r_1, r_2, \dots\}$ be an enumeration of the set of all rational numbers, and for each $j \in \mathbb{N}$, let $(s_{j,n})$ be a sequence of irrational numbers which converges to r_j. Then it can be easily seen that $\{s_{j,n} : j, n \in \mathbb{N}\}$ is a countable dense subset of the set of irrational numbers. The idea involved in this argument can be used to prove something more about separability. We leave the proof of the following proposition to the reader.

Proposition 2.34 (i) *Every subset of a separable metric space is separable.*

(ii) *Suppose Ω_1 and Ω_2 are metric spaces, and $f : \Omega_1 \to \Omega_2$ is a continuous function. If Ω_1 is separable space, then $f(\Omega_1)$ is separable as a subspace of Ω_2.*

Now we prove two more results which help us to infer separability or non-separability of certain spaces which we are already familiar with. We use the notation $\mathbb{Q}_{\mathbb{K}}$ for the set $\mathbb{Q}$ of all rational numbers if $\mathbb{K} = \mathbb{R}$, and the set of all complex numbers with real and imaginary parts as rational numbers if $\mathbb{K} = \mathbb{C}$, i.e.,

$$\mathbb{Q}_{\mathbb{K}} = \begin{cases} \mathbb{Q} & \text{if } \mathbb{K} = \mathbb{R}, \\ \mathbb{Q} + i\,\mathbb{Q} & \text{if } \mathbb{K} = \mathbb{C}. \end{cases}$$

Theorem 2.35 *Every finite dimensional normed linear space is separable.*

Proof. Let X be a finite dimensional normed linear space and $\{u_1, \ldots, u_n\}$ be a basis of X. Since $\mathbb{Q}_{\mathbb{K}}$ is a countable subset of $\mathbb{K}$, and since the cartesian product of a finite number of countable sets is countable, it follows that the set

$$D := \left\{ \sum_{j=1}^{n} \alpha_j u_j : \alpha_j \in \mathbb{Q}_{\mathbb{K}} \right\}$$

is a countable subset of $\operatorname{span}\{u_1, \ldots, u_n\} = X$. Also, D is dense in X. To see this, let $x \in X$ and $\varepsilon > 0$. Let $\beta_1, \ldots, \beta_n$ be scalars such that $x = \sum_{j=1}^{n} \beta_j u_j$. By the denseness of $\mathbb{Q}_{\mathbb{K}}$ in $\mathbb{K}$, there exist $\alpha_1, \ldots, \alpha_n$ in $\mathbb{Q}_{\mathbb{K}}$ such that $|\beta_j - \alpha_j| < \varepsilon$ for all $j \in \{1, \ldots, n\}$. Then it follows that

$$\left\| x - \sum_{j=1}^{n} \alpha_j u_j \right\| \leq \sum_{j=1}^{n} |\beta_j - \alpha_j| \|u_j\| \leq \varepsilon \sum_{j=1}^{n} \|u_j\|.$$

This shows that D is dense in X. ∎

Theorem 2.36 *Every normed linear space with a Schauder basis is separable. In particular, every normed linear space with a denumerable basis is separable.*

Proof. Let X be a normed linear space with a Schauder basis $\{u_1, u_2, \ldots\}$. Since

$$Y := \bigcup_{n=1}^{\infty} Y_n \quad \text{with} \quad Y_n = \operatorname{span}\{u_1, \ldots, u_n\} \quad \forall n \in \mathbb{N}$$

is dense in X, it is enough to show that Y has a countable dense subset. For this, let

$$D_n = \left\{ \sum_{j=1}^{n} \alpha_j u_j : \alpha_j \in \mathbb{Q}_{\mathbb{K}} \right\}, \qquad n \in \mathbb{N}.$$

Then, as in the proof of Theorem 2.35, D_n is a countable dense subset of Y_n. Thus, $\cup_{n=1}^{\infty} D_n$ is a countable dense subset of Y. ∎

Now making use of Proposition 2.34 and Theorem 2.36, we can derive the conclusions in the examples that follow.

EXAMPLE 2.10 (i) The spaces c_{00} and $\mathcal{P}[a,b]$ are separable with respect to any norm, as these spaces have denumerable bases.

(ii) The space ℓ^p for $1 \leq p < \infty$ is separable, since it has the Schauder basis $\{e_1, e_2, \ldots\}$, where $e_j(i) = \delta_{ij}$, $(i,j) \in \mathbb{N} \times \mathbb{N}$.

(iii) The space c_0 with $\|\cdot\|_\infty$ is separable since c_{00} is dense in c_0, and c_{00} is separable.

(iv) The space $C[a,b]$ with $\|\cdot\|_p$ for $1 \leq p \leq \infty$ is separable since $\mathcal{P}[a,b]$ is dense in $C[a,b]$, and $\mathcal{P}[a,b]$ is separable.

(v) The space $L^p[a,b]$ for $1 \leq p < \infty$ is separable since $C[a,b]$ is dense in $L^p[a,b]$, and $C[a,b]$ is separable.

The following result helps us to infer that certain spaces are not separable.

Proposition 2.37 *Let Ω be a metric space with metric d. If there exists an uncountable family of disjoint nonempty open sets in Ω, then Ω is not separable. In particular, if there exists an uncountable set E and $r > 0$ such that $d(x,y) \geq r$ for every distinct $x, y \in E$, then Ω is not separable.*

Proof Suppose $\mathcal{G}$ is an uncountable family of disjoint nonempty open sets in Ω, and D is a dense subset of Ω. Then each $G \in \mathcal{G}$ contains at least one element from D, say $x_G \in D \cap G$. Since $\mathcal{G}$ is an uncountable set, the set $\{x_G : G \in \mathcal{G}\}$ is also uncountable subset of D. Thus, no dense subset of Ω can be countable. ∎

EXAMPLE 2.11 (i) The space $\ell^\infty(\mathbb{N})$ is not separable since the set

$$E = \{x \in \ell^\infty : x(j) \in \{0, 1\},\ j \in \mathbb{N}\}$$

is an uncountable set, and for every distinct $x, y \in E$, $\|x - y\|_\infty = 1$.

(ii) The spaces $\ell^\infty([a, b])$ is not separable since the set

$$E = \{\chi_{[a,s]} : a < s \leq b\}$$

is an uncountable subset of $\ell^\infty([a, b])$ such that, for every distinct $x, y \in E$, $\|x - y\|_\infty = 1$.

(iii) The set E in the above example also serves to show that $L^\infty[a, b]$ is not separable.

Remark 2.9 We have seen that every normed linear space with a Schauder basis is separable. Most of the separable spaces which one comes across, including $C[a, b]$ and $L^p[a, b]$ for $1 \leq p < \infty$, have Schauder bases (cf. [20], Problems 8.14 and 8.15). It was a longstanding open problem whether every separable Banach space has a Schauder basis. This question has been settled in negative in the year 1973 by Enflo [12] by giving an example of a separable Banach space which does not have a Schauder basis.

2.4 Heine-Borel Theorem and Riesz Lemma

Recall that a subset of a metric space is compact if it is complete and totally bounded. We have already observed that every compact subset of a metric space is closed and bounded, and compact subsets of $\mathbb{K}$ are precisely closed and bounded sets. The latter is the well-known *Heine-Borel theorem*.

The following characterizations of compactness and totally boundedness are useful in the sequel:

(1) A subset S of a metric space is *compact* if and only if every sequence in S has a subsequence which converges to a point in S.

(2) A subset S of a metric space is *totally bounded* if and only if every sequence in S has a Cauchy subsequence.

Using the above characterization of compactness and the Heine-Borel theorem on $\mathbb{K}$, we obtain the *Bolzano-Weierstrass property* of $\mathbb{K}$: Every bounded sequence in $\mathbb{K}$ has a convergent subsequence.

One may ask: Is every closed and bounded set in a normed linear space X is compact? Certainly, not for all normed linear spaces. For example, consider $X = c_{00}$ with the norm $\|\cdot\|_\infty$, and the unit sphere $S = \{x \in X : \|x\|_\infty = 1\}$ in X. Clearly, S is closed and bounded in X. But it is not compact. To see this, consider the sequence (e_n) in S, where $e_n(j) = \delta_{nj}$ for all $n, j \in \mathbb{N}$. We note that

$$\|e_n - e_m\|_\infty = 1 \quad \forall n \neq m$$

so that the sequence (e_n) has no Cauchy subsequence. Thus, we have shown that S is not totally bounded in X; consequently, not compact in X.

We show that the answer to the above raised question is affirmative for every finite dimensional space, and negative for every infinite dimensional space.

We shall make use of the following easily verifiable results: If $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent norms on a linear space X, and $S \subseteq X$, then

(1) S is bounded in $(X, \|\cdot\|) \iff S$ is bounded in $(X, \|\cdot\|_*)$,

(2) S is closed in $(X, \|\cdot\|) \iff S$ is closed in $(X, \|\cdot\|_*)$, and

(3) S is compact in $(X, \|\cdot\|) \iff S$ is compact in $(X, \|\cdot\|_*)$.

Theorem 2.38 (Heine-Borel theorem) *Every closed and bounded subset of a finite dimensional normed linear space is compact.*

Proof. Let X be a normed linear space of dimension k and $E = \{u_1, \ldots, u_k\}$ be a basis of X. By Theorem 2.25 and the observations (1), (2), (3) made above, it is enough to prove the result with respect to the norm

$$\|x\|_E = \max\{|f_1(x)|, \ldots, |f_k(x)|\}, \qquad x \in X,$$

where $f_1, \ldots, f_n$ are the coordinate functionals corresponding to the basis E.

Now, let $S \subseteq X$ be a closed and bounded subset of X (with respect to the norm $\|\cdot\|_E$). In order to show that it is compact, it is enough to show that every sequence in S has a subsequence which converges to a point in S. So, let (x_n) be a sequence in S. Since S is bounded, by the definition of the norm $\|\cdot\|_E$, $(f_j(x_n))$ is a bounded sequence in $\mathbb{K}$ for each $j \in \{1, \dots, k\}$. By the Heine-Borel theorem on $\mathbb{K}$, $(f_1(x_n))$ has a convergent subsequence, that is, (x_n) has a subsequence $(x_{1,n})$ such that $(f_1(x_{1,n}))$ converges. Since $(f_2(x_{1,n}))$, which is a subsequence of $(f_2(x_n))$, is a bounded sequence in $\mathbb{K}$, it has a convergent subsequence. Thus, $(x_{1,n})$ has a subsequence $(x_{2,n})$ such that $(f_2(x_{2,n}))$ converges. Continuing this procedure for k times, the sequence $(x_{j-1,n})$ has a subsequence $(x_{j,n})$ such that $(f_j(x_{j,n}))$ converges. Let

$$\alpha_j := \lim_{n\to\infty} f_j(x_{j,n}), \qquad j = 1, \dots, k.$$

Since for each i with $1 < i \leq k$, $(x_{i,n})$ is a subsequence of $(x_{i-1,n})$, it follows that

$$\lim_{n\to\infty} f_j(x_{k,n}) = \alpha_j, \qquad j = 1, \dots, k.$$

Therefore,

$$\max\{|f_j(x_{k,n}) - \alpha_j| : j = 1, \dots, k\} \to 0 \quad \text{as} \quad n \to \infty.$$

Hence, by taking

$$\tilde{x}_n = f_1(x_{k,n})u_1 + \cdots + f_k(x_{k,n})u_k, \qquad x = \alpha_1 u_1 + \cdots + \alpha_k u_k,$$

we see that $(\tilde{x}_n)$ is a subsequence of (x_n) which converges to x with respect to the norm $\|\cdot\|_E$. Since S is closed, $x \in S$. Thus, the proof of the theorem is complete. ∎

Now we show that the Heine-Borel theorem is, in fact, a characterization of finite dimensional normed linear spaces.

Theorem 2.39 *Let X be a normed linear space. Then the following are equivalent:*

(i) *X is finite dimensional.*

(ii) *The unit sphere $\{x \in X : \|x\| = 1\}$ is compact.*

(iii) *The unit sphere $\{x \in X : \|x\| = 1\}$ is totally bounded.*

For the proof of this theorem, we make use of the following result known as the *Riesz lemma.*

Lemma 2.40 (Riesz lemma) *Let X_0 be a proper closed subspace of a normed linear space X. Then for every r with $0 < r < 1$, there exists $x_r \in X$ such that*

$$\|x_r\| = 1, \qquad \text{dist}\,(x_r, X_0) \geq r.$$

Proof. Let X_0 be a proper closed subspace of a normed linear space X, and let $x \in X \setminus X_0$. Since X_0 is closed, $d = \text{dist}\,(x, X_0) > 0$. Let $0 < r < 1$. Then there exists $u \in X_0$ such that $\|x - u\| \leq d/r$. Now, let

$$x_r = \frac{x - u}{\|x - u\|}.$$

Then we have $\|x_r\| = 1$ and

$$\text{dist}\,(x_r, X_0) = \frac{\text{dist}\,(x, X_0)}{\|x - u\|} \geq \frac{d}{d/r} = r. \qquad \blacksquare$$

Another Proof. Suppose the result is not true. Then there exists $r \in (0, 1)$ such that

$$\text{dist}\,(x, X_0) < r \quad \forall x \in S, \tag{2.11}$$

where $S = \{u \in X : \|u\| = 1\}$. Since X_0 is a proper subspace of X, there exists $x_0 \in S$ such that $x_0 \notin X_0$, and since X_0 is a closed subspace of X, we have

$$0 < d := \text{dist}\,(x_0, X_0) < r.$$

Hence, there exists $u_0 \in X_0$ such that $\|x_0 - u_0\| < r$. Then by taking

$$x_1 = \frac{x_0 - u_0}{\|x_0 - u_0\|},$$

we have $x_1 \in S$ so that by the relation (2.11) above, $\text{dist}\,(x_1, X_0) < r$. Since

$$\text{dist}\,(x_1, X_0) = \frac{1}{\|x_0 - u_0\|}\,\text{dist}\,(x_0 - u_0, X_0) = \frac{d}{\|x_0 - u_0\|},$$

it follows that

$$d < r\|x_0 - u_0\| < r^2.$$

Again, since $d := \text{dist}\,(x_0, X_0) < r^2$, there exists $u_1 \in X_0$ such that $\|x_0 - u_1\| < r^2$. Taking

$$x_2 = \frac{x_0 - u_1}{\|x_0 - u_1\|},$$

we have $x_2 \in S$ so that by (2.11), $\operatorname{dist}(x_2, X_0) < r$. Thus, as above,

$$d < r\|x_0 - u_1\| < r^3.$$

Continuing this process, we get a sequence (u_n) in X_0 such that

$$d < r\|x_0 - u_n\| < r^{n+2} \quad \forall n \in \mathbb{N}.$$

Since $0 < r < 1$, it follows, by letting n tend to infinity, that $d = 0$ which is a contradiction. ∎

Proof of Theorem 2.39. If the space X is finite dimensional, then by Theorem 2.38, the closed unit ball of X is compact and, in particular, it is totally bounded. Thus, (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii).

Next, we show that (iii) $\Longrightarrow$ (i).

Suppose that X is infinite dimensional and $\{u_1, u_2, \ldots\}$ is a denumerable linearly independent subset of X. For each $j \in \mathbb{N}$, let

$$X_j = \operatorname{span}\{u_1, \ldots, u_j\}.$$

Then each X_j is a proper closed subspace of X_{j+1} so that by Riesz lemma 2.40, there exists $x_{j+1} \in X_{j+1}$ such that

$$\|x_{j+1}\| = 1, \qquad \operatorname{dist}(x_{j+1}, X_j) \geq \frac{1}{2}.$$

Since $X_i \subseteq X_j$ for every $i \leq j$, it follows that

$$\|x_n - x_m\| \geq \frac{1}{2} \quad \forall n, m \in \mathbb{N}, \qquad n \neq m.$$

Thus, the sequence (x_n) in the unit sphere $S = \{x \in X : \|x\| = 1\}$ of X cannot have a Cauchy subsequence, showing that the unit sphere S is not totally bounded. ∎

Remark 2.10 In Theorem 2.39, the set $\{x \in X : \|x\| = 1\}$ can be replaced by $S_r := \{x \in X : \|x\| = r\}$ for any $r > 0$, because it is easily seen that S_1 is compact (resp. totally bounded) if and only if S_r is compact (resp. totally bounded). Thus, we can also deduce (*How?*) that a normed linear space X is finite dimensional if and only if $B(x_0, r)$ is totally bounded for some $x_0 \in X$ and $r > 0$.

Recall that we have proved Riesz Lemma for $0 < r < 1$. Can the result be true for $r = 1$? Not always. For example, consider the linear space

$$X = \{x \in C[0,1] : x(0) = 0\}$$

with $\|\cdot\|_\infty$, and let

$$X_0 = \left\{x \in X : \int_0^1 x(t)\,dt = 0\right\}.$$

It is easily seen (*Verify*) that X_0 is a closed proper subspace of X.

Now suppose that Riesz Lemma 2.40 holds for X_0 with $r = 1$ so that there exists $x \in X$ such that

$$\|x\|_\infty = 1 \quad \text{and} \quad \operatorname{dist}(x, X_0) = 1.$$

We shall arrive at a contradiction. First we observe the following facts:

(a) $\left|\int_0^1 x(t)\,dt\right| < 1$.

(b) $\left|\int_0^1 u(t)\,dt\right| \leq \left|\int_0^1 x(t)\,dt\right|$ for all $u \in X$ with $\|u\|_\infty = 1$.

Clearly,

$$\left|\int_0^1 x(t)\,dt\right| \leq \int_0^1 |x(t)|\,dt \leq \|x\|_\infty = 1.$$

Hence, if $\left|\int_0^1 x(t)\,dt\right| = 1$, then we get $\int_0^1 (1 - |x(t)|)\,dt = 0$, implying that $|x(t)| = 1$ for every $t \in [0,1]$, a contradiction to the fact that $x(0) = 0$. This proves (a).

Now, to see (b), first we observe that the inequality is true for all $u \in X_0$ with $\|u\|_\infty = 1$. Next suppose that $u \in X \setminus X_0$ is such that $\|u\|_\infty = 1$. Then

$$\tilde{u} := x - \frac{\int_0^1 x(t)\,dt}{\int_0^1 u(t)\,dt}\, u$$

belongs to X_0 and

$$\|x - \tilde{u}\|_\infty = \left|\frac{\int_0^1 x(t)\,dt}{\int_0^1 u(t)\,dt}\right|.$$

By our assumption, $\|x - \tilde{u}\| \geq \operatorname{dist}(x, X_0) = 1$. Thus, (b) is also proved.

But, (a) and (b) together lead to a contradiction, because, for every $\varepsilon > 0$, we can find $u_\varepsilon \in X$ such that

$$\|u_\varepsilon\|_\infty = 1, \qquad \left|\int_0^1 u(t)\,dt\right| > 1 - \varepsilon.$$

We now show that if we impose certain additional conditions on the subspace X_0 or the space X, then the conclusion in Riesz lemma (Theorem 2.40) holds for $r = 1$ as well. The Riesz lemma for $r = 1$ is closely associated with best approximation theorems to be considered in Section 2.5. First we consider the case of X_0 a finite dimensional subspace.

Theorem 2.41 *Let X_0 be a finite dimensional proper subspace of a normed linear space X. Then there exists $x \in X$ such that*

$$\|x\| = 1, \qquad \operatorname{dist}(x, X_0) = 1.$$

Proof. Let $x_0 \in X \setminus X_0$ and $Y = \operatorname{span}\{x_0; X_0\}$. Applying Riesz Lemma 2.40 with Y in place of X, we obtain a sequence (u_n) in Y such that

$$\|u_n\| = 1, \qquad 1 \geq \operatorname{dist}(u_n, X_0) \geq 1 - \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Since the closed unit ball in the finite dimensional space Y is compact (see Theorem 2.39), there is a convergent subsequence (v_n) of (u_n). Suppose $v_n \to v$ in Y as $n \to \infty$. Then it follows (*How?*) that

$$\|v\| = 1, \qquad \operatorname{dist}(v, X_0) = 1.$$

This completes the proof. ∎

In the following section we shall give another proof of Theorem 2.41 by making use of a *best approximation property* of finite dimensional subspaces.

2.5 Best Approximation Theorems

Let X be a normed linear space and X_0 be a subspace of X. For $x \in X$, an element $x_0 \in X_0$ is called a **best approximation** of x, if

$$\|x - x_0\| = \operatorname{dist}(x, X_0).$$

A subspace X_0 of X is said to have the **best approximation property** if every $x \in X$ has a best approximation in X_0.

Theorem 2.42 *Suppose X_0 is a proper closed subspace of a normed linear space X. If X_0 has the best approximation property, then the Riesz lemma holds for X_0 with $r = 1$.*

Proof. Suppose X_0 has the best approximation property. Let $x \in X \setminus X_0$, and let $x_0 \in X_0$ be such that $\|x - x_0\| = \text{dist}(x, X_0)$. Then by taking

$$\tilde{x} := \frac{x - x_0}{\|x - x_0\|},$$

it follows that

$$\|\tilde{x}\| = 1, \qquad \text{dist}(\tilde{x}, X_0) = 1.$$

This completes the proof. ∎

Now we consider a few cases of the spaces having the best approximation property.

Theorem 2.43 *Finite dimensional subspaces of normed linear spaces have the best approximation property, i.e., if X_0 is a finite dimensional subspace of a normed linear space X, then for every $x \in X$, there exists $x_0 \in X_0$ such that*

$$\|x - x_0\| = \text{dist}(x, X_0).$$

Proof. Let X be a normed linear space and X_0 be a finite dimensional subspace of X. Let $x \in X$. For a fixed $u_0 \in X_0$, consider the set

$$S = \{u \in X_0 : \|x - u\| \leq \|x - u_0\|\}$$

and the map

$$f : u \mapsto \|x - u\|, \qquad u \in S.$$

Since S is a closed and bounded subset of a finite dimensional space X_0 and since f is a continuous function, there is some $x_0 \in S$ such that

$$f(x_0) = \inf \{f(u) : u \in S\}.$$

From this it follows (*How?*) that

$$\|x - x_0\| = \inf \{\|x - u\| : u \in X_0\}.$$

Thus, the proof is complete. ∎

Another proof of Theorem 2.41. Suppose X_0 is a finite dimensional subspace of a normed linear space X. By Theorem 2.43, X_0 has the best approximation property, and hence by Theorem 2.42, the Riesz lemma (Theorem 2.40) holds for X_0 with $r = 1$. ∎

Remark 2.11 (a) In view of Theorem 2.42, and the fact that the Riesz lemma need not hold for every closed subspace, we can assert that every closed subspace of a normed linear space need not have the best approximation property.

(b) We shall prove in Chapter 8 that, if X belongs to a particular class of Banach spaces, namely, *reflexive Banach spaces*, then Theorem 2.43 holds for every closed subspace X_0 of X. We shall see that Hilbert spaces, and the spaces $\ell^p(\mathbb{N})$ and $L^p[a,b]$ for $1 < p < \infty$ belong to this class.

We observe that, in proving Theorem 2.43, we made use of a characterizing property of finite dimensionality of the subspace X_0. Next, we show that if the norm on X is induced by an inner product, then the finite dimensionality of X_0 can be replaced by assuming that it is complete. In this case we can also infer that the best approximation is unique.

Theorem 2.44 *Let X be an inner product space and X_0 be a complete subspace of X. Then for every $x \in X$, there exists a unique $x_0 \in X_0$ such that*

$$\|x - x_0\| = \text{dist}(x, X_0).$$

Moreover,

(i) $x - x_0 \in X_0^{\perp}$, *and*

(ii) *if X_0 is a proper subspace of X, then there exists $v \in X$ such that*

$$\|v\| = 1, \qquad \text{dist}\,(v, X_0) = 1$$

Proof. Let $x \in X$ and let $d = \text{dist}\,(x, X_0)$. If $x \in X_0$, we may take $x_0 = x$. Now, let $x \notin X_0$, and let (u_n) in X_0 be such that $\|x - u_n\| \to d$ as $n \to \infty$. We shall show that (u_n) converges to an element $u \in X_0$. Assuming this for the time being, we obtain $\|x - u_n\| \to \|x - u\|$ as $n \to \infty$ so that $\|x - u\| = d$. Since X_0 is a complete space, the convergence of (u_n) to an element in X_0 will follow once we show that it is a Cauchy sequence. Observe that, by parallelogram law,

$$\begin{aligned}\|u_n - u_m\|^2 &= \|(u_n - x) + (x - u_m)\|^2 \\ &= 2\left(\|u_n - x\|^2 + \|x - u_m\|^2\right) \\ &\quad -\|(u_n - x) - (x - u_m)\|^2,\end{aligned}$$

where $\|u_n - x\| \to d$ as $n \to \infty$, and

$$\|(u_n - x) - (x - u_m)\| = \|(u_n + u_m) - 2x\| \geq 2d$$

for all $n, m \in \mathbb{N}$. Hence, it follows that $\|u_n - u_m\| \to 0$ as $n, m \to \infty$.

To see the uniqueness of u, let $v \in X_0$ be another element such that $\|x - v\| = d$. Again, by parallelogram law, we have

$$\begin{aligned}\|u - v\|^2 &= \|(u - x) + (x - v)\|^2 \\ &= 2\left(\|u - x\|^2 + \|x - v\|^2\right) - \|(u - x) - (x - v)\|^2.\end{aligned}$$

Since

$$\|x - u\| = d = \|x - v\|, \qquad \left\|x - \frac{u + v}{2}\right\| \geq d,$$

it follows that $u = v$.

Now we prove the additional results (i) and (ii). To establish (i), we have to show that $\langle x - u, y\rangle = 0$ for every $y \in X_0$. Clearly, it is enough to show this for $y \in X_0$ with $\|y\| = 1$. So, let $y \in X_0$ be such that $\|y\| = 1$. First we observe that

$$\|x - u\| \leq \|x - z\| \quad \forall z \in X_0.$$

In particular, for $z = u + \alpha y$ for some $\alpha \in \mathbb{K}$,

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

Now taking $\alpha = \langle x - u, y\rangle$, we get

$$\|x - u\|^2 \leq \|x - u\|^2 - |\langle x - u, y\rangle|^2,$$

showing that

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

The result in (ii) follows by taking $y = (x - x_0)/\|x - x_0\|$. ∎

In the proof of the first part of the above theorem, the fact that X_0 is a subspace is used only to have the relation

$$\frac{1}{2}(u+v) \in X_0 \quad \text{whenever} \quad u, v \in X_0.$$

Thus, the first part of the theorem remains unchanged if X_0 is a *convex* subset of X which is complete with respect to the induced metric.

A subset S of a linear space Y is said to be a **convex set** if for every $x, y \in Y$,

$$\lambda x + (1-\lambda)y \in S, \qquad 0 < \lambda < 1.$$

Clearly, every subspace of a linear space is convex.

Thus, we have the following result.

Theorem 2.45 *Let X be a Hilbert space and X_0 be a closed convex subset of X. Then for every $x \in X$, there exists a unique $x_0 \in X_0$ such that*

$$\|x - x_0\| = \text{dist}\,(x, X_0).$$

If, in addition, X_0 is a subspace of X, then $x - x_0 \in X_0^{\perp}$.

The uniqueness of the best approximation, if it exists, can be proved for any strictly convex space X (see Remark 2.5 for the definition).

Theorem 2.46 *Let X be a strictly convex normed linear space, X_0 be a subspace of X, and $x \in X$. Then a best approximation of x from X_0, if it exists, is unique.*

Proof. Suppose $u, v \in X_0$ are such that

$$\|x - u\| = d = \|x - v\|, \qquad d := \text{dist}\,(x, X_0).$$

Then we have

$$d \leq \left\| x - \frac{u+v}{2} \right\| = \left\| \frac{x-u}{2} + \frac{x-v}{2} \right\| \leq d$$

so that by taking $x_1 = (x-u)/d$ and $x_2 = (x-v)/d$ we have

$$\|x_1\| = 1 = \|x_2\|, \qquad \left\| \frac{x_1 + x_2}{2} \right\| = 1.$$

Since X is strictly convex, we have $x_1 = x_2$; consequently, $u = v$. ∎

Recall (cf. Theorem 2.16) that every inner product space is strictly convex.

Exercise 2.15 Let X be a normed linear space and X_0 be a subspace of X with the best approximation property. Show that the set of best approximations of an element $x \in X$ is a convex subset of X_0. □

Exercise 2.16 If S is a closed convex subset of a Hilbert space, then show that there is an element in S of minimal norm, i.e., there is an $x_0 \in S$ such that $\|x_0\| = \inf\{\|u\| : u \in S\}$. □

2.6 Projection Theorem

We use Theorem 2.44 to derive the following result, known as *projection theorem.*

Theorem 2.47 (Projection theorem) *Let X be an inner product space and X_0 be a complete subspace of X. Then,*

$$X = X_0 + X_0^\perp, \qquad (X_0^\perp)^\perp = X_0.$$

Proof. By Theorem 2.44, for every $x \in X$, there exists a unique $u \in X_0$ such that $x - u \in X_0^\perp$. Thus,

$$x = u + (x - u) \quad \text{with} \quad u \in X_0, \quad x - u \in X_0^\perp$$

so that $X = X_0 + X_0^\perp$. It remains to show that $(X_0^\perp)^\perp = X_0$.

If $x \in X_0$, then for every $u \in X_0^\perp$, we have $\langle x, u\rangle = 0$ so that we have $x \in (X_0^\perp)^\perp$, showing that $X_0 \subseteq (X_0^\perp)^\perp$. Next, let $x \in (X_0^\perp)^\perp$. Again, by Theorem 2.44, there exists $v \in X_0$ such that $x - v \in X_0^\perp$. Now, since $\langle x, x - v\rangle = 0$, $\langle v, x - v\rangle = 0$, we have $\langle x - v, x - v\rangle = 0$ so that $x = v \in X_0$. ∎

Usually, in books on functional analysis, the following particular case of the above theorem is known as the *projection theorem.*

Corollary 2.48 *Let X be a Hilbert space and X_0 be a closed subspace of X. Then*

$$X = X_0 + X_0^\perp, \qquad (X_0^\perp)^\perp = X_0.$$

We have seen in Section 1.2.5 that if X_1 and X_2 are subspaces of a linear space X satisfying

$$X = X_1 + X_2, \qquad X_1 \cap X_2 = \{0\},$$

then there is a unique projection operator $P : X \to X$ such that

$$R(P) = X_1 \quad \text{and} \quad N(P) = X_2.$$

Thus, by Theorem 2.47, if X is an inner product space and X_0 is a complete subspace of X, then there is a unique projection $P : X \to X$ with

$$R(P) = X_0, \qquad N(P) = X_0^{\perp}.$$

In particular, the above projection has the property that $R(P) \perp N(P)$. In fact, the projection P associates every $x \in X$ the unique element $x_0 \in X_0$ such that $x - x_0 \in X_0^{\perp}$, and

$$\|x - x_0\| = \inf\{\|x - u\| : u \in X_0\}.$$

A projection operator $P : X \to X$ on an inner product space X is said to be an **orthogonal projection** if

$$R(P) \perp N(P).$$

Remark 2.12 (a) Suppose $P : X \to X$ is an orthogonal projection on an inner product space X. Then, by the Pythagoras theorem (Theorem 2.17), we have

$$\|x\|^2 = \|Px\|^2 + \|x - Px\|^2 \quad \forall x \in X.$$

In particular,

$$\|Px\| \le \|x\| \quad \forall x \in X.$$

Thus, it is immediate that P is a continuous function as well.

In fact, if $A : X \to Y$ is a linear operator between normed linear spaces X and Y such that there exists $c > 0$, and

$$\|Ax\| \le c\,\|x\| \quad \forall x \in X,$$

then it follows that A is a continuous function as well. We shall show in the next chapter that the above requirement is a characterizing property of continuity of a linear operator A.

(b) Now we show that the conclusion of Corollary 2.48 need not hold if X is not a Hilbert space. For this, consider $X = c_{00}$ with the norm $\|\cdot\|_2$, and let

$$X_0 := \Big\{x \in X : \sum_{j=1}^{\infty} \frac{x(j)}{j} = 0\Big\}.$$

Note that $X_0 = N(f)$, where $f : X \to \mathbb{K}$ is defined by

$$f(x) = \sum_{j=1}^{\infty} \frac{x(j)}{j}, \qquad x \in X.$$

Since

$$\begin{aligned} |f(x)| &\leq \Big(\sum_{j=1}^{\infty} \frac{1}{j^2}\Big)^{1/2} \Big(\sum_{j=1}^{\infty} |x(j)|^2\Big)^{1/2} \\ &= \frac{\pi}{\sqrt{6}} \|x\|_2 \end{aligned}$$

for all $x \in X$, it follows from the remark (a) above that f is a continuous function, and hence $X_0 := N(f)$ is a closed subspace of X.

We now show that $X_0^{\perp} = \{0\}$ so that the conclusion of Corollary 2.48 does not hold for this example: Suppose $u \in X_0^{\perp}$. Let $k \in \mathbb{N}$ be such that $u(j) = 0$ for all $j > k$. For each $n \in \{1, \ldots, k\}$, let $x_n \in X$ be defined by

$$x_n(j) = \begin{cases} n & \text{if } j = n, \\ -k-1 & \text{if } j = k+1, \\ 0 & \text{if } j \notin \{n, k+1\}. \end{cases}$$

Clearly, $x_n \in X_0$ so that $\langle u, x_n \rangle = 0$ for $n = 1, \ldots, k$. But, $\langle u, x_n \rangle = n\,u(n)$, $n = 1, \ldots, k$. Hence, $u(j) = 0$ for all $j \in \mathbb{N}$, showing that $u = 0$. Thus, we have shown that $X_0^{\perp} = \{0\}$.

Exercise 2.17 Let X be a Hilbert space. Prove the following:

(i) There is a one to one correspondence between the set of all closed subspaces of X and the set of all orthogonal projections on X.

(ii) If S is a nonempty subset of X, then $S^{\perp\perp} = \overline{\operatorname{span} S}$.

(iii) Let S be a nonempty subset of X. If $S^{\perp} = \{0\}$, then $\operatorname{span} S$ is dense in X. $\square$

PROBLEMS

1. Let X be a normed linear space with norm $\|\cdot\|$, and for nonzero $\beta \in \mathbb{K}$, let

$$\|x\|_\beta = \|\beta x\| \quad \forall x \in X.$$

Show that $\|\cdot\|_\beta$ is a norm on X.

2. Show that there is only one norm $\|\cdot\|$ on $\mathbb{K}$ such that $\|1\| = 1$.

3. Let X be a normed linear space with norm $\|\cdot\|$, $x_0 \in X$ and $r > 0$. Show that

(a) $\mathrm{cl}\,\{x \in X : \|x - x_0\| < r\} = \{x \in X : \|x - x_0\| \leq r\}$,

(b) $B(x_0, r) + x = B(x_0 + x, r)$ for all $x \in X$.

4. Let $t_1, \ldots, t_k$ be distinct points in $[a, b]$, and for $f \in \mathcal{P}_n[a, b]$, define

$$\nu(f) = \sum_{j=1}^{k} |f(t_j)|.$$

Show that ν is a norm on $\mathcal{P}_n[a, b]$ if and only if $k \geq n + 1$.

5. Suppose X is a linear space and $E \subseteq X$ is such that it is

(a) *convex*, i.e., for every $x, y \in E$, $\lambda x + (1-\lambda)y \in E$ for $0 < \lambda < 1$,

(b) *balanced*, i.e., $\alpha x \in E$ whenever $x \in E$ and $|\alpha| \leq 1$,

(c) *absorbing*, i.e., $\forall x \in X$, $\exists \lambda > 0$ such that $\lambda x \in E$, and

(d) there is no subspace $Y \neq \{0\}$ such that $Y \subseteq E$.

Show that $x \mapsto \|x\|_E := \inf\{\lambda > 0 : \lambda^{-1} x \in E\}$ is a norm on X.

6. Let $\mathcal{N}$ be the set of all norms on a linear space X such that $\sup_{\nu \in \mathcal{N}} \nu(x) < \infty$ for every $x \in X$. Show that $x \mapsto \sup_{\nu \in \mathcal{N}} \nu(x)$ is a norm on X.

7. For $1 \leq p \leq \infty$ and $1 \leq r \leq \infty$, find $c_1 > 0$, $c_2 > 0$ such that

$$c_1 \|x\|_p \leq \|x\|_r \leq c_2 \|x\|_p \quad \forall x \in \mathbb{K}^n.$$

Also, show that, for every $x \in \mathbb{K}^n$, $\|x\|_p \to \|x\|_\infty$ as $p \to \infty$.

8. For $1 \leq p \leq r \leq \infty$, show that

$$\ell^p(\mathbb{N}) \subseteq \ell^r(\mathbb{N}), \qquad L^r[a, b] \subseteq L^p[a, b].$$

Also, show that, if $p < r$, then the inclusions above are strict.

9. Give an example of a sequence in $C[a,b]$ to illustrate that there does not exist $c > 0$ such that

$$\|x\|_\infty \leq c\|x\|_1 \quad \forall x \in C[a,b].$$

10. Let $\mathbb{K}^{m\times n}$ be the set of all $m \times n$ matrices with entries in $\mathbb{K}$. Show that

$$\langle A, B\rangle := \operatorname{trace}(B^*A), \qquad A, B \in \mathbb{K}^{m\times n},$$

defines an inner product on $\mathbb{K}^{m\times n}$.

Also show that

$$\|AB\|_{\mathrm{F}} \leq \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}} \quad \forall A, B \in \mathbb{K}^{n\times n},$$

where $\|\cdot\|_{\mathrm{F}}$ is the norm induced by the above inner product, and it is called the *Frobeneus norm* on $\mathbb{K}^{m\times n}$.

11. For every linear space X, does there exist a norm on X?

12. Let X and Y be normed linear spaces and (A_n) be a sequence of linear operators from X to Y. Prove that, if there exists $x_0 \in X$ such that $(A_n x_0)$ does not converge, then there exists a dense subset D of X such that $(A_n x)$ does not converge for any $x \in D$.

[*Hint*: Use the fact that $\{x \in X : (A_n x) \text{ converges}\}$ is a proper subspace, and use Lemma 2.32 and the exercise following it.]

13. Let X be a normed linear space. Consider the norms

$$(x, y) \mapsto \|x\| + \|y\|, \qquad (x, y) \in X \times X,$$

$$(\alpha, x) \mapsto |\alpha| + \|x\|, \qquad (\alpha, x) \in \mathbb{K} \times X$$

on $X \times X$ and $\mathbb{K} \times X$, respectively. Show that the maps

$$(x, y) \mapsto x + y, \qquad x \mapsto -x, \qquad (\alpha, x) \mapsto \alpha x$$

are continuous functions. Show also that for each $u \in X$ and nonzero $\alpha \in \mathbb{K}$, the maps

$$x \mapsto x + u, \qquad x \mapsto \alpha x$$

are homeomorphisms from X onto itself, i.e., these maps are bijective, continuous, and their inverses are also continuous.

14. Let X be a normed linear space and U and V be subsets of X. Show that

(a) if one of U and V is an open set, then $U + V$ is an open set, and

(b) if U is compact and V is closed, then $U + V$ is closed.

15. Let X and Y be finite dimensional normed linear spaces with the same dimension, say n. Let $\{u_1, \ldots, u_n\}$ be a basis of X and $\{v_1, \ldots, v_n\}$ be a basis of Y. For $x = \sum_{j=1}^{n} \alpha_j u_j \in X$, define

$$T(x) = \sum_{j=1}^{n} \alpha_j v_j.$$

Show that $T : X \to Y$ is a linear homeomorphism from X onto Y.

16. Let X be an inner product space over $\mathbb{R}$. If $x, y \in X$ are such that $\|x + y\| = \|x - y\|$, then show that $x \perp y$.

17. Let X be an inner product space over $\mathbb{C}$ and $x, y \in X$. Show that $x \perp y$ if and only if $\|\alpha x + \beta y\|^2 = \|\alpha x\|^2 + \|\beta y\|^2$ for every $\alpha, \beta \in \mathbb{C}$.

[*Hint*: Look at the relation when $\alpha = 1$, $\beta = \langle x, y \rangle$.]

18. Let X be an inner product space and (x_n) be a sequence in X. For $x \in X$, show that $\|x_n - x\| \to 0$ as $n \to \infty$ whenever $\|x_n\| \to \|x\|$ and $\langle x_n, x \rangle \to \|x\|^2$ as $n \to \infty$.

19. Let X be a Hilbert space and Y be a closed subspace of X. If Z is a subspace of X such that $Y \subseteq Z$ and $Y \neq Z$, then show that there exists $x \in Z$ such that $x \perp Y$.

20. If $\psi(\cdot, \cdot)$ is a semi-inner product on a linear space X and if

$$Y = \{x \in X : \psi(x, x) = 0\},$$

then show that Y is a subspace of X and

$$\|x + Y\|_* := \sqrt{\psi(x, x)}, \qquad x + Y \in X/Y,$$

defines a norm on X/Y.

21. Let X be an inner product space and $S \subseteq X$.

(a) Show that $S^\perp = (\operatorname{span} S)^\perp = \overline{\operatorname{span} S}^\perp$.

(b) If $X = C[a,b]$ with L^2-inner product and $S = \mathcal{P}[a,b]$, then find $S^\perp$?

22. Let $X = C[-1,1]$ be equipped with L^2-inner product,

$$S_1 = \{f \in X : f(-t) = f(t) \quad \forall t \in [-1,1]\},$$
$$S_2 = \{f \in X : f(-t) = -f(t) \quad \forall t \in [-1,1]\}.$$

Show that $S_1 \perp S_2$, i.e., $\langle f, g \rangle = 0$ for every $(f,g) \in S_1 \times S_2$. Also, show that

$$S_1^\perp = S_2, \qquad S_2^\perp = S_1, \qquad X = S_1 + S_2.$$

23. Let $t_1, \ldots, t_n$ be distinct points in $[a,b]$. For $x \in C[a,b]$, let

$$\nu_p = \begin{cases} \left(\sum_{j=1}^n |x(t_j)|^p\right)^{1/p} & \text{if } 1 \le p < \infty, \\ \max\{|x(t_j)| : j = 1, \ldots, n\} & \text{if } p = \infty. \end{cases}$$

Show that ν_p is a seminorm on $C[a,b]$, and ν_p is not a norm on $C[a,b]$

24. Let $X = C[a,b]$ with $\|\cdot\|_\infty$ and $t_1, \ldots, t_n$ be distinct points in $[a,b]$. Let

$$X_n = \{x \in X : x(t_j) = 0, \; j = 1, \ldots, n\}.$$

Show that X_n is a closed subspace of X. What is the dimension of X/X_n?

25. Give an example to illustrate that if X_1 and X_2 are closed subspaces of a normed linear space X, then

$$X_1 + X_2 := \{x_1 + x_2 : x_1 \in X_1, \; x_2 \in X_2\}$$

need not be a closed subspace.

26. Let X be a linear space, Y be a normed linear space, and $A : X \to Y$ be a linear operator. Suppose $y_0 \in R(A)$ and $r > 0$ are such that the equation $Ax = y$ has a solution for every $y \in B_Y(y_0, r)$. Show that $Ax = y$ has a solution for every $y \in Y$.

27. Show that $C^1[a,b]$ is a Banach space with respect to the norms

$$\|x\| = |x(a)| + \|x'\|_\infty, \qquad \|x\|_* = \|x\|_\infty + \|x'\|_\infty \quad \forall x \in x \in C^1[a,b].$$

28. If $1 \le p < \infty$, then there does not exist $c > 0$ such that

$$\|x\|_\infty \le c\|x\|_p \quad \forall x \in C[a,b].$$

Why?

3

Operators on Normed Linear Spaces

In applications, it is very important to know whether a given linear operator is continuous or not. For instance, suppose a physical problem is modelled in the form of an operator equation

$$Ax = y,$$

where $A : X \to Y$ is a linear operator between normed linear spaces X and Y. The problem may be to find $x \in X$ from the knowledge of $y \in Y$, or to find $y \in Y$ from the knowledge of $x \in X$. The given information is often known as 'data' and the information which is to be obtained is known as a 'solution'. The questions of existence and uniqueness of a solution are purely set-theoretic or algebraic ones. But, in applications, exact data are not usually available. Also, in numerical approximation procedures, one normally considers approximate data. In such situations, 'stability' of the solution is of great importance. It amounts to knowing whether the operator A or its inverse A^{-1} is continuous.

In this chapter our main study is about *continuous linear operators* between normed linear spaces and their various properties. We shall also discuss briefly a class of discontinuous operators which are closely related to continuous operators, namely, the class of *closed operators.*

3.1 Bounded Operators

Suppose (Ω_1, d_1) and (Ω_2, d_2) are metric spaces. Recall from Section 2.1 that a function $f : \Omega_1 \to \Omega_2$ is continuous at a point $t \in \Omega_1$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$s \in \Omega_1,\ d_1(s,t) < \delta \implies d_2(f(s), f(t)) < \varepsilon.$$

One of the important characterizations of continuity of a function $f : \Omega_1 \to \Omega_2$ is in terms of convergence of sequences (*Verify*): A function $f : \Omega_1 \to \Omega_2$ is continuous at $t \in \Omega_1$ if and only if for every sequence (t_n) in Ω_1 that converges to t, the sequence $(f(t_n))$ converges to $f(t)$.

Now, let X and Y be normed linear spaces and $A : X \to Y$ be a linear operator, i.e., for all $x, y \in X$ and for all $\alpha \in \mathbb{K}$,

$$A(x+y) = Ax + Ay, \qquad A(\alpha x) = \alpha Ax.$$

As in Chapter 2, we may use the same notation $\|\cdot\|$ for the norm on any normed linear space. Thus, a linear operator $A : X \to Y$ is **continuous** at $x \in X$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$u \in X, \quad \|u - x\| < \delta \implies \|Au - Ax\| < \varepsilon.$$

3.1.1 Some Basic Results and Examples

Linearity of a function $A : X \to Y$ provides us additional information which facilitates the study of continuity of a function. Recall from Remark 2.12(b) that if a linear operator $A : X \to Y$ satisfies

$$\|Ax\| \le c\,\|x\| \quad \forall x \in X$$

for some $c > 0$, then A is a continuous function as well.

In the following theorem we list some of the characterizations of continuity of a linear operator.

Theorem 3.1 *Let X and Y be normed linear spaces and $A : X \to Y$ be a linear operator. Then the following are equivalent:*

(i) *A is continuous at 0.*

(ii) *A is continuous at every $x \in X$.*

(iii) *A is uniformly continuous.*

(iv) *There exists $c > 0$ such that $\|Ax\| \le c\|x\|$ for all $x \in X$.*

(v) *$\{Ax : x \in X, \; \|x\| = 1\}$ is a bounded subset of Y.*

(vi) *For every bounded set $E \subseteq X$, the set $\{Ax : x \in E\}$ is bounded in Y.*

Proof. The following implications are quite clear:

$$\text{(iii)} \implies \text{(ii)} \implies \text{(i)}, \quad \text{(iv)} \implies \text{(i)}, \quad \text{(iv)} \implies \text{(vi)} \implies \text{(v)}.$$

Hence, it is enough to prove the implications:

$$\text{(i)} \Longrightarrow \text{(iii)}, \quad \text{(i)} \Longrightarrow \text{(iv)}, \quad \text{(v)} \Longrightarrow \text{(iv)}.$$

Let $\varepsilon > 0$ and assume (i). Then there exists $\delta > 0$ such that

$$x \in X, \; \|x\| < \delta \Longrightarrow \|Ax\| < \varepsilon.$$

Hence, for every $u \in X$, taking $x - u$ in place of x above,

$$\|x - u\| < \delta \Longrightarrow \|Ax - Au\| = \|A(x - u)\| < \varepsilon.$$

Therefore, A is continuous at u. Since the above δ is independent of u, A is uniformly continuous. Thus, (i) implies (ii) and (iii). Also, for every $u \neq 0$, $v := \delta u/2\|u\|$ is of norm less than δ, so that by (i), we have $\|Av\| < \varepsilon$. Therefore,

$$\|Au\| = \left\|A\left(\frac{2\|u\|}{\delta}v\right)\right\| \leq \frac{2\,\varepsilon}{\delta}\|u\| \quad \forall u \in X.$$

This proves (iv). Thus, we have proved that (i) implies (ii), (iii), and (iv).

Next, assume (v), and let $x \neq 0$. Since $\tilde{x} := x/\|x\|$ is of norm 1, by (v), there is $c > 0$ such that $\|A\tilde{x}\| \leq c$. Consequently,

$$\|Ax\| \leq c\|x\|,$$

proving (iv). Thus, (v) implies (iv). ∎

We have mentioned in the beginning of this section that an important characterization of continuity of a function is in terms of convergence of sequences. Thus, in view of the above theorem, a linear operator $A : X \to Y$ is continuous if and only if for every sequence (x_n) in X that converges to 0, the sequence (Ax_n) also converges to 0.

We observe that the statements (i) and (vi) in the above theorem can be rewritten as

$$\forall \varepsilon > 0, \, \exists \delta > 0 \;\text{ such that }\; \|x\| < \delta \Longrightarrow \|Ax\| < \varepsilon,$$

$$\forall \alpha > 0, \, \exists \beta > 0 \;\text{ such that }\; \|x\| \leq \alpha \Longrightarrow \|Ax\| \leq \beta,$$

respectively.

In the above theorem, the implications

$$\text{(iii)} \Longrightarrow \text{(ii)} \Longrightarrow \text{(i)}, \quad \text{(iv)} \Longrightarrow \text{(v)}, \quad \text{(vi)} \Longrightarrow \text{(v)}$$

hold even if $A : X \to Y$ is not linear. But, reverse inequalities need not hold if A is not linear. (*Give examples to justify this statement.*)

In view of the equivalence of (ii) and (vi) in Theorem 3.1, a continuous linear operator is often called a **bounded linear operator** or simply a **bounded operator**.

If $A : X \to Y$ is a bounded linear operator, then we may say that A is a bounded operator between X and Y, and if $Y = X$, then we may say that A is a bounded operator on X.

A linear operator which is not a bounded operator is called an **unbounded operator**. Thus, a linear operator $A : X \to Y$ is an unbounded operator, i.e., a discontinuous operator, if and only if there exists a bounded sequence (x_n) in X such that (Ax_n) is unbounded in Y; equivalently, there exists a sequence (x_n) in X and $\delta > 0$ such that $x_n \to 0$, but $\|Ax_n\| \geq \delta$ for all $n \in \mathbb{N}$.

It has to be borne in mind that a bounded operator $A : X \to Y$ is not a bounded function on X, unless A is the zero operator. Thus, if $A : X \to Y$ is a nonzero bounded operator, then *it is not true* that there exists $c > 0$ such that $\|Ax\| \leq c$ for all $x \in X$ (*Why?*).

Exercise 3.1 Show that a linear operator $A : X \to Y$ is a bounded operator if and only if, for any $r > 0$, its restriction to any of the sets

$$\{x \in X : \|x\| = r\}, \quad \{x \in X : \|x\| \leq r\}, \quad \{x \in X : \|x\| < r\}$$

is a bounded function. □

A characterization of continuity of the inverse operator

Before giving specific examples of bounded operators, let us observe a result about the inverse of a linear operator.

Recall from Section 1.2.4 that if $A : X \to Y$ is a linear operator between linear spaces X and Y and if it is injective, then there exists a unique linear operator $B : R(A) \to X$ such that

$$(BA)(x) = x \quad \forall x \in X, \quad (AB)(y) = y \quad \forall y \in R(A),$$

and this operator B is called the **inverse** of A, denoted by A^{-1}.

Proposition 3.2 *Let X and Y be normed linear spaces, and $A : X \to Y$ be a linear operator. Then there exists $\gamma > 0$ such that*

$$\|Ax\| \geq \gamma\|x\| \quad \forall x \in X$$

if and only if A is injective and $A^{-1} : R(A) \to X$ is continuous, and in that case,

$$\|A^{-1}y\| \leq \frac{1}{\gamma}\|y\| \quad \forall y \in R(A).$$

Proof. Suppose there exists $\gamma > 0$ such that $\|Ax\| \geq \gamma\|x\|$ for all $x \in X$. Then, clearly, A is injective. To see that $A^{-1} : R(A) \to X$ is continuous, let $y \in R(A)$, and let $x \in X$ be such that $y = Ax$. Then we have

$$\|y\| = \|Ax\| \geq \gamma\|x\| = \gamma\|A^{-1}y\|.$$

Hence, by the equivalence of (ii) and (iv) in Theorem 3.1, the operator $A^{-1} : R(A) \to X$ is continuous, and $\|A^{-1}y\| \leq (1/\gamma)\|y\|$ for all $y \in R(A)$.

Conversely, suppose A is injective and $A^{-1} : R(A) \to X$ is continuous. Again, by the equivalence of (ii) and (iv) in Theorem 3.1, there exists $c > 0$ such that

$$\|A^{-1}y\| \leq c\|y\| \quad \forall y \in R(A).$$

For $x \in X$, let $y = Ax$. Then, by the above relation, we have $\|x\| \leq c\|Ax\|$ so that we may take $\gamma = 1/c$. ∎

Now a definition: A linear operator $A : X \to Y$ between normed linear spaces X and Y is said to be **bounded below**, if there exists $\gamma > 0$ such that

$$\|Ax\| \geq \gamma\|x\| \quad \forall x \in X.$$

In other words, $\{\|Ax\| : \|x\| = 1\}$ is bounded below by a positive number γ.

Recall that a linear operator $A : X \to Y$ between normed linear spaces X and Y is a bounded operator if and only if there exists $c > 0$ such that

$$\|Ax\| \leq c\|x\| \quad \forall x \in X,$$

i.e., $\{\|Ax\| : \|x\| = 1\}$ is *bounded above* by a positive number c.

A linear operator which is also a homeomorphism is called a **linear homeomorphism**, and a linear operator which is also an isometry is called a **linear isometry**.

Thus, by Proposition 3.2, a bijective linear operator $A : X \to Y$ is a linear homeomorphism if and only if there exists $c_1 > 0$ and $c_2 > 0$ such that

$$c_1\|x\| \leq \|Ax\| \leq c_2\|x\| \quad \forall x \in X.$$

In particular, a surjective linear isometry is a linear homeomorphism. Recall that a linear operator $A : X \to Y$ is a linear isometry if and only if $\|Ax\| = \|x\|$ for all $x \in X$.

Normed linear spaces X and Y are said to be **linearly homeomorphic** if there is a linear homeomorphism from X onto Y, and X and Y are said to be **linearly isometric** if there is a linear isometry from X onto Y.

The following result will be of use in later chapters.

Proposition 3.3 *Suppose X is a Banach space, Y is a normed linear space, and $A : X \to Y$ is a bounded linear operator. If A is bounded below, then $R(A)$ is a closed subspace of Y.*

Proof. Suppose A is bounded below. Let (y_n) be a sequence in $R(A)$ which converges in Y to an element $y \in Y$. We show that $y \in R(A)$. For this, let $x_n \in X$ be such that $y_n = Ax_n$ for all $n \in \mathbb{N}$. Since (y_n) is a Cauchy sequence and A is bounded below, it follows that (x_n) is a Cauchy sequence in X. Since X is a Banach space, (x_n) converges to some $x \in X$. Hence, by using the continuity of A, (Ax_n) converges to Ax. Thus, we have $y = Ax \in R(A)$. ∎

EXAMPLE 3.1 In all the following examples, a linear operator $A : X \to Y$ is shown to be continuous, i.e., a bounded operator, by producing a constant $c > 0$ such that

$$\|Ax\| \leq c\|x\| \quad \forall x \in X.$$

(i) Let $A : \ell^p(n) \to \ell^r(m)$ be a linear operator. Then we have

$$Ax = \sum_{j=1}^{n} x(j)Ae_j, \qquad x \in \ell^p(n),$$

where $e_j(i) = \delta_{ij}$ for all $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. Hence, by triangle inequality and Hölder's inequality,

$$\|Ax\|_r \leq \sum_{j=1}^{n} |x(j)|\,\|Ae_j\|_r \leq \|u_r\|_q\|x\|_p \quad \forall x \in \ell^p(n),$$

where $u_r \in \mathbb{K}^n$ is such that $u_r(j) = \|Ae_j\|_r$ for $j = 1, \ldots, n$, and q is the conjugate exponent of p. Thus, A is a bounded linear operator.

Note that if A is represented by an $m \times n$ matrix (a_{ij}), then

$$(Ae_j)(i) = a_{ij}, \qquad i \in \{1, \ldots, m\};\ j \in \{1, \ldots, n\}.$$

Thus, as a particular case of the above, we also have

$$\|Ax\|_1 \le \sum_{j=1}^{n} |x(j)|\, \|Ae_j\|_1 \le \left(\max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| \right) \|x\|_1,$$

$$\|Ax\|_\infty \le \sum_{j=1}^{n} |x(j)|\, \|Ae_j\|_\infty \le \left(\max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| \right) \|x\|_\infty.$$

Is a linear operator $A : \mathbb{K}^n \to \mathbb{K}^m$ continuous with respect to any norm on the spaces $\mathbb{K}^n$ and $\mathbb{K}^m$?

The answer is affirmative due to the following general result.

Theorem 3.4 *Let X and Y be normed linear spaces. If X is finite dimensional, then every linear operator $A : X \to Y$ is continuous.*

Proof. Let X be a finite dimensional space and $A : X \to Y$ be a linear operator. Recall from Section 2.1.4 that

$$x \mapsto \|x\|_* := \|x\| + \|Ax\|, \qquad x \in X,$$

is also a norm on the linear space X, called the graph norm induced by A. Since X is finite dimensional, by Theorem 2.24, there exists $a > 0$ and $b > 0$ such that

$$a\,\|x\| \le \|x\|_* \le b\,\|x\| \quad \forall x \in X.$$

In particular,

$$\|Ax\| \le b\,\|x\| \quad \forall x \in X.$$

This completes the proof. ∎

Corollary 3.5 *Any two finite dimensional normed linear spaces of the same dimension are linearly homeomorphic.*

Proof. Let X and Y be normed linear spaces of dimension n. By the above theorem, it is enough to find a bijective linear operator from X to Y. Now, let $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$ be bases of X and Y, respectively. Then it is easily seen that the map $T : X \to Y$ defined by

$$Tx = \sum_{j=1}^{n} \alpha_j v_j, \qquad x := \sum_{j=1}^{n} \alpha_j u_j \in \mathbb{K}^n,$$

is a bijective linear operator. ∎

In examples (ii) and (iii) below, we consider bounded operators between certain sequence spaces induced by infinite matrices.

Suppose (a_{ij}) is an infinite matrix of scalars, i.e., $a_{ij} \in \mathbb{K}$ for $i, j \in \mathbb{N}$. For $x \in \mathcal{F}(\mathbb{N}, \mathbb{K})$, consider the series

$$(Ax)(i) := \sum_{j=1}^{\infty} a_{ij} x(j). \qquad i \in \mathbb{N}, \tag{3.1}$$

Given certain sequence spaces X and Y, i.e., $X, Y \subseteq \mathcal{F}(\mathbb{N}, \mathbb{K})$, with some norms on them, we shall impose some conditions on (a_{ij}), so that $A : X \to Y$ defines a bounded linear operator. The first requirement on (a_{ij}) is that the series in (3.1) converges. For this, we assume that

$$\sum_{j=1}^{\infty} |a_{ij}|\, |x(j)| < \infty \quad \forall i \in \mathbb{N}, \forall x \in X. \tag{3.2}$$

EXAMPLE 3.1(*cont.*) (ii) Let (a_{ij}) be an infinite matrix of scalars such that $\alpha := \sup_j \sum_{i=1}^{\infty} |a_{ij}| < \infty$. Then, for all $x \in \ell^1$, we have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|\, |x(j)| = \sum_{j=1}^{\infty} \Big(\sum_{i=1}^{\infty} |a_{ij}| \Big) |x(j)|$$

so that $(Ax)(i)$ in (3.1) is well defined for all $x \in \ell^1$ and for all $i \in \mathbb{N}$, and

$$\sum_{i=1}^{\infty} |(Ax)(i)| \leq \alpha \|x\|_1.$$

Hence,

$$Ax \in \ell^1, \quad \|Ax\|_1 \leq \alpha \|x\|_1 \quad \forall x \in \ell^1,$$

so that $A : \ell^1 \to \ell^1$ is a bounded linear operator.

(iii) Again, let (a_{ij}) be an infinite matrix of scalars such that $\beta := \sup_i \sum_{j=1}^\infty |a_{ij}| < \infty$. Then

$$\sum_{j=1}^\infty |a_{ij}|\,|x(j)| \leq \beta\,\|x\|_\infty \quad \forall x \in \ell^\infty$$

so that $(Ax)(i)$ in (3.1) is well defined for all $x \in \ell^\infty$ and for all $i \in \mathbb{N}$, and

$$Ax \in \ell^\infty, \quad \|Ax\|_\infty \leq \beta\|x\|_\infty \quad \forall x \in \ell^\infty.$$

Thus, $A : \ell^\infty \to \ell^\infty$ is a bounded linear operator.

(iv) Let $X = C[a,b]$ with $\|\cdot\|_\infty$. Let

$$f(x) = \int_a^b x(t)\,dt, \qquad g(x) = x(t_0), \qquad (Ax)(s) = \int_a^s x(t)\,dt$$

for all $x \in X$, for some $t_0 \in [a,b]$. Then, f and g are linear functionals on X and $A : X \to X$ is a linear operator. Moreover, they are continuous. This is seen by observing that for every $x \in X$,

$$|f(x)| \leq (b-a)\|x\|_\infty, \qquad |g(x)| \leq \|x\|_\infty, \qquad \|Ax\|_\infty \leq (b-a)\|x\|_\infty.$$

Fredholm integral operator

(v) Let $X = C[a,b]$ with $\|\cdot\|_\infty$. For $k(\cdot,\cdot) \in C\left([a,b] \times [a,b]\right)$, let

$$(Ax)(s) = \int_a^b k(s,t)x(t)\,dt, \qquad x \in X, \; s \in [a,b].$$

We show that $Ax \in C[a,b]$ for all $x \in C[a,b]$ and $A : X \to X$ is a bounded linear operator: Observe that for every $s_0, s \in [a,b]$ and $x \in C[a,b]$,

$$\begin{aligned}
|(Ax)(s) - (Ax)(s_0)| &= \left|\int_a^b [k(s,t) - k(s_0,t)]x(t)\,dt\right| \\
&\leq \int_a^b |k(s,t) - k(s_0,t)|\;|x(t)|\,dt \\
&\leq \|x\|_\infty \int_a^b |k(s,t) - k(s_0,t)|\,dt.
\end{aligned}$$

Now, since the continuous function $k(\cdot,\cdot)$ defined on the compact set $[a,b]\times[a,b]$ is uniformly continuous, it follows that, for $\varepsilon>0$ there exists $\delta>0$ such that

$$|s-s_0|<\delta \implies |k(s,t)-k(s_0,t)|<\varepsilon \quad \forall\, t\in[a,b].$$

Consequently,

$$|(Ax)(s)-(Ax)(s_0)|\le \varepsilon(b-a)\|x\|_\infty.$$

Thus, $Ax\in C[a,b]$ for all $x\in C[a,b]$. The linearity of $A: X\to X$ can be seen easily. To show its continuity, we again observe that

$$|(Ax)(s)|\le \int_a^b |k(s,t)|\,|x(t)|\,dt \le \left(\int_a^b |k(s,t)|\,dt\right)\|x\|_\infty$$

so that

$$\|Ax\|\le \left(\sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt\right)\|x\|_\infty.$$

Since the function

$$s\mapsto \int_a^b |k(s,t)|\,dt, \qquad s\in[a,b]$$

is continuous, we have

$$c:=\sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt<\infty$$

and hence,

$$\|Ax\|\le c\|x\|_\infty \quad \forall\, x\in X.$$

The operator in the above example is called a *Fredholm integral operator* with kernel $k(s,t)$.

Orthogonal projection

(vi) Let X be an inner product space and $P: X\to X$ be an *orthogonal projection*, i.e., P is a linear operator such that

$$P^2=P, \qquad R(P)\perp N(P).$$

Recall from Remark 2.12 that

$$\|Px\|\le\|x\| \quad \forall\, x\in X,$$

showing that P is continuous.

(vii) Let X be an inner product space. For $u \in X$, consider $f_u : X \to \mathbb{K}$ defined by

$$f_u(x) = \langle x, u\rangle, \qquad x \in X.$$

Obviously, f_u is a linear functional on X. By Schwarz inequality (Theorem 2.14), we have

$$|f_u(x)| = |\langle x, u\rangle| \le \|u\|\,\|x\| \quad \forall x \in X,$$

so that f_u is continuous.

We shall soon show that if X is a Hilbert space, then every continuous linear functional is of the above form (Theorem 3.9). This result is known as the *Riesz representation theorem.*

Lagrange interpolatory projection

(viii) For distinct points $t_{1,n}, \ldots, t_{n,n}$ in $[a,b]$ and for $x \in C[a,b]$, let $L_n x$ be the *Lagrange interpolation* of x, i.e.,

$$L_n x = \sum_{j=1}^{n} x(t_{j,n})\ell_{j,n},$$

where

$$\ell_{j,n}(t) = \prod_{\substack{i=1\\ i\ne j}}^{n} \frac{t - t_{i,n}}{t_{j,n} - t_{i,n}}, \qquad t \in [a,b].$$

Clearly, $L_n x \in \mathcal{P}_{n-1}[a,b]$ for each $x \in C[a,b]$, and L_n is a linear operator on $C[a,b]$. Moreover, since $\ell_{i,n}(t_{j,n}) = \delta_{ij}$,

$$L_n(L_n x) = L_n x \quad \forall x \in C[a,b].$$

Hence, L_n is a projection operator on $C[a,b]$. We also note that

$$\begin{aligned} |(L_n x)(t)| &\le \sum_{j=1}^{n} |x(t_{j,n})|\,|\ell_{j,n}(t)| \\ &\le \left(\sum_{j=1}^{n} |\ell_{j,n}(t)|\right)\|x\|_\infty \end{aligned}$$

so that

$$\|L_n x\|_\infty \le \left(\sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)|\right)\|x\|_\infty$$

for all $x \in C[a,b]$. Thus, $L_n : C[a,b] \to C[a,b]$ is continuous with respect to the norm $\|\cdot\|_\infty$ on $C[a,b]$.

The projections L_n are called *Lagrange interpolatory projections.*

(ix) For $u \in C[a,b]$, let $A : C[a,b] \to C[a,b]$ be defined by

$$(Ax)(t) = u(t)x(t) \quad \forall x \in C[a,b],\ t \in [a,b].$$

Then we have

$$\|Ax\|_\infty \leq \|u\|_\infty \|x\|_\infty \quad \forall x \in C[a,b].$$

Hence A is a bounded linear operator on $C[a,b]$ with respect to the norm $\|\cdot\|_\infty$.

Examples of unbounded operators

Now we give some examples of *unbounded linear operators*, i.e., linear operators that are not continuous.

EXAMPLE 3.2 In the following, a linear operator $A : X \to Y$ is shown to be discontinuous by showing that there exists a bounded set $E \subseteq X$ such that $\{Ax : x \in E\}$ is not bounded in Y (cf. Theorem 3.1). In fact, we produce a bounded sequence (x_n) in X such that the sequence (Ax_n) is not bounded in Y.

(i) Let $X = C^1[0,1]$ with $\|\cdot\|_\infty$, and let $f : X \to \mathbb{K}$ be defined by

$$f(x) = x'(1), \qquad x \in X.$$

Clearly, f is a linear functional. It is not continuous. To show this, consider the sequence (x_n) in X defined by $x_n(t) = t^n$ for $t \in [a,b]$ and $n \in \mathbb{N}$. Then it is seen that

$$\|x_n\|_\infty = 1, \quad |f(x_n)| = |x_n'(1)| = n \quad \forall n \in \mathbb{N}$$

so that $(f(x_n))$, the image of the bounded sequence (x_n), is not bounded in $\mathbb{K}$.

(ii) Let $X = C^1[0,1]$ and $Y = C[0,1]$, both with $\|\cdot\|_\infty$, and let $A : X \to Y$ be defined by

$$(Ax)(t) = x'(t), \qquad x \in X,\ t \in [0,1].$$

Clearly, A is a linear operator. It is not continuous. To see this, consider the sequence (x_n) in X defined as in example (i) above.

Then again we have

$$\|x_n\|_\infty = 1, \qquad \|Ax_n\|_\infty = n \quad \forall n \in \mathbb{N}$$

so that (Ax_n), the image of the bounded sequence (x_n), is not bounded in Y.

Another sequence for which the above conclusion holds is defined by

$$x_n(t) = \frac{\sin(nt)}{\sqrt{n}}, \qquad t \in [0,1],\ n \in \mathbb{N}.$$

In this case, we have

$$\|x_n\|_\infty = \frac{1}{\sqrt{n}}, \qquad \|Ax_n\|_\infty = \sqrt{n} \quad \forall n \in \mathbb{N}.$$

(iii) Let $X = c_{00}$ with $\|\cdot\|_\infty$. Consider the linear functional $f : X \to \mathbb{K}$ defined by

$$f(x) = \sum_{j=1}^{\infty} x(j), \qquad x \in X.$$

Clearly, f is linear. To see that it is not continuous, consider the sequence (x_n) defined by

$$x_n(j) = \begin{cases} 1 & \text{if } j \le n, \\ 0 & \text{if } j > n. \end{cases}$$

Then we have $\|x_n\|_\infty = 1$ and $|f(x_n)| = n$ for all $n \in \mathbb{N}$. Thus, (x_n) is bounded in X but $(f(x_n))$ is not bounded in $\mathbb{K}$.

Exercise 3.2 Let $X = c_{00}$ with $\|\cdot\|_p$, and $f : X \to \mathbb{K}$ be defined by $f(x) = \sum_{j=1}^{\infty} x(j)$, $x \in X$. Show that f is continuous if and only if $p = 1$. $\square$

Remark 3.1 Examples 3.2((i) and (iii)) show that the assumption of finite dimensionality of X in Theorem 3.4 cannot be replaced by finite dimensionality of Y.

While discussing the continuity of a linear operator, an important point one should keep in mind is that continuity of an operator depends very much on the norms on the spaces under consideration. That is, a linear operator $A : X \to Y$ may not be continuous with respect to certain norms on X and Y, but it can be continuous with respect to certain other norms. Recall from Example 3.2(ii) that $A : C^1[0,1] \to C[0,1]$, defined by $Ax = x'$, $x \in C^1[0,1]$, is not continuous with respect to the norm $\|\cdot\|_\infty$. But, if we endow $C^1[0,1]$ with the norm $\|x\| := \|x\|_\infty + \|x'\|_\infty$, $x \in C^1[0,1]$, then the operator A is continuous. More generally, suppose that X and Y are normed linear spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $A : X \to Y$ is a linear operator, possibly discontinuous. Let $\|\cdot\|_*$ be the graph norm on X with respect to A, i.e.,

$$\|x\|_* = \|x\|_X + \|Ax\|_Y, \qquad x \in X.$$

Since $\|Ax\|_Y \leq \|x\|_*$ for all $x \in X$, $A : X \to Y$ is continuous with respect to the norms $\|\cdot\|_*$ on X and $\|\cdot\|_Y$ on Y. In application, the procedure of replacing a norm by another norm so as to make the operator continuous may not be desirable, as usually one would like to compute the size of certain quantities in a certain specific manner.

3.2 The Space $\mathcal{B}(X,Y)$

We use the notation $\mathcal{B}(X,Y)$ to denote the set of all bounded (linear) operators, that is, continuous linear operators from a normed linear space X to a normed linear space Y. If $Y = X$, then we write $\mathcal{B}(X)$ instead of $\mathcal{B}(X,X)$.

Theorem 3.6 *The set $\mathcal{B}(X,Y)$ is a subspace of $\mathcal{L}(X,Y)$.*

Proof. Clearly, $\mathcal{B}(X,Y) \subseteq \mathcal{L}(X,Y)$. Hence, it is enough to show that for $A, B \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{K}$,

$$A + B \in \mathcal{B}(X,Y), \qquad \alpha A \in \mathcal{B}(X,Y).$$

So let $A, B \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{K}$. By Theorem 3.1, there exists $c_A, c_B > 0$ such that

$$\|Ax\| \leq c_A\|x\|, \quad \|Bx\| \leq c_B\|x\| \quad \forall x \in X.$$

Hence,

$$\|(A+B)x\| = \|Ax+Bx\| \leq \|Ax\|+\|Bx\| \leq (c_A+c_B)\|x\| \quad \forall x \in X,$$

$$\|(\alpha A)x\| = \|\alpha(Ax)\| = |\alpha|\,\|Ax\| \leq |\alpha|c_A\|x\| \quad \forall x \in X.$$

Thus, we have $A + B \in \mathcal{B}(X,Y)$ and $\alpha A \in \mathcal{B}(X,Y)$. ∎

The space $\mathcal{B}(X,\mathbb{K})$ of all continuous linear functionals or bounded linear functionals is called the **dual space** of X or, simply, the **dual** of X. Dual of X is denoted by X'.

3.2.1 Norm on $\mathcal{B}(X,Y)$

Recall that if $A \in \mathcal{B}(X,Y)$, then there exists $c > 0$ such that

$$\|Ax\| \le c\|x\| \quad \forall x \in X.$$

In particular, we have

$$\sup\left\{\frac{\|Ax\|}{\|x\|} : x \in X,\ x \neq 0\right\} \le c.$$

Let

$$\nu_A = \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in X,\ x \neq 0\right\}$$

Then it follows that

$$\|Ax\| \le \nu_A\|x\| \quad \forall x \in X.$$

Also, we have

$$\nu_A \le c \quad \text{whenever} \quad \|Ax\| \le c\|x\| \quad \forall x \in X.$$

Therefore, for $A \in \mathcal{B}(X,Y)$, we have

$$\nu_A = \inf\{c > 0 : \|Ax\| \le c\|x\|, \forall x \in X\}.$$

Note that the quantity ν_A depends on the norms of both X and Y.

Theorem 3.7 *Let X and Y be normed linear spaces. Then the map $A \mapsto \nu_A$ is a norm on $\mathcal{B}(X,Y)$.*

Proof. Clearly, for every $A \in \mathcal{B}(X,Y)$,

$$\nu_A = 0 \iff \|Ax\| = 0 \quad \forall x \in X \iff A = 0.$$

For every $A, B \in \mathcal{B}(X,Y)$ and $x \in X$, we have

$$\|(A+B)(x)\| \le \|Ax\| + \|Bx\| \le \nu_A\|x\| + \nu_B\|x\| = (\nu_A + \nu_B)\|x\|$$

so that

$$\nu_{A+B} \le \nu_A + \nu_B.$$

Also, for every $A \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{K}$,

$$\begin{aligned} \nu_{\alpha A} &= \sup\left\{\frac{\|(\alpha A)x\|}{\|x\|} : x \in X,\ x \neq 0\right\} \\ &= |\alpha| \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in X,\ x \neq 0\right\} \\ &= |\alpha|\nu_A. \end{aligned}$$

This completes the proof. ∎

Exercise 3.3 If $A \in \mathcal{B}(X,Y)$, then show that each of the quantities

$$\sup_{\|x\|\leq 1} \|Ax\|, \qquad \sup_{\|x\|<1} \|Ax\|, \qquad \sup_{\|x\|=1} \|Ax\|$$

is equal to ν_A. □

Having proved that $A \mapsto \nu_A$ is a norm on $\mathcal{B}(X,Y)$, we denote ν_A by $\|A\|$, i.e., for $A \in \mathcal{B}(X,Y)$,

$$\|A\| = \sup\{\|Ax\| : x \in X, \|x\| \leq 1\},$$

and the quantity $\|A\|$ is called the **norm of** A. Thus, if $A \in \mathcal{B}(X,Y)$, then

$$\|Ax\| \leq \|A\| \, \|x\| \quad \forall x \in X.$$

As a particular case of the above consideration, the dual X' of a normed linear space X is a normed linear space with the norm

$$\|f\| = \sup\{|f(x)| : x \in X, \|x\| \leq 1\}, \quad f \in X'.$$

Thus, from Corollary 3.5, it follows that, if X is finite dimensional, then X' is linearly homeomorphic with X.

It can be easily seen that if $A \in \mathcal{B}(X,Y)$ and X_0 is a subspace of X, then the restriction operator $A_0 : A|_{X_0} : X_0 \to Y$ is also continuous, and

$$\|A_0\| \leq \|A\|.$$

One may also observe that a continuous linear operator remains continuous if its codomain is restricted to a subspace which contains its range. That is, if $A : X \to Y$ is a continuous linear operator and if Y_1 is a subspace of Y such that $R(A) \subseteq Y_1$, then the operator $A_1 : X \to Y_1$ defined by $A_1 x = Ax$ for all $x \in X$ is a continuous linear operator. More over, it is seen that $\|A_1\| = \|A\|$. In the above two situations, we may use the same notation for the operator, but specify the domain and codomain accordingly. That is, the operators $A_0 : X_0 \to Y$ and $A_1 : X \to Y_1$ will be represented by $A : X_0 \to Y$ and $A : X \to Y_1$, respectively.

Recall from Section 1.2.3 that if X, Y and Z are linear spaces and $A \in \mathcal{L}(X,Y)$ and $B \in \mathcal{L}(Y,Z)$, then the linear operator $BA : X \to Z$ is defined as the composition of B and A, i.e.,

$$(BA)(x) = B\,(Ax) \quad \forall x \in X.$$

In case X, Y, Z are normed linear spaces, and $A \in \mathcal{B}(X,Y)$ and $B \in \mathcal{B}(Y,Z)$, then it is seen that $BA \in \mathcal{B}(X,Z)$. In fact, we have

$$\|BA\| \le \|B\|\,\|A\|.$$

Transpose of an operator

If $A \in \mathcal{B}(X,Y)$, then there is a natural way of associating it with an operator in $\mathcal{B}(Y',X')$. For $f \in Y'$ and $x \in X$, we define

$$(A'f)(x) = f(Ax).$$

Note that

$$|(A'f)(x)| = |f(Ax)| \le \|f\|\,\|Ax\| \le \|f\|\,\|A\|\,\|x\|$$

for all $x \in X$ and $f \in Y'$ so that $A'f \in X'$ for all $f \in Y'$. Clearly, $A' : Y' \to X'$ is a linear operator. It also follows that $\|A'f\| \le \|f\|\,\|A\|$, and hence,

$$A' \in \mathcal{B}(Y',X'), \quad \|A'\| \le \|A\|.$$

Later, as a consequence of *Hahn-Banach extension theorem*, we shall see that $\|A'\| = \|A\|$.

The operator A' defined above is called the **transpose** of A.

Exercise 3.4 Suppose X and Y are finite dimensional spaces with ordered bases U and V, respectively and $A : X \to Y$ is a linear operator. If U' and V' are the dual bases of X' and Y' associated with U and V, respectively, then show that the matrix representation $[A']_{V',U'}$ of A' is the *transpose* of the matrix representation $[A]_{U,V}$ of A. □

Recall that, for a linear operator $A : X \to Y$, if there exists $c > 0$ such that $\|Ax\| \le c\|x\|$ for all $x \in X$, then $A \in \mathcal{B}(X,Y)$ and $\|A\| \le c$. In case we can find an $x_0 \ne 0$ such that $\|Ax_0\| = c\|x_0\|$, then we have

$$c\|x_0\| = \|Ax_0\| \le \|A\|\,\|x_0\|$$

and hence,

$$\|A\| = c.$$

We observe that the existence of such x_0 is equivalent to the fact that the (continuous) function $x \mapsto \|Ax\|$ defined on the closed unit ball attains its supremum at some element.

Sometimes the above procedure helps us to compute the norm of an operator in $\mathcal{B}(X,Y)$. But not always, since the closed unit ball in an infinite dimensional space is not compact (cf. Theorem 2.39).

Let us give some examples wherein the norm can be computed exactly.

NOTATION: If $\alpha \in \mathbb{K}$, and $f : \Omega \to \mathbb{K}$ is a function defined on a set Ω, then $\operatorname{sgn}(\alpha) \in \mathbb{K}$ and the function $\operatorname{sgn}(f) : \Omega \to \mathbb{K}$ are defined by

$$\operatorname{sgn}(\alpha) = \begin{cases} |\alpha|/\alpha & \text{if } \alpha \neq 0, \\ 0 & \text{if } \alpha = 0, \end{cases}$$

and $\operatorname{sgn}(f)(t) := \operatorname{sgn}(f(t))$ for all $t \in \Omega$, respectively.

EXAMPLE 3.3 (i) Consider the operators f, g, A as in Example 3.1(iv): Let $X = C[a,b]$ with $\|\cdot\|_\infty$ and

$$f(x) = \int_a^b x(t)\,dt, \quad g(x) = x(t_0), \quad (Ax)(s) = \int_a^s x(t)\,dt$$

for all $x \in X$ and for some $t_0 \in [a,b]$. We have seen that

$$|f(x)| \leq (b-a)\|x\|_\infty, \quad |g(x)| \leq \|x\|_\infty, \quad \|Ax\|_\infty \leq (b-a)\|x\|_\infty$$

for all $x \in X$. If we take $x_0(t) = 1$ for every $t \in [a,b]$, we get

$$|f(x_0)| = (b-a)\|x_0\|_\infty, \quad |g(x_0)| = \|x_0\|_\infty, \quad \|Ax_0\|_\infty = (b-a)\|x_0\|_\infty.$$

Hence, it follows that

$$\|f\| = b-a, \qquad \|g\| = 1, \qquad \|A\| = b-a.$$

(ii) Let $X = C^1[0,1]$ with norm $x \mapsto \|x\|_* = \|x\|_\infty + \|x'\|_\infty$, $Y = C[0,1]$ with $\|\cdot\|_\infty$ and $A : X \to Y$ be defined by

$$Ax = x', \qquad x \in X.$$

Clearly,

$$\|Ax\| = \|x'\|_\infty \leq \|x\|_* \quad \forall x \in X.$$

Therefore, A is a bounded linear operator and $\|A\| \leq 1$. Now, consider the sequence (x_n) defined by

$$x_n(t) = \frac{t^n}{n+1}, \qquad t \in [0,1],\ n \in \mathbb{N}.$$

Then we see that $\|x_n\|_* = 1$ and

$$\frac{n}{n+1} = \|Ax_n\|_\infty \le \|A\|\ \|x_n\|_* = \|A\| \le 1 \quad \forall n \in \mathbb{N}.$$

Letting $n \to \infty$, we obtain $\|A\| = 1$.

In this example, it is easily seen that there is no nonzero x_0 such that $\|Ax_0\|_\infty = \|x_0\|_*$. This is not surprising since the closed unit ball in the space X is not compact.

(iii) Let X be an inner product space and $u \in X$. Let $f_u : X \to \mathbb{K}$ be defined by

$$f_u(x) = \langle x, u \rangle, \qquad x \in X.$$

We have seen in Example 3.1(vii) that f_u is a continuous linear functional and $|f_u(x)| \le \|u\|\ \|x\|$ for all $x \in X$. Thus, $\|f_u\| \le \|u\|$. Since

$$|f_u(u)| = |\langle u, u \rangle| = \|u\|^2,$$

it follows that $\|f_u\| = \|u\|$.

From the above observation it also follows that the map $u \mapsto f_u$ from X into X' is one-one. We shall show, in Theorem 3.9, that if X is a Hilbert space, then the above map is onto as well.

(iv) Let X be an inner product space and $P : X \to X$ be an orthogonal projection. We have seen in Example 3.1(vi) that $\|Px\| \le \|x\|$ for all $x \in X$. Hence, $\|P\| \le 1$. Now we show that

$$\|P\| = \begin{cases} 0 & \text{if } P = 0, \\ 1 & \text{if } P \ne 0. \end{cases}$$

Clearly, $P = 0$ implies $\|P\| = 0$, and $P = I$ implies $\|Px\| = \|x\|$ for all $x \in X$ so that $\|P\| = 1$. Now let $0 \ne P \ne I$. Then there exists $0 \ne x_0 \in R(P)$, and we have

$$\|Px_0\| = \|x_0\|.$$

This, together with the fact that $\|P\| \le 1$, implies that $\|P\| = 1$.

(v) Let (a_{ij}) be an $m \times n$ matrix and $A : \mathbb{K}^n \to \mathbb{K}^m$ be defined by

$$(Ax)(i) = \sum_{j=1}^{n} a_{ij} x(j), \qquad x \in X;\ i \in \{1, \ldots, m\}.$$

We have seen in Example 3.1(i) that for every $x \in \mathbb{K}^n$,

$$\|Ax\|_1 \leq a\,\|x\|_1, \qquad a = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|.$$

We show that there is a nonzero $x_0 \in X$ such that $\|Ax_0\|_1 = a\|x_0\|_1$, which would imply that the norm of $A : \ell^1(n) \to \ell^1(m)$ is a. For this, let $k \in \{1, \ldots, n\}$ be such that

$$a = \sum_{i=1}^{m} |a_{ik}|,$$

and let $x_0 \in \mathbb{K}^n$ be defined by $x_0(j) = \delta_{jk}$ for $j = 1, \ldots, n$. Then it follows that $\|x_0\|_1 = 1$ and

$$\|Ax_0\|_1 = \sum_{i=1}^{m} |a_{ik}| = a\|x_0\|_1.$$

Also, from Example 3.1(i), we have, for every $x \in \mathbb{K}^n$,

$$\|Ax\|_\infty \leq b\,\|x\|_\infty, \qquad b = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|.$$

We show that the operator $A : \ell^\infty(n) \to \ell^\infty(m)$ is of norm b, by showing that there is a nonzero $u_0 \in X$ such that $\|Au_0\|_\infty \geq b\|u_0\|_\infty$. For this, let $\ell \in \{1, \ldots, m\}$ be such that $b = \sum_{j=1}^{n} |a_{\ell j}|$, and let $u_0 \in \mathbb{K}^n$ be defined by $u_0(j) = \operatorname{sgn}(a_{\ell j})$ for all $j = 1, \ldots, n$. Then it follows that $\|u_0\|_\infty = 1$ and

$$\|Au_0\|_\infty \geq |(Au_0)(\ell)| = \sum_{j=1}^{n} |a_{\ell j}| = b\,\|u_0\|_\infty.$$

We already know that $\|Au_0\|_\infty \leq b\,\|u_0\|_\infty$. Thus, $\|Au_0\|_\infty = b\,\|u_0\|_\infty$.

(vi) Let A be as in Example 3.1(ii). We saw there that $A \in \mathcal{B}(\ell^1)$ and for every $x \in \ell^1$,

$$\|Ax\|_1 \leq \alpha\|x\|_1, \qquad \alpha = \sup_j \sum_{i=1}^{\infty} |a_{ij}| < \infty.$$

Thus, $\|A\| \le \alpha$. Now we show that $\|A\| \ge \alpha$ so that $\|A\| = \alpha$. For this, let (j_n) be a sequence in $\mathbb{N}$ such that

$$\alpha_n := \sum_{i=1}^{\infty} |a_{ij_n}| \to \alpha \quad \text{as} \quad n \to \infty.$$

Let $x_n(j) = \delta_{jj_n}$ for $j \in \mathbb{N}$. Then we have $\|x_n\|_1 = 1$, and

$$\|Ax_n\|_1 = \sum_{i=1}^{\infty} |a_{ij_n}| = \alpha_n$$

so that $\alpha_n \le \|A\|$ for all $n \in \mathbb{N}$. Hence, taking limit as $n \to \infty$, $\alpha \le \|A\|$.

(vii) Let A be as in Example 3.1(iii). We have seen there that $A \in \mathcal{B}(\ell^\infty)$ and for every $x \in \ell^\infty$,

$$\|Ax\|_\infty \le \beta \|x\|_\infty, \qquad \beta = \sup_i \sum_{j=1}^{\infty} |a_{ij}|.$$

Thus, $\|A\| \le \beta$. Now we show that $\|A\| \ge \beta$ so that $\|A\| = \beta$. For this, let (i_n) be a sequence in $\mathbb{N}$ such that

$$\beta_n := \sum_{j=1}^{\infty} |a_{i_n j}| \to \beta \quad \text{as} \quad n \to \infty.$$

Let $u_n(j) = \operatorname{sgn}(a_{i_n j})$ for all $j \in \mathbb{N}$. Then it follows that $\|u_n\|_\infty \le 1$ and

$$|(Au_n)(i_n)| = \sum_{j=1}^{\infty} |a_{i_n j}| = \beta_n$$

so that $\beta_n \le \|Au_n\|_\infty \le \|A\|$ for all $n \in \mathbb{N}$. Hence, it follows that $\beta \le \|A\|$.

Exercise 3.5 Let (λ_n) be a bounded sequence of scalars. For $x \in \ell^\infty$, let

$$(Ax)(i) = \lambda_i x(i), \quad i \in \mathbb{N}.$$

Show that, for any $p \in [1, \infty]$, $Ax \in \ell^p$ for all $x \in \ell^p$, and $A : \ell^p \to \ell^p$ is a bounded linear operator with

$$\|A\| = \sup_{n \in \mathbb{N}} |\lambda_n|. \quad \square$$

EXAMPLE 3.3 (*cont.*)(viii) Consider the operator A in Example 3.1(v). We have seen that

$$\|Ax\| \le \left(\sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt\right)\|x\|_\infty \quad \forall x\in X.$$

Hence,

$$\|A\| \le \sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt.$$

Now we show that

$$\|A\| = \sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt.$$

Since the function $s \mapsto \int_a^b |k(s,t)|\,dt$ is continuous on the compact interval $[a,b]$, there exists $s_0\in[a,b]$ such that

$$\sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt = \int_a^b |k(s_0,t)|\,dt.$$

Then for every $\varepsilon > 0$, we have

$$\begin{aligned}
\int_a^b (|k(s_0,t)|-\varepsilon)\,dt &= \int_a^b \frac{|k(s_0,t)|^2-\varepsilon^2}{|k(s_0,t)|+\varepsilon}\,dt\\
&\le \int_a^b k(s_0,t)\frac{\overline{k(s_0,t)}}{|k(s_0,t)|+\varepsilon}\,dt\\
&= |(Ax_\varepsilon)(s_0)|\\
&\le \|A\|,
\end{aligned}$$

where $x_\varepsilon(t) = \overline{k(s_0,t)}/\left(|k(s_0,t)|+\varepsilon\right)$. Hence,

$$\int_a^b |k(s_0,t)|\,dt \le \|A\| + \varepsilon(b-a) \quad \forall \varepsilon>0.$$

Letting $\varepsilon\to 0$, we get

$$\sup_{a\le s\le b}\int_a^b |k(s,t)|\,dt = \int_a^b |k(s_0,t)|\,dt \le \|A\|.$$

Note that the above example includes the example of f in (i), obtained by taking $k(s,t)=1$ for every $s,t\in[a,b]$. In that example, it was easy to find $x_0 \neq 0$ such that $|f(x_0)| = \|f\| \, \|x_0\|$. This may not be possible if we change the space, as the following example shows.

(ix) Let $X = \{x \in C[0,1] : x(0) = 0\}$ with $\|\cdot\|_\infty$, and

$$f(x) = \int_0^1 x(t)\,dt, \qquad x \in X.$$

Note that X is a closed subspace of $C[0,1]$ with $\|\cdot\|_\infty$. Clearly,

$$|f(x)| \leq \|x\|_\infty \quad \forall x \in X,$$

so that $\|f\| \leq 1$. It is seen that there is no $x_0 \in X$ such that $\|x_0\|_\infty = 1$ and $|f(x_0)| = 1$. However, we can still show that $\|f\| = 1$. For this, consider the sequence (x_n) defined by

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

for $t \in [0,1]$ and $n \in \mathbb{N}$. Then we have $x_n \in X$, $\|x_n\|_\infty = 1$, and

$$\|f\| \geq |f(x_n)| = 1 - \frac{1}{2n} \quad \forall n \in \mathbb{N}.$$

Letting $n \to \infty$, it follows that $\|f\| \geq 1$. Thus, $\|f\| = 1$.

(x) Recall from Example 3.1(viii) that the Lagrange interpolatory projection $L_n : C[a,b] \to C[a,b]$ associated with distinct points $t_{1,n}, \ldots, t_{n,n}$ in $[a,b]$ is given by

$$L_n x = \sum_{j=1}^{n} x(t_{j,n}) \ell_{j,n}, \qquad x \in C[a,b],$$

where

$$\ell_{j,n}(t) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{t - t_{i,n}}{t_{j,n} - t_{i,n}}, \qquad t \in [a,b];\ j = 1, \ldots, n.$$

We have seen that, with respect to the norm $\|\cdot\|_\infty$ on $C[a,b]$, L_n is a bounded linear operator on $C[a,b]$ and

$$|(L_n x)(t)| \leq \left(\sup_{a \leq t \leq b} \sum_{j=1}^{n} |\ell_{j,n}(t)| \right) \|x\|_\infty \quad \forall x \in C[a,b].$$

Thus,

$$\|L_n\| \le \sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)|.$$

We now show that $\|L_n\| = \sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)|$. Since each $\ell_{j,n}$ is continuous on $[a,b]$, there exists $\tau \in [a,b]$ such that

$$\sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)| = \sum_{j=1}^{n} |\ell_{j,n}(\tau)|.$$

Choose $x_0 \in C[a,b]$ such that

$$\|x_0\|_\infty = 1, \qquad x_0(a) = x_0(t_{1,n}), \qquad x_0(b) = x_0(t_{n,n}),$$

$$x_0(t_{j,n}) = \operatorname{sgn}(\ell_{j,n}(\tau)), \qquad j \in \{1,\ldots,n\}.$$

Then it follows that

$$(L_n x)(\tau) = \sum_{j=1}^{n} |\ell_{j,n}(\tau)|$$

so that we have

$$\sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)| = |(L_n x)(\tau)| \le \|L_n x\|_\infty \le \|L_n\|.$$

Thus,

$$\|L_n\| = \sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)|.$$

(xi) For $t_1,\ldots,t_k$ in $[a,b]$ and scalars $w_1,\ldots,w_k$, let

$$Q(x) = \sum_{j=1}^{k} x(t_j)w_j, \quad x \in C[a,b].$$

Then it follows that $Q : C[a,b] \to \mathbb{K}$ is a linear functional. We show that it is continuous with respect to the norm $\|\cdot\|_\infty$ on $C[a,b]$ and $\|Q\| = \sum_{j=1}^{k} |w_j|$.

Note that for every $x \in C[a,b]$,

$$|Q(x)| \le \sum_{j=1}^{k} |x(t_j)|\,|w_j| \le \left(\sum_{j=1}^{k} |w_j|\right) \|x\|_\infty$$

so that Q is continuous. Now, let $x_0 \in C[a,b]$ be such that

$$\|x_0\|_\infty = 1, \; x_0(t_j) = \operatorname{sgn}(w_j), \qquad j \in \{1,\ldots,k\}.$$

Then we see that $|Q(x_0)| = \sum_{j=1}^k |w_j|$ so that $\|Q\| = \sum_{j=1}^k |w_j|$.

The above functional Q is called a **quadrature formula** for approximating the integral

$$f(x) := \int_a^b x(t)\,dt.$$

The points $t_1,\ldots,t_k$ are the *nodes* and the scalars $w_1,\ldots,w_k$ are the *weights* of the quadrature formula Q. Many of the standard numerical integration procedures such as composite rectangular rule, composite trapezoidal rule and composite Simpson's rule can be put in the above general form. For example, taking $t_1,\ldots,t_k$ to be equidistant points and $w_1,\ldots,w_k$ to be equal to the length of the sub-interval, i.e.,

$$t_j = a + \frac{j}{n}(b-a), \qquad w_j = \frac{b-a}{n}, \qquad j = 1,\ldots,n,$$

the quadrature formula which is a composite rectangular rule represents the total area of the rectangles over the sub-intervals $[t_{i-1}, t_i]$, $i = 1,\ldots,k$.

(xii) Let $A : C[a,b] \to C[a,b]$ be defined by

$$(Ax)(t) = u(t)x(t), \qquad x \in C[a,b],\, t \in [a,b],$$

where u is a given function in $C[a,b]$. We have seen in Example 3.1(ix) that A is a bounded operator with respect to the norm $\|\cdot\|_\infty$ on $C[a,b]$, and $\|Ax\|_\infty \le \|u\|_\infty \|x\|_\infty$ for all $x \in C[a,b]$. Hence, $\|A\| \le \|u\|_\infty$. Taking $x_0(t) = 1$ for all $t \in [a,b]$, we have $\|x_0\|_\infty = 1$, $\|Ax_0\|_\infty = \|u\|_\infty$ so that $\|A\| = \|u\|_\infty$.

Some estimates for norms of certain operators

In Example 3.3, we computed norms of certain operators. In the following example we consider a few cases in which we give only some estimates for the norm of the operators.

EXAMPLE 3.4 (i) Let $a_{ij} \in \mathbb{K}$ for $(i,j) \in \mathbb{N} \times \mathbb{N}$, and let

$$\alpha_{p,q} = \sum_{i=1}^\infty \Big(\sum_{j=1}^\infty |a_{ij}|^q\Big)^{p/q}, \quad \beta = \sup_i \sum_{j=1}^\infty |a_{ij}|, \quad \gamma = \sup_j \sum_{i=1}^\infty |a_{ij}|,$$

where $1 < p < \infty$ and q is the conjugate exponent of p. We show that if

$$\min\{\alpha_{p,q}^{1/p},\ \beta^{1/q}\gamma^{1/p}\} < \infty,$$

then

$$(Ax)(i) = \sum_{j=1}^{\infty} a_{ij}x(j), \qquad i \in \mathbb{N},$$

defines a bounded operator on ℓ^p and

$$\|A\| \le \min\{\alpha_{p,q}^{1/p},\ \beta^{1/q}\gamma^{1/p}\}.$$

By Hölder's inequality, for every $x \in \ell^p$, we have

$$\sum_{j=1}^{\infty} |a_{ij}|\ |x(j)| \le \left(\sum_{j=1}^{\infty} |a_{ij}|^{1/q}\right)^{1/q} \|x\|_p \le \alpha_{pq}^{1/p}\|x\|_p$$

so that

$$\sum_{i=1}^{\infty} |(Ax)(i)|^p \le \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} |a_{ij}|\ |x(j)|\right)^p \le \alpha_{p,q}\|x\|_p^p.$$

Again, by using Hölder's inequality, for every $x \in \ell^p$, we have

$$\begin{aligned}
\sum_{j=1}^{\infty} |a_{ij}|\ |x(j)| &= \sum_{j=1}^{\infty} |a_{ij}|^{1/q}|a_{ij}|^{1/p}|x(j)| \\
&\le \left(\sum_{j=1}^{\infty} |a_{ij}|\right)^{1/q}\left(\sum_{j=1}^{\infty} |a_{ij}|\ |x(j)|^p\right)^{1/p}
\end{aligned}$$

so that

$$\begin{aligned}
\left(\sum_{j=1}^{\infty} |a_{ij}|\ |x(j)|\right)^p &\le \left(\sum_{j=1}^{\infty} |a_{ij}|\right)^{p/q}\left(\sum_{j=1}^{\infty} |a_{ij}||x(j)|^p\right) \\
&\le \beta^{p/q}\sum_{j=1}^{\infty} |a_{ij}||x(j)|^p.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\sum_{i=1}^{\infty} |(Ax)(i)|^p &\le \beta^{p/q}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}||x(j)|^p \\
&= \beta^{p/q}\sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} |a_{ij}|\right)|x(j)|^p \\
&\le \beta^{p/q}\gamma\|x\|_p^p.
\end{aligned}$$

Thus, $Ax \in \ell^p$ for every $x \in \ell^p$ whenever $\min\{\alpha_{p,q}^{1/p},\ \beta^{1/q}\gamma^{1/p}\} < \infty$ and, in that case, $A : \ell^p \to \ell^p$ is a bounded linear operator and

$$\|A\| \le \min\{\alpha_{p,q}^{1/p},\ \beta^{1/q}\gamma^{1/p}\}.$$

A continuous analogue of the above example is the following (We leave the verification to the reader):

(ii) Let $J = [a,b]$, and $X = L^p(J)$ with $1 < p < \infty$. Let $k(\cdot,\cdot) \in C(J \times J)$, and let

$$(Ax)(s) = \int_a^b k(s,t)x(t)\,dt, \quad x \in X.$$

We have already seen that $A : X \to X$ is a bounded linear operator. Then, as in the above example, with summations replaced by integrals, it can be seen that

$$\|A\| \le \min\{\alpha_{p,q}^{1/p},\ \beta^{1/q}\gamma^{1/p}\},$$

where

$$\begin{aligned} \alpha_{p,q} &= \int_a^b \left(\int_a^b |k(s,t)|^q\,dt\right)^{p/q} ds, \\ \beta &= \sup_{s\in[a,b]} \int_a^b |k(s,t)|\,dt, \\ \gamma &= \sup_{t\in[a,b]} \int_a^b |k(s,t)|ds. \end{aligned}$$

(iii) Let $X = L^2[a,b]$, and for $u \in L^\infty[a,b]$, let $A : X \to X$ be defined by $Ax = u.x$, $x \in X$. Then, for every $x \in X$, we have

$$\|Ax\|_2^2 = \int_a^b |Ax|^2\,d\mu = \int_a^b |ux|^2\,d\mu \le \|u\|_\infty \|x\|_2^2$$

so that $A \in \mathcal{B}(X)$ and $\|A\| \le \|u\|_\infty$.

Exercise 3.6 (i) Let $a_{ij} \in \mathbb{K}$ for $(i,j) \in \mathbb{N}\times\mathbb{N}$. For $x \in \ell^\infty$, let

$$(Ax)(i) = \sum_{j=1}^{\infty} a_{ij}x(j), \qquad i \in \mathbb{N}.$$

(a) Let $1 \leq p \leq \infty$ and $1 \leq r \leq \infty$, and assume that $\alpha_i := \sum_{j=1}^{\infty} |a_{ij}| < \infty$ for all $i \in \mathbb{N}$. Show that if $u = (\alpha_1, \alpha_2, \ldots) \in \ell^r$, then $Ax \in \ell^r$ for all $x \in \ell^p$ and

$$\|Ax\|_r \leq \|u\|_r \|x\|_p \quad \forall x \in \ell^p.$$

(b) Let $1 < p < \infty$ and assume that

$$u_q(i) := \left(\sum_{j=1}^{\infty} |a_{ij}|^q \right)^{1/q} < \infty \quad \forall i \in \mathbb{N}.$$

Show that, if $1 \leq r \infty$ and $u_q \in \ell^r$, then

$$\|Ax\|_r \leq \|u_q\|_r \|x\|_p \quad \forall x \in \ell^p.$$

(c) Let $1 < p \leq \infty$ and $1 \leq r \leq \infty$. Let

$$c_{r,q} = \begin{cases} \max\limits_i \left(\sum_{j=1}^{\infty} |a_{ij}|^q \right)^{1/q} & \text{if } r = \infty \\ \left[\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} |a_{ij}|^q \right)^{r/q} \right]^{1/r} & \text{if } 1 \leq r < \infty. \end{cases}$$

Show that if $c_{r,q} < \infty$, then

$$\|Ax\|_r \leq c_{r,q} \|x\|_p \quad \forall x \in \ell^p.$$

(ii) Let $k(\cdot, \cdot) \in C([a, b] \times [a, b])$, and for $x \in L^1[a, b]$, define

$$(Ax)(s) = \int_a^b k(s, t) x(t)\, dt \quad s \in [a, b].$$

(a) For $1 \leq p \leq \infty$ and $1 \leq r \leq \infty$, show that

$$\|Ax\|_r \leq \|\phi\|_r \|x\|_p \quad \forall x \in L^p[a, b],$$

where $\phi(s) = \int_a^b |k(s, t)|\, dt$, $s \in [a, b]$.

(b) For $1 < p < \infty$ and $1 \leq r \leq \infty$, show that

$$\|Ax\|_r \leq \|\psi_q\|_r \|x\|_p \quad \forall x \in L^p[a, b],$$

where $\psi_q(s) = \left(\int_a^b |k(s,t)|^q \, dt \right)^{1/q}$, $s \in [a,b]$.

(c) For $1 < p \leq \infty$ and $1 \leq r \leq \infty$, show that

$$\|Ax\|_r \leq c_{r,q}\|x\|_p \quad \forall x \in L^p[a,b],$$

where

$$c_{r,q} = \begin{cases} \max_i \left(\int_a^b |k(s,t)|^q \, dt \right)^{1/q} & \text{if } r = \infty \\ \left[\int_a^b \left(\int_a^b |k(s,t)|^q \, dt \right)^{r/q} ds \right]^{1/r} & \text{if } 1 \leq r < \infty. \end{cases} \quad \square$$

A Characterization of Bounded Linear Functionals

Let $A : X \to Y$ be a linear operator between normed linear spaces X and Y. It can be easily seen that if $A \in \mathcal{B}(X,Y)$, then its null space $N(A)$ is a closed subspace of X. But the converse need not be true. For example, Let $X = C^1[a,b]$, $Y = C[a,b]$ both with $\|\cdot\|_\infty$. Let $A : X \to Y$ be defined by $Ax = x'$, $x \in X$. In this case, we see that $N(A)$ is the set of all constant functions, a one-dimensional subspace of X. Therefore, $N(A)$ is a closed subspace. But we know that A is not a bounded operator.

Now we show that such a situation will not arise for linear functionals, i.e., if $Y = \mathbb{K}$.

Theorem 3.8 *Let X be a normed linear space and $f : X \to \mathbb{K}$ be a nonzero linear functional with its null space $N(f)$ closed in X. Then f is continuous, and for any $x_0 \in X \setminus N(f)$,*

$$\|f\| = \frac{|f(x_0)|}{\text{dist}\,(x_0, N(f))}.$$

Proof. Let $x_0 \in X$ be such that $f(x_0) \neq 0$. Then, for every $x \in X$, $x = y + \alpha x_0$, where

$$y = x - \frac{f(x)}{f(x_0)}x_0, \qquad \alpha = \frac{f(x)}{f(x_0)}.$$

Since $y \in N(f)$, we have

$$\text{dist}\,(x, N(f)) = \text{dist}\,(\alpha x_0, N(f)) = \left|\frac{f(x)}{f(x_0)}\right| \text{dist}\,(x_0, N(f)),$$

and since $x_0 \notin N(f)$, dist $(x_0, N(f)) > 0$. Hence,

$$|f(x)| = \frac{|f(x_0)|}{\text{dist }(x_0, N(f))} \text{dist }(x, N(f)) \leq \frac{|f(x_0)|}{\text{dist }(x_0, N(f))}\|x\|.$$

Thus, f is continuous and

$$\|f\| \leq \frac{|f(x_0)|}{\text{dist }(x_0, N(f))}.$$

Moreover, for every $u \in N(f)$,

$$|f(x_0)| = |f(x_0 - u)| \leq \|f\| \, \|x_0 - u\|$$

so that by taking infimum over all $u \in N(f)$, we get

$$|f(x_0)| \leq \|f\| \text{ dist }(x_0, N(f)).$$

Hence,

$$\|f\| \geq \frac{|f(x_0)|}{\text{dist }(x_0, N(f))}.$$

Thus, we have

$$\|f\| = \frac{|f(x_0)|}{\text{dist }(x_0, N(f))}. \quad \blacksquare$$

Exercise 3.7 If f is a nonzero continuous linear functional on a normed linear space X, then show that $\|f\| = 1/d$, where $d = \inf\{\|x\| : f(x) = 1\}$. $\square$

Now, we illustrate Theorem 3.8 by examples.

EXAMPLE 3.5 (i) Let $X = C^1[0,1]$ with $\|\cdot\|_\infty$, and let $f : X \to \mathbb{K}$ be defined by $f(x) = x'(1)$, $x \in X$. From Example 3.2(i), we know that f is a discontinuous linear functional on X. By Theorem 3.8, we know that $N(f)$ is not closed. To see this explicitly, consider

$$x(t) = t \quad \forall t \in [0,1], \qquad x_n(t) = t - \frac{t^n}{n} \quad \forall t \in [0,1]\, \forall n \in \mathbb{N}.$$

Observe that $f(x_n) = 0$ for all $n \in \mathbb{N}$ and $\|x_n - x\|_\infty = 1/n \to 0$ as $n \to \infty$, but $f(x) \neq 0$.

(ii) Let $X = c_{00}$ with $\|\cdot\|_\infty$, and for $x \in X$, let $f(x) = \sum_{j=1}^{\infty} x(j)$. From Example 3.2(iii), we know that f is a discontinuous linear functional on X. To see explicitly that its null space is not closed, consider

$$x_n = \left(-1, \frac{1}{n}, \ldots, \frac{1}{n}, 0, 0, \ldots\right), \qquad n \in \mathbb{N},$$

with $1/n$ occurring n times following -1. Then it is seen that $x_n \in N(f)$ for all $n \in \mathbb{N}$. Also, with $x = (-1, 0, 0, \ldots)$,

$$\|x_n - x\|_\infty = 1/n \to 0 \quad \text{as} \quad n \to \infty,$$

but $x \notin N(f)$.

Remark 3.2 (a) Theorem 3.8 shows that closed hyperspaces in a normed linear space are precisely the null spaces of continuous linear functionals.

(b) We have seen examples of discontinuous linear functionals ((i) and (iii)) in Example 3.2). In those examples, the space under consideration are not Banach spaces. One may ask whether discontinuous linear functionals exist on infinite dimensional Banach spaces. The answer is in affirmative.

In fact, discontinuous linear functionals exist on every infinite dimensional normed linear space. To see this, let X be an infinite dimensional normed linear space and let E_0 be a denumerable linearly independent subset of X, say $E_0 = \{u_1, u_2, \ldots\}$. Let E be a basis of X such that $E_0 \subseteq E$. Assume, without loss of generality that $\|u_n\| = 1$ for every $n \in \mathbb{N}$. Clearly, every $x \in X$ is of the form

$$x = \sum_{j=1}^{n} \alpha_j u_j + z$$

for some scalars $\alpha_1, \ldots, \alpha_n$ and $z \in \operatorname{span}(E \setminus E_0)$. For x having the above representation, define

$$f(x) = \sum_{i=1}^{n} j\alpha_j.$$

Then it can be easily seen that f defines a linear functional on X. Note that, for each $k \in \mathbb{N}$,

$$x_n^{(k)} := \frac{u_k}{k} - \frac{u_n}{n} \in N(f) \quad \forall n \in \mathbb{N}, \quad \text{and} \quad x_n^{(k)} \to \frac{u_k}{k} \quad \text{as} \quad n \to \infty.$$

But $\dfrac{u_k}{k} \notin N(f)$. Thus, f is not continuous and $N(f)$ is not a closed subspace of X.

Exercise 3.8 (i) Let X be a normed linear space and $f \in X'$. Show that, if $N(f)$ is a nonzero complete subspace, then X is also a complete space, i.e., a Banach space. $\square$

3.3 Riesz Representation Theorem

Recall from Example 3.3(iii) that if X is an inner product space and $u \in X$, then the map $f_u : X \to \mathbb{K}$ defined by

$$f_u(x) = \langle x, u \rangle, \qquad x \in X,$$

is a continuous linear functional on X, and $\|f_u\| = \|u\|$.

Does every continuous linear functional on X have the above form?

If the space is finite dimensional, then we can give an affirmative answer rather easily. To see this, suppose X is a finite dimensional inner product space and $\{u_1, \ldots, u_n\}$ is a basis which is also an orthonormal set. Let $f : X \to \mathbb{K}$ be a linear functional on X. Since

$$x = \sum_{j=1}^{n} \langle x, u_j \rangle u_j$$

for every $x \in X$, we have

$$f(x) = \sum_{j=1}^{n} \langle x, u_j \rangle f(u_j) = \sum_{j=1}^{n} \langle x, \overline{f(u_j)} u_j \rangle.$$

Thus,

$$f(x) = \langle x, u \rangle \quad \text{with} \quad u = \sum_{j=1}^{n} \overline{f(u_j)}\, u_j.$$

What about the situation in infinite dimensional inner product spaces? Unfortunately, in this case, we may not be able to find a $u \in X$ such that

$$f(x) = \langle x, u \rangle \quad \forall x \in X.$$

For example, consider the space $X = c_{00}$ with the inner product

$$\langle x, y \rangle = \sum_{j=1}^{\infty} x(j)\overline{y(j)}, \qquad x, y \in X.$$

We know that for every $u \in X$, there exists $N \in \mathbb{N}$ such that

$$\langle u, e_n \rangle = u(n) = 0 \quad \forall n \geq N,$$

where $e_n(j) = \delta_{jn}$ for $j, n \in \mathbb{N}$. Therefore, if f is a continuous linear functional on X such that $f(e_n) \neq 0$ for every n, then there is no

$u \in X$ satisfying $f(e_n) = \langle e_n, u\rangle$ for all $n \in \mathbb{N}$. For instance, let (λ_n) be a sequence of scalars with $\sum_{j=1}^{\infty} |\lambda_j|^2 < \infty$, and define

$$f(x) = \sum_{j=1}^{\infty} \lambda_j x(j), \qquad x \in X.$$

Clearly, $f : X \to \mathbb{K}$ is a linear functional. Also, by using Schwarz inequality, we have

$$|f(x)| \le \sum_{j=1}^{\infty} |\lambda_j|\,|x(j)| \le \Big(\sum_{j=1}^{\infty} |\lambda_j|^2\Big)^{1/2} \|x\| \quad \forall\, x \in X.$$

Thus, f is a continuous linear functional on X, and $f(e_n) = \lambda_n$ for all $n \in \mathbb{N}$. Hence, if (λ_n) is such that $\lambda_n \ne 0$ for infinitely many n's, then there cannot exist $u \in c_{00}$ such that $f(x) = \langle x, u\rangle$ for all $x \in X$. For example, we may take $\lambda_n = 1/n^p$ for any $p \ge 1$.

In fact, what we have shown is that if $(\lambda_n) \in \ell^2 \setminus c_{00}$, then the functional f defined by

$$f(x) = \sum_{j=1}^{\infty} \lambda_j x(j), \qquad x \in X,$$

belongs to X', but there is no $u \in X$ such that $f(x) = \langle x, u\rangle$ for all $x \in X$.

However, if X is a Hilbert space, then we do have an affirmative answer.

Theorem 3.9 (Riesz representation theorem) *If X is a Hilbert space, then for every continuous linear functional f on X, there exists a unique $v \in X$ such that*

$$f(x) = \langle x, v\rangle \quad \forall\, x \in X.$$

Moreover, $\|v\| = \|f\|$.

Proof. Let X be a Hilbert space, and let $f : X \to \mathbb{K}$ be a continuous linear functional on X. If $f = 0$, then $v = 0$ satisfies $f(x) = \langle x, v\rangle$ for all $x \in X$.

Now assume that $f \ne 0$. Then $N(f)$ is a proper closed subspace of X. Hence, by projection theorem (Theorem 2.47), $N(f)^\perp \ne \{0\}$.

Let $x_0 \in N(f)^{\perp}$ be such that $\|x_0\| = 1$. Note that, for every $x \in X$, $x = y + \alpha x_0$, where

$$y = x - \frac{f(x)}{f(x_0)} x_0 \in N(f), \qquad \alpha = \frac{f(x)}{f(x_0)}.$$

Since

$$\langle x, x_0 \rangle = \langle y, x_0 \rangle + \alpha \langle x_0, x_0 \rangle = \alpha,$$

it follows that

$$f(x) = \alpha f(x_0) = \langle x, x_0 \rangle f(x_0) = \langle x, \overline{f(x_0)} x_0 \rangle$$

for all $x \in X$. Thus, taking $v = \overline{f(x_0)} x_0$,

$$f(x) = \langle x, v \rangle \quad \forall x \in X.$$

The other properties of u can be easily verified. ∎

Remark 3.3 From the discussion preceding the theorem, it is clear that the conclusion in the Riesz representation theorem need not hold if the space is not complete.

If we observe the proof of Theorem 3.9 closely, we find that the completeness of the space is used only to make use of Theorem 2.47. Thus, essentially, what we have used is the completeness of $N(f)$. In view of this, one may ask whether it is possible to have continuous linear functional f on an incomplete inner product space such that $N(f)$ is complete so that the conclusion of Theorem 3.9 would be true for all such f. Unfortunately, the answer is in the negative (see Exercise 3.8).

After developing some more theory on Hilbert spaces, it is possible to give a specific expression for the representer v of $f \in X'$ in terms of an *orthonormal basis* of the space, analogous to the situation in finite dimensional inner product spaces (Chapter 4, Problem 12).

By the Riesz representation theorem, it is apparent that the dual space X' of a Hilbert space X is in one-to-one correspondence with the space X itself. In fact, we can say something more about this correspondence.

Let X be a Hilbert space, and for $f \in X'$, let u_f be the unique element in X which represents f as in Theorem 3.9. Consider the map $T : X' \to X$ defined by

$$T(f) = u_f, \quad f \in X'.$$

Then we have

$$\|T(f)\| = \|u_f\| = \|f\| \quad \forall f \in X'.$$

Moreover, it can be easily seen (*Verify*) that

$$T(f+g) = T(f) + T(g), \qquad T(\alpha f) = \bar{\alpha} T(f)$$

for all $f, g \in X'$ and $\alpha \in \mathbb{K}$. Thus, $T : X' \to X$ is a *conjugate linear* isometry between X' and X.

Because of this isometry, one often identifies X' with X. In fact, the above correspondence can be used to define an inner product on X', namely,

$$\langle f, g \rangle = \langle u_g, u_f \rangle, \qquad f, g \in X'.$$

We may observe that the norm induced by this inner product is the same as the norm on X' (*Verify*).

As examples, duals of ℓ^2 (respectively, $L^2[a,b]$) can be identified with ℓ^2 (respectively, $L^2[a,b]$) itself. More precisely, we have the following:

(a) For every $f \in (\ell^2)'$, there exists a unique $y \in \ell^2$ such that $\|y\|_2 = \|f\|$, and

$$f(x) = \sum_{j=1}^{\infty} x(j)\overline{y(j)} \quad \forall x \in \ell^2.$$

(b) For every $f \in (L^2(E))'$, there exists a unique $y \in L^2(E)$ such that $\|y\|_2 = \|f\|$, and

$$f(x) = \int_E x(t)\overline{y(t)}d\mu(t) \quad \forall x \in L^2(E).$$

In Chapter 8, we shall show that if $1 \le p < \infty$, then there is a natural isometry between the dual of ℓ^p (respectively, $L^p[a,b]$) and the space ℓ^q (respectively, $L^q[a,b]$), given as in (a) and (b) above with f belongs to $(\ell^p)'$ (respectively, $(L^p[a,b])'$) and y belongs to ℓ^q (respectively, $L^q[a,b]$).

Exercise 3.9 Let X^* be the space of all continuous conjugate linear functionals on a Hilbert space X, i.e., $f \in X^*$ if and only if $f : X \to \mathbb{K}$ is continuous and satisfies

$$f(x+y) = f(x) + f(y), \qquad f(\alpha x) = \bar{\alpha} f(x)$$

for all $x, y \in X$ and $\alpha \in \mathbb{K}$. Show the following:

(i) X^* is a normed linear space with the norm

$$\|f\| := \sup\{|f(x)| : x \in X, \|x\| = 1\}.$$

(ii) For every $f \in X^*$, there exists a unique $v_f \in X$ such that

$$f(x) = \langle v_f, x\rangle \quad \forall x \in X,$$

and, in that case, $\|f\| = \|v_f\|$.

(iii) Show that

$$\langle f, g\rangle = \langle v_f, v_g\rangle, \qquad f, g \in X^*,$$

defines a complete inner product on X^*.

The space X^* above is called the **adjoint space** of X. □

3.4 Convergence of Sequence of Operators

In Section 3.3, we observed that the dual of a Hilbert space is a Hilbert space. One may ask whether the dual of every Banach space is a Banach space. We answer this question affirmatively. More generally, we show that if Y is a Banach space, then $\mathcal{B}(X, Y)$ is a Banach space. For this purpose, as also to discuss more results, we prove a result concerning convergence of a sequence of bounded operators.

Suppose X and Y are normed linear spaces and (A_n) is a sequence of linear operators from X to Y. It is easy to see that, if $(A_n x)$ converges for every $x \in X$, then the function $A : X \to Y$ defined by

$$Ax = \lim_{n\to\infty} A_n x, \qquad x \in X,$$

is a linear operator. If, in addition, each A_n is a bounded operator, one may ask whether A is also a bounded operator. The answer is in the negative, as we see in the following example.

EXAMPLE 3.6 Let $X = c_{00}$ with $\|\cdot\|_\infty$ and $Y = \mathbb{K}$, and for $n \in \mathbb{N}$, let

$$f_n(x) = \sum_{j=1}^{n} x(j), \qquad x \in X.$$

Then it can be seen that f_n is a continuous linear functional on X and $\|f_n\| = n$ for all $n \in \mathbb{N}$. Also, we have

$$f_n(x) \to f(x) := \sum_{j=1}^{\infty} x(j) \quad \forall x \in X.$$

Recall from Example 3.2(iii)) that $f : X \to \mathbb{K}$ is linear but not continuous.

By imposing boundedness of $(\|A_n\|)$, we do obtain the continuity of A, as the following theorem shows.

Theorem 3.10 *Let X and Y be normed linear spaces and (A_n) be a sequence in $\mathcal{B}(X, Y)$ such that $(A_n x)$ converges in Y for every $x \in X$. If $(\|A_n\|)$ is a bounded sequence, then the operator $A : X \to Y$ defined by*

$$Ax = \lim_{n\to\infty} A_n x, \qquad x \in X,$$

belongs to $\mathcal{B}(X, Y)$ and $\|A\| \le \liminf_n \|A_n\|$.

Proof. Suppose $(A_n x)$ converges in Y for every $x \in X$. Define $A : X \to Y$ by

$$Ax = \lim_{n\to\infty} A_n x, \qquad x \in X.$$

The linearity of A follows from its definition. Also, for every $x \in X$, we have

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| = \liminf_n \|A_n x\| \le \liminf_n \|A_n\| \, \|x\|$$

so that $A \in \mathcal{B}(X, Y)$ and $\|A\| \le \liminf_n \|A_n\|$. ∎

Remark 3.4 Example 3.6 shows that the boundedness of $(\|A_n\|)$ in Theorem 3.10 cannot be dropped. We shall see in Chapter 6, as a consequence of the *Uniform boundedness principle*, that the condition of boundedness of $(\|A_n\|)$ in the above theorem would be redundant if X is a Banach space. Note that the space X in Example 3.6 is not a Banach space.

We know from real analysis that if (f_n) is a sequence in $C[a, b]$ such that $f_n(t) \to f(t)$ for every $t \in [a, b]$ for some $f \in C[a, b]$, then it is not necessary that

$$\sup_{t\in[a,b]} |f_n(t) - f(t)| \to 0.$$

One may ask whether it is true that, if (A_n) is a sequence in $\mathcal{B}(X,Y)$ such that $A_n x \to Ax$ for every $x \in X$ for some $A \in \mathcal{B}(X,Y)$, then

$$\sup\{\|A_n x - Ax\| : \|x\| \le 1\} \to 0.$$

The answer is again in the negative. To see this, consider the following example.

EXAMPLE 3.7 Let $X = \ell^2$, and for $n \in \mathbb{N}$, let $A_n : X \to X$ be defined by

$$(A_n x)(j) = \begin{cases} x(j) & \text{if } j \le n, \\ 0 & \text{if } j > n. \end{cases}$$

Then for every $x \in \ell^2$, we have

$$\|A_n x - x\|_2^2 = \sum_{j=n+1}^{\infty} |x(j)|^2 \to 0$$

as $n \to \infty$, but

$$\sup_{\|x\| \le 1} \|A_n x - x\|_2 \ge \|A_n e_{n+1} - e_{n+1}\|_2 = 1 \quad \forall n \in \mathbb{N}.$$

We may observe that for each $n \in \mathbb{N}$, the operators A_n and $I - A_n$ are orthogonal projections, and hence, we already know that

$$\sup_{\|x\| \le 1} \|A_n x - x\|_2 = \|I - A_n\| = 1 \quad \forall n \in \mathbb{N}.$$

In Theorem 3.10, it was required that $(A_n x)$ converges for every $x \in X$. One may ask whether this condition can be guaranteed by knowing the convergence of $(A_n x)$ for x in some subset E of X. It can be easily seen that if $E \subseteq X$ is such that $\operatorname{span} E = X$ or if E contains any of the sets $\{x \in X : \|x\| < r\}$, $\{x \in X : \|x\| = r\}$ for some $r > 0$, then the convergence of $(A_n x)$ for $x \in E$ implies the convergence of $(A_n x)$ for all $x \in X$. Now we show that if Y is a Banach space, then E can be any set such that $\operatorname{span} E$ is dense in X.

Theorem 3.11 *Let X be a normed linear space, Y be a Banach space and (A_n) be a sequence in $\mathcal{B}(X,Y)$ such that $(\|A_n\|)$ is bounded. Suppose $E \subseteq X$ is such that* $\operatorname{span} E$ *is dense in X. If $(A_n x)$ converges for every $x \in E$, then $(A_n x)$ converges for every $x \in X$, and the conclusions of Theorem 3.10 hold.*

Proof. Suppose $(A_n x)$ converges for every $x \in E$. Then it is obvious that $(A_n x)$ converges for every $x \in D := \operatorname{span} E$.

Now let $x \in X$ and let $\varepsilon > 0$ be given. Since D is dense in X, there is $u \in D$ such that $\|x - u\| < \varepsilon$. Now, for any $n, m \in \mathbb{N}$,

$$\begin{aligned} \|A_n x - A_m x\| &\le \|A_n x - A_n u\| + \|A_n u - A_m u\| + \|A_m u - A_m x\| \\ &\le \|A_n\|\,\|x - u\| + \|A_n u - A_m u\| + \|A_m\|\,\|u - x\| \\ &\le (\|A_n\| + \|A_m\|)\,\|x - u\| + \|A_n u - A_m u\|. \end{aligned}$$

Since $(A_n u)$ converges, there is $N \in \mathbb{N}$ such that $\|A_n u - A_m u\| < \varepsilon$ for all $n, m \ge N$. Now, let $c \ge \sup_n \|A_n\|$. Then we have

$$\|A_n x - A_m x\| \le (2c + 1)\varepsilon \quad \forall n, m \ge N.$$

Thus, $(A_n x)$ is a Cauchy sequence in Y. Since Y is a Banach space, $(A_n x)$ converges. ▮

Exercise 3.10 Suppose X and Y are normed linear spaces and (A_n) is a sequence of linear operators from X to Y. If X is finite dimensional, and $(A_n x)$ converges for every $x \in X$, then show that $\|A_n - A\| \to 0$ as $n \to \infty$, where $Ax := \lim_{n\to\infty} A_n x$, $x \in X$. □

Now we prove the result which we promised in the beginning of this section.

3.4.1 Completeness of $\mathcal{B}(X, Y)$

Theorem 3.12 *Let X and Y be normed linear spaces. If Y is a Banach space, then $\mathcal{B}(X, Y)$ is a Banach space. In particular, the dual of a normed linear space is a Banach space.*

Proof. Suppose Y is a Banach space, and (A_n) is a Cauchy sequence in $\mathcal{B}(X, Y)$. We show that (A_n) converges in $\mathcal{B}(X, Y)$. For this, let $\varepsilon > 0$ be given, and let $N \in \mathbb{N}$ be such that

$$\|A_n - A_m\| < \varepsilon \quad \forall n, m \ge N.$$

Then for each $x \in X$, we have

$$\|(A_n - A_m)x\| < \|A_n - A_m\|\,\|x\| < \varepsilon\|x\| \quad \forall n, m \ge N.$$

Thus, for every $x \in X$, $(A_n x)$ is a Cauchy sequence in Y. Since Y is complete, $(A_n x)$ converges. Also, since (A_n) is a Cauchy sequence in

$\mathcal{B}(X, Y)$, $(\|A_n\|)$ is bounded. Hence, by Theorem 3.10, $A : X \to Y$ defined by

$$Ax = \lim_{n\to\infty} A_n x, \quad x \in X,$$

belongs to $\mathcal{B}(X, Y)$. Now, for each $x \in X$ and $m \geq N$,

$$\|(A - A_m)x\| = \lim_{n\to\infty} \|(A_n - A_m)x\| \leq \limsup_{n,m} \|A_n - A_m\| \, \|x\| < \varepsilon \|x\|.$$

Thus, $\|A - A_m\| < \varepsilon$ for all $m \geq N$, showing that (A_n) converges to A in $\mathcal{B}(X, Y)$.

The particular case is obvious since $\mathbb{K}$ is a Banach space. ∎

What about the converse of the above theorem? Is it true that, if $\mathcal{B}(X, Y)$ is a Banach space, then Y is also a Banach space? We shall answer this question affirmatively by making use of the *Hahn-Banach extension theorem* which we shall prove in Chapter 5.

Now we consider an application of Theorem 3.11 to discuss the question of convergence of quadrature formulas.

3.4.2 Convergence of Quadrature Formulas

In numerical analysis, it is important to have nodes $t_{1,n}, \ldots, t_{n,n}$ and weights $w_{1,n}, \ldots, w_{n,n}$ such that the sequence (Q_n) of quadrature formulas given by

$$Q_n(x) := \sum_{j=1}^{n} x(t_{j,n}) w_{j,n}, \qquad x \in C[a, b],$$

converges to the integral of x, for every $x \in C[a, b]$, i.e.,

$$Q_n(x) \to \phi(x) := \int_a^b x(t)\, dt \quad \text{as} \quad n \to \infty$$

for every $x \in C[a, b]$. We have already seen that the functions Q_n and ϕ defined above are continuous linear functionals on $C[a, b]$ with $\|\cdot\|_\infty$. In fact, we have seen in Example 3.3 that

$$\|Q_n\| = \sum_{j=1}^{n} |w_{j,n}|, \qquad \|\phi\| = b - a.$$

Hence, as an immediate consequence of Theorem 3.11, we have the following result.

Theorem 3.13 *Let Q_n and ϕ be as above, and let E be a subset of $C[a,b]$ such that* $\operatorname{span} E$ *is dense in $C[a,b]$ with respect to the norm $\|\cdot\|_\infty$. If there exists $c>0$ such that*

$$\sum_{j=1}^{n} |w_{j,n}| \le c \quad \forall n \in \mathbb{N},$$

and $(Q_n(x))$ converges to $\phi(x)$ for every $x \in E$, then $(Q_n(x))$ converges to $\phi(x)$ for every $x \in C[a,b]$.

In the above theorem, we may take $E = \{x_0, x_1, x_2, \ldots\}$, where $x_k(t) = t^{k-1}$ for $k \in \mathbb{N}$.

We remark that if $w_{j,n}$ are real and non-negative, then with $x_0(t) = 1$ for all $t \in [a,b]$, we have

$$Q_n(x_0) = \sum_{j=1}^{n} w_{j,n} = \sum_{j=1}^{n} |w_{j,n}|$$

so that the convergence $Q_n(x_0) \to \int_a^b x_0(t)\,dt$ implies the boundedness of the sequence $\left(\sum_{j=1}^{n} |w_{j,n}|\right)$. Thus, we have proved the following result.

Corollary 3.14 *Let $x_k(t) = t^{k-1}$, $k \in \mathbb{N}$ and $w_{j,n} \ge 0$ for all $j = 1, \ldots, n$; $n \in \mathbb{N}$. If $(Q_n(x_k))$ converges to $\int_a^b x_k(t)\,dt$ for every $k \in \mathbb{N}$, then $(Q_n(x))$ converges to $\int_a^b x(t)\,dt$ for every $x \in C[a,b]$.*

Gaussian quadrature formula

We illustrate the above corollary by giving an example.

EXAMPLE 3.8 Let $\{p_0(t), p_1(t), p_2(t), \ldots\}$ be the set of orthogonal polynomials obtained by orthogonalizing $\{1, t, t^2, \ldots\}$ with respect to the L^2-inner product on $\mathcal{P}[a,b]$. These polynomials are known as the *Legendre polynomials* (cf. Example 2.5(iv)). Since, for every $n \in \mathbb{N}$,

$$\operatorname{span}\{1, t, \ldots, t^n\} = \operatorname{span}\{p_0(t), p_1(t), \ldots, p_n(t)\},$$

it follows that, for every polynomial $q(t)$ of degree atmost $n-1$, $\langle p_n, q\rangle = 0$.

It is known (cf. Wendroff [34]) that the zeros $t_{1,n}, \ldots, t_{n,n}$ of the n-th degree polynomial $p_n(t)$ are real, distinct and lie in the open interval (a, b). Define the quadrature formula

$$Q_n(x) = \sum_{j=1}^{n} x(t_{j,n}) w_{j,n}, \qquad x \in C[a, b],$$

where, for $j = 1, \ldots, n$,

$$w_{j,n} = \int_a^b \ell_{j,n}(t)\, dt, \qquad \ell_{j,n}(t) = \prod_{i=1, i \neq j}^{n} \frac{t - t_{i,n}}{t_{j,n} - t_{i,n}} \quad \text{for } t \in [a, b].$$

Note that

$$Q_n(x) = \int_a^b (L_n x)(t)\, dt,$$

where $L_n : C[a, b] \to C[a, b]$ is the Lagrange interpolatory projection

$$L_n x = \sum_{j=1}^{n} x(t_{j,n}) \ell_{j,n}, \qquad x \in C[a, b].$$

The above quadrature formula is called the **Gaussian quadrature formula.** We show that, for each $n \in \mathbb{N}$,

$$Q_n(p) = \int_a^b p(t)\, dt \quad \forall p \in \mathcal{P}_{n-1} \tag{3.3}$$

$$w_{j,n} > 0 \quad \forall j \in \{1, \ldots, n\}$$

so that, by Theorem 3.13, the Gaussian quadrature formula converges.

To prove (3.3) above, let $p(t)$ be a polynomial of degree atmost $n-1$. Since $(L_n p)(t)$ is also a polynomial of degree atmost $n-1$, and $(L_n p)(t_{i,n}) = p(t_{i,n})$ for $i = 1, \ldots, n$, it follows that $(L_n p)(t) = p(t)$ for every $t \in [a, b]$. Hence, we have

$$Q_n(p) = \int_a^b (L_n p)(t)\, dt = \int_a^b p(t)\, dt,$$

proving (3.3). Now, to obtain $w_{j,n} > 0$ for $j = 1, \ldots, n$; $n \in \mathbb{N}$, we make use of a stronger result than the one given in (3.3), namely,

$$Q_n(f) = \int_a^b f(t)\, dt \quad \forall f \in \mathcal{P}_{2n-1}. \tag{3.4}$$

First, let us assume (3.4). Since $(\ell_{j,n}(t))^2$ is a polynomial of degree atmost $2n-1$, and it satisfies $(\ell_{j,n}(t_{i,n}))^2 = \delta_{ij}$, it follows from (3.4) that

$$0 < \int_a^b (\ell_{j,n}(t))^2\,dt = Q_n\left((\ell_{j,n})^2\right) = w_{j,n}, \qquad j = 1, \ldots, n.$$

Now, we prove (3.4): Let $f(t)$ be a polynomial of degree atmost $2n-1$. Since $p_n \in \mathcal{P}_n$, we know from algebra that there are polynomials $q(t)$ and $r(t)$, each of degree atmost $n-1$, such that

$$f(t) = q(t)p_n(t) + r(t).$$

Since $\int_a^b q(t)p_n(t)\,dt = 0$, we have

$$\int_a^b f(t)\,dt = \int_a^b r(t)\,dt.$$

Now using (3.3) and the fact that

$$f(t_{i,n}) = q(t_{i,n})p_n(t_{i,n}) + r(t_{i,n}) = r(t_{i,n}),$$

we have

$$\int_a^b f(t)\,dt = \int_a^b r(t)\,dt = \sum_{j=1}^n r(t_{j,n})w_{j,n} = \sum_{j=1}^n f(t_{j,n})w_{j,n}.$$

Thus, the proof of (3.4) is complete.

Next, we consider a special class of (not necessarily bounded) operators, namely, the closed operators, which are very useful in applications.

3.5 Closed Operators

Recall the following example of an unbounded linear operator (Example 3.2(ii)):

$$(Ax)(t) = x'(t), \qquad x \in C^1[0,1];\ t \in [0,1],$$

defined from the space $X_0 = C^1[0,1]$ to the space $Y = C[0,1]$, both with $\|\cdot\|_\infty$. Although this operator is not continuous, by a result from real analysis, we know that for every (x_n) in X_0, if

$$\|x_n - x\|_\infty \to 0, \qquad \|Ax_n - y\|_\infty \to 0$$

as $n \to \infty$ for some $x, y \in C[0,1]$, then

$$x \in C^1[0,1], \qquad Ax = y.$$

This example and many more such examples of differential operators provide the motivation to consider a class of operators, called *closed linear operators*, having the above property.

Let X and Y be normed linear spaces and X_0 be a subspace of X. A linear operator $A : X_0 \to Y$ is said to be a **closed linear operator** or simply a **closed operator** if for every (x_n) in X_0 that satisfies $x_n \to x$ and $Ax_n \to y$ for some $x \in X$, $y \in Y$, we have $x \in X_0$ and $Ax = y$.

The nomenclature "closed" is used in the above definition because of the following result regarding the *graph* of A:

$$G(A) := \{(x, Ax) : x \in X_0\}.$$

Note that the graph of A is a subspace of the product space $X \times Y$. In the following, we assume, without explicitly mentioning, that the norm on the product space $X \times Y$ is

$$(x,y) \mapsto \|x\|_X + \|y\|_Y, \quad (x,y) \in X \times Y$$

or any equivalent norm.

Proposition 3.15 *Let X and Y be normed linear spaces and X_0 be a subspace of X. A linear operator $A : X_0 \to Y$ is a closed linear operator if and only if its graph $G(A) := \{(x, Ax) : x \in X_0\}$ is a closed subspace of the product space $X \times Y$.*

Proof. We have already observed that the graph $G(A)$ is a subspace of $X \times Y$. Suppose $A : X_0 \to Y$ is a closed linear operator. We prove that $G(A)$ is a closed subset of $X \times Y$. For this, suppose (x, y) belongs to the closure of $G(A)$. Then there exists a sequence $(x_n, y_n) \in G(A)$ such that $\|x_n - x\| + \|y_n - y\| \to 0$ as $n \to \infty$. But, $y_n = Ax_n$ for all $n \in \mathbb{N}$. Thus, we have $x_n \to x$ in X and $Ax_n \to y$ in Y. Since A is a closed operator, we have $x \in X_0$ and $Ax = y$. Consequently, $(x, y) \in G(A)$. Thus we have shown that $G(A)$ is a closed subset of $X \times Y$.

Conversely, suppose that $G(A)$ is a closed subspace of $X \times Y$. We show that A is a closed operator. For this, let (x_n) in X_0 be such that $x_n \to x$ and $Ax_n \to y$ for some $x \in X$ and $y \in Y$. Then, clearly,

$\|x_n - x\| + \|Ax_n - y\| \to 0$ as $n \to \infty$. Since $G(A)$ is a closed subset of $X \times Y$, it follows that $(x, y) \in G(A)$, i.e., $x \in X_0$ and $y = Ax$. Thus, we have showed that A is a closed operator. ∎

In the sequel, we may write

$$A : X_0 \subseteq X \to Y$$

to specify that A is a linear operator defined from a subspace X_0 of a normed linear space X to a normed linear space Y.

Although most of the differential operators which occur in applications can be shown to be closed operators, in this book we intend to discuss only a few properties of closed operators.

Closed operators arise naturally while studying bounded operators. Look at the following simple example.

EXAMPLE 3.9 Let $X = C^1[0,1]$, with the norm $x \mapsto \|x\| := \|x\|_\infty + \|x'\|_\infty$, and $Y = C^1[0,1]$ with $\|\cdot\|_\infty$. Let $A : X \to Y$ be defined by $Ax = x$ for every $x \in X$. Clearly, A is a bijective bounded linear operator. Since X is a Banach space and Y is not a Banach space, the inverse of A is not continuous. However, it is easy to see that $A^{-1} : Y \to X$ is a closed operator.

The situation described in the above example prevails for all injective bounded operators defined on the whole space, as the following theorem shows.

Theorem 3.16 *Let $A : X \to Y$ be an injective bounded operator. Then $A^{-1} : R(A) \subseteq Y \to X$ is a closed operator.*

Proof. Let (y_n) in $R(A)$ be such that $y_n \to y$ in Y and $A^{-1}y_n \to x$ in X for some $y \in Y$, $x \in X$. Taking $x_n = A^{-1}y_n$, $n \in \mathbb{N}$, it follows that $x_n \to x$ and $Ax_n \to y$ as $n \to \infty$. But, by continuity of A, we have $Ax_n \to Ax$ so that $y = Ax \in R(A)$ and $A^{-1}y = x$. Thus, $A^{-1} : R(A) \subseteq Y \to X$ is a closed operator. ∎

We have seen in the beginning of this section that a closed operator need not be a bounded operator. A question that naturally arises is whether every bounded operator is a closed operator. The answer is again in the negative, as the following example shows.

EXAMPLE 3.10 Let $X = C[a,b]$ with $\|\cdot\|_\infty$, $X_0 = C^1[a,b]$, and $A : X_0 \to X$ be the inclusion operator, i.e., $Ax = x$ for every $x \in X_0$. Clearly, A is a bounded operator. Since X_0 is a proper dense subspace of X, it is easily seen that the graph of A is not closed in $X \times X$ so that by Proposition 3.15, A is not a closed operator.

However, if the domain of a bounded operator is a closed subspace, then it is a closed operator, as the following theorem shows.

Theorem 3.17 *Let $A : X_0 \subseteq X \to Y$ be a bounded operator.*

(i) *If X_0 is closed in X, then A is a closed operator.*

(ii) *If Y is a Banach space and A is a closed operator, then X_0 is a closed subspace of X.*

Proof. Let $A : X_0 \subseteq X \to Y$ be a bounded operator.

(i) Suppose X_0 is a closed subspace of X. To see that A is a closed operator, let (x_n) be a sequence in X_0 such that $x_n \to x$ and $Ax_n \to y$ for some $x \in X$, $y \in Y$. Since X_0 is closed in X, we have $x \in X_0$, and then by the continuity of A, we have $Ax = y$. This proves that A is a closed operator.

(ii) Suppose $A : X_0 \subseteq X \to Y$ is a closed as well as a bounded operator, and Y is a Banach space. In order to show that X_0 is a closed subspace of X, let (x_n) in X_0 be such that $x_n \to x$ for some $x \in X$. We show that $x \in X_0$. Since A is continuous, (Ax_n) is a Cauchy sequence in Y. Now, by the completeness of Y, there exists $y \in Y$ such that $Ax_n \to y$. Therefore, by using the fact that A is a closed operator, $x \in X_0$ and $Ax = y$.

This completes the proof. ∎

Exercise 3.11 Let $A : X_0 \subseteq X \to Y$ be a closed operator. If Y is a Banach space and X_0 is not closed in X, then A is an unbounded operator. Why? □

In view of the above exercise, one may ask whether every closed operator $A : X_0 \subseteq X \to Y$ with closed X_0 and complete Y is a bounded operator.

Again, the answer is in the negative, as the following example shows.

EXAMPLE 3.11 Let $X_0 = X = C^1[0,1]$ and $Y = C[0,1]$, both with $\|\cdot\|_\infty$ and $Ax = x'$, $x \in X$. Then $A : X_0 \to Y$ is a closed operator with domain X_0 closed and codomain Y a Banach space, but A is not a bounded operator.

We shall prove in Chapter 7 that if both the domain X_0 and the codomain Y of a closed operator $A : X_0 \subseteq X \to Y$ are Banach spaces, then it is a bounded operator. This result is called the *closed graph theorem*. As a consequence of closed graph theorem, we can infer that if X and Y are Banach spaces, X_0 is a proper dense subspace of X, then a closed operator $A : X_0 \subseteq X \to Y$ does not have a closed extension to all of X (*Why?*). This is in contrast with bounded operators as the following theorem shows.

Theorem 3.18 *Let $A_0 : X_0 \subseteq X \to Y$ be a bounded operator, where X_0 is dense in X, and Y is a Banach space. Then there exists a unique $A \in \mathcal{B}(X,Y)$ such that A is an extension of A_0. Moreover, $\|A\| = \|A_0\|$, and for $x \in X$, $Ax = \lim_{n\to\infty} A_0x_n$, where (x_n) is a sequence in X_0 such that $x_n \to x$.*

Proof. Let $x \in X$ and (x_n) in X_0 be such that $x_n \to x$ as $n \to \infty$ (such a sequence exists, since X_0 is dense in X). As

$$\|A_0x_n - A_0x_m\| \leq \|A_0\| \, \|x_n - x_m\| \quad \forall n, m \in \mathbb{N},$$

the sequence (A_0x_n) is Cauchy in Y. Since Y is a Banach space, $\lim_{n\to\infty} A_0x_n$ exists. Define $A : X \to Y$ by

$$Ax = \lim_{n\to\infty} A_0x_n, \qquad x \in X.$$

It is easily seen (*Verify*) that A is well defined, and it is a linear operator that extends A_0 to the whole space X. Clearly, $\|A_0\| \leq \|A\|$. Also, for $x \in X$, let (x_n) be a sequence in X_0 such that $x_n \to x$. Then, since $\|A_0x_n\| \leq \|A_0\| \, \|x_n\|$ for all $n \in \mathbb{N}$, and $\|x_n\| \to \|x\|$ as $n \to \infty$, we have

$$\|Ax\| = \lim_{n\to\infty} \|A_0x_n\| \leq \|A_0\| \, \|x\|.$$

Hence,

$$A \in \mathcal{B}(X,Y), \qquad \|A\| = \|A_0\|.$$

Now suppose that $B \in \mathcal{B}(X,Y)$ is another extension of A_0. For $x \in X$, let (x_n) in X_0 be such that $\lim_{n\to\infty} x_n = x$. Then, by continuity of B and by the definition of A, we have

$$Bx = \lim_{n\to\infty} Bx_n = \lim_{n\to\infty} A_0x_n = Ax.$$

Therefore, the operator A defined above is the unique operator in $\mathcal{B}(X,Y)$ satisfying the required conditions. ∎

We know that if $A \in \mathcal{B}(X,Y)$, then $N(A)$ is a closed subspace of X, and if A is injective, then by Theorem 3.16 that $A^{-1} : R(A) \to X$ is a closed operator. Now we show that these results hold for a closed operator $A : X_0 \subseteq X \to Y$ as well.

Theorem 3.19 *Suppose $A : X_0 \subseteq X \to Y$ is a closed operator. Then we have the following.*

(i) *$N(A)$ is a closed subspace of X.*

(ii) *If A is injective, then $A^{-1} : R(A) \subseteq Y \to X$ is a closed operator.*

Proof. (i) To see that $N(A)$ is a closed subspace of X, let (x_n) be in $N(A)$ such that $x_n \to x$ in X. We show that $x \in N(A)$. Since $Ax_n = 0$ for all n, by the definition of closedness of A, it follows that $x \in X_0$ and $Ax = 0$. Thus, $x \in N(A)$.

(ii) Next, suppose that A is injective. To see that A^{-1} is a closed operator, let (y_n) in $R(A)$ be such that $y_n \to y$ in Y and $A^{-1}y_n \to x$ in X. We show that $y \in R(A)$ and $A^{-1}y = x$. Let $x_n = A^{-1}y_n$. Then we have

$$x_n \to x, \quad Ax_n \to y \quad \text{as} \quad n \to \infty.$$

By closedness of A, $x \in X_0$ and $Ax = y$. Therefore, $y \in R(A)$ and $A^{-1}y = x$. ∎

As a consequence of Theorems 3.17 and 3.19, we have the following result which is very important in the context of solving operator equations.

Theorem 3.20 *Suppose X is a Banach space and $A : X_0 \subseteq X \to Y$ is an injective closed operator. If $R(A)$ is not closed in Y, then $A^{-1} : R(A) \subseteq Y \to X$ is an unbounded operator.*

Proof. By Theorem 3.19, $A^{-1} : R(A) \subseteq Y \to X$ is a closed operator. If $A^{-1} : R(A) \subseteq Y \to X$ is a bounded operator, then since X is a Banach space, by Theorem 3.17, it follows that $R(A)$, which is the domain of A^{-1}, is a closed subspace. ∎

Exercise 3.12 (i) Prove that a linear functional on a normed linear space is a closed operator if and only if it is continuous.

(ii) Give a direct proof for Theorem 3.19 (ii) when $X_0 = X$ and A is a bounded operator.

(iii) Give a direct proof for Theorem 3.20 if $X_0 = X$ and A is a bounded operator. Have you seen this result earlier?

(iv) Suppose X is a Banach space and $A : X_0 \subseteq X \to Y$ is a closed operator which is bounded below. Show that $R(A)$ is a closed subspace of X.

(v) Let X be Hilbert space and $A : X_0 \subseteq X \to X$ be a linear operator such that for some $c > 0$,

$$|\langle Ax, x\rangle| \geq c\,\|x\|^2 \quad \forall x \in X.$$

Show that A is injective, $R(A)$ dense in X, and $A^{-1} : R(A) \to X$ has a unique continuous extension to all of X. [*Hint:* See that A is bounded below, $R(A)^{\perp} = \{0\}$, and use Theorem 3.18.]

(vi) Suppose A in (v) is a closed operator. Then show that A is bijective. [*Hint:* Use (v) and show that $R(A)$ is closed.] □

Theorem 3.19(ii) can be used for showing that certain operators are closed, as in the following example.

EXAMPLE 3.12 Let $X = \ell^2$ and (λ_n) be a bounded sequence of positive real numbers which converges to 0. Let

$$X_0 = \left\{x \in \ell^2 : \sum_{j=1}^{\infty} \frac{|x(j)|^2}{\lambda_j^2} < \infty\right\}$$

and $A : X_0 \subseteq X \to X$ be defined by

$$Ax = \left(\frac{x(1)}{\lambda_1}, \frac{x(2)}{\lambda_2}, \ldots\right), \qquad x \in X_0.$$

Then A is a closed unbounded operator: Note first that A is an injective linear operator. Also, for every $y \in \ell^2$,

$$x = (\lambda_1 y(1), \lambda_2 y(2), \ldots) \in X_0, \qquad Ax = y,$$

so that $A : X_0 \to X$ is surjective as well. Moreover, if $c > 0$ is such that $|\lambda_j| \leq c$ for all $j \in \mathbb{N}$, then

$$\|A^{-1}y\|^2 = \sum_{j=1}^{\infty} |\lambda_j y(j)|^2 \leq c^2\, \|y\|_2^2 \quad \forall y \in \ell^2.$$

Thus, $A^{-1} \in \mathcal{B}(X)$. Hence, by Theorems 3.17 and 3.19, it follows that

$$A^{-1} : X \to X, \qquad A = \left(A^{-1}\right)^{-1} : X_0 \subseteq X \to X$$

are closed operators. To see that A is not a bounded operator, consider the sequence (x_n) defined by $x_n(j) = \delta_{nj}$ for $j, n \in \mathbb{N}$. Then (x_n) is a bounded sequence in X_0, but $\|Ax_n\| = 1/|\lambda_n| \to \infty$ as $n \to \infty$. Therefore, A is not a bounded operator.

The above example also shows that the inverse of an injective bounded operator need not be a bounded operator. In that example, $A^{-1} : X \to X$ is an injective bounded linear operator, but its inverse, namely the operator $A : X_0 \subseteq X \to X$, is not continuous.

Exercise 3.13 Show that X_0 in Example 3.12 is not a closed subspace of ℓ^2. □

Now we show how a closed operator defined on a subspace X_0 of a Banach space can be used for defining a new norm on X_0 which makes it a Banach space.

Theorem 3.21 *Let $A : X_0 \subseteq X \to Y$ be a closed operator, where X and Y are Banach spaces. Define*

$$\|x\|_A = \|x\| + \|Ax\|, \qquad x \in X_0.$$

Then $x \mapsto \|x\|_A$ is a complete norm on X_0.

Proof. Recall that $x \mapsto \|x\|_A := \|x\| + \|Ax\|$, $x \in X_0$ is the graph norm on X_0. To see that it is complete, let (x_n) be a Cauchy sequence in X_0 with respect to $\|\cdot\|_A$. Then it follows that (x_n) and (Ax_n) are Cauchy sequences in X and Y, respectively. Since both X and Y are Banach spaces, there exist $x \in X$, $y \in Y$ such that $\|x_n - x\| \to 0$ and $\|Ax_n - y\| \to 0$ as $n \to \infty$. Now, using the fact that A is a closed operator, we have $x \in X_0$, $Ax = y$. Thus,

$$\|x_n - x\| + \|A(x_n - x)\| \to 0 \quad \text{as} \quad n \to \infty,$$

showing that (x_n) converges in X_0 with respect to $\|\cdot\|_A$. ∎

PROBLEMS

1. Let X and Y be normed linear spaces, and $A : X \to Y$ be a linear operator. Show that A is continuous if and only if for every Cauchy sequence (x_n) in X, (Ax_n) is a Cauchy sequence in Y.

2. Let $X = c$, the space of all convergent scalar sequences with norm $\|\cdot\|_\infty$. Show that $f : X \to \mathbb{K}$ defined by $f(x) = \lim_{n\to\infty} x(n)$, $x \in X$, is a continuous linear functional and $\|f\| = 1$.

3. For a given $u \in \mathbb{K}^n$, let $f : \mathbb{K}^n \to \mathbb{K}$ be defined by

$$f(x) = u(1)x(1) + \cdots + u(n)x(n), \qquad x \in \mathbb{K}^n.$$

If $\mathbb{K}^n$ is endowed with the norm $\|\cdot\|_p$ for $1 \le p \le \infty$, then show that $\|f\| = \|u\|_q$, where $q \in [1, \infty]$ is the conjugate exponent of p.

4. Let X be an n-dimensional inner product space and $A : X \to X$ be a linear operator. If $\{u_1, \ldots, u_n\}$ is an orthonormal subset X, then show that

$$\|A\| \le \left(\sum_{j=1}^{n} \|Au_j\|^2\right)^{1/2}.$$

If there exist $\lambda_1, \ldots, \lambda_n$ in $\mathbb{K}$ such that $Au_j = \lambda_j u_j$ for $j = 1, \ldots, n$, then show that

$$\|A\| = \max\{\,|\lambda_j| : j = 1, \ldots, n\}.$$

5. Let X be $C[a,b]$ with $\|\cdot\|_\infty$ or $L^2[a,b]$. For $u \in C[a,b]$, define $A : X \to X$ by

$$(Ax)(t) = u(t)x(t), \qquad t \in [a,b],\ x \in X.$$

Show that $A \in \mathcal{B}(X)$ and $\|A\| = \|u\|_\infty$.

[*Hint*: The relation $\|A\| \ge \|u\|_\infty$, in the case of $X = L^2[a,b]$, can be proved by *reductio ad absūrdum* by making use of the continuity of u: Assuming that $\|A\| < \|u\|_\infty$, obtain an interval $J \subseteq [a,b]$ such that $\|A\| < |u(t)|$ for all $t \in J$, and an $x_0 \in C[a,b]$ satisfying $\|Ax_0\|_2 > \|A\|\,\|x_0\|_2$.]

6. For $1 \le p \le \infty$, let $X = \ell^p(\mathbb{N})$, and let $A : X \to X$ and $B : X \to X$ be defined by

$$Ax = (0, x(1), x(2), \ldots), \qquad Bx = (x(2), x(3), \ldots) \quad \text{for } x \in X.$$

Show that $A, B \in \mathcal{B}(X)$ and $\|A\| = 1 = \|B\|$.

The operators A and B above are called the *right shift operator* and *left shift operator*, respectively.

7. Let $X = C[a,b]$ with $\|\cdot\|_\infty$. Let $t_1, \ldots, t_k$ be distinct points in $[a,b]$ and $u_1, \ldots, u_k$ in X be such that $u_j(t_i) = \delta_{ij}$ for $i, j = 1, \ldots, k$. Let $P : X \to X$ be defined by

$$Px = \sum_{j=1}^{k} x(t_j)u_j, \qquad x \in X.$$

Then show that P is a projection operator onto span$\{u_1, \ldots, u_k\}$, $P \in \mathcal{B}(X)$, and

$$\|P\| = \sup_{a \le t \le b} \sum_{j=1}^{k} |u_j(t)|.$$

The projection operator P above is called the *interpolatory projection* onto the span of $\{u_1, \ldots, u_k\}$ corresponding to the 'nodes' $\{t_1, \ldots, t_k\}$.

8. For each $n \in \mathbb{N}$, let $t_{1,n}, \ldots, t_{n,n}$ be points in $[a,b]$ such that

$$a = t_{0,n} < t_{1,n} < t_{2,n} < \ldots < t_{n,n} = b,$$

$$\max_{1 \le j \le n} (t_{j,n} - t_{j-1,n}) \to 0 \quad \text{as} \quad n \to \infty.$$

For each $j \in \{1, \ldots, n\}$, let $u_{j,n} \in C[a,b]$ be such that $u_{j,n}(t_{i,n}) = \delta_{ij}$ and the restriction of u_j to the interval $[t_{i-1,n}, t_{i,n}]$, $i = 1, \ldots, n$, is a polynomial of degree 1. Let P_n be the interpolatory projection onto span$\{u_{j,n} : j = 1, \ldots, n\}$ corresponding to the nodes $t_{1,n}, \ldots, t_{n,n}$, and let Q_n be the quadrature formula with the nodes $t_{1,n}, \ldots, t_{n,n}$ and weights $w_{1,n}, \ldots, w_{n,n}$, where $w_{j,n} = \int_a^b u_{j,n}(t)\,dt$, $j = 1, \ldots, n$. Show that

(a) $\|P_n x - x\|_\infty \to 0$ as $n \to \infty$ for all $x \in C[a,b]$.

(b) $Q_n(x) \to \int_a^b x(t)\,dt$ as $n \to \infty$ for all $x \in C[a,b]$.

9. If $A \in \mathcal{B}(X, Y)$ is injective, then show that

$$\|A^{-1}y\| \ge \frac{\|y\|}{\|A\|} \quad \forall y \in R(A).$$

10. If $A \in \mathcal{B}(X)$ is such that $\|A\| < 1$, then show that $I - A$ injective and

$$\|(I - A)^{-1}y\| \le \frac{\|y\|}{1 - \|A\|} \quad \forall y \in R(I - A).$$

11. Let X be a normed linear space and X^* be the set of all continuous conjugate linear functionals on X, i.e., $f \in X^*$ if and only if $f : X \to \mathbb{K}$ is continuous and

$$f(\alpha x + \beta y) = \bar{\alpha} f(x) + \bar{\beta} f(y) \quad \forall x, y \in X; \forall \alpha, \beta \in \mathbb{K}.$$

Show that X^* is a normed linear space with norm defined by

$$\|f\| = \sup\{|f(x)| : x \in X, \|x\| \le 1\}, \qquad f \in X^*,$$

and with respect to this norm, X^* is a Banach space.

12. Let $A \in \mathcal{B}(X, Y)$. For $f \in Y^*$, define

$$(A^* f)(x) = f(Ax), \quad f \in Y^*, \ x \in X.$$

Show that $A^* f \in X^*$ for all $f \in Y^*$, and $A^* \in \mathcal{B}(Y^*, X^*)$.

13. Let X be a Banach space, Y be a normed linear space, X_0 be a subspace of X and $A : X_0 \subseteq X \to Y$ be a closed operator which is bounded below, i.e., there exists $c > 0$ such that $\|Ax\| \ge c\|x\|$ for every $x \in X_0$. Show the following:

(a) $R(A)$ is a closed subspace of Y.

(b) A is injective, and $A^{-1} : R(A) \subseteq Y \to X$ is a closed and bounded operator.

14. Let X be a normed linear space and $P : X \to X$ be a projection operator. Show that P is a closed operator if and only if $R(P)$ and $N(P)$ are closed subspaces.

15. Every finite dimensional normed linear space is linearly homeomorphic with $\mathbb{K}^n$ for some $n \in \mathbb{N}$ – Why? Here, $\mathbb{K}^n$ may be endowed with any norm.

16. Show that a nonzero linear functional on a normed linear space is discontinuous if and only if its null space is dense.

17. Let (λ_n) be a sequence of positive real numbers such that $\lambda_n \to 0$ as $n \to \infty$, and let

$$X = \left\{ x \in \ell^2(\mathbb{N}) : \sum_{j=1}^{\infty} \frac{|x(j)|^2}{\lambda_j^2} < \infty \right\}.$$

(a) Show that X is a Hilbert space with respect to the inner product

$$\langle x, y \rangle_* = \sum_{j=1}^{\infty} \frac{x(j)\overline{y(j)}}{\lambda_j^2}, \qquad x, y \in X.$$

(b) Let X be with the above inner product $\langle \cdot, \cdot \rangle_*$. For $u \in \ell^2$, let $f_u : X \to \mathbb{K}$ be defined by $f_u(x) = \langle x, u \rangle$, $x \in X$. Show that

$$f_u \in X', \qquad \|f_u\| = \left(\sum_{j=1}^{\infty} \lambda_j^2 |u(j)|^2 \right)^{1/2}.$$

[*Hint*: For $u \in \ell^2$, take x_0 such that $x_0(j) = \lambda_j^2 \overline{u(j)}$, $j \in \mathbb{N}$, and observe that $x_0 \in X$, and $f_u(x_0) = \|x_0\|_*^2$.]

18. Let p and q be conjugate exponents with $1 < p < \infty$. For $y \in \ell^q$, let

$$f(x) = \sum_{j=1}^{\infty} x(j) y(y), \qquad x \in \ell^p.$$

Show that $f \in (\ell^p)'$ and $\|f\| \leq \|y\|_q$.

19. Let p and q be conjugate exponents with $1 < p < \infty$. For $y \in L^q[a, b]$, let

$$f(x) = \int_a^b x(t) y(t)\, d\mu(t), \qquad x \in L^p[a, b].$$

Then show that f defines a continuous linear functional on $L^p[a, b]$ and $\|f\| \leq \|y\|_q$.

20. Let X be an inner product space. Show that for every $A \in \mathcal{B}(X)$,

$$\|A\| = \sup \{ |\langle Ax, y \rangle| : x, y \in X, \|x\| = 1 = \|y\| \}.$$

21. A Hilbert space $\mathcal{H} \subseteq \mathcal{F}([a, b], \mathbb{K})$, is said to have a *reproducing kernel* if for every $s \in [a, b]$, there exists $h_s \in \mathcal{H}$ such that

$$x(s) = \langle x, h_s \rangle_{\mathcal{H}} \quad \forall\, x \in \mathcal{H}.$$

The function h defined by $h(s,t) = h_s(t)$, $s,t \in [a,b]$ is called the *reproducing kernel* of the Hilbert space $\mathcal{H}$.

Show that a Hilbert space $\mathcal{H} \subseteq \mathcal{F}([a,b], \mathbb{K})$ has a reproducing kernel if the evaluation function $t \mapsto x(t)$, $t \in [a,b]$ is a continuous linear functional on $\mathcal{H}$.

22. Let $k(\cdot,\cdot) \in C([a,b]\times[a,b])$ and $A : L^2[a,b] \to L^2[a,b]$ be defined by

$$(Ax)(s) = \int_a^b k(s,t)x(t)\,d\mu(t), \qquad x \in L^2[a,b],\ s \in [a,b].$$

Show that, if A is an injective linear operator, then $R(A)$ with the inner product

$$\langle f,g\rangle = \langle A^{-1}f, A^{-1}g\rangle_2, \qquad f,g \in R(A),$$

is a Hilbert space with reproducing kernel $h(\cdot,\cdot)$ defined by

$$h(s,t) = \int_a^b \overline{k(s,\tau)}k(t,\tau)\,d\tau, \qquad s,t \in [a,b].$$

23. Let X be a linear space, Y be a Hilbert space, and $A : X \to Y$ be a linear operator. For $y \in R(A) + R(A)^\perp$, let

$$S_y = \{x \in X : \|Ax - y\| \leq \|Au - y\| \quad \forall u \in X\}.$$

Show that S_y is nonempty if and only if $y \in R(A) + R(A)^\perp$.

24. In the above problem, suppose X is also a Hilbert space and $N(A)$ is a closed subspace of X. Then, show that, for every $y \in R(A) + R(A)^\perp$, there exists a unique $x_y \in S_y$ such that

$$\|x_y\| = \inf\{\|x\| : x \in S_y\}.$$

The above x_y is called the *generalized solution* of the equation $Ax = y$, and the map $y \mapsto x_y$ is called the *generalized inverse* of A or the *Moore-Penrose inverse* of A, denoted by $A^\dagger$.

25. Let X and Y be Hilbert spaces and $A \in \mathcal{B}(X,Y)$. Show that the generalized inverse $A^\dagger : R(A) + R(A)^\perp \subseteq Y \to X$ is a closed linear operator.

4

More about Hilbert Spaces

In this chapter we consider more results on inner product spaces and Hilbert spaces which are essentially based on the inner product structure of the space. Such results may not be available for a general normed linear space or a Banach space.

4.1 Orthonormal Sets and Orthonormal Bases

Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$, and let $\| \cdot \|$ be the corresponding norm, i.e.,

$$\|x\| = \langle x, x \rangle^{1/2} \quad \forall x \in X.$$

Unless otherwise stated explicitly, the inner product spaces that we consider are nontrivial, i.e., of dimension at least one.

Recall from Chapter 2 that a subset S of X is an *orthogonal set* in X if $\langle x, y \rangle = 0$ for every distinct $x, y \in S$, and an orthogonal set S is called an *orthonormal set* if $\|x\| = 1$ for every $x \in S$. Recall also that every orthonormal set is linearly independent.

By an **orthonormal sequence** we mean a sequence (u_n) in X whose terms form an orthonormal set.

Using the concept of orthonormality, we define a new concept in an inner product space, namely, an *orthonormal basis*.

Let E be an orthonormal set in an inner product space X. Then E is said to be an **orthonormal basis** of X or a **complete orthonormal system** for X if it is a maximal orthonormal set in X, that is, E is an orthonormal set, and for every orthonormal set $\widetilde{E}$ satisfying $E \subseteq \widetilde{E}$, we have $\widetilde{E} = E$.

Proposition 4.1 *If E is an orthonormal set in an inner product space X, then X has an orthonormal basis containing E.*

Proof. Let X be an inner product space and let E be an orthonormal set in X. Consider the family $\mathcal{E}$ of all orthonormal sets in X containing E. Clearly, $\mathcal{E}$ is nonempty, as $E \in \mathcal{E}$. It is seen that $\mathcal{E}$ is a partially ordered set with respect to set inclusion. If $\mathcal{T}$ is a totally ordered subset of $\mathcal{E}$, then the union of all elements of $\mathcal{T}$ is an upper bound of $\mathcal{T}$. Therefore, by Zorn's lemma, $\mathcal{E}$ has a maximal element which is an orthonormal basis of X. ∎

Since every orthonormal set is linearly independent, it is clear that every basis which is an orthonormal set is an orthonormal basis as well. But, an orthonormal basis need not be a basis. Here is an example to this effect.

EXAMPLE 4.1 Let $X = \ell^2$, and $E = \{e_1, e_2, \ldots\}$, with $e_j(i) = \delta_{ij}$; $i, j \in \mathbb{N}$. Clearly, E is an orthonormal set. Since $\langle x, e_n\rangle = x(n)$ for all $x \in \ell^2$ and for all $n \in \mathbb{N}$,

$$x \in \ell^2, \qquad \langle x, e_n\rangle = 0 \quad \forall n \in \mathbb{N} \implies x = 0.$$

This shows that, there cannot be an orthonormal set $\widetilde{E}$ which properly contains E. Thus, E is an orthonormal basis. But, since ℓ^2 is a complete space, by Theorem 2.30, E is not a basis of ℓ^2.

In fact, the property of the set E that we used in the above example is a characterization property of any orthonormal basis, as we shall see in the following theorem. We shall show, in Section 4.4, that an orthonormal basis of a Hilbert space is a basis if and only if the space is finite dimensional .

Recall from Chapter 2 that for a set $S \subseteq X$, we write

$$S^\perp = \{u \in X : \langle u, x\rangle = 0 \quad \forall x \in S\}.$$

Theorem 4.2 *Let E be an orthonormal set in an inner product space X. Then E is an orthonormal basis if and only if $E^\perp = \{0\}$, i.e.,*

$$x \in X, \quad \langle x, u\rangle = 0 \quad \forall u \in E \implies x = 0.$$

Proof. Suppose E is an orthonormal set in X. We show that $E^\perp \neq \{0\}$ if and only if E is not an orthonormal basis.

If $E^\perp \neq \{0\}$, and $x \in E^\perp$ is a nonzero element, then taking $u = x/\|x\|$, we see that the set $\tilde{E} := E \cup \{u\}$ is an orthonormal set which properly contains E so that E is not an orthonormal basis.

Conversely, suppose that E is not an orthonormal basis. Then there exists an orthonormal set $\widetilde{E} \neq E$ such that $E \subset \widetilde{E}$. Let $v \in \widetilde{E} \setminus E$. Then we have $v \neq 0$ and $v \in E^{\perp}$. Thus, $E^{\perp} \neq \{0\}$. ∎

We have already mentioned that every orthonormal set which is also a basis is an orthonormal basis. From Theorem 4.2, we can derive something more.

Corollary 4.3 *Let E be an orthonormal set in an inner product space X. Then E is an orthonormal basis of* $\operatorname{span} E$ *and its closure.*

Proof. Since E is a basis of $\operatorname{span} E$, it is an orthonormal basis of $\operatorname{span} E$ as well. Now, to show that E is an orthonormal basis of $\overline{\operatorname{span} E}$, by Theorem 4.2, it is enough to prove that

$$x \in \overline{\operatorname{span} E}, \quad \langle x, u\rangle = 0 \quad \forall u \in E \implies x = 0.$$

For this, suppose $x \in \overline{\operatorname{span} E}$ such that $\langle x, u\rangle = 0$ for every $u \in E$. Then it follows that $\langle x, y\rangle = 0$ for every $y \in \operatorname{span} E$. Now, let (x_n) be a sequence in $\operatorname{span} E$ such that $x = \lim_{n\to\infty} x_n$. Then, using the continuity of inner product and the above observed fact, we get

$$\langle x, x\rangle = \lim_{n\to\infty} \langle x, x_n\rangle = 0,$$

proving $x = 0$. ∎

EXAMPLE 4.2 (i) Let $X = \mathbb{K}^n$ with the standard inner product

$$\langle x, y\rangle = \sum_{j=1}^{n} \alpha_j \bar{\beta}_j, \qquad x, y \in \mathbb{K}^n.$$

Clearly, $\{e_1, \ldots, e_n\}$ with $e_j(i) = \delta_{ij}$ for $i, j = 1, \ldots, n$, is an orthonormal set which is also a basis of X. Hence, it is an orthonormal basis of X.

(ii) Let $X = c_{00}$ with the ℓ^2-inner product, i.e.,

$$\langle x, y\rangle = \sum_{j=1}^{\infty} x(j)\overline{y(j)}, \qquad x, y \in c_{00}.$$

Then $E = \{e_1, e_2, \ldots\}$ with $e_j(i) = \delta_{ij}$ for $i, j = 1, 2, \ldots$, is an orthonormal set, and it is also a basis of X. Therefore, E is an orthonormal basis of X.

We have already observed in Example 4.1 that E is an orthonormal basis of ℓ^2. This is a consequence of Corollary 4.3 as well, since c_{00} is dense in ℓ^2 and E is an orthonormal basis of c_{00}.

(iii) Let $X_0 = \mathcal{P}[a,b]$ with the L^2-inner product

$$\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}\,dt, \qquad x, y \in \mathcal{P}[a,b].$$

Let $x_j(t) = t^{j-1}$ for $t \in [a,b]$, $j = 1, 2, \ldots$, and let $E = \{u_1, u_2, \ldots\}$ be the orthonormal set obtained from $\{x_1, x_2, \ldots\}$ by Gramm-Schmidt orthogonalization procedure (see the Remark following Theorem 2.19). Since E is a basis of $\mathcal{P}[a,b]$, by Corollary 4.3, it is an orthonormal basis of X_0.

Recall that if $[a,b] = [-1,1]$, then the polynomials $u_1, u_2, \ldots$ are known as *Legendre polynomials* (Example 2.5 (iv)).

(iv) Let $X = C[a,b]$ with L^2-inner product, i.e.,

$$\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}\,dt, \qquad x, y \in C[a,b],$$

and let X_0 and E be as in (iii) above. Since E is an orthonormal basis of X_0 and since $\mathcal{P}[a,b]$ is dense in X (Example 2.8), by Corollary 4.3, E is an orthonormal basis of X.

(v) Let X and E be as in (iv). Since E is an orthonormal basis of X and since X is dense in $L^2[a,b]$ (cf. Theorem 2.28), by Corollary 4.3, E is an orthonormal basis of $L^2[a,b]$.

We give another orthonormal basis of $L^2[a,b]$ using the definition of Fourier coefficients. We take $\mathbb{K} = \mathbb{C}$ and for convenience of representation we consider $[a,b] = [0, 2\pi]$. Let

$$u_n(t) = \frac{e^{int}}{\sqrt{2\pi}}, \qquad t \in [0, 2\pi],\ n \in \mathbb{Z}.$$

Then it can be seen that

$$\langle u_n, u_m\rangle = \delta_{nm} \quad \forall n, m \in \mathbb{Z},$$

so that $E = \{u_1, u_2, \ldots\}$ is an orthonormal set. It is known by Weierstrass approximation theorem (cf. [28], viz. Theorem 4.25), that span E, the set of all *trigonometric polynomials*, is dense in $C[0, 2\pi]$ with respect to the norm $\|\cdot\|_\infty$. Since

$$\|x\|_2 \le \sqrt{2\pi}\|x\|_\infty \quad \forall x \in C[0, 2\pi],$$

it follows that span(E) is dense in $C[0,2\pi]$ with respect to $\|\cdot\|_2$ as well. Now since $C[0,2\pi]$ is dense in $L^2[a,b]$ (cf. Theorem 2.28), by Corollary 4.3, E is an orthonormal basis of $L^2[0,2\pi]$.

(vi) Let $u_1(t) = 1/\sqrt{2\pi}$, and for $n \in \mathbb{N}$,

$$u_{2n}(t) = \frac{\sin(nt)}{\sqrt{\pi}}, \qquad u_{2n+1}(t) = \frac{\cos(nt)}{\sqrt{\pi}}.$$

Then it can be seen that $E = \{u_1, u_2, \ldots\}$ is an orthonormal set in $L^2[0,2\pi]$. By using the same arguments as in the last example, it also follows that E is an orthonormal basis of $L^2[0,2\pi]$.

We note that the orthonormal bases for all the spaces $\mathbb{K}^n$, c_{00}, ℓ^2, $C[a,b]$, $L^2[a,b]$ considered in Example 4.2 are countable. Recall that these spaces are separable. Now we see that this is, in fact, the case for every separable inner product space.

Proposition 4.4 *Every orthonormal set in a separable inner product space is countable.*

Proof. Let X be an inner product space and E be an orthonormal set in X. We observe that for every $u, v \in E$ with $u \neq v$,

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 = 2.$$

Therefore, if E is uncountable, then there is an uncountable family of disjoint open sets, namely, $\{B(u, \sqrt{2}) : u \in E\}$ in X. Hence, by Proposition 2.37, the space X is not separable. ∎

Exercise 4.1 Show that every orthonormal set in an inner product space is a closed set. □

We know that the spaces $C[a,b]$, ℓ^2 and $L^2[a,b]$ cannot have denumerable bases (see the remarks following the statement of Theorem 2.30). Hence, the orthonormal bases considered in Example 4.2 for these spaces are not bases.

But if the space is finite dimensional, then every orthonormal basis is a basis (*Why?*). We have also observed in Chapter 2 that if X is an inner product space of dimension n, and if $\{u_1, \ldots, u_n\}$ is an orthonormal set in X, then for every $x \in X$,

$$x = \sum_{j=1}^{n} \langle x, u_j \rangle u_j, \qquad \|x\|^2 = \sum_{j=1}^{n} |\langle x, u_j \rangle|^2.$$

Next, suppose that X is an inner product space with a denumerable, i.e., countably infinite, orthonormal basis $E = \{u_1, u_2, \ldots\}$. One would like to know whether the relations

$$x = \sum_{j=1}^{\infty} \langle x, u_j \rangle u_j, \qquad \|x\|^2 = \sum_{j=1}^{\infty} |\langle x, u_j \rangle|^2 \tag{4.1}$$

hold for every $x \in X$.

We may observe that if the series $\sum_{j=1}^{\infty} \langle x, u_j \rangle u_j$ converges to an element $y \in X$, then $\sum_{j=1}^{\infty} |\langle x, u_j \rangle|^2 = \|y\|^2$. Indeed, if

$$s_n := \sum_{j=1}^{n} \langle x, u_j \rangle u_j \to y \quad \text{as} \quad n \to \infty,$$

then

$$\|s_n\|^2 = \sum_{j=1}^{n} |\langle x, u_j \rangle|^2 \to \|y\|^2 \quad \text{as} \quad n \to \infty.$$

Thus, what one would like to know is whether the series $\sum_j \langle x, u_j \rangle u_j$ converges to x.

Clearly, the equalities in (4.1) hold if E is a basis of X. For, then corresponding to every $x \in X$, there are scalars $\alpha_1, \ldots, \alpha_k$, where k depending on x, such that

$$x = \sum_{j=1}^{k} \alpha_j u_j$$

so that $\alpha_j = \langle x, u_j \rangle$ for $j = 1, \ldots, k$, and $\langle x, u_j \rangle = 0$ for all $j \geq k$. Thus, equalities in (4.1) hold if E is a basis of X.

What about for the case when E is not a basis of X? Let us look at the example of $X = \ell^2$ and $E = \{e_1, e_2, \ldots\}$. In this case, we know that $\langle x, e_j \rangle = x(j)$ for all $j \in \mathbb{N}$, so that

$$\|x\|^2 = \sum_{j=1}^{\infty} |x(j)|^2 = \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 \quad \forall x \in \ell^2.$$

For $x \in \ell^2$, if we take $s_n = \sum_{j=1}^{n} x(j) e_j$, then it follows that

$$\|x - s_n\|^2 = \sum_{j=n+1}^{\infty} |x(j)|^2 \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, we also have

$$x = \lim_{n\to\infty} s_n = \sum_{j=1}^{\infty} \langle x, e_j\rangle e_j.$$

We shall show that equalities do hold in (4.1) for every Hilbert space.

Before going further, we shall observe a few facts concerning a series of the form $\sum_{j=1}^{\infty}\langle x, u_j\rangle u_j$, where $\{u_1, u_2, \ldots\}$ is a denumerable orthonormal set. Suppose a series $\sum_{j=1}^{\infty}\langle x, u_j\rangle u_j$ converges. Can we rearrange the series and still have its convergence? The answer is in the affirmative, as the following proposition shows.

Proposition 4.5 *Let $E = \{u_1, u_2, \ldots\}$ be a denumerable orthonormal set in an inner product space X, and let $\psi : E \to \mathbb{K}$ be a function. Suppose (v_n) is a sequence obtained by rearranging the terms of (u_n). If the series $\sum_{j=1}^{\infty} \psi(u_j)u_j$ converges, say to $s \in X$, then the series $\sum_{j=1}^{\infty} \psi(v_j)v_j$ also converges to s.*

Proof. Suppose that the series $\sum_{j=1}^{\infty} \psi(u_j)u_j$ converges, say $s = \sum_{j=1}^{\infty} \psi(u_j)u_j$. For $n \in \mathbb{N}$, let

$$s_n = \sum_{j=1}^{n} \psi(u_j)u_j, \qquad \zeta_n = \sum_{j=1}^{n} \psi(v_j)v_j \quad \forall n \in \mathbb{N}.$$

We show that (ζ_n) converges to s. For this, first we observe that

$$\|s - \zeta_n\|^2 = \|s\|^2 + \|\zeta_n\|^2 - \langle s, \zeta_n\rangle - \langle \zeta_n, s\rangle,$$

where

$$\|\zeta_n\|^2 = \langle \zeta_n, \zeta_n\rangle = \sum_{j=1}^{n} |\psi(v_j)|^2, \qquad \langle s, \zeta_n\rangle = \lim_{m\to\infty} \langle s_m, \zeta_n\rangle.$$

Taking m large enough such that $\{v_1, \ldots, v_n\} \subseteq \{u_1, \ldots, u_m\}$, it follows that

$$\langle s_m, \zeta_n\rangle = \sum_{j=1}^{n} |\psi(v_j)|^2 = \langle \zeta_n, s_m\rangle.$$

Thus,

$$\|s - \zeta_n\|^2 = \|s\|^2 - \|\zeta_n\|^2 \quad \forall n \in \mathbb{N}.$$

Now, since

$$\lim_{n\to\infty} \|\zeta_n\|^2 = \sum_{j=1}^{\infty} |\psi(v_j)|^2 = \sum_{j=1}^{\infty} |\psi(u_j)|^2 = \lim_{n\to\infty} \|s_n\|^2 = \|s\|^2,$$

we have $\lim_{n\to\infty} \zeta_n = s$. ∎

NOTATION: Suppose E is a countable orthonormal set in an inner product space X and $\psi : E \to \mathbb{K}$ be a function. If E is a finite orthonormal set, then the meaning of $\sum_{u\in E} \psi(u)u$ is obvious. Now, suppose that E is a denumerable orthonormal set and $x \in X$. Suppose $u_1, u_2, \ldots$ and $v_1, v_2, \ldots$ are two different enumerations of E. Then by Proposition 4.5, we know that $\sum_{j=1}^{\infty} \psi(u_j)u_j$ converges if and only if $\sum_{j=1}^{\infty} \psi(v_j)v_j$ converges. Thus, we say that the expression $\sum_{u\in E} \psi(u)u$ converges or well defined provided the series $\sum_{j=1}^{\infty} \psi(u_j)u_j$ converges for a particular enumeration, say $u_1, u_2, \ldots$, of elements E.

Next, suppose that E is an arbitrary, but not necessarily countable, orthonormal set in X, and let $\psi : E \to \mathbb{K}$. If the set $E_0 := \{u \in E : \psi(u) \neq 0\}$ is countable, and if $\sum_{u\in E_0} \psi(u)u$ is convergent, then we denote this sum by $\sum_{u\in E} \psi(u)u$, and say that $\sum_{u\in E} \psi(u)u$ is convergent or well defined.

4.2 Bessel's Inequality

The following inequality, which is a generalization of Schwarz inequality, is very useful to establish certain important results. For instance, it is effectively used for establishing equalities of the form (4.1) when the space is complete.

Theorem 4.6 (Bessel's inequality) *Let E be a countable orthonormal set in an inner product space X. Then for every $x \in X$,*

$$\sum_{u\in E} |\langle x, u\rangle|^2 \leq \|x\|^2.$$

Moreover, for $x \in X$,

$$\sum_{u\in E} |\langle x, u\rangle|^2 = \|x\|^2 \iff \sum_{u\in E} \langle x, u\rangle u = x.$$

Proof. Suppose $\{u_1, \ldots, u_n\} \subseteq E$. For $x \in X$, let

$$s_n = \sum_{j=1}^{n} \langle x, u_j \rangle u_j.$$

Then, using the orthonormality of E, we have

$$\langle x, s_n \rangle = \langle s_n, x \rangle = \|s_n\|^2 = \sum_{j=1}^{n} |\langle x, u_j \rangle|^2$$

so that

$$0 \leq \|x - s_n\|^2 = \|x\|^2 - \sum_{j=1}^{n} |\langle x, u_j \rangle|^2.$$

From this the conclusions in theorem follow. ∎

An obvious consequence of Theorem 4.6 is that if $\{u_1, u_2, \ldots\}$ is a denumerable orthonormal set in an inner product space X, then

$$\langle x, u_n \rangle \to 0 \quad \text{as} \quad n \to \infty$$

for every $x \in X$. This implies, in view of Example 4.2(vi), for every $x \in L^2[0, 2\pi]$,

$$\int_0^{2\pi} x(t) \cos(nt)\, d\mu(t) \to 0, \qquad \int_0^{2\pi} x(t) \sin(nt)\, d\mu(t) \to 0$$

as $n \to \infty$. These results are known as the **Riemann-Lebesgue lemma.**

Here is an important consequence of Theorem 4.6.

Theorem 4.7 *Let E be an orthonormal set in an inner product space X. Then for every $x \in X$, the set*

$$E_x := \{u \in E : \langle x, u \rangle \neq 0\}$$

is a countable set.

Proof. For $x \in X$, we note that

$$E_x := \{u \in E : \langle x, u \rangle \neq 0\} = \cup_{n=1}^{\infty} E_{x,n},$$

where

$$E_{x,n} = \left\{u \in E : |\langle x, u \rangle| > \frac{1}{n}\right\}.$$

Then it is enough to show that $E_{x,n}$ is finite for each n. For this, let $u_1, \ldots, u_k$ in $E_{x,n}$. Then, by Theorem 4.6, it follows that

$$\frac{k}{n^2} \le \sum_{j=1}^{k} |\langle x, u_j \rangle|^2 \le \|x\|^2$$

so that $k \le n^2\|x\|^2$. This proves (*How?*) that $E_{x,n}$ is a finite set for each n, and hence, E_x is a countable set. ∎

Exercise 4.2 Let X be an inner product space and E be an orthonormal set.

(i) If E is a finite set, then show that equality occurs in Bessel's inequality if and only if $x \in \operatorname{span} E$.

(ii) If $E = \{u_1, u_2, \ldots\}$, a denumerable set, then show that the function

$$T : x \mapsto (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots), \qquad x \in X,$$

is a bounded linear operator from X into ℓ^2, and it is injective if and only if E is an orthonormal basis. Give an example to show that T need not be surjective even if E is an orthonormal basis. □

4.3 Fourier Expansion and Parseval's Formula

In this section, we prove the equalities in (4.1) whenever X is a Hilbert space and $\{u_n : n \in \mathbb{N}\}$ is an orthonormal basis. In fact, in view of Theorem 4.7, we prove generalized forms of (4.1) with an arbitrary (not necessarily countable) orthonormal set. To this end, first we prove a general result.

Proposition 4.8 *Let X be a Hilbert space and E be an orthonormal set in X. Then for every $x \in X$, the series $\sum_{u \in E} \langle x, u \rangle u$ is well defined, and*

$$x - \sum_{u \in E} \langle x, u \rangle u \in E^{\perp}.$$

Proof. Let $x \in X$ and $E_x = \{u \in E : \langle x, u \rangle \neq 0\}$. By Theorem 4.7, we know that E_x is a countable set. Suppose E_x is a finite set, say $E_x = \{u_1, \ldots, u_n\}$. Then $y := \sum_{j=1}^{n} \langle x, u_j \rangle u_j = \sum_{u \in E} \langle x, u \rangle u$, and for every $v \in E$,

$$\langle y, v \rangle = \sum_{j=1}^{n} \langle x, u_j \rangle \langle u_j, v \rangle = \langle x, v \rangle.$$

Hence, $x - y \in E^\perp$. Next, suppose that $E_x = \{u_1, u_2, \ldots\}$, a denumerable set. Let $s_n = \sum_{j=1}^{n} \langle x, u_j \rangle u_j$ for $n \in \mathbb{N}$. Then for $n > m$, we have

$$\|s_n - s_m\|^2 = \sum_{j=m+1}^{n} |\langle x, u_j \rangle|^2$$

Now, by Bessel's inequality, it follows that (s_n) is a Cauchy sequence. Since X is a Hilbert space, the sequence (s_n) converges. Let y be its limit, i.e., $y := \lim_{n\to\infty} s_n = \sum_{j=1}^{\infty} \langle x, u_j \rangle u_j$. Then, using the continuity of the inner product and the orthonormality of $\{u_1, u_2, \ldots\}$, it follows that

$$\langle y, v \rangle = \sum_{j=1}^{\infty} \langle x, u_j \rangle \langle u_j, v \rangle = \langle x, v \rangle \quad \forall v \in E,$$

so that $x - y \in E^\perp$. This completes the proof. ∎

As a corollary to the above proposition, we prove the *Fourier expansion theorem* and *Parseval's formula*.

Theorem 4.9 *Let X be a Hilbert space and E be an orthonormal set in X. Then the following are equivalent:*

(i) *E is an orthonormal basis of X.*

(ii) **(Fourier expansion)** *For every $x \in X$, $x = \sum_{u \in E} \langle x, u \rangle u$.*

(iii) **(Parseval's formula)** *For every $x \in X$, $\|x\|^2 = \sum_{u \in E} |\langle x, u \rangle|^2$.*

Proof. Let $x \in X$ and $E_x = \{u \in E : \langle x, u \rangle \neq 0\}$. By Bessel's inequality (Theorem 4.7), we know that E_x is a countable set. Then by Theorem 4.6, we have

$$x = \sum_{u \in E_x} \langle x, u \rangle u \iff \|x\|^2 = \sum_{u \in E_x} |\langle x, u \rangle|^2.$$

Thus, we have the equivalence of (ii) and (iii).

To see the equivalence of (i) and (ii), recall from Theorem 4.2 that E is an orthonormal basis if and only if $E^\perp = \{0\}$. But, by Proposition 4.8, $E^\perp = \{0\}$ if and only if $x = \sum_{u \in E} \langle x, u \rangle u$. Thus, (i) and (ii) are equivalent. ∎

Remark 4.1 If E is a countable orthonormal set, then in the proof of the above theorem, we can take E itself in place of E_x. Thus, if $\{u_1, u_2, \ldots\}$ is an orthonormal set in a Hilbert space, then it is an orthonormal basis if and only if every $x \in X$ can be written as

$$x = \sum_n \langle x, u_n \rangle u_n.$$

In particular, we can conclude that every countable orthonormal basis of a Hilbert space is a *Schauder basis.*

EXAMPLE 4.3 (i) We have observed earlier that if $X = \ell^2$ and the orthonormal basis $E = \{e_1, e_2, \ldots\}$, then

$$x = \sum_{j=1}^{\infty} x(j) e_j \quad \forall x \in \ell^2.$$

Parseval's formula in this case is nothing but the definition of the norm in ℓ^2.

(ii) Consider the Hilbert space $X = L^2[0, 2\pi]$, and orthonormal basis as in Example 4.2(v). By Theorem 4.9, we have

$$x = \sum_{n=-\infty}^{\infty} \hat{x}(n) u_n, \qquad \|x\|^2 = \sum_{n=-\infty}^{\infty} |\hat{x}(n)|^2,$$

where $\hat{x}(n)$ is the n-th Fourier coefficient, i.e.,

$$\hat{x}(n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} x(t) e^{-int}\, dt.$$

Exercise 4.3 Let X be a Hilbert space and E be an orthonormal set.

(i) If E is a countable set, then prove that $\sum_{u \in E} |f(u)|^2 \leq \|f\|^2$ for every $f \in X'$.

[*Hint*: Use the Riesz representation theorem and Bessel's inequality.]

(ii) Show that, for every $f \in X'$, $E_f := \{u \in E : f(u) \neq 0\}$ is countable, and $\sum_{u \in E} \overline{f(u)} u$ converges.

[*Hint*: Use the Riesz representation theorem and Proposition 4.8.]

(iii) Prove that for every $f \in X'$, there exists a countable set $E_0 \subseteq E$ such that $v := \sum_{u \in E} \overline{f(u)} u$ is well defined and $f(x) = \langle x, v \rangle$ for all $x \in X$.

[*Hint*: The Riesz representation theorem and Theorem 4.9(ii).] □

In view of the above exercise, one may ask whether one can show the convergence of $\sum_{u \in E} \overline{f(u)} u$ for every $f \in X'$ without resorting to the Riesz representation theorem. This can be answered affirmatively, after proving *Riesz-Fischer theorem* in Section 4.4 (see Problem 12).

4.4 Riesz-Fischer Theorem

We have already observed in Proposition 4.4 that every orthonormal set in a separable inner product space is countable.

If X is a Hilbert space, then the existence of a countable orthonormal basis is one of the characterizing properties of separable Hilbert spaces.

Theorem 4.10 *Let X be a Hilbert space and E be an orthonormal basis of X. Then E is countable if and only if X is separable.*

Proof. Suppose X is a separable Hilbert space. Then we already know, by Proposition 4.4, that every orthonormal basis of X is countable.

Conversely, suppose E is a countable orthonormal basis of X. By Theorem 4.9, we know that $\operatorname{span} E$ is dense in X. In fact, E is a Schauder basis of X. Hence, by Theorem 2.36, X is a separable space. ∎

We already know (*Do you?*) that every finite dimensional inner product space X is linearly isometric with $\ell^2(n)$ for some $n \in \mathbb{N}$, with isometry given by

$$x \mapsto (\langle x, u_1 \rangle, \ldots, \langle x, u_n \rangle), \qquad x \in X,$$

where $\{u_1, \ldots, u_n\}$ is an orthonormal basis of X.

Also, if X is an infinite dimensional separable Hilbert space and $\{u_1, u_2, \ldots\}$ is an orthonormal basis of X, then by Theorem 4.9, we know that the map

$$x \mapsto (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots), \qquad x \in X,$$

is a linear isometry from X into ℓ^2. A question that naturally arises is whether this map is onto. We answer this affirmatively, by making use of the following theorem.

Theorem 4.11 (Riesz-Fischer theorem) *Let $\{u_1, u_2, \ldots\}$ be an orthonormal set in a Hilbert space X and let (α_n) be a sequence of scalars. Then*

$$\sum_{n=1}^{\infty} |\alpha_n|^2 \quad \text{converges if and only if} \quad \sum_{n=1}^{\infty} \alpha_n u_n \quad \text{converges,}$$

and, in that case,

$$\alpha_n = \langle x, u_n \rangle \quad \forall n \in \mathbb{N}, \quad \text{where} \quad x = \sum_{n=1}^{\infty} \alpha_n u_n.$$

Proof. Let $s_n = \sum_{j=1}^{n} \alpha_j u_j$ for $n \in \mathbb{N}$. Then for $n > m$, it follows that

$$\|s_n - s_m\|^2 = \sum_{j=m+1}^{n} |\alpha_n|^2.$$

From this it is clear that $\sum_{n=1}^{\infty} |\alpha_n|^2$ converges if and only if (s_n) is a Cauchy sequence. Since X is complete, (s_n) is a Cauchy sequence if and only if $\sum_{n=1}^{\infty} \alpha_n u_n$ converges. Now the results follow by using the properties of an inner product. ∎

As a consequence of the above theorem, we show that every infinite dimensional separable Hilbert space is linearly isometric with ℓ^2. More generally, we have the following.

Theorem 4.12 *Let X be an infinite dimensional Hilbert space and (u_n) be an orthonormal sequence in X. Consider the linear operator $T : X \to \ell^2$ defined by*

$$Tx = (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots), \qquad x \in X.$$

Then we have the following:

(i) *T is surjective.*

(ii) *T is an isometry if and only if E is an orthonormal basis.*

Proof. Clearly, by Bessel's inequality, the map T is well defined. By Riesz-Fischer Theorem 4.11, it is seen that T is surjective. By Theorem 4.9, T is a linear isometry if and only if E is an orthonormal basis. ∎

We have observed earlier that an orthonormal basis of an inner product space need not be a (Hamel) basis. We have also given examples of infinite dimensional inner product spaces with orthonormal bases which are bases as well. The following theorem shows that, in an infinite dimensional Hilbert space, an orthonormal basis can never be a basis.

Theorem 4.13 *Let X be a Hilbert space and E be an orthonormal basis of X. Then E is a basis of X if and only if X is finite dimensional.*

Proof. Clearly, if X is finite dimensional, then every orthonormal basis is a basis.

Conversely, suppose that X is an infinite dimensional Hilbert space and E is an orthonormal basis of X. Let $\{u_1, u_2, \ldots\}$ be a denumerable subset of E. Let (α_n) be a sequence of nonzero scalars such that $\sum_{j=1}^{\infty} |\alpha_j|^2 < \infty$. Then, by Riesz-Fischer Theorem 4.11, it follows that the series $\sum_{j=1}^{\infty} \alpha_j u_j$ converges.

Let $x = \sum_{j=1}^{\infty} \alpha_j u_j$. If E is a basis of X, then there exists $\{v_1, \ldots, v_k\} \subseteq E$ and scalars $\beta_1, \ldots, \beta_k$ such that $x = \sum_{j=1}^{k} \beta_j v_j$. Suppose n is large enough such that $u_n \notin \{v_1, \ldots, v_k\}$. Then, using the orthonormality of $\{u_1, u_2, \ldots\}$, we have

$$\alpha_n = \left\langle \sum_{j=1}^{\infty} \alpha_j u_j, u_n \right\rangle = \left\langle \sum_{j=1}^{k} \beta_j v_j, u_n \right\rangle = 0$$

a contradiction to the assumption that $\alpha_j \neq 0$ for all $j \in \mathbb{N}$. Thus, if X is infinite dimensional, then E cannot be a basis of X. ∎

Next, we give examples of nonseparable inner product spaces so that by Proposition 2.34 (i), completions of such spaces are nonseparable Hilbert spaces.

EXAMPLE 4.4 (i) For $r \in \mathbb{R}$, let $u_r(t) = e^{irt}$, $t \in \mathbb{R}$. Let X be the set of all finite linear combinations of members of $\{u_r : r \in \mathbb{R}\}$ with coefficients from $\mathbb{C}$, i.e., $x \in X$ if and only if there exists a finite subset Δ_x of $\mathbb{R}$ and $a_r \in \mathbb{K}$, $r \in \Delta_x$ such that $x = \sum_{r \in \Delta_x} a_r u_r$. Clearly, X is a linear space over $\mathbb{C}$. For $p = \sum_{j=1}^{n} c_j u_{r_j}$ and $q = \sum_{j=1}^{m} d_j u_{s_j}$ in X, define

$$\langle p, q \rangle = \sum_{(i,j) \in D_{p,q}} c_i \bar{d}_j,$$

where $D_{p,q} = \{(i,j) : r_i = s_j, i = 1, \ldots, n; j = 1, \ldots, m\}$. Then we see that the map $(p,q) \mapsto \langle p,q \rangle$ is an inner product on X, and $\{u_r : r \in \mathbb{R}\}$ is an orthonormal basis of X.

We may observe that

$$\lim_{\tau\to\infty} \frac{1}{2\tau}\int_{-\tau}^{\tau} e^{irt}\,dt = \begin{cases} 1 & \text{if } r = 0, \\ 0 & \text{if } r \neq 0. \end{cases}$$

Hence, we see that

$$\langle p,q \rangle = \lim_{\tau\to\infty} \frac{1}{2\tau}\int_{-\tau}^{\tau} p(t)\overline{q(t)}\,dt.$$

The completion of the above space X is known as the space of *almost periodic functions*.

In fact, on every linear space with an uncountable basis, we can define an inner product under which the basis is an orthonormal basis so that the space and its completion are not separable. This is described in the following:

(ii) Let X be a linear space having an uncountable basis, say $E = \{u_\alpha : \alpha \in \Lambda\}$. (Any underlying linear space of an infinite dimensional Banach space is a candidate for this X.) For $x = \sum_{\alpha\in\Lambda_1} a_\alpha u_\alpha$, $y = \sum_{\beta\in\Lambda_2} b_\beta u_\beta$ in X, define

$$\langle x,y \rangle = \sum_{\alpha,\beta\in\Lambda_1\cap\Lambda_2} a_\alpha \bar{b}_\beta\, g(\alpha,\beta),$$

where Λ_1, Λ_2 are finite subsets of Λ; $a_\alpha, b_\beta \in \mathbb{K}$ for $\alpha \in \Lambda_1$, $\beta \in \Lambda_2$, and

$$g(\alpha,\beta) = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta. \end{cases}$$

Then it is seen that $\langle \cdot,\cdot \rangle$ is an inner product on X, and the uncountable set E is an orthonormal basis of X. Thus, the completion of X is a non-separable Hilbert space.

PROBLEMS

In the following, X is a Hilbert space and E is an orthonormal set in X.

1. Show that for every $x \in X$, $\|x\| = \sup\{|\langle x, u\rangle| : u \in X, \|u\| = 1\}$. Deduce that, for every $A \in \mathcal{B}(X)$,

$$\|A\| = \sup\{|\langle Ax, y\rangle| : x, y \in X, \|x\| = 1 = \|y\|\}.$$

2. Let $\{u_1, \ldots, u_n\} \subseteq E$, $f \in X'$, and $y_n := \sum_{j=1}^{n} \overline{f(u_j)} u_j$. Show that for every $x \in X$,

$$\langle x, y_n\rangle = f(x_n), \qquad x_n := \sum_{j=1}^{n} \langle x, u_j\rangle u_j, \qquad \sum_{j=1}^{n} |f(u_j)|^2 \leq \|f\|^2.$$

3. Show that E is an orthonormal basis of X if and only if $\operatorname{span} E$ is dense in X.

4. If X is a Hilbert space, then show that E is an orthonormal basis if and only if

$$\langle x, y\rangle = \sum_{u \in E} \langle x, u\rangle \overline{\langle y, u\rangle} \quad \forall x, y \in X.$$

5. Suppose X is a Hilbert space and E is an orthonormal basis of X. For each $x \in X$, let $\hat{x} : E \to \mathbb{K}$ be defined by $\hat{x}(u) = \langle x, u\rangle$, $u \in E$. Show that

(a) $\widehat{X} := \{\hat{x} : x \in X\}$ is a subspace of $\mathcal{F}(E, \mathbb{K})$,

(b) $(\hat{x}, \hat{y}) \mapsto \sum_{u \in E} \hat{x}(u)\overline{\hat{y}(u)}$ is an inner product on $\widehat{X}$,

(c) $\widehat{X}$ with the above inner product is a Hilbert space, and

(d) $\{\hat{u} : u \in E\}$ is an orthonormal basis of $\widehat{X}$.

Show also that the map $x \mapsto \hat{x}$ is a linear isometry from X onto $\widehat{X}$.

6. Let S be a nonempty set and $\mathcal{H}(S)$ be the subset of $\mathcal{F}(S, \mathbb{K})$ consisting of all functions $f : S \to \mathbb{K}$ such that $\{s \in S : f(s) \neq 0\}$ is a countable set, and $\sum_{s \in S} |f(s)|^2 < \infty$. Show the following:

(a) $\mathcal{H}(S)$ is a subspace of $\mathcal{F}(S, \mathbb{K})$.

(b) $(f, g) \mapsto \sum_{s \in S} f(s)\overline{g(s)}$ is an inner product on $\mathcal{H}(S)$.

(c) $\mathcal{H}(S)$ with the above inner product is a Hilbert space.

(d) $\{\chi_{\{s\}} : s \in S\}$ is an orthonormal basis of $\mathcal{H}(S)$.

(e) If X is a Hilbert space and E is an orthonormal basis of X, then the map $T : f \mapsto \sum_{u \in E} f(u)u$ is a linear isometry from $\mathcal{H}(E)$ onto X, and $\mathcal{H}(E)$ is the space $\widehat{X}$ as in Problem 5.

7. Suppose Λ_1, Λ_2 are nonempty sets such that there is a bijective function T from Λ_1 onto Λ_2. Let $E_1 = \{u_\lambda : \lambda \in \Lambda\}$ and $E_2 = \{v_\lambda : \lambda \in \Lambda\}$ be orthonormal bases of Hilbert spaces X_1 and X_2, respectively. Prove the following:

(a) For each $x \in X_1$, $\widehat{T}(x) := \sum_{u \in E_1} \langle x, u \rangle T(u)$ is a convergent series in X_2.

(b) $\langle x, u \rangle = \langle \widehat{T}(x), T(u) \rangle$ for all $u \in E_1$.

(c) $\widehat{T} : X_1 \to X_2$ is a surjective linear isometry.

8. If X is a Hilbert space, then show that $P : X \to X$ defined by

$$Px = \sum_{u \in E} \langle x, u \rangle u, \quad x \in X,$$

is an orthogonal projection onto $X_0 := \overline{\operatorname{span} E}$.

9. (Projection theorem) Suppose X is a Hilbert space and X_0 is a closed subspace of X. Let E_0 be an orthonormal basis of X_0. Then show that for every $x \in X$,

$$y := \sum_{u \in E_0} \langle x, u \rangle u \in X_0, \qquad x - y \in X_0^\perp.$$

Show also that $P : X \to X$ defined by

$$Px = \sum_{u \in E_0} \langle x, u \rangle u \in X_0, \quad x \in X,$$

is an orthogonal projection onto X_0.

10. Let X_1 and X_2 be closed subspaces of a Hilbert space X, and let P_1 and P_2 be orthogonal projections onto X_1 and X_2, respectively. If $\langle x, y \rangle = 0$ for all $x \in X_1$, $y \in X_2$, then show that

(a) $X_1 + X_2$ is a closed subspace of X,

(b) $P_1 + P_2$ is the orthogonal projection onto $X_1 + X_2$, and

(c) $P_1 P_2 = 0 = P_2 P_1$.

11. Show that, for $f \in X'$, the set $E_f := \{u \in E : f(u) \neq 0\}$ is countable. [*Hint*: Use Problem 2 above and the arguments in Theorem 4.7.]

12. (Riesz representation theorem) If X is a Hilbert space and $f \in X'$, then show that $v := \sum_{u \in E} \overline{f(u)} u$ is well defined and

$$f(x) = \langle x, v \rangle \quad \forall x \in X.$$

[*Hint*: Use Problems 2, 11 above and Theorem 4.11.]

13. Let (σ_n) be a sequence of positive real numbers which converge to 0. Show that

(a) $X := \{x \in \ell^2 : \sum_{n=1}^{\infty} |x(n)|^2/\sigma_n^2 < \infty\}$ is a subspace of ℓ^2,

(b) $\langle x, y \rangle_* := \sum_{n=1}^{\infty} x(n)\overline{y(n)}/\sigma_n^2$ defines an inner product on X,

(c) X with $\langle \cdot, \cdot \rangle_*$ is a Hilbert space.

Also, find an orthonormal basis for X.

14. Suppose X is a separable Hilbert space and $E = \{u_n : n \in \mathbb{N}\}$ is a denumerable orthonormal basis. Let (x_n) be a sequence in X and $x \in X$. Show that $f(x_n) \to f(x)$ for each $f \in X'$ as $n \to \infty$ if and only if $\langle x_n, u_m \rangle \to \langle x, u_m \rangle$ as $n \to \infty$ for each $m \in \mathbb{N}$.

15. Let X be a Hilbert space, and $\{u_1, u_2, \ldots\}$ and $\{v_1, v_2, \ldots\}$ are orthonormal sets in X. If $\{u_1, u_2, \ldots\}$ is an orthonormal basis of X and if $\sum_{j=1}^{\infty} \|u_j - v_j\|^2 < \infty$, then show that $\{v_1, v_2, \ldots\}$ is also an orthonormal basis of X.

16. Using the fact that $X_0 = \{x \in C[\pi, \pi] : x(-\pi) = x(\pi)\}$ is dense in $L^1[-\pi, \pi]$, show that

$$\hat{x}(n) := \frac{1}{2\pi} \int_{-\pi}^{\pi} x(t) e^{-int} d\mu(t) \to 0 \quad \text{as} \quad |n| \to \infty.$$

17. Suppose E and F are measurable subsets of $\mathbb{R}$, and $\{u_1, u_2, \ldots\}$ and $\{v_1, v_2, \ldots\}$ are orthonormal bases of $L^2(E)$ and $L^2(F)$, respectively. Let

$$\phi_{jk}(s,t) = u_j(s) v_k(t), \qquad (s,t) \in E \times F, \qquad (j,k) \in \mathbb{N} \times \mathbb{N}.$$

Show that $\{\phi_{jk} : j, k \in \mathbb{N}\}$ is an orthonormal basis of $L^2(E \times F)$.

18. Let $\{u_1, u_2, \ldots\}$ be an orthonormal basis of $\ell^2(\mathbb{N})$, and for k, ℓ in $\mathbb{N}$, let $w_{k\ell} \in \mathcal{F}(\mathbb{N} \times \mathbb{N}, \mathbb{K})$ be defined by

$$w_{k\ell}(i,j) = u_k(i) \bar{u}_\ell(j), \qquad (i,j) \in \mathbb{N} \times \mathbb{N}.$$

Show that $\{w_{k\ell} : (k, \ell) \in \mathbb{N} \times \mathbb{N}\}$ is an orthonormal basis of $\ell^2(\mathbb{N} \times \mathbb{N})$.

5

Hahn-Banach Theorem and Its Consequences

Recall from Chapter 3 that if X is a normed linear space, then its dual X', the space of all bounded linear functionals on X, is a normed linear space with the norm given by

$$\|f\| = \sup\{\,|f(x)| : x \in X,\ \|x\| = 1\}, \qquad f \in X',$$

and X' is a Banach space with respect to this norm (Theorem 3.12).

If X is finite dimensional, then we know that X' is of the same dimension as of X. What about the situation if X is an infinite dimensional space?

If X is an inner product space, then we know that every $y \in X$ gives rise to a bounded linear functional f_y on X, defined by $f_y(x) = \langle x, y\rangle$ for all $x \in X$ (Example 3.1(vii)). Moreover, in this case, the map $y \mapsto f_y$ is injective (Example 3.3(iii)). We also know that, if X is a Hilbert space, then the above map is a conjugate linear isometry from X onto X' (Theorem 3.9).

But the theory developed so far in this book does not give any clue for the existence of any bounded linear functional on a general normed linear space or a Banach space. Of course, in specific examples, we may be able to give plenty of bounded linear functionals (see Exercise 5.1 below).

Now, let us attempt to find some bounded linear functionals on a general normed linear space.

It is often possible to find bounded linear functionals on subspaces. For instance, if x_0 is a nonzero element in X, then the function $f_0 : \operatorname{span}\{x_0\} \to \mathbb{K}$ defined by

$$f_0(\alpha x_0) = \alpha, \qquad \alpha \in \mathbb{K},$$

is a bounded linear functional on $\operatorname{span}\{x_0\}$. Thus, one would like to get an affirmative answer to the following question: Suppose X

is a normed linear space, X_0 is a subspace of X, and $g : X_0 \to \mathbb{K}$ is a bounded linear functional. Does there exist a bounded linear functional $f : X \to \mathbb{K}$ such that f is an extension of g? We answer this question affirmatively. We, in fact, show that a norm preserving extension is possible, i.e., with $\|f\| = \|g\|$. Before proving this in a general setting, let us look at certain special cases wherein the proofs are obtained rather easily.

(1) Suppose X_0 is a dense subspace of X and $g \in X_0'$. Then by Theorem 3.18, g has a unique extension $f \in X'$, and $\|f\| = \|g\|$. In fact, this f is defined by

$$f(x) = \lim_{n\to\infty} g(x_n), \qquad x \in X,$$

where (x_n) is a sequence in X_0 such that $x = \lim_{n\to\infty} x_n$.

(2) Suppose X is an inner product space, X_0 is a complete subspace of X, and $g \in X_0'$. Since X_0 is a Hilbert space, by the Riesz representation theorem (Theorem 3.9), there exists $v \in X_0$ such that

$$g(x) = \langle x, v\rangle \quad \forall x \in X_0.$$

Then $f : X \to \mathbb{K}$ defined by

$$f(x) = \langle x, v\rangle, \qquad x \in X,$$

is a bounded linear extension of g. Moreover,

$$\|f\| = \|v\| = \|g\|.$$

Obtaining a linear extension of $g \in X_0'$ is an easy task. In fact, every linear functional on X_0 can be extended linearly to all of X (see Exercise 5.2 below). But every linear extension need not be continuous. For instance, suppose X is an infinite dimensional Banach space and X_0 is a finite dimensional subspace of X. Let $f : X \to \mathbb{K}$ be a discontinuous linear functional. We know that such a functional exists (see Remark 3.2(b)). Since every linear functional on a finite dimensional space is continuous (Theorem 3.4), it follows that the restriction $g = f|_{X_0}$ of f is a continuous linear functional on X_0. Clearly, the discontinuous linear functional f is a linear extension of $g \in X_0'$.

The crux of the matter is to obtain an extension which is continuous as well as linear, and which preserves the norm. We shall do this in the next section. This is the *Hahn-Banach extension theorem* (Theorem 5.1).

In Section 5.2, we deduce some important consequences of the extension theorem.

In Section 5.4, we prove another theorem concerning continuous linear functionals, namely, the *Hahn-Banach separation theorem* (Theorem 5.14), which essentially shows the existence of a continuous linear functional that separates two disjoint convex sets, one of which is an open set.

Exercise 5.1 (i) Let X be the space $\mathcal{P}[a,b]$ or $C[a,b]$ with the norm $\|\cdot\|_\infty$. Show that for each $\tau \in [a,b]$, the map $f_\tau : x \mapsto x(\tau)$ is a bounded linear functional on X, and the map $\tau \mapsto f_\tau$ is injective from $[a,b]$ into X'.

(ii) Let $1 \le p \le \infty$. Show that for each $y \in \ell^q$, $\psi \in L^q[a,b]$, the maps

$$f_y : x \mapsto \sum_{i=1}^{\infty} x(i)y(i), \qquad g_\psi : \phi \mapsto \int_a^b \phi(t)\psi(t)\,dt$$

are bounded linear functionals on ℓ^p and $L^p[a,b]$, respectively. Show also that the maps $y \mapsto f_y$ and $\psi \mapsto g_\psi$ are injective. □

Exercise 5.2 Let X_0 be a subspace of a normed linear space X. Show that every linear functional on X_0 can be extended linearly to all of X. □

5.1 The Extension Theorem

Theorem 5.1 (Hahn-Banach extension theorem) *Let X_0 be a subspace of a normed linear space X and let $g \in X_0'$. Then there exists $f \in X'$ such that*

$$f|_{X_0} = g, \qquad \|f\| = \|g\|.$$

We deduce the above theorem from the following generalized version. For its statement, we recall from Chapter 2 that a function $\nu : X \to \mathbb{R}$ defined on a linear space X is a *seminorm* on X if

$$\nu(x+y) \le \nu(x) + \nu(y), \qquad \nu(\alpha x) = |\alpha|\nu(x)$$

for all $x, y \in X$ and for all $\alpha \in \mathbb{K}$. We have already observed that the conditions on ν imply that $\nu(x) \ge 0$ for all $x \in X$ and $\nu(0) = 0$.

Theorem 5.2 *Let X be a linear space and $\nu : X \to \mathbb{R}$ be a seminorm on X. Suppose X_0 is a subspace of X and $g : X_0 \to \mathbb{K}$ is a linear functional satisfying*

$$|g(x)| \le \nu(x) \quad \forall x \in X_0.$$

Then there exists a linear functional $f : X \to \mathbb{K}$ such that

$$f|_{X_0} = g \quad \text{and} \quad |f(x)| \le \nu(x) \quad \forall x \in X.$$

Let us first deduce Theorem 5.1 from Theorem 5.2.

Proof of Theorem 5.1. Define $\nu : X \to \mathbb{K}$ by

$$\nu(x) = \|g\|\,\|x\| \quad \forall x \in X.$$

Clearly, ν is a seminorm on X and

$$|g(x)| \le \nu(x) \quad \forall x \in X_0.$$

Hence, by Theorem 5.2, there exists a linear extension $f : X \to \mathbb{K}$ of g such that

$$|f(x)| \le \nu(x) \quad \forall x \in X.$$

Since $\nu(x) = \|g\|\,\|x\|$, the above relation, in particular, implies that f is continuous and $\|f\| \le \|g\|$. Since f is an extension of g, we have $\|g\| \le \|f\|$. Thus, we have a continuous norm preserving extension of g. ∎

Now we proceed to prove Theorem 5.2. The idea is to prove the theorem first for the case $\mathbb{K} = \mathbb{R}$ and then use the following lemma to obtain the proof for the case $\mathbb{K} = \mathbb{C}$.

We shall introduce certain terminology.

Suppose X is a linear space over $\mathbb{C}$. Then X can be considered as a linear space over $\mathbb{R}$ as well. A function $f : X \to \mathbb{R}$ is called a **real-linear functional** if it is linear considering X as a linear space over $\mathbb{R}$, and a function $f : X \to \mathbb{C}$ is called a **complex-linear functional** if it is linear considering X as a linear space over $\mathbb{C}$.

Lemma 5.3 *Let X be a linear space over $\mathbb{C}$. Then we have the following:*

(i) *Let f be a complex-linear functional on X and $\phi : X \to \mathbb{R}$ be defined by*

$$\phi(x) = \operatorname{Re} f(x), \qquad x \in X.$$

Then ϕ is a real-linear functional on X, and

$$f(x) = \phi(x) - i\phi(ix), \qquad x \in X.$$

(ii) *Let ψ be a real-linear functional on X and $f : X \to \mathbb{C}$ be defined by*

$$f(x) = \psi(x) - i\psi(ix), \qquad x \in X.$$

Then f is a complex-linear functional on X.

(iii) *Let $\nu : X \to \mathbb{R}$ be a seminorm on X and f be a complex-linear functional on X. Then*

$$|f(x)| \le \nu(x) \quad \forall x \in X \iff |\mathrm{Re}\, f(x)| \le \nu(x) \quad \forall x \in X.$$

Proof. (i) Clearly, $\phi : X \to \mathbb{R}$ is real-linear. Let $h(x) := \mathrm{Im}\, f(x)$, $x \in X$. Then $f(x) = \phi(x) + ih(x)$ so that

$$f(ix) = \phi(ix) + i\,h(ix),$$

$$if(x) = i\phi(x) - h(x).$$

Since $f(ix) = if(x)$, it follows that $h(x) = -\phi(ix)$ for all $x \in X$. Hence,

$$f(x) = \phi(x) - i\phi(ix), \qquad x \in X.$$

(ii) It can be easily seen that f is real-linear, and $\mathrm{Re}\, f(x) = \psi(x)$ for every $x \in X$. Also,

$$\begin{aligned} f(ix) &= \psi(ix) - i\psi(-x) \\ &= \psi(ix) + i\psi(x) \\ &= i[\psi(x) - i\psi(ix)] \\ &= if(x). \end{aligned}$$

From this, it can be seen (*How?*) that f is a complex-linear functional.

(iii) Since $|\mathrm{Re}\, f(x)| \le |f(x)|$ for every $x \in X$, we have

$$|f(x)| \le \nu(x) \quad \forall x \in X \implies |\mathrm{Re}\, f(x)| \le \nu(x) \quad \forall x \in X.$$

Next, suppose $|\mathrm{Re}\, f(x)| \le \nu(x)$ for all $x \in X$. Now, let $x \in X$, and let $\alpha \in \mathbb{C}$ be such that $|\alpha| = 1$ and $|f(x)| = \alpha f(x)$. Then

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

Thus, $|f(x)| = \operatorname{Re} f(\alpha x)$. Therefore, using the hypothesis that $|\operatorname{Re} f(x)| \leq \nu(x)$ for all $x \in X$, we have

$$|f(x)| = |\operatorname{Re} f(\alpha x)| \leq \nu(\alpha x) = |\alpha|\nu(x) = \nu(x).$$

This completes the proof. ∎

From the above lemma it follows that the imaginary part of a linear functional is completely determined by its real part. Thus, if f_1 and f_2 are real-linear functionals on a complex-linear space, then the function f, defined by $f(x) = f_1(x) + if_2(x)$, $x \in X$, need not be a complex-linear functional. But the function does satisfy the relations

$$f(x+y) = f(x) + f(y), \qquad f(\alpha x) = \alpha f(x)$$

for all $x, y \in X$ and $\alpha \in \mathbb{R}$.

Proof of Theorem 5.2. Case (i): Suppose $\mathbb{K} = \mathbb{R}$. If $X_0 = X$ or $g = 0$, then we may take $f = g$. Now, suppose that $X_0 \neq X$ and $g \neq 0$. Let $u \in X \setminus X_0$, and let

$$X_u := \operatorname{span}\{u; X_0\} = \{x + \alpha u : x \in X_0, \ \alpha \in \mathbb{R}\}.$$

First we find a linear extension $f_u : X_u \to \mathbb{R}$ of g such that

$$|f_u(x)| \leq \nu(x) \quad \forall x \in X_u.$$

Assume for a moment that such an extension exists. Then its value at $x + \alpha u$ for $x \in X_0$ and $\alpha \in \mathbb{R}$ would be $g(x) + \alpha f_u(u)$. Now, we define $f_u : X_u \to \mathbb{R}$ by

$$f_u(x + \alpha u) = g(x) + \alpha c \quad \forall x \in X_0, \qquad \alpha \in \mathbb{R},$$

where $c \in \mathbb{R}$ is yet to be specified. Clearly, this f_u is a linear extension of g. The scalar $c \in \mathbb{R}$ is to be chosen in such a way that

$$|f_u(x + \alpha u)| \leq \nu(x + \alpha u) \quad \forall x \in X_0, \qquad \alpha \in \mathbb{R},$$

i.e.,

$$|g(x) + \alpha c| \leq \nu(x + \alpha u) \quad \forall x \in X_0, \qquad \alpha \in \mathbb{R}. \tag{5.1}$$

Clearly, (5.1) is satisfied for $\alpha = 0$. We note that (5.1) is satisfied for a c with $\alpha \neq 0$ if and only if

$$|g(x) - c| \leq \nu(x - u) \quad \forall x \in X_0. \tag{5.2}$$

Indeed, (5.2) follows from (5.1) by taking $\alpha = -1$, and (5.1) follows from (5.2) by replacing x by $x/(-\alpha)$. Now, (5.2) is satisfied if and only if

$$g(x) - \nu(x-u) \le c \quad \forall x \in X_0, \qquad c \le g(x) + \nu(x-u) \quad \forall x \in X_0,$$

i.e. if and only if

$$g(x) - \nu(x-u) \le c \le g(y) + \nu(y-u) \quad \forall x, y \in X_0. \tag{5.3}$$

Thus, it is enough to find $c \in \mathbb{R}$ such that the relation (5.3) is satisfied.

We note that for every $x, y \in X_0$,

$$g(x) - g(y) = g(x-y) \le \nu(x-y) \le \nu(x-u) + \nu(y-u)$$

so that

$$g(x) - \nu(x-u) \le g(y) + \nu(y-u) \quad \forall x, y \in X_0.$$

From this we have

$$\sup\{g(x) - \nu(x-u) : x \in X_0\} \le \inf\{g(x) + \nu(x-u) : x \in X_0\}.$$

Thus, if $c \in \mathbb{R}$ is such that

$$\sup\{g(x) - \nu(x-u) : x \in X_0\} \le c \le \inf\{g(x) + \nu(x-u) : x \in X_0\},$$

then for such c, (5.3) is satisfied.

We shall use the above result, along with Zorn's lemma, to obtain a linear extension $f : X \to \mathbb{R}$ of g such that $|f(x)| \le \nu(x)$ for every $x \in X$. For this purpose, consider the family $\mathcal{F}$ of all pairs (h, X_h), where X_h is a subspace of X such that $X_0 \subseteq X_h$ and $h : X_h \to \mathbb{R}$ is a linear extension of g such that $|h(x)| \le \nu(x)$ for all $x \in X_h$. For (h_1, X_{h_1}), (h_2, X_{h_2}) in $\mathcal{F}$, define

$$(h_1, X_{h_1}) \preccurlyeq (h_2, X_{h_2}) \quad \text{whenever} \quad X_{h_1} \subseteq X_{h_2},\ h_{2|X_{h_1}} = h_1.$$

It can be seen that $\mathcal{F}$ is a partially ordered set with partial order $\preccurlyeq$. Suppose $\mathcal{T}$ is a totally ordered subset of $\mathcal{F}$. Then consider

$$Y = \cup\{X_h : (h, X_h) \in \mathcal{T}\},$$

and define $\phi : Y \to \mathbb{R}$ such that

$$\phi(x) = h(x) \text{ whenever } x \in X_h, (h, X_h) \in \mathcal{T}.$$

Clearly, $(\phi, Y) \in \mathcal{F}$, and (ϕ, Y) is an upper bound of $\mathcal{T}$. Therefore, by Zorn's lemma, $\mathcal{F}$ has a maximal element, say (f, X_f). Now, it is enough to show that $X_f = X$.

Suppose $X_f \neq X$, and let $x_0 \in X \setminus X_f$. Then, by the first part of the proof, f has a linear extension, say $\tilde{f}$ to $\widetilde{X}_f := \operatorname{span}\{x_0; X_f\}$ satisfying $|\tilde{f}(x)| \leq \nu(x)$ for all $x \in \widetilde{X}_f$. Thus, we have

$$(f, X_f) \preccurlyeq (\tilde{f}, \tilde{X}_f) \in \mathcal{F}, \qquad (f, X_f) \neq (\tilde{f}, \tilde{X}_f)$$

contradicting the maximality of (f, X_f). Therefore, $X_f = X$, and f is a linear extension of g satisfying $|f(x)| \leq \nu(x)$ for all $x \in X$.

Case (ii): Suppose $\mathbb{K} = \mathbb{C}$. Let $g_0 : X_0 \to \mathbb{R}$ be defined by $g_0(x) = \operatorname{Re} g(x)$, $x \in X_0$. Then by Lemma 5.3, g_0 is a real-linear functional on X satisfying $|g_0(x)| \leq \nu(x)$ for all $x \in X_0$. Therefore, by what we have already proved in case (i), g_0 has a linear extension $f_0 : X \to \mathbb{R}$ satisfying $|f_0(x)| \leq \nu(x)$ for all $x \in X$. Again, by Lemma 5.3, the function $f : X \to \mathbb{C}$ defined by

$$f(x) = f_0(x) - i f_0(ix), \qquad x \in X,$$

is a complex-linear functional satisfying $|f(x)| \leq \nu(x)$ for all $x \in X$. This f is an extension of g since, for every $x \in X_0$,

$$f(x) = f_0(x) - i f_0(ix) = g_0(x) - i g_0(ix) = g(x).$$

This completes the proof of the theorem. ∎

Let X be a linear space. A function $p : X \to \mathbb{R}$ is said to be a **sublinear functional** or a **convex functional** if, for all $x, y \in X$ and $t \geq 0$,

$$p(x + y) \leq p(x) + p(y), \qquad p(tx) = tp(x).$$

If the linear space is over $\mathbb{R}$, then the following result can be proved by a slight modification of the proof of Theorem 5.2 (see Problem 10 at the end of the chapter).

Theorem 5.4 *Let X be a linear space over $\mathbb{R}$ and $p : X \to \mathbb{R}$ be a sublinear functional on X. Suppose X_0 is a subspace of X and $g : X_0 \to \mathbb{R}$ is a linear functional such that*

$$g(x) \le p(x) \quad \forall x \in X_0.$$

Then there exists a linear extension $f : X \to \mathbb{R}$ of g such that

$$f(x) \le p(x) \quad \forall x \in X.$$

5.2 Consequences

Now we consider some of the immediate corollaries and important consequences of Theorem 5.1.

Corollary 5.5 *Suppose X_0 is a closed subspace of a normed linear space X and $x_0 \in X \setminus X_0$. Then there exists $f \in X'$ such that $f(x) = 0$ for all $x \in X_0$, and*

$$f(x_0) = \operatorname{dist}(x_0, X_0), \qquad \|f\| = 1.$$

Proof. Let $X_1 = \operatorname{span}\{x_0; X_0\}$, i.e.,

$$X_1 = \{x + \alpha x_0 : x \in X_0, \alpha \in \mathbb{K}\}.$$

Define $g : X_1 \to \mathbb{K}$ by

$$g(x + \alpha x_0) = \alpha \operatorname{dist}(x_0, X_0), \qquad x \in X_0,\ \alpha \in \mathbb{K}.$$

Clearly, g is linear, $g(x) = 0$ for all $x \in X_0$, and $g(x_0) = \operatorname{dist}(x_0, X_0)$. Since $x_0 \notin X_0$ and X_0 is closed, we have $\operatorname{dist}(x_0, X_0) > 0$. Also, observe that the null space of g is X_0. Hence, by Theorem 3.8, g is a bounded linear functional on X_1 and

$$\|g\| = \frac{g(x_0)}{\operatorname{dist}(x_0, X_0)} = 1.$$

Now by Theorem 5.1, there exists $f \in X'$ satisfying $f(x) = g(x)$ for all $x \in X_1$ and $\|f\| = \|g\|$. In particular, $f(x) = 0$ for all $x \in X_0$,

$$f(x_0) = g(x_0) = \operatorname{dist}(x_0, X_0), \qquad \|f\| = \|g\| = 1.$$

This completes the proof. ∎

Corollary 5.6 *Let X be a normed linear space and $x \in X$. Then there exists $f \in X'$ such that*

$$f(x) = \|x\|, \qquad \|f\| = 1.$$

Proof. If $x = 0$, then we may take f to be any element in X' of norm 1. If $x \neq 0$, then the existence of $f \in X'$ satisfying the requirements is guaranteed by the last corollary with $X_0 = \{0\}$. ∎

Exercise 5.3 Show that if $x \in X$ is such that $f(x) = 0$ for every $f \in X'$, then $x = 0$. □

Corollary 5.7 *Let X be a normed linear space and $\{x_1, \ldots, x_n\}$ be linearly independent in X. Then then there exist $f_1, \ldots, f_n$ in X' such that*

$$f_j(x_i) = \delta_{ij} \quad \forall\, i, j = 1, \ldots, n.$$

Proof. Let $X_0 := \operatorname{span}\{x_1, \ldots, x_n\}$. Then $E_0 := \{x_1, \ldots, x_n\}$ is a basis of X_0. Let $g_1, \ldots, g_n$ be the coordinate functionals on X_0 corresponding to the basis E_0. By their definition, we have $g_j(x_i) = \delta_{ij}$ for $i, j = 1, \ldots, n$. Since X_0 is a finite dimensional space, each g_j is a continuous. Hence, by Theorem 5.1, there exist $f_1, \ldots, f_n$ in X' such that $f_j(x) = g_j(x)$ for every $x \in X_0$, $j = 1, \ldots, n$. In particular, we have $f_j(x_i) = g_j(x_i) = \delta_{ij}$ for all $i, j = 1, \ldots, n$. ∎

Another Proof. (Derived from Corollary 5.5) For $j \in \{1, \ldots, n\}$, let

$$X_j = \operatorname{span}\{x_i : i = 1, \ldots, n,\ i \neq j\}.$$

Since each subspace X_j is a closed subspace of X (Theorem 2.25) and $x_j \notin X_j$, by Corollary 5.5, there exists $g_j \in X'$ such that

$$g_j(x_j) = \operatorname{dist}(x_j, X_j) \quad \text{and} \quad g_j(x) = 0 \quad \forall\, x \in X_j,\ j = 1, \ldots, n.$$

The functionals

$$f_j = \frac{g_j}{g_j(x_j)}, \qquad j = 1, \ldots, n,$$

satisfy the requirements. ∎

Exercise 5.4 For every normed linear space X, the dual space X' is nontrivial and $\dim X' \geq \dim X$. □

Corollary 5.8 *Let X be a normed linear space and X_0 be a finite dimensional subspace of X. Then there exists a closed subspace X_1 of X such that*

$$X = X_0 + X_1, \qquad X_0 \cap X_1 = \{0\}.$$

In fact, if $\{x_1, \ldots, x_n\}$ is a basis of X_0, and if $f_1 \ldots, f_n$ in X' are as in Corollary 5.7, then

$$X_1 = \bigcap_{j=1}^{n} N(f_j),$$

and for every $x \in X$, $x = y + z$, where

$$y = \sum_{j=1}^{n} f_j(x)x_j, \qquad z = x - y.$$

Proof. The proof involves only verifications, and hence it is left as an exercise for the reader. ∎

Remark 5.1 Suppose that X, X_0 and X_1 are as in the above corollary. Consider $P : X \to X$ defined by

$$P(x) = \sum_{j=1}^{n} f_j(x)x_j, \qquad x \in X.$$

Then we see that P is a bounded linear operator satisfying $P^2 = P$ with

$$R(P) = X_0, \qquad N(P) = X_1.$$

In other words, $P \in \mathcal{B}(X, Y)$ and is a projection onto X_0 along X_1.

A closed subspace X_0 of a normed linear space X is said to have a **complementary subspace** if there exists a closed subspace X_1 of X such that

$$X = X_0 + X_1, \qquad X_0 \cap X_1 = \{0\},$$

and, in that case, X_0 and X_1 are called **complementary subspaces**.

What we have seen in Corollary 5.8 is that every finite dimensional subspace of a normed linear space has a complementary subspace. We have also seen in Theorem 2.47 that every complete subspace of an inner product space has a complementary subspace. But, in general, a closed subspace of a Banach space need not have a complementary subspace. It has been proved by Phillips [25] that the closed subspace c_0 of ℓ^∞ does not have a complementary subspace in ℓ^∞. For a straightforward proof of this result, one may refer Whitely [35]. In fact, every infinite dimensional Banach space which is not linearly homeomorphic with a Hilbert space has at least one closed subspace having no complementary subspace (cf. Lindenstrauss and Tzafriri [22]).

Theorem 5.9 *For every x in a normed linear space X,*

$$\|x\| = \sup\{|f(x)| : f \in X', \ \|f\| \le 1\}.$$

Proof. Clearly, if $f \in X'$ is such that $\|f\| = 1$, then $|f(x)| \le \|x\|$. Therefore,

$$\sup\{|f(x)| : f \in X', \|f\| \le 1\} \le \|x\|.$$

On the other hand, by Corollary 5.6, for every $x \in X$, there exists $f_x \in X'$ such that $\|f_x\| = 1$ and $f_x(x) = \|x\|$. Therefore,

$$\|x\| = |f_x(x)| \le \sup\{|f(x)| : f \in X', \|f\| \le 1\}.$$

This completes the proof. ∎

We recall from Chapter 3 that if $A \in \mathcal{B}(X,Y)$, then the *transpose* of A is the operator $A' : Y' \to X'$ defined by

$$(A'f)(x) = f(Ax) \quad \forall f \in Y', \quad x \in X.$$

We have already seen that $A' \in \mathcal{B}(Y', X')$ and $\|A'\| \le \|A\|$.

Theorem 5.10 *Let X and Y be a normed linear spaces, and $A \in \mathcal{B}(X,Y)$. Then $\|A'\| = \|A\|$.*

Proof. We have already observed that $\|A'\| \le \|A\|$. Now by Theorem 5.9, for every $x \in X$, we have

$$\|Ax\| = \sup\{|f(Ax)| : f \in Y', \|f\| \le 1\}.$$

Hence, using the definition of A', we have

$$\begin{aligned}\|Ax\| &= \sup\{|f(Ax)| : f \in Y', \|f\| \le 1\}\\ &= \sup\{|(A'f)(x)| : f \in Y', \|f\| \le 1\}\\ &\le \|A'\|\,\|x\|\end{aligned}$$

for every $x \in X$, showing that $\|A\| \le \|A'\|$. Thus, $\|A'\| = \|A\|$. ∎

Completion of a normed linear space revisited

Let X be a normed linear space. For each $x \in X$, consider the *evaluation functional* φ_x defined on X', i.e.,

$$\varphi_x(f) = f(x), \qquad f \in X'.$$

It is clear that φ_x is a linear functional on X'. Also we see that

$$|\varphi_x(f)| = |f(x)| \leq \|x\| \, \|f\| \quad \forall f \in X',$$

so that

$$\varphi_x \in (X')', \qquad \|\varphi_x\| \leq \|x\|.$$

NOTATION: We shall denote the space $(X')'$, the dual of the dual space X', by X'', and call it the **second dual** of X. Similarly, we can define the third dual, the fourth dual, and so on.

Now let $J : X \to X''$ be defined by

$$J(x) = \varphi_x, \qquad x \in X.$$

Then, from the above observation, we see that J is a linear operator and $\|J(x)\| \leq \|x\|$ for all $x \in X$. Now we show that J is, in fact, an isometry.

Exercise 5.5 Let X be a normed linear space. Show that the map $J : X \to X''$ defined above is injective. □

Theorem 5.11 *Let X be a normed linear space and $J : X \to X''$ be defined by $J(x) = \varphi_x$, where*

$$\varphi_x(f) = f(x) \quad \forall f \in X', \; \forall x \in X.$$

Then J is a linear isometry.

Proof. We have already observed that $J : X \to X''$ defined by $J(x) = \varphi_x$, where

$$\varphi_x(f) = f(x) \quad \forall f \in X', \; \forall x \in X,$$

is a bounded linear operator and $\|J(x)\| \leq \|x\|$ for every $x \in X$. Hence, it is enough to show that $\|x\| \leq \|J(x)\|$ for every $x \in X$. Now, by Corollary 5.6, for every $x \in X$, there exists $f_x \in X'$ such that $\|f_x\| = 1$ and $f_x(x) = \|x\|$. Hence,

$$\|x\| = |f_x(x)| = |(Jx)(f_x)| \leq \|Jx\| \quad \forall x \in X.$$

This completes the proof. ∎

The isometry $J : X \to X''$ obtained in Theorem 5.11 is known as the **canonical linear isometry** from X into X''.

A normed linear space X is said to be a **reflexive space**, if the canonical linear isometry $J : X \to X''$ is surjective, i.e., if for every $\varphi \in X''$ there exists $x \in X$ such that

$$\varphi(f) = f(x) \quad \forall f \in X'.$$

Clearly, every reflexive space is a Banach space. Reflexive spaces are considered in more detail in Section 8.2.

Since X'' is a Banach space for every normed linear space X, the closure of the range of the canonical linear isometry $J : X \to X''$ is a *completion* of X.

Exercise 5.6 Let X be a normed linear space and $\Omega := \{f \in X' : \|f\| = 1\}$. For each $x \in X$, let $\phi_x : \Omega \to \mathbb{K}$ be defined by $\phi_x(f) = f(x)$, $f \in \Omega$. Show that $\phi_x \in C(\Omega)$ for every $x \in X$, and the map $T : x \to \phi_x$ is a linear isometry from X into $C(\Omega)$. $\square$

The above exercise gives a completion of a normed linear space X as a closed subset of $C(\Omega)$. The next theorem gives another completion of a normed linear space.

Theorem 5.12 *Let X be a normed linear space and Ω be a dense subset of X. Then X is linearly isometric with a subspace of $\ell^\infty(\Omega)$.*

Proof. For each $u \in \Omega$, by Corollary 5.6, there exists $f_u \in X'$ such that

$$f_u(u) = \|u\|, \qquad \|f_u\| = 1.$$

Now, for $x \in X$, let $\phi_x : \Omega \to \mathbb{K}$ be defined by

$$\phi_x(u) = f_u(x), \qquad u \in \Omega.$$

Note that for every $x \in X$,

$$|\phi_x(u)| = |f_u(x)| \le \|f_u\| \, \|x\| = \|x\| \quad \forall u \in \Omega$$

so that $\phi_x \in \ell^\infty(\Omega)$ and $\|\phi_x\|_\infty \le \|x\|$ for every $x \in X$. Next, consider the function $T : X \to \ell^\infty(\Omega)$ defined by

$$Tx = \phi_x, \qquad x \in X.$$

Then it is easily seen that T is a linear operator. Moreover,

$$\|Tx\| = \|\phi_x\|_\infty \leq \|x\| \quad \forall x \in X.$$

Thus, the proof will be completed once we show that $\|\phi_x\|_\infty = \|x\|$ for every $x \in X$. For this let $x \in X$. We observe that, for every $u \in \Omega$,

$$\begin{aligned}
| \|x\| - |\phi_x(u)| | &\leq | \|x\| - |\phi_u(u)| | + | |\phi_u(u)| - |\phi_x(u)| | \\
&\leq | \|x\| - |f_u(u)| | + |f_u(u) - f_x(u)| \\
&= | \|x\| - \|u\| | + |f_u(u) - f_u(x)| \\
&= \|x - u\| + |f_u(u - x)| \\
&\leq \|x - u\| + \|f_u\| \, \|u - x\| \\
&= 2\,\|x - u\|.
\end{aligned}$$

Therefore, for every $u \in \Omega$, $\|x\| \leq |\phi_x(u)| + 2\,\|x - u\|$ so that

$$\|x\| \leq \|\phi_x\|_\infty + 2\,\|x - u\| \quad \forall u \in \Omega.$$

Since Ω is dense in X, it follows that $\|x\| \leq \|\phi_x\|_\infty$. ∎

Exercise 5.7 Every separable normed linear space is linearly isometric with a subspace of $\ell^\infty(\mathbb{N})$. □

The following theorem is a converse of Theorem 3.12.

Theorem 5.13 *Let X and Y be normed linear spaces such that $X \neq \{0\}$, and let $\mathcal{B}(X, Y)$ be a Banach space. Then Y is a Banach space.*

Proof. Let (y_n) be a Cauchy sequence in Y and let $0 \neq x_0 \in X$. By Corollary 5.6, there exists $f \in X'$ such that $f(x_0) = \|x_0\|$ and $\|f\| = 1$. Now consider $F_n : X \to Y$ defined by

$$F_n(x) = f(x)y_n, \qquad x \in X.$$

Clearly, F_n is linear and $\|F_n(x)\| \leq \|x\|\,\|y_n\|$ for every $x \in X$ so that $F_n \in \mathcal{B}(X, Y)$ for each n. Moreover,

$$\|(F_n - F_m)(x)\| \leq \|x\|\,\|y_n - y_m\| \quad \forall x \in X,$$

$$\|(F_n - F_m)(x_0)\| = \|f(x_0)(y_n - y_m)\| = \|x_0\|\,\|y_n - y_m\|.$$

Thus, we get $\|F_n - F_m\| = \|y_n - y_m\|$ for all $n, m \in \mathbb{N}$. This shows that (F_n) is a Cauchy sequence in $\mathcal{B}(X, Y)$. Since $\mathcal{B}(X, Y)$ is a Banach space, there exists $F \in \mathcal{B}(X, Y)$ such that $\|F_n - F\| \to 0$ as $n \to \infty$. In particular, $(F_n(x_0))$ converges to $F(x_0)$. That is,

$$f(x_0)y_n = F_n(x_0) \to F(x_0) \quad \text{as} \quad n \to \infty.$$

Since $f(x_0) \neq 0$, it follows that (y_n) converges in Y. ∎

5.3 On Uniqueness of Extension

We have not addressed so far the question whether the extension obtained in the Hahn-Banach Theorem is unique. In fact, the Hahn-Banach extension need not be unique even for a finite dimensional normed linear space as the following example shows. By a *Hahn-Banach extension* of $g \in X_0'$, we mean a norm preserving continuous linear extension $f : X \to \mathbb{K}$ of g.

EXAMPLE 5.1 Let $X = \mathbb{K}^2$ with $\|\cdot\|_1$, and

$$X_0 = \{x \in X : x(2) = 0\}.$$

Let $g : X_0 \to \mathbb{K}$ be defined by

$$g(x) = x(1), \qquad x \in X_0.$$

Then it is clear that $g \in X_0'$ and $\|g\| = 1 = g(x_0)$, where $x_0 = (1, 0)$. Let $f \in X'$. Now, f is of the form

$$f(x) = ax(1) + bx(2), \qquad x \in X$$

for some $a, b \in \mathbb{K}$ (*Why?*), and its norm is given by (see Example 3.3(v))

$$\|f\| = \max\{|a|, |b|\}.$$

For the above f to be a Hahn-Banach extension of g, it is necessary and sufficient that

$$f(x_0) = g(x_0) = 1, \qquad \|f\| = \|g\| = 1.$$

Thus, it follows that for any b with $|b| \leq 1$, $f : X \to \mathbb{K}$ defined by

$$f(x) = x(1) + bx(2), \qquad x \in X,$$

is a Hahn-Banach extension of g.

Next, we give examples where the Hahn-Banach extension is unique.

EXAMPLE 5.2 (i) We have remarked in the beginning of this chapter that if X_0 is a dense subspace of a normed linear space X, then the Hahn-Banach extension is unique.

(ii) If X is a Hilbert space, then a Hahn-Banach extension of every bounded linear functional on a subspace is unique:

Suppose X is a Hilbert space. Assume first that X_0 is a closed subspace of X. Let $g : X_0 \to \mathbb{K}$ be a bounded linear functional on X_0. We have already observed in the beginning of the chapter that g has a continuous norm preserving extension to all of X.

Suppose f_1 and f_2 are continuous norm preserving extensions of g. Then, by the Riesz representation theorem (Theorem 3.9), there exist $v_1, v_2 \in X$ such that

$$f_1(x) = \langle x, v_1 \rangle, \qquad f_2(x) = \langle x, v_2 \rangle$$

for all $x \in X$, and

$$\|v_1\| = \|f_1\| = \|g\|, \qquad \|v_2\| = \|f_2\| = \|g\|.$$

It is clear that the function $f : X \to \mathbb{K}$ defined by

$$f(x) = \left\langle x, \frac{v_1 + v_2}{2} \right\rangle, \qquad x \in X,$$

is also a continuous linear extension of g. Moreover,

$$\|g\| \leq \|f\| = \left\| \frac{v_1 + v_2}{2} \right\| \leq \frac{1}{2} (\|v_1\| + \|v_2\|) = \|g\|.$$

In particular,

$$\left\| \frac{v_1 + v_2}{2} \right\| = \|f\| = \|g\|.$$

Now, by the parallelogram law, we have

$$\|v_1 + v_2\|^2 + \|v_1 - v_2\|^2 = 2 \left(\|v_1\|^2 + \|v_2\|^2 \right)$$

so that

$$4 \|g\|^2 + \|v_1 - v_2\|^2 = 4 \|g\|^2$$

showing that $v_1 = v_2$. Consequently, $f_1 = f_2$.

In case X_0 is not closed in X, we may extend g to the closure of X_0 so as to obtain a unique norm preserving continuous linear extension $\tilde{g} \in (\overline{X_0})'$ as in example (i) above, and then $\tilde{g}$ can be uniquely extended to all of X.

In example (ii) above, we may observe that

$$f = \frac{f_1 + f_2}{2}, \qquad \|f\| = \|f_1\| = \|f_2\|.$$

If we take, without loss of generality, $\|g\| = 1$, then what we have essentially used is the fact that the dual of a Hilbert space is strictly convex.

Exercise 5.8 Show that Hilbert spaces and the spaces $\ell^p(n)$, ℓ^p and $L^p[a,b]$ for $1 < p < \infty$ are strictly convex spaces. □

EXAMPLE 5.2 (*cont.*) (iii) Let X be a normed linear space such that the dual space X' is strictly convex, X_0 be a subspace of X, and $g : X_0 \to \mathbb{K}$ be a continuous linear functional on X_0. We show that g has a unique Hahn-Banach extension.

Suppose f_1 and f_2 are Hahn-Banach extensions of g. Without loss of generality, assume that $\|g\| = 1$ so that $\|f_1\| = 1 = \|f_2\|$. Then it follows that $f := (f_1 + f_2)/2$ is an extension of g and it satisfies

$$1 = \|g\| \leq \|f\| \leq \frac{\|f_1\| + \|f_2\|}{2} \leq 1$$

so that $\|f\| = 1$. Therefore, by the strict convexity of X', we have $f_1 = f_2$.

We shall show in Chapter 8 that the duals of the spaces $\ell^p(n)$, ℓ^p and $L^p[a,b]$ for $1 < p < \infty$ are linearly isometric with $\ell^q(n)$, ℓ^q and $L^q[a,b]$, respectively. Thus, in view of Exercise 5.8, in each of these spaces, we have a unique Hahn-Banach extension.

5.4 Separation Theorem

Now we consider the separation theorem.

Theorem 5.14 (Hahn-Banach separation theorem) *Let X be a normed linear space and E_1 and E_2 be nonempty disjoint convex subsets of X with E_1 being an open set. Then there exists $f \in X'$ and a real number γ such that*

$$E_1 \subseteq \{x \in X : \mathrm{Re} f(x) < \gamma\}, \qquad E_2 \subseteq \{x \in X : \mathrm{Re} f(x) \geq \gamma\}.$$

Before giving its proof, let us see how the nomenclature "separation theorem" is given to the above theorem.

Recall from Section 1.2.3 that a proper subspace X_0 of a linear space X is called a *hyperspace* if there is $u \in X \setminus X_0$ such that $\operatorname{span}\{u; X_0\} = X$.

A subset H of a linear space X is called a **hyperplane** if it is a translation of a hyperspace by a vector, i.e., if H is of the form

$$H = x + X_0,$$

where X_0 is a hyperspace and $x \in X$.

We may also recall a characterization (Theorem 1.12): A proper subspace X_0 of a linear space X is a hyperspace if and only if it is the null space of a nonzero linear functional on X. From this we obtain a characterization of a hyperplane: A subset $H \subseteq X$ is a hyperplane in X if and only if there exists a nonzero linear functional f and a scalar α such that

$$H = H_{\alpha,f} := \{x \in X : f(x) = \alpha\}$$

since

$$\{x \in X : f(x) = \alpha\} = x_\alpha + N(f)$$

for some $x_\alpha \in X$ with $f(x_\alpha) = \alpha$. Thus, hyperplanes are of the form $H_{\alpha,f}$ for some nonzero linear functional f and for some $\alpha \in \mathbb{K}$.

If the scalar field is $\mathbb{R}$, then we may say that a set E is on the *left side of* a hyperplane $H_{\alpha,f}$ if

$$E \subseteq \{x \in X : f(x) \leq \alpha\},$$

and it is *strictly on the left side* of $H_{\alpha,f}$ if

$$E \subseteq \{x \in X : f(x) < \alpha\}.$$

Similarly, we may say that E is on the *right side* (respectively, *strictly on the right side*) of the hyperplane $H_{\alpha,f}$ if $\leq$ (respectively, $<$) above is replaced by $\geq$ (respectively, $>$).

Thus, the conclusion of Theorem 5.14 can be interpreted as: there exists a hyperplane H which *separates* E_1 and E_2 such that E_1 is *strictly on the left side* of H and E_2 is *on the right side* of H.

For proving Theorem 5.14, we shall make use of Theorem 5.4 and a lemma.

We use a few terminologies in the lemma: Let X be a linear space. A set $E \subseteq X$ is said to be an **absorbing set** if for every $x \in X$, there exists $t > 0$ such that $t^{-1}x \in E$.

Let $E \subseteq X$ be a convex, absorbing set. Then $\mu_E : X \to \mathbb{R}$ defined by

$$\mu_E(x) := \inf\{t > 0 : t^{-1}x \in E\}$$

is called a **Minkowski functional** of E.

Since E is absorbing, it is clear that $\mu_E(x) < \infty$ for every $x \in X$.

We may observe (*Verify*) that, if E is an absorbing set, then $0 \in E$, and if X is a normed linear space, then every open set containing 0 is an absorbing set.

Lemma 5.15 *Let E be a convex, absorbing subset of a linear space X and let μ_E be the corresponding Minkowski functional. Then μ_E is a convex functional, and*

$$\{x \in X : \mu_E(x) < 1\} \subseteq E \subseteq \{x \in X : \mu_E(x) \leq 1\}.$$

Proof. For μ_E to be a convex functional, it has to satisfy the properties

$$\mu_E(x+y) \leq \mu_E(x) + \mu_E(y), \qquad \mu_E(\tau x) = \tau\mu_E(x)$$

for all $x, y \in X$ and for all $\tau \geq 0$.

Let $x, y \in X$. Let $s > 0$, $t > 0$ be such that $s^{-1}x \in E$, $t^{-1}y \in E$. Then, using the convexity of E, we have

$$(s+t)^{-1}(x+y) = \left(\frac{s}{s+t}\right)s^{-1}x + \left(\frac{t}{s+t}\right)t^{-1}y \in E.$$

Hence, $\mu_E(x+y) \leq s+t$. Now, taking infimum over all such s and t, it follows that

$$\mu_E(x+y) \leq \mu_E(x) + \mu_E(y).$$

Now, we show that $\mu_E(\tau x) = \tau\mu_E(x)$ for all $x \in X$ and for all $\tau \geq 0$. For this, let $x \in X$ and $\tau > 0$. Let $t > 0$ be such that $t^{-1}x \in E$. Since $t^{-1}x = (\tau t)^{-1}(\tau x)$, it follows that $\mu_E(\tau x) \leq \tau t$. Taking infimum over all $t > 0$ with $t^{-1}x \in E$, we get

$$\mu_E(\tau x) \leq \tau\mu_E(x).$$

Next, taking τx in place of x, let $s > 0$ be such that $s^{-1}(\tau x) \in E$. Since $s^{-1}(\tau x) = (\tau^{-1}s)^{-1}x$, we have $\mu_E(x) \leq \tau^{-1}s$. Taking infimum over all such $s > 0$, we obtain

$$\mu_E(x) \leq \tau^{-1}\mu_E(\tau x),$$

i.e., $\tau\mu_E(x) \leq \mu_E(\tau x)$. Thus, we have proved that

$$\mu_E(\tau x) = \tau\mu_E(x).$$

Now, we prove the last part of the theorem. For this, suppose $x \in E$. Then we have $1 \in \{t > 0 : t^{-1}x \in E\}$ so that $\mu_E(x) \leq 1$. Thus, $E \subseteq \{x \in X : \mu_E(x) \leq 1\}$.

Next, suppose that $x \in X$ is such that $\mu_E(x) < 1$. Then there exists $t_0 > 0$ such that $t_0 < 1$ with $t_0^{-1}x \in E$. Since E is convex and $0 \in E$, we have

$$x = t_0\left(t_0^{-1}x\right) + (1 - t_0)\, 0 \in E.$$

Thus, $\mu_E(x) < 1$ implies $x \in E$. This completes the proof. ∎

Proof of Theorem 5.14. We prove for the case of $\mathbb{K} = \mathbb{R}$. The case of $\mathbb{K} = \mathbb{C}$ follows by using Lemma 5.3.

Let $u_1 \in E_1$, $u_2 \in E_2$. Take $E = E_1 - E_2 + x_0$, where $x_0 := u_2 - u_1$. Then it can be seen that E is an open convex set with $0 \in E$ (see Exercise 5.9(i) below). Hence, E is an absorbing set as well. Let μ_E be the Minkowski functional of E.

In order to obtain the required functional, we apply Theorem 5.4 by taking $X_0 = \operatorname{span}\{x_0\}$, $p := \mu_E$, and the linear functional $f_0 : X_0 \to \mathbb{R}$ defined by

$$f_0(\alpha x_0) = \alpha, \qquad \alpha \in \mathbb{R}.$$

Since $E_1 \cap E_2 = \varnothing$ and $x_0 \notin E$, by Lemma 5.15, we have $\mu_E(x_0) \geq 1$, and hence,

$$f_0(\alpha x_0) = \alpha \leq \alpha\mu_E(x_0) = \mu_E(\alpha x_0) \quad \forall\, \alpha \in \mathbb{R}.$$

Thus, by Theorem 5.4, f_0 has a linear extension $f : X \to \mathbb{R}$ such that

$$f(x) \leq \mu_E(x) \quad \forall\, x \in X.$$

Recall that, by Lemma 5.15, $\mu_E(x) \leq 1$ for every $x \in E$. Hence, $f(x) \leq 1$ for every $x \in E$. From this, it also follows that $f(x) \geq -1$ for every $x \in (-E)$. Thus, we have

$$|f(x)| \leq 1 \quad \forall x \in E \cap (-E).$$

Since $E \cap (-E)$ is an open set containing 0, it follows (*How?*) that f is continuous.

Now, we show that there exists $\gamma \in \mathbb{R}$ such that

$$f(x) < \gamma \leq f(y) \quad \forall x \in E_1,\, y \in E_2.$$

For this, we may first observe (see Exercise 5.9(ii) below) that $f(E_1)$ and $f(E_2)$ are intervals in $\mathbb{R}$ and $f(E_1)$ is an open interval. Hence, it is enough to show that $f(x) \leq f(y)$ for every $x \in E_1$ and for every $y \in E_2$. So, let $x \in E_1$, $y \in E_2$. Since $x - y + x_0 \in E$, by Lemma 5.15, we have

$$f(x) - f(y) + 1 = f(x - y + x_0) \leq \mu_E(x - y + x_0) \leq 1.$$

Thus, we have proved that $f(x) \leq f(y)$ for all $x \in E_1$, $y \in E_2$. ∎

Exercise 5.9 Let E_1, E_2 be as in Theorem 5.14.

(i) For $u_1 \in E_1$, $u_2 \in E_2$, let $E = E_1 - E_2 + u_2 - u_1$. Show that the set E is open, convex, absorbing, and $0 \in E$.

(ii) If $f \in X'$, then show that $f(E_1)$ and $f(E_2)$ are intervals in $\mathbb{R}$ and $f(E_1)$ is an open interval. □

PROBLEMS

In the following problems, X denotes a normed linear space.

1. Let X_0 be a subspace of X. Give reasons to assertions in the following statements:

(a) If X_0 is a proper closed subspace, then there exists a closed hyperspace H of X such that $X_0 \subseteq H$.

(b) If $x_0 \in X \setminus X_0$ and $S = \{f \in X' : X_0 \subseteq N(f)\}$, then $x_0 \in \overline{X_0}$ if and only if $f(x_0) = 0$ for every $f \in S$.

2. Let X_0 be a closed subspace of X. Then show that

$$X_0 = \cap\{N(f) : f \in X_0^a\},$$

where $X_0^a = \{f \in X' : f(x) = 0 \ \forall x \in X_0\}$.

3. For the case when X is an inner product space, derive the conclusion of Corollary 5.8 without using Corollary 5.7.

4. Show that for $x, y \in X$, $f(x) = f(y)$ for every $f \in X'$ implies $x = y$.

5. Let X and Y be normed linear spaces. Show that the map $A \mapsto A'$ is a linear isometry from $\mathcal{B}(X, Y)$ into $\mathcal{B}(Y', X')$.

6. Show that every normed linear space is linearly isometric with a subspace of $(C(\Omega), \|\cdot\|_\infty)$, where $\Omega = \{f \in X' : \|f\| = 1\}$.

7. Show that every separable normed linear space is linearly isometric with a subspace of $\ell^\infty(\mathbb{N})$.

8. Prove Theorem 5.4.
[*Hint*: Modify the arguments used in the proof of Theorem 5.1.]

6

Uniform Boundedness Principle

Let X and Y be normed linear spaces, and (A_n) be a sequence in $\mathcal{B}(X, Y)$ such that $(A_n x)$ converges in Y for every $x \in X$. If $(\|A_n\|)$ is bounded, then we know by Theorem 3.10, that $A : X \to Y$ defined by

$$Ax = \lim_{n\to\infty} A_n x, \qquad x \in X,$$

is a bounded linear operator and

$$\|A\| \leq \liminf_n \|A_n\|.$$

Also, we have seen (Example 3.6) that, for the above conclusion, the assumption of boundedness of the sequence $(\|A_n\|)$ was not redundant. We now show that, if X is a Banach space, then the assumption of boundedness of $(\|A_n\|)$ is, in fact, a consequence of the boundedness of $(A_n x)$ for each $x \in X$.

First we introduce a few terminologies.

Let X and Y be normed linear spaces, $E \subseteq X$, and $\mathcal{A}$ be a family of linear operators from X to Y. We say that $\mathcal{A}$ is **pointwise bounded** on E if for each $x \in E$, there exists $M_x > 0$ such that $\|Ax\| \leq M_x$ for all $A \in \mathcal{A}$. The family $\mathcal{A}$ is said to be **uniformly bounded** on E if there exists $M > 0$ such that $\|Ax\| \leq M$ for all $A \in \mathcal{A}$ and for all $x \in E$.

Let E_0 be any of the sets $\{x \in X : \|x\| \leq 1\}$, $\{x \in X : \|x\| < 1\}$ or $\{x \in X : \|x\| = 1\}$. Then it can be easily seen that

(1) $\mathcal{A}$ is pointwise bounded on E_0 if and only if it is pointwise bounded on X, and

(2) $\mathcal{A}$ is uniformly bounded on E_0 if and only if $\mathcal{A} \subseteq \mathcal{B}(X, Y)$ and $\{\|A\| : A \in \mathcal{A}\}$ is a bounded subset of $\mathbb{R}$.

If E_0 is as above, then, in the definition of pointwise boundedness and uniform boundedness, we drop the special mention of the set

E_0. Thus, a family $\mathcal{A}$ of linear operators from X to Y is **pointwise bounded** if and only if $\mathcal{A}$ is pointwise bounded on X, and $\mathcal{A}$ is **uniformly bounded** if $\mathcal{A} \subseteq \mathcal{B}(X,Y)$ and $\{\|A\| : A \in \mathcal{A}\}$ is a bounded set.

Clearly, if $\mathcal{A}$ is uniformly bounded, then it is pointwise bounded on X. The *uniform boundedness principle* asserts that the converse is also true, provided $\mathcal{A} \subseteq \mathcal{B}(X,Y)$ and X is a Banach space. In case X is a finite dimensional space, the proof of the above assertion is very easy: Suppose X is a finite dimensional normed linear space and $\mathcal{A}$ is a pointwise bonded subset of $\mathcal{B}(X,Y)$. Let $\{u_1,\ldots,u_k\}$ be a basis of X. Then for every $x \in X$ and $A \in \mathcal{A}$, we have

$$Ax = \sum_{j=1}^{k} f_j(x)Au_j, \qquad x \in X,$$

where $\{f_1,\ldots,f_k\}$ is the dual basis associated with $\{u_1,\ldots,u_k\}$. Hence,

$$\|Ax\| \leq \sum_{j=1}^{k} \|f_j\|\,\|Au_j\|\,\|x\|$$

for all $x \in X$ and $A \in \mathcal{A}$. Since $\mathcal{A}$ is pointwise bounded on X, for each $j = 1,\ldots,k$, there exists $\beta_j > 0$ such that $\|Au_j\| \leq \beta_j$ for all $A \in \mathcal{A}$. Thus,

$$\|Ax\| \leq \alpha\beta\|x\| \quad \forall x \in X, \forall A \in \mathcal{A},$$

where $\beta = \max\{\beta_j : j = 1,\ldots,k\}$ and $\alpha = \sum_{j=1}^{k} \|f_j\|$. Thus, $\|A\| \leq \alpha\beta$ for all $A \in \mathcal{A}$.

6.1 The Theorem and Its Consequences

Theorem 6.1 (Uniform boundedness principle) *Let X be a Banach space, Y be a normed linear space, and $\mathcal{A} \subseteq \mathcal{B}(X,Y)$. If $\mathcal{A}$ is pointwise bounded, then $\mathcal{A}$ is uniformly bounded.*

Proof. Suppose $\mathcal{A}$ is pointwise bounded. Then, for each $x \in X$ there exists $M_x > 0$ such that $\|Ax\| \leq M_x$ for all $A \in \mathcal{A}$. Hence, $X = \cup_{n=1}^{\infty} E_n$, where

$$E_n = \{x \in X : \|Ax\| \leq n \quad \forall A \in \mathcal{A}\}, \qquad n \in \mathbb{N}.$$

Since $E_n = \cap_{A\in\mathcal{A}}\{x \in X : \|Ax\| \leq n\}$, and $\{x \in X : \|Ax\| \leq n\}$ is a closed set for every $n \in \mathbb{N}$, E_n is a closed subset of X. Since X is a Banach space, by Baire Category Theorem 2.31, there exists $k \in \mathbb{N}$ such that E_k has nonempty interior.

Let $u \in E_k$ and let $r > 0$ be such that $B(u,r) \subseteq E_k$. Then for every nonzero $x \in X$, we have

$$u + \frac{rx}{2\|x\|} \in B(u,r) \subseteq E_k,$$

i.e.,

$$\left\|A\Big(u + \frac{rx}{2\|x\|}\Big)\right\| \leq k \quad \forall A \in \mathcal{A}.$$

Consequently, since $\|Au\| \leq k$,

$$\left\|A\Big(\frac{rx}{2\|x\|}\Big)\right\| = \left\|A\Big(u + \frac{rx}{2\|x\|}\Big) - Au\right\| \leq 2k \quad \forall A \in \mathcal{A}.$$

From this, we get

$$\|Ax\| \leq \frac{4k}{r}\|x\| \quad \forall x \in X, \forall A \in \mathcal{A}.$$

Consequently, $\|A\| \leq 4k/r$ for all $A \in \mathcal{A}$, and hence $\mathcal{A}$ is uniformly bounded. ∎

Let X be a Banach space, Y be a normed linear space and $\mathcal{A}$ be a subset of $\mathcal{B}(X,Y)$. If $\mathcal{A}$ is not uniformly bounded, then we know by the above theorem that there exists $x_0 \in X$ such that the set $\{Ax_0 : A \in \mathcal{A}\}$ is not bounded in Y. In fact, we have the following.

Corollary 6.2 *Let X be a Banach space, Y be a normed linear space, and $\mathcal{A}$ be a subset of $\mathcal{B}(X,Y)$. If $\mathcal{A}$ is not uniformly bounded, then there exists a dense subset D of X such that for every $x \in D$, $\{Ax : A \in \mathcal{A}\}$ is not bounded in Y.*

Proof. Suppose $\mathcal{A}$ is not uniformly bounded. Then by Theorem 6.1, there exists $x_0 \in X$ such that the set $\{Ax_0 : A \in \mathcal{A}\}$ is not bounded in Y. Therefore, the set

$$X_0 = \{x \in X : \{Ax : A \in \mathcal{A}\} \text{ bounded in } Y\}$$

is a proper subset of X. It can be easily seen (*Verify*) that X_0 is also a subspace of X. Hence, by Lemma 2.32, interior of X_0 is empty. Therefore, the complement of X_0, namely, the set

$$\{x \in X : \{Ax : A \in \mathcal{A}\} \text{ not bounded in } Y\}$$

is dense in X (Exercise 2.11). This completes the proof. ∎

Banach-Steinhaus theorem

The following result is immediate from Theorems 6.1 and 3.10.

Corollary 6.3 (Banach-Steinhaus theorem) *Let X be a Banach space, Y be a normed linear space, and (A_n) be a sequence in $\mathcal{B}(X, Y)$ such that $(A_n x)$ converges for every $x \in X$. Let $A : X \to Y$ be defined by*

$$A(x) = \lim_{n\to\infty} A_n(x), \qquad x \in X.$$

Then $(\|A_n\|)$ is bounded and $A \in \mathcal{B}(X, Y)$.

Combining the above corollary with Theorem 3.11, we obtain the following result.

Theorem 6.4 *Let X and Y be Banach spaces, and (A_n) be a sequence in $\mathcal{B}(X, Y)$. Then $(A_n x)$ converges for every $x \in X$ if and only if $(\|A_n\|)$ is bounded and there exists a dense subset D of X such that $(A_n u)$ converges for every $u \in D$.*

Corollary 6.5 *Let X be a normed linear space and $E \subseteq X$. Then E is bounded in X iff $f(E)$ is bounded in $\mathbb{K}$ for every $f \in X'$.*

Proof. Suppose E is bounded in X, and let $c > 0$ be such that $\|x\| \le c$ for every $x \in E$. Then, for every $f \in X'$,

$$|f(x)| \le \|f\| \, \|x\| \le c\|f\| \quad \forall x \in E.$$

Thus, $f(E)$ is bounded in $\mathbb{K}$ for every $f \in X'$.

Conversely, suppose that $f(E)$ is bounded for every $f \in X'$. For $x \in E$, consider $\varphi_x : X' \to \mathbb{K}$ defined by

$$\varphi_x(f) = f(x), \qquad f \in X'.$$

Then, by Corollary 5.11, we have

$$\varphi_x \in \mathcal{B}(X', \mathbb{K}), \qquad \|\varphi_x\| = \|x\| \quad \forall x \in E.$$

By hypothesis, $\{|\varphi_x(f)| : x \in E\}$ is bounded for every $f \in X'$. Therefore, by uniform boundedness principle, $\{\varphi_x : x \in X\}$ is uniformly bounded, i.e., $\{\|\varphi_x\| : x \in E\}$ is bounded. Thus, $\{\|x\| : x \in E\}$ is bounded. In other words, E is bounded in X. ∎

If the scalar field $\mathbb{K}$ is the set of all real numbers, we can give some geometric meaning to the above result.

Let X be a normed linear space over $\mathbb{R}$ and f be a linear functional on X. Corresponding to $\alpha \in \mathbb{R}$, consider the hyperplane

$$H_{f,\alpha} := \{x \in X : f(x) = \alpha\} = x_\alpha + N(f),$$

where $x_\alpha \in X$ is such that $f(x_\alpha) = \alpha$.

Now, for $E \subseteq X$, the set $f(E)$ is bounded if and only if there exist $\alpha, \beta \in \mathbb{R}$ such that $\alpha \le f(x) \le \beta$ for every $x \in E$. Hence, using the terminology in Section 5.4, $f(E)$ is bounded if and only if there are $\alpha, \beta \in \mathbb{R}$ such that E is on the left side of $H_{f,\beta}$ and on the right side of $H_{f,\alpha}$. Thus, Corollary 6.5 states that a subset $E \subseteq X$ is bounded in X if and only if for every $f \in X'$ there are $\alpha, \beta \in \mathbb{R}$ such that E *lies between* the hyperplanes $H_{f,\alpha}$ and $H_{f,\beta}$.

Exercise 6.1 Let X and Y be normed linear spaces and $A : X \to Y$ be a linear operator. Then prove that A is a bounded linear operator if and only if the composition map $gA : X \to \mathbb{K}$ belongs to X' for every $g \in Y'$. □

A consequence of the uniform boundedness principle, which is very useful in *numerical functional analysis*, is the following.

Corollary 6.6 *Let X be a Banach space, Y be a normed linear space, and (A_n) be a sequence in $\mathcal{B}(X,Y)$ such that $(A_n x)$ converges for every $x \in X$. Let $A : X \to Y$ be defined by $Ax = \lim_{n\to\infty} A_n x$, $x \in X$. Then for every totally bounded subset $S \subseteq X$,*

$$\sup_{x \in S} \|A_n x - Ax\| \to 0 \quad \text{as } n \to \infty.$$

Proof. Let S be a totally bounded subset of X, and let $\varepsilon > 0$ be given. Then there exist $x_1, \ldots, x_k$ in S such that

$$S \subseteq \bigcup_{j=1}^{k} \{x \in X : \|x - x_j\| < \varepsilon\}.$$

Let $x \in S$ and let $j \in \{1, \ldots, k\}$ be such that $\|x - x_j\| < \varepsilon$. Then we have

$$\begin{aligned}\|A_n x - Ax\| &\leq \|A_n x - A_n x_j\| + \|A_n x_j - Ax_j\| + \|Ax_j - Ax\| \\ &\leq \|A_n\| \, \|x - x_j\| + \|A_n x_j - Ax_j\| + \|A\| \, \|x - x_j\|.\end{aligned}$$

Now, by the hypothesis on (A_n), for each $i \in \{1, \ldots, k\}$, there exists $N_i \in \mathbb{N}$ such that $\|A_n x_i - Ax_i\| < \varepsilon$ for all $n \geq N_i$. In particular,

$$\|A_n x_j - Ax_j\| < \varepsilon \quad \forall n \geq N := \max\{N_i : i = 1, \ldots, k\}.$$

Therefore, using the fact that $\|A_n\| \leq c$ for every n and for some $c > 0$ (Corollary 6.3), we have

$$\|A_n x - Ax\| \leq (c + 1 + \|A\|)\, \varepsilon \quad \forall n \geq N.$$

Since this N is independent of the element x, we can conclude that

$$\sup_{x \in S} \|A_n x - Ax\| \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, the proof is complete. ∎

Exercise 6.2 Let X, Y and (A_n) be as in the above corollary. If $T \in \mathcal{B}(X)$ is a finite rank operator, then show that $\|(A - A_n)T\| \to 0$ as $n \to \infty$. □

Arzela-Ascoli's theorem

In Corollary 6.6, we obtained *uniform convergence* of a sequence of operators in $\mathcal{B}(X, Y)$ on totally bounded subsets of X provided we have *pointwise convergence* on X. We know that if X is a finite dimensional normed linear space, then every bounded subset of X is totally bounded, and this is not the case if X is infinite dimensional (cf. Theorem 2.39). For a compact metric space Ω, *Arzela-Ascoli theorem* proved below specifies certain sufficient conditions for a pointwise bounded subset S of $C(\Omega)$ to be a totally bounded set. We shall make use of this result on many other occasions in the due course. For stating the result, we recall the following definitions from analysis:

Let Ω be a nonempty set. A set $\mathcal{S}$ of $\mathbb{K}$-valued functions on Ω, i.e., $\mathcal{S} \subseteq \mathcal{F}(\Omega, \mathbb{K})$, is said to be **pointwise bounded** on Ω if for each $t \in \Omega$, there exists $M_t > 0$ such that

$$|\phi(t)| \leq M_t \quad \forall \phi \in \mathcal{S}.$$

Now, let Ω be a metric space with metric d. Then a subset S of $\mathcal{F}(\Omega, \mathbb{K})$ is said to be **equicontinuous** on Ω if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$s, t \in \Omega, \ d(s,t) < \delta \implies |x(s) - x(t)| < \varepsilon \quad \forall x \in S.$$

Theorem 6.7 (Arzela-Ascoli) *Let (Ω, d) be a compact metric space and let X be the space $C(\Omega)$ with the norm $\|\cdot\|_\infty$. Then a subset S of X is totally bounded in X if and only if it is pointwise bounded and equicontinuous on Ω.*

Proof. Suppose $S \subseteq X$ is totally bounded in X. It is easy to see that S is a bounded subset of X. In particular, S is pointwise bounded on Ω. To see that it is equicontinuous, let $\varepsilon > 0$. By the total boundedness of S, there exist $x_1, \ldots, x_k$ in S such that $S \subseteq \cup_{j=1}^k B(x_j, \varepsilon)$. Note that for every $x \in S$, $j \in \{1, \ldots, k\}$ and $s, t \in \Omega$,

$$\begin{aligned} |x(s) - x(t)| &\leq |x(s) - x_j(s)| + |x_j(s) - x_j(t)| + |x_j(t) - x(t)| \\ &\leq \|x - x_j\|_\infty + |x_j(s) - x_j(t)| + \|x_j - x\|_\infty. \end{aligned}$$

Since each x_j, $j \in \{1, \ldots, k\}$ is uniformly continuous on Ω, there exists $\delta_j > 0$ such that

$$|x_j(s) - x_j(t)| < \varepsilon \quad \text{whenever} \quad s, t \in \Omega, \quad d(s,t) < \delta_j.$$

Now, let $x \in S$, and let $i \in \{1, \ldots, k\}$ be such that $\|x - x_i\|_\infty < \varepsilon$. Then we have

$$\begin{aligned} |x(s) - x(t)| &\leq \|x - x_i\|_\infty + |x_i(s) - x_i(t)| + \|x_i - x\|_\infty \\ &\leq 3\varepsilon, \end{aligned}$$

for every $s, t \in \Omega$ with $d(s,t) < \delta = \min\{\delta_j : j = 1, \ldots, k\}$. Note that δ does not depend on $x \in S$. Thus, we have shown that S is equicontinuous.

Conversely, suppose that S is pointwise bounded and equicontinuous, and $\varepsilon > 0$. By equicontinuity of S, there exists $\delta > 0$ such that

$$s, t \in \Omega, \ d(s,t) < \delta \implies |x(s) - x(t)| < \varepsilon \quad \forall x \in S. \qquad (6.1)$$

Clearly, $\Omega = \cup_{t\in\Omega} B_\Omega(t,\delta)$. Since Ω is a compact metric space, there exists $t_1,\ldots,t_k$ such that

$$\Omega = \bigcup_{j=1}^{k} B_\Omega(t_j,\delta). \tag{6.2}$$

Since S is pointwise bounded on Ω, it follows that the set

$$E = \{(x(t_1),\ldots,x(t_k)) : x \in S\}$$

is a bounded subset of $\ell^\infty(k)$. (Recall that $\ell^\infty(k)$ is the space $\mathbb{K}^k$ with $\|\cdot\|_\infty$.) In fact, if $M_1,\ldots,M_k$ are such that $|x(t_j)| \le M_j$ for all $x \in S$ and $j = 1,\ldots,k$, then

$$\|u\|_\infty \le \max\{M_j : j = 1,\ldots,k\} \quad \forall u \in E.$$

Since $\ell^\infty(k)$ is a finite dimensional normed linear space, the bounded set E is totally bounded in $\ell^\infty(k)$. Hence, there exists $u_1,\ldots,u_n$ in E such that

$$E \subseteq \bigcup_{j=1}^{n} \{u \in \ell^\infty(k) : \|u - u_j\|_\infty < \varepsilon\}.$$

For each $j = 1,\ldots,n$, let $x_j \in S$ be such that $u_j = (x_j(t_1),\ldots,x_j(t_k))$.

Now, let $x \in S$. Then we have $u := (x(t_1),\ldots,x(t_k)) \in E$, so that by the above property of E, there exists $i \in \{1,\ldots,n\}$ such that $\|u - u_i\|_\infty < \varepsilon$, i.e., there exists $i \in \{1,\ldots,n\}$ such that

$$|x(t_j) - x_i(t_j)| < \varepsilon \quad \forall j \in \{1,\ldots,k\}.$$

Let $t \in \Omega$. By (6.2), there exist $j \in \{1,\ldots,k\}$ such that $d(t,t_j) < \delta$. Then, using (6.1), we have

$$\begin{aligned} |x(t) - x_i(t)| &\le |x(t) - x(t_j)| + |x(t_j) - x_i(t_j)| + |x_i(t_j) - x_i(t)| \\ &< 3\varepsilon. \end{aligned}$$

Consequently, $\|x - x_i\|_\infty < 3\varepsilon$. Thus, $S \subseteq \cup_{j=1}^{n} B_X(x_i, 3\varepsilon)$, showing that S is totally bounded in X. ∎

6.2 Some Applications

6.2.1 On Divergence of Lagrange Interpolation

We recall from Example 3.1(viii) that the Lagrange interpolation of a function $x \in C[a, b]$, associated with a partition

$$a = t_{0,n} < t_{1,n} < \ldots < t_{n,n} \leq b$$

of $[a, b]$, is defined as

$$L_n x = \sum_{j=1}^{n} x(t_{j,n}) \ell_{j,n},$$

where

$$\ell_{j,n} = \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{t - t_{i,n}}{t_{j,n} - t_{i,n}}, \qquad j = 1, \ldots, n.$$

One may ask whether $L_n x$ is close to x in some sense. It is a well known result in numerical analysis (see, for example, Wendroff [34]) that if $x \in C[a, b]$ is differentiable n times at every point in (a, b), then for every $t_{1,n}, \ldots, t_{n,n}$ in $[a, b]$, there exists $\xi_n \in [a, b]$ such that

$$x(t) - (L_n x)(t) = \frac{x^{(n)}(\xi_n)}{n\,!} \prod_{i=1}^{n} (t - t_{i,n}).$$

Thus, one can have a pointwise error estimate provided one knows a bound for $|x^{(n)}(t)|$. What about uniform approximation? That is, do we have $\|x - L_n x\|_\infty \to 0$ as $n \to \infty$ for every $x \in C[a, b]$? The following theorem shows that the answer is negative for *many* functions x in $C[a, b]$.

Theorem 6.8 *There exists a dense set $D \subseteq C[a, b]$ such that for every $x \in D$, the sequence $(\|L_n x\|_\infty)$ is unbounded.*

For proving the above theorem, we make use of the following result proved by Faber [14].

Lemma 6.9 *Given distinct points $\tau_1, \ldots, \tau_k$ in $[a, b]$, there exists a polynomial $p(t)$ of degree atmost $k - 1$ and $s \in [a, b]$ such that*

$$p(\tau_j) \leq 8\sqrt{\pi} \quad \forall j = 1, \ldots, k; \quad p(s) > \ln k.$$

Proof of Theorem 6.8. By Corollary 6.2, it is enough to show that $(\|L_n\|)$ is unbounded. Recall from Example 3.3(x) that

$$\|L_n\| = \sup_{a\le t\le b} \sum_{j=1}^{n} |\ell_{j,n}(t)|.$$

By the above lemma, there exists a polynomial $p_n(t)$ of degree at most $n-1$ and $s_n \in [a,b]$ such that

$$p_n(t_{j,n}) \le 8\sqrt{\pi} \quad \forall j = 1,\ldots,n, \qquad p_n(s_n) > \ln n.$$

Hence,

$$|p_n(t)| = |(L_n p_n)(t)| \le \sum_{j=1}^{n} |p_n(t_{j,n})|\, |\ell_{j,n}(t)| \le 8\sqrt{\pi}\|L_n\|$$

so that

$$\|L_n\| \ge \frac{|p_n(t)|}{8\sqrt{\pi}} \quad \forall t \in [a,b].$$

In particular,

$$\|L_n\| \ge \frac{|p_n(s_n)|}{8\sqrt{\pi}} > \frac{\ln n}{8\sqrt{\pi}}.$$

Thus, $(\|L_n\|)$ is unbounded. This completes the proof. ∎

Theorem 6.8 shows that we cannot expect uniform convergence of $(L_n x)$ to x for a large class of functions. Can we have pointwise convergence for all $x \in C[a,b]$? Again, the answer is in the negative. This was proved by Bernstein, way back in 1912 by showing that for $x_0 \in C[-1,1]$ defined by $x_0(t) = |t|$, $-1 \le t \le 1$, $(L_n x_0(t))$ converges to $x_0(t)$ only for $t \in \{-1,0,1\}$. (For a detailed discussion of the above mentioned divergence results, refer Baker and Mills [7].)

6.2.2 A Necessary condition for the Convergence of Quadrature Formulas

As an application of Theorem 3.11, we have shown that a sequence $(Q_n(x))$ of quadrature formulas given by

$$Q_n(x) = \sum_{j=1}^{n} x(t_{j,n}) w_{j,n}, \qquad x \in C[a,b],$$

converges to $\int_a^b x(t)\,dt$ for every $x \in C[a,b]$ provided the sequence $\left(\sum_{j=1}^{n} |w_{j,n}|\right)$ is bounded, and for some set $E \subseteq C[a,b]$ with span E dense in $C[a,b]$,

$$Q_n(x) \to \int_a^b x(t)\,dt \quad \forall x \in E.$$

The question one would like to ask is whether the boundedness of the sequence $\left(\sum_{j=1}^{n} |w_{j,n}|\right)$ is necessary for the convergence of $(Q_n(x))$ for every $x \in C[a,b]$. We have already seen (Example 3.3) that, for each $n \in \mathbb{N}$, Q_n is a continuous linear functional on $C[a,b]$ with $\|\cdot\|_\infty$, and $\|Q_n\| = \sum_{j=1}^{n} |w_{j,n}|$. Thus, Theorem 6.4 shows that the answer is in the affirmative.

6.2.3 On Divergence of Fourier Series

We recall from real analysis that the *Fourier series* of a function $x \in L^1[-\pi,\pi]$ is defined by

$$\sum_{n=-\infty}^{\infty} \hat{x}(n) e^{int},$$

where

$$\hat{x}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t) e^{-int}\,d\mu(t), \qquad n \in \mathbb{Z}.$$

For $x \in L^1[-\pi,\pi]$, $t \in [-\pi,\pi]$, $n \in \mathbb{Z}$, let

$$s_n(x,t) = \sum_{k=-n}^{n} \hat{x}(k) e^{ikt}$$

It is important to know for what kind of functions x, we have the convergence

$$s_n(x,t) \to x(t) \quad \text{as} \quad n \to \infty \quad \forall t \in [-\pi,\pi].$$

This is not true for all $x \in L^1[-\pi,\pi]$; not true even for all $x \in C[-\pi,\pi]$, as the following theorem shows.

Theorem 6.10 *Let $X = \{x \in C[-\pi,\pi] : x(-\pi) = x(\pi)\}$ with norm $\|\cdot\|_\infty$. Then there exists a dense subset D of X such that for every $x \in D$, the sequence $(s_n(x,0))$ diverges.*

Proof. First we observe that X is a closed subspace of the Banach space $C[-\pi,\pi]$ with $\|\cdot\|_\infty$, so that it is a Banach space. For each $n \in \mathbb{Z}$, let $f_n : X \to \mathbb{K}$ be defined by

$$f_n(x) = s_n(x,0), \qquad x \in X.$$

Clearly, f_n is a linear functional on X. We show that $f_n \in X'$ for every $n \in \mathbb{Z}$ and $(\|f_n\|)$ is unbounded so that the result follows from Theorem 6.4. Let

$$y_n(t) = \sum_{k=-n}^{n} e^{ikt}, \qquad n \in \mathbb{Z}.$$

Using the fact that $y_n(-t) = y_n(t)$ for all $n \in \mathbb{Z}$ and $t \in [-\pi,\pi]$, we have

$$f_n(x) = s_n(x,0) = \sum_{k=-n}^{n} \hat{x}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)y_n(t)\,d\mu(t).$$

Hence,

$$|f_n(x)| \le \frac{1}{2\pi}\|x\|_\infty\|y_n\|_1 \quad \forall x \in X \quad \forall n \in \mathbb{Z}$$

so that

$$f_n \in X', \qquad \|f_n\| \le \frac{1}{2\pi}\|y_n\|_1 \quad \forall n \in \mathbb{Z}.$$

We show that $\|f_n\| = \|y_n\|_1/2\pi$ for every $n \in \mathbb{Z}$ and $(\|y_n\|)$ is unbounded.

Let $\phi_n : [-\pi,\pi] \to \mathbb{K}$ be defined by

$$\phi_n(t) = \begin{cases} 1 & \text{if } y_n(t) \ge 0 \\ -1 & \text{if } y_n(t) < 0, \end{cases}$$

and let $\psi_{n,m} \in C[-\pi,\pi]$ be such that for each $n \in \mathbb{Z}$, $\psi_{n,m}(t) \to \phi_n(t)$ as $m \to \infty$ for every $t \in [\pi,\pi]$ and $\|\psi_{n,m}\|_\infty \le 1$ for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$. Such a sequence can be easily constructed as $y_n(t)$ has only a finite number of zeroes in $[-\pi,\pi]$. Then by dominated convergence theorem (Theorem 2.11), it follows that

$$f_n(\psi_{n,m}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi_{n,m}(t)y_n(t)\,d\mu(t) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_n(t)y_n(t)\,d\mu(t)$$

as $m \to \infty$. Thus,

$$\lim_{m\to\infty} f_n(\psi_{n,m}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |y_n(t)|\, d\mu(t) = \frac{1}{2\pi}\|y_n\|_1.$$

Consequently, we have $\|f_n\| = \|y_n\|_1/2\pi$ for every $n \in \mathbb{Z}$. It remains to show that $(\|y_n\|)$ is unbounded. To see this, first we observe that

$$y_n(t) = 1 + \sum_{k=1}^{n}\left(e^{ikt} + e^{-ikt}\right) = 1 + 2\sum_{k=1}^{n}\cos(kt)$$

so that

$$y_n(t)\sin\frac{t}{2} = \sin\frac{t}{2} + \sum_{k=1}^{n}\left[\sin\left(k+\frac{1}{2}\right)t - \sin\left(k-\frac{1}{2}\right)t\right] = \sin\left(n+\frac{1}{2}\right)t.$$

Hence,

$$y_n(t) = \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}} \quad \text{for} \quad t \neq 0; \;\; y_n(0) = 2n+1.$$

Therefore, since $|\sin t| \leq |t|$ for every $t \in [-\pi, \pi]$, we have

$$\begin{aligned}\|y_n\|_1 &\geq 2\int_{-\pi}^{\pi}\left|\frac{\sin(n+1/2)t}{t}\right| dt = 4\int_0^{(2n+1)\pi/2}\left|\frac{\sin t}{t}\right| dt\\ &\geq 4\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi}\frac{|\sin t|}{k\pi}\, dt = \frac{8}{\pi}\sum_{k=1}^{n}\frac{1}{k}.\end{aligned}$$

Therefore, $(\|y_n\|)$ is unbounded. ∎

PROBLEMS

1. Let X and Y be Banach spaces and (A_n) be a sequence in $\mathcal{B}(X,Y)$. If $(A_n x)$ is a Cauchy sequence in Y for every $x \in X$, then prove that there exists $A \in \mathcal{B}(X,Y)$ such that $A_n x \to Ax$ for every $x \in X$.

2. Let X be a Banach space, Y be a normed linear space and $\mathcal{A} \subseteq \mathcal{B}(X,Y)$ be such that $\{Ax : A \in \mathcal{A}\}$ is bounded in Y for every

$x \in X$. Then $\sup\{\|Ax\| : x \in E, A \in \mathcal{A}\} < \infty$ for every bounded set $E \subseteq X$ – Why?

3. Let X be a Banach space, Y be a normed linear space and $\mathcal{A} \subseteq \mathcal{B}(X, Y)$. Show that $\{\|A\| : A \in \mathcal{A}\}$ is not bounded if and only if there exists a dense subset $D \subseteq X$ such that for every $x \in D$, $\{\|Ax\| : A \in \mathcal{A}\}$ is not bounded.

4. Let X be a Banach space, Y be a normed linear space and (A_n) be a sequence in $\mathcal{B}(X, Y)$. Show that $(\|A_n\|)$ is bounded if and only if $(|f(A_n x)|)$ is bounded for every $x \in X$ and for every $f \in Y'$.

5. Let $X = \mathcal{P}[a, b]$ with $\|\cdot\|_\infty$. Using uniform boundedness principle, show that X is not a Banach space.

6. Let (α_n) be a sequence of scalars such that $\sum_{j=0}^{\infty} \alpha_j \beta_j$ converges for every (β_n) which converges to zero. Show that $\sum_{j=0}^{\infty} |\alpha_j| < \infty$.

7. Let X be a Banach space and (P_n) be a sequence of projections in $\mathcal{B}(X)$ such that $P_n x \to x$ as $n \to \infty$ for every $x \in X$. If $A \in \mathcal{B}(X)$ is such that $\{Ax : x \in X, \|x\| \le 1\}$ is totally bounded in X, then show that $\|A - P_n A\| \to 0$ as $n \to \infty$.

8. Let X be Hilbert space and (P_n) be a sequence of bounded projections such that $\|x - P_n x\| \to 0$ for every $x \in X$. Show that there exists $c > 0$ such that $\|x - P_n x\| \le c \inf\{\|x - u\| : u \in R(P_n)\}$ for every $x \in X$.

7

Closed Graph Theorem and Its Consequences

Let X and Y be normed linear spaces and X_0 be a subspace of X. Recall from Section 3.5 that a linear operator $A : X_0 \subseteq X \to Y$ is said to be a *closed operator* if for every (x_n) in X_0 which satisfies $x_n \to x$ and $Ax_n \to y$ for some $x \in X$ and $y \in Y$, we have $x \in X_0$ and $Ax = y$. Also, we have seen in Section 3.6 that a bounded operator $A : X_0 \subseteq X \to Y$ with its domain X_0 closed in X is a closed operator, whereas a closed operator $A : X_0 \subseteq X \to Y$ need not be a bounded operator even if its domain is closed.

Can we impose additional assumptions on the spaces involved so that a closed operator is a bounded operator?

Closed graph theorem asserts that if the spaces X_0 and Y are Banach spaces, then a closed operator $A : X_0 \subseteq X \to Y$ is a bounded operator.

7.1 Closed Graph Theorem

Let X and Y be normed linear spaces. Recall from Proposition 3.15 that, if X_0 is a subspace of X, then a linear operator $A : X_0 \subseteq X \to Y$ is a closed operator if and only if its graph

$$G(A) := \{(x, Ax) : x \in X_0\}$$

is closed in $X_0 \times Y$. Thus, if our discussion is about a closed operator defined on a closed subspace, then, without loss of generality, we can assume that the operator is defined on the whole space X. The following theorem should be seen in this perspective.

Theorem 7.1 (Closed graph theorem) *If X and Y are Banach spaces, then every closed linear operator $A : X \to Y$ is continuous.*

Proof. Let X and Y be Banach spaces, and $A : X \to Y$ be a closed linear operator. Let $B_0 = \{x \in X : \|x\| < 1\}$. We show that

$$B_0 \subseteq \{x \in X : \|Ax\| \le c\}$$

for some $c > 0$, so that by Theorem 3.1, A is continuous.

Note that $X = \cup_{j=1}^{\infty} V_j$, where

$$V_\alpha := \{x \in X : \|Ax\| \le \alpha\}, \qquad \alpha > 0.$$

Since X is complete, by the Baire category theorem (Theorem 2.31), there is some $k \in \mathbb{N}$ such that the interior of $\overline{V_k}$ is nonempty. Thus, there is some $x_0 \in X$ and $r > 0$ such that $B(x_0, r) \subseteq \overline{V_k}$. This implies that

$$B_0 \subseteq \overline{V_{2k/r}}.$$

To see this, let $x \in B_0$. Then, $u := x_0 + rx \in B(x_0, r) \subseteq \overline{V_k}$. Let (u_n) and (v_n) be in V_k such that $u_n \to u$ and $v_n \to x_0$. Then we have

$$x = \frac{1}{r}(u - x_0) = \lim_{n\to\infty} \frac{1}{r}(u_n - v_n),$$

$$\left\| A\left(\frac{u_n - v_n}{r}\right)\right\| \le \frac{1}{r}(\|Au_n\| + \|Av_n\|) \le \frac{2k}{r}.$$

Therefore, $(u_n - v_n)/r$ belongs to $V_{2k/r}$ so that $x \in \overline{V_{2k/r}}$.

Let us denote $V_{2k/r}$ by W. Let $x \in B_0$ and $0 < \varepsilon < 1$. Since $B_0 \subseteq \overline{W}$, there exists $x_1 \in W$ such that $\|x - x_1\| < \varepsilon$. Hence, $\varepsilon^{-1}(x - x_1) \in B_0$. By the same argument, there exists $x_2 \in W$ such that $\|\varepsilon^{-1}(x - x_1) - x_2\| < \varepsilon$, i.e.,

$$\|x - (x_1 + \varepsilon x_2)\| < \varepsilon^2.$$

Having obtained $x_1, x_2, \ldots, x_n \in W$ such that

$$\left\|x - (x_1 + \varepsilon x_2 + \varepsilon^2 x_3 + \cdots + \varepsilon^{n-1} x_n)\right\| < \varepsilon^n,$$

we have

$$\left\|\varepsilon^{-n}[x - (x_1 + \varepsilon x_2 + \varepsilon^2 x_3 + \cdots + \varepsilon^{n-1} x_n)]\right\| < 1.$$

Hence, again using the fact that $B_0 \subseteq \overline{W}$, there exists $x_{n+1} \in W$ satisfying

$$\left\|\varepsilon^{-n}[x - (x_1 + \varepsilon x_2 + \varepsilon^2 x_3 + \cdots + \varepsilon^{n-1} x_n)] - x_{n+1}\right\| < \varepsilon.$$

Thus, by induction, we obtain a sequence (x_n) in W with

$$\left\|x - (x_1 + \varepsilon x_2 + \varepsilon^2 x_3 + \cdots + \varepsilon^{n-1} x_n)\right\| < \varepsilon^n$$

for every positive integer n. In particular, taking $s_n := \sum_{j=1}^{n} \varepsilon^{j-1} x_j$, $n \in \mathbb{N}$, the sequence (s_n) converges to x. Recall that $x_j \in W$ implies $\|Ax_j\| \le 2k/r$. Hence, for $n > m$, we have

$$\|As_n - As_m\| \le \sum_{j=m+1}^{n} \left\|A(\varepsilon^{j-1} x_j)\right\| \le \frac{2k}{r} \sum_{j=m+1}^{n} \varepsilon^{j-1}.$$

Thus, (As_n) is a Cauchy sequence in Y. Since Y is also a Banach space, the sequence (As_n) converges to some $y \in Y$. Since A is a closed linear operator, it follows that $y = Ax = \lim_{n\to\infty} As_n$. Note that

$$\|As_n\| \le \sum_{j=1}^{n} \left\|A(\varepsilon^{j-1} x_j)\right\| \le \frac{2k}{r} \sum_{j=1}^{n} \varepsilon^{j-1} \le \frac{2k}{r(1-\varepsilon)}.$$

Hence,

$$\|Ax\| = \lim_{n\to\infty} \|As_n\| \le \frac{2k}{r(1-\varepsilon)}.$$

Thus, $B_0 \subseteq V_c$ with $c = 2k/r(1-\varepsilon)$, and the proof is complete. ∎

Now, we state a seemingly generalized form of the closed graph theorem. One part of it is nothing but the above theorem, and the other part is contained in Theorem 3.17.

Theorem 7.2 *Let X, Y be Banach spaces and $A : X_0 \subseteq X \to Y$ be a closed operator. Then A is continuous if and only if its domain X_0 is closed in X.*

Next, we give examples to show that the completeness assumptions on the space X and Y cannot be dropped.

We have already seen examples to the effect that if X_0 is not complete, then a closed operator $A : X_0 \subseteq X \to Y$ need not be continuous. Recall the following examples from Section 3.6.

EXAMPLE 7.1 (i) Let $X = Y = C[0,1]$ with $\|\cdot\|_\infty$, $X_0 = C^1[0,1]$ and $A : X_0 \subseteq X \to Y$ be defined by $Ax = x'$, $x \in X_0$. Then A is a closed operator which is not continuous. Note that Y is a Banach space but X_0 is not a Banach space.

(ii) Let $X = \ell^2$, (λ_n) be a sequence of positive scalars such that $\lambda_n \to 0$ as $n \to \infty$, and

$$X_0 = \left\{ x \in \ell^2 : \sum_{j=1}^{\infty} \frac{|x(j)|^2}{\lambda_j^2} < \infty \right\}.$$

Let $A : X_0 \subseteq X \to X$ be defined by

$$(Ax)(j) = \frac{x(j)}{\lambda_j}, \quad j \in \mathbb{N};\ x \in X.$$

We have seen in Example 3.12 that A is a closed, but an unbounded operator. Here, Y is a Banach space but X_0 is not a Banach space.

The following example shows that the completeness assumption on the codomain space also cannot be dropped.

EXAMPLE 7.2 Let X be an infinite dimensional Banach space. By Theorem 2.30, we know that every basis of X is an uncountable set. Let $\{u_\lambda : \lambda \in \Lambda\}$ be a basis of X with $\|u_\lambda\| = 1$, $\lambda \in \Lambda$. Then every $x \in X$ can be written as

$$x = \sum_{\lambda \in \Lambda} \hat{x}(\lambda) u_\lambda,$$

where $\hat{x} : \Lambda \to \mathbb{K}$ is a function such that $\hat{x}(\lambda) = 0$ for all but a finite number of λ's. Define

$$\|x\|_* = \sum_{\lambda \in \Lambda} |\hat{x}(\lambda)|.$$

Then it is easily seen that $\|\cdot\|_*$ is also a norm on X. Moreover,

$$\|x\|_* \geq \|x\| \quad \forall x \in X.$$

We show that the norm $\|\cdot\|_*$ is not complete. For this, consider a denumerable subset $\{\lambda_1, \lambda_2, \ldots\}$ of Λ, and a sequence (x_n) in X defined by

$$x_n = \sum_{\lambda \in \Lambda} \hat{x}_n(\lambda) u_\lambda$$

with

$$\hat{x}_n(\lambda) = \begin{cases} 1/j^2 & \text{if } \lambda = \lambda_j,\ 1 \leq j \leq n, \\ 0 & \text{if } \lambda \notin \{\lambda_1, \ldots, \lambda_n\}. \end{cases}$$

Let $\Delta_n = \{\lambda_1, \ldots, \lambda_n\}$, $n \in \mathbb{N}$. Then for $n > m$, we have

$$\begin{aligned}\|x_n - x_m\|_* &= \sum_{\lambda\in\Delta_n} |\hat{x}_n(\lambda) - \hat{x}_m(\lambda)| + \sum_{\lambda\notin\Delta_n} |\hat{x}_n(\lambda) - \hat{x}_m(\lambda)| \\ &= \sum_{j=1}^{n} |\hat{x}_n(\lambda_j) - \hat{x}_m(\lambda_j)| = \sum_{j=m+1}^{n} \frac{1}{j^2}.\end{aligned}$$

Thus, (x_n) is a Cauchy sequence with respect to $\|\cdot\|_*$. We show that (x_n) does not converge to any $x \in X$ with respect to $\|\cdot\|_*$.

Let $x \in X$. Then we observe that for all $j \leq n$,

$$\|x - x_n\|_* = \sum_{\lambda\in\Lambda} |\hat{x}(\lambda) - \hat{x}_n(\lambda)| \geq \sum_{j=1}^{n} \left|\hat{x}(\lambda_j) - \frac{1}{j^2}\right| \geq \left|\hat{x}(\lambda_j) - \frac{1}{j^2}\right|.$$

Since $\hat{x}(\lambda) = 0$ for all but a finite number of λ's, it follows that $\|x - x_n\|_* \not\to 0$.

Now let $X_0 = X$ with $\|\cdot\|$ and $Y = X$ with norm $\|\cdot\|_*$. Recall that $\|x\| \leq \|x\|_*$ for all $x \in X$. Hence, the identity operator $B : Y \to X$, i.e., $Bx = x$ for $x \in Y$, is a continuous bijective operator. Hence, by Theorems 3.17 and 3.19, its inverse $A := B^{-1} : X \to Y$ is a closed operator. But A is not continuous. If it were continuous, then there would exist $c > 0$ such that $\|x\|_* \leq c\|x\|$ for every $x \in X$, giving

$$\|x\| \leq \|x\|_* \leq c\|x\| \quad \forall x \in X.$$

This is not possible since $\|\cdot\|$ is complete but $\|\cdot\|_*$ is not complete.

Continuity of projection operators

The closed graph theorem gives a criterion for continuity of projection operators. Recall that a linear operator $P : X \to X$ on a linear space X is called a projection if $P^2 = P$.

Suppose X is a normed linear space, and $P : X \to X$ is a continuous projection. Since $R(P) = N(I - P)$, it follows that both $N(P)$ and $R(P)$ are closed subspaces of X.

Corollary 7.3 *Let X be a Banach space and $P : X \to X$ be a projection. If $N(P)$ and $R(P)$ are closed subspaces of X, then P is continuous.*

Proof. Suppose $N(P)$ and $R(P)$ are closed subspaces. To see that P is continuous, by closed graph theorem, it is enough to prove that P is a closed linear operator. For this, let (x_n) in X be such that $x_n \to x$ and $Px_n \to y$. We have to show that $y = Px$. Since $R(P)$ is closed, it follows that $y \in R(P)$ so that $Py = y$. Also, since $x_n - Px_n \in R(I-P) = N(P)$, $x_n - Px_n \to x - y$ and $N(P)$ is closed, we have $x - y \in N(P)$. Hence, $Px = Py = y$. Thus, the proof is complete. ∎

7.2 Bounded Inverse Theorem

We have seen in Section 3.5 that the inverse of an injective closed operator is a closed operator, whereas inverse of an injective bounded operator need not be a bounded operator.

But as an immediate consequence of closed graph theorem, we have the following theorem.

Theorem 7.4 (Bounded inverse theorem) *Let X and Y be Banach spaces, and $A : X_0 \subseteq X \to Y$ be an injective closed operator. Then $A^{-1} : R(A) \to X$ is continuous if and only if $R(A)$ is closed in Y.*

Proof. By Theorem 3.19, $A^{-1} : R(A) \subseteq Y \to X$ is a closed operator. Hence, by Theorem 7.2, $A^{-1} : R(A) \subseteq Y \to X$ is continuous if and only if $R(A)$ is closed in Y. ∎

As a corollary to the above result, we obtain the following, which is also known as bounded inverse theorem.

Corollary 7.5 (Bounded inverse theorem) *Let X and Y be Banach spaces, and $A \in \mathcal{B}(X,Y)$ be bijective. Then $A^{-1} \in \mathcal{B}(Y,X)$.*

Exercise 7.1 Let X be a linear space such that it is a Banach space with respect to norms $\|\cdot\|$ and $\|\cdot\|_*$. Suppose that there exists $c > 0$ such that $\|x\|_* \le c\|x\|$ for all $x \in X$. Show that there exists $c_1 > 0$ such that $\|x\| \le c_1\|x\|_*$ for all $x \in X$; consequently, $\|\cdot\|$ and $\|\cdot\|_*$ are equivalent. □

Recall that a function $f : \Omega_1 \to \Omega_2$ between two metric spaces Ω_1 and Ω_2 is a homeomorphism if and only if f is bijective, continuous and its inverse $f^{-1} : \Omega_2 \to \Omega_1$ is also continuous. Thus, the bounded inverse theorem states that a bijective bounded linear operator between two Banach spaces is a linear homeomorphism.

One may ask whether the completeness condition on X and Y in the above two results is necessary for the inverse operator to be continuous. Of course not, as the identity operator on a normed linear space has always continuous inverse. A more general situation is when the operator $A : X \to Y$ is a linear isometry, i.e., when $\|Ax\| = \|x\|$ for every $x \in X$. In fact, if $A : X \to Y$ is a linear isometry, then the inverse operator $A^{-1} : R(A) \to X$ is also an isometry, in particular, the inverse is continuous.

The above remark does not mean that the completeness condition on the spaces X and Y in Theorem 7.4 can be dropped. Recall that in Example 7.2 we have a closed operator with its domain space not a Banach space, and its inverse not continuous. The following example shows that the completeness assumption on the codomain space also cannot be dropped.

EXAMPLE 7.3 Let $X = C[a,b]$ with $\|\cdot\|_\infty$, and $Y = C[a,b]$ with $\|.\|_1$. Consider the map $A : X \to Y$ defined by $Ax = x$, $x \in X$. Then we see that A is bijective, linear and continuous. But its inverse is not continuous since X is a Banach space, but Y is not a Banach space.

7.3 Open Mapping Theorem

Recall that a function $f : \Omega_1 \to \Omega_2$ between two metric spaces Ω_1 and Ω_2 is an *open map* if for every open subset G of Ω_1, its image $f(G)$ is an open subset of Ω_2. It is easily seen that

(1) if $f : \Omega_1 \to \Omega_2$ and $g : \Omega_2 \to \Omega_3$ are open maps, then their composition $g \circ f : \Omega_1 \to \Omega_3$ is an open map, and

(2) if $f : \Omega_1 \to \Omega_2$ is bijective, then f is open if and only if its inverse $f^{-1} : \Omega_2 \to \Omega_1$ is continuous.

In the context of normed linear spaces and linear operators, the following result is worth noting.

Proposition 7.6 *Let X and Y be normed linear spaces, and $A : X \to Y$ be a linear operator. If A is an open map, then A is surjective.*

Proof. If A is an open map, then, in particular, the subspace $R(A)$ is an open set. Then by Lemma 2.32, $R(A)$ cannot be a proper subspace of Y. Thus, A is surjective. ∎

What about the converse of the above result? The converse holds good, provided X and Y are Banach spaces, and A is a bounded linear operator. This is the *open mapping theorem.* For its proof, we shall make use of the following two lemmas.

Recall from Example 1.6(vi) that if Z is a subspace of a linear space X, then the *quotient map* $\eta : X \to X/Z$ defined by $\eta(x) = x + Z$, $x \in X$, is a linear operator.

Lemma 7.7 *Let Z be a closed subspace of a normed linear space X. Then the quotient map $\eta : X \to X/Z$ is a linear, surjective, continuous and open map.*

Proof. Linearity and surjectivity of η is clear from its definition. Since

$$\|\eta(x)\| = \|x + Z\| \leq \|x\| \quad \forall x \in X,$$

the map η is continuous as well. To see that η is open, let G be an open subset of X. It is enough to show that for every $x_0 \in G$, there exists $r > 0$ such that

$$x \in X, \quad \|(x + Z) - (x_0 + Z)\| < r \implies x + Z \in \eta(G).$$

Let $x_0 \in G$. Since G is open in X, there exists $r > 0$ such that

$$x \in X, \qquad \|x - x_0\| < r \implies x \in G.$$

Now for $x \in X$, the relation $\|(x + Z) - (x_0 + Z)\| < r$ implies that there exists $z \in Z$ such that $\|x - x_0 + z\| < r$. Therefore, $x + z \in G$, and hence, $x + Z = x + z + Z \in \eta(G)$. ∎

Suppose $A : X \to Y$ is a linear operator, and Z is a closed subspace of X such that $Z \subseteq N(A)$. Then it is easily seen (*Verify*) that

$$\widetilde{A}(x + Z) = Ax, \qquad x \in X$$

defines a linear operator $\widetilde{A} : X/Z \to Y$. We may call this operator as **quotient operator** induced by A. Note that $\widetilde{A} \circ \eta = A$. We may also observe that $\widetilde{A}$ is injective if and only if $Z = N(A)$.

Lemma 7.8 *Let X and Y be normed linear spaces, $A \in \mathcal{B}(X, Y)$, and Z be a closed subspace of X such that $Z \subseteq N(A)$. Then $\widetilde{A}$ defined above, is a bounded linear operator, and $\|\widetilde{A}\| = \|A\|$.*

Proof. The operator $\widetilde{A} : X/Z \to Y$ is continuous because, for every $x \in X$,

$$\|\widetilde{A}(x + Z)\| = \|Ax\| = \|A(x + u)\| \leq \|A\| \, \|x + u\| \quad \forall u \in Z,$$

so that

$$\|\widetilde{A}(x+Z))\| \leq \|A\| \, \|x+Z\|.$$

This also shows that $\|\widetilde{A}\| \leq \|A\|$. To see the reverse inequality, we observe that for every $x \in X$,

$$\|Ax\| = \|\widetilde{A}(x+Z)\| \leq \|\widetilde{A}\| \, \|x+Z\| \leq \|\widetilde{A}\| \, \|x\|.$$

Thus, $\|\widetilde{A}\| = \|A\|$. ∎

Theorem 7.9 (Open mapping theorem) *Let X and Y be Banach spaces, and $A \in \mathcal{B}(X,Y)$. If A is surjective, then it is an open map.*

Proof. Suppose A is surjective. Taking $Z = N(A)$ in Lemma 7.8, it follows that the operator $\widetilde{A} : X/N(A) \to Y$ defined by $\widetilde{A}(x + N(A)) = Ax$, $x \in X$, is a bijective bounded linear operator. Therefore, by Corollary 7.5, the inverse of $\widetilde{A}$ is also continuous. In particular, $\widetilde{A}$ is an open map. Since $A = \widetilde{A} \circ \eta$ and, by Lemma 7.7, the quotient map η is an open map, it follows that A is an open map. ∎

Corollary 7.10 *Every nonzero continuous linear functional on a Banach space is an open map.*

Proof. We know that every nonzero linear functional is surjective. Therefore, the result follows by the open mapping theorem. ∎

Exercise 7.2 Give a proof for the above corollary without using the completeness of X and continuity of f. □

Examples 7.2 and 7.3 also serve (*How?*) to show that completeness assumptions in the open mapping theorem cannot be dropped.

We now consider a consequence of the open mapping theorem, which is useful while solving an operator equation.

7.3.1 A Stability Result for Operator Equations

Consider an operator equation

$$Ax = y,$$

where $A : X \to Y$ is a bounded linear operator between Banach spaces X and Y. As we have already pointed out in the beginning of Chapter 3, one of the important issues to be considered while solving the above equation, in practical situations, is to see whether the solution is *stable* under perturbation in the *data* y. That is, if x and $\tilde{x}$ are solutions corresponding to data y and $\tilde{y}$, respectively, then we would like to have $\|x - \tilde{x}\|$ to be small whenever $\|y - \tilde{y}\|$ is small. Of course, if A is bijective, then by bounded inverse theorem, the above requirement will be satisfied. What can we say, if we only know that A is surjective? Here is an answer.

Theorem 7.11 *Let X and Y be Banach spaces, and $A \in \mathcal{B}(X, Y)$. If A is surjective, then there exists $c > 0$ such that for every $y \in Y$, there exists $x \in X$ such that*

$$Ax = y, \qquad \|x\| \leq c\|y\|.$$

Proof. Suppose $A \in \mathcal{B}(X, Y)$ is surjective. Then, by open mapping theorem, $A : X \to Y$ is an open map. In particular, the image of the open unit ball in X is open in Y and it contains 0. Hence, there exists $r > 0$ such that

$$\{v \in Y : \|v\| < r\} \subseteq \{Ax : \|x\| < 1\}.$$

Now, let $0 \neq y \in Y$. Then $v := ry/2\|y\|$ satisfies $\|v\| < r$, so that there exists $u \in X$ such that $\|u\| < 1$ and $Au = v$. Thus, for $y \in Y$, the result holds with $x = (2\|y\|/r)u$ and $c = 2/r$. ∎

PROBLEMS

1. Let X be a normed linear space, and Y and Z be Banach spaces. Let $A : X_0 \subseteq X \to Y$ be a closed operator and $B \in \mathcal{B}(Z, X)$ such that $R(B) \subseteq X_0$. Show that $AB \in \mathcal{B}(Z, Y)$.

2. Let X and Y be normed linear spaces and $A : X_0 \subseteq X \to Y$ be a closed operator. If $B : X \to Y$ is a bounded operator, then show that $A + B : X_0 \to Y$ is a closed linear operator.

3. Using bounded inverse theorem, prove the closed graph theorem.

[*Hint*: Observe that, if $A : X_0 \subseteq X \to Y$ is a closed operator, where X and Y are Banach spaces, then $A = \pi_2\pi_1^{-1}$, where $\pi_1 : G(A) \to X_0$ and $\pi_2 : X_0 \times Y \to Y$ are defined by $\pi_1(x, Ax) = x$, $\pi_2(x, y) = y$ for $x \in X_0$, $y \in Y$.]

4. Let X and Y be Banach spaces and $\mathcal{F}$ be a pointwise bounded subset of $\mathcal{B}(X,Y)$. Let $\ell^\infty(\mathcal{F},Y)$ be the set of all bounded functions from $\mathcal{F}$ into Y, and $T : X \to \ell^\infty(\mathcal{F},Y)$ be defined by $(Tx)(A) = Ax$ for every $x \in X$, $A \in \mathcal{F}$. Show the following:

(a) $\ell^\infty(\mathcal{F},Y)$ is a linear space, and it is Banach space with respect to the norm $f \mapsto \|f\| = \sup_{A\in\mathcal{F}} \|f(A)\|$.

(b) T is a closed linear operator.

5. Using the above problem, give another proof for the uniform boundedness principle.

6. Suppose that X and Y are Banach spaces, and $A : X \to Y$ is a bijective bounded operator. For $y, y_n \in Y$, let $x, x_n \in X$ be such that $Ax = y$ and $Ax_n = y_n$ for $n \in \mathbb{N}$. Show that $y_n \to y$ as $n \to \infty$ implies $x_n \to x$ as $n \to \infty$.

7. If X is a Banach space with respect to norms $\|\cdot\|_1$ and $\|\cdot\|_2$, and if there exists $c > 0$ such that $\|x\|_2 \le c\|x\|_1$ for all $x \in X$, then show that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.

8. Let X and Y be Banach spaces and $A \in \mathcal{B}(X,Y)$. If there exists a linear operator $B : Y \to X$ such that AB and BA are identity operators on Y and X, respectively, then show that A is bijective, $B \in \mathcal{B}(Y,X)$ and $A^{-1} = B$.

9. Let X be a Banach space with a Schauder basis $\{u_1, u_2, \ldots\} \subseteq X$. For $x \in X$, let $\alpha_j(x) \in \mathbb{K}$ for $j \in \mathbb{N}$, be such that $x = \sum_{j=1}^\infty \alpha_j u_j$. Show that for each $j \in \mathbb{N}$, the map $f_j : X \to \mathbb{K}$ defined by $f_j(x) = \alpha_j(x)$, $x \in X$, is a continuous linear functional.

[*Hint*: For $x \in X$, let $\|x\|_* = \sup_{n\in\mathbb{N}} \left\|\sum_{j=1}^n \alpha_j u_j\right\|$, where (α_j) is such that $x = \sum_{j=1}^\infty \alpha_j u_j \in X$. Show that $\|\cdot\|_*$ is a complete norm on X, and then use the fact that $|f_j(x)| \le 2\|x\|_*$.]

10. Let X be a Hilbert space and $A : X \to X$ be a linear operator such that $\langle Ax, y\rangle = \langle x, Ay\rangle$ for every $x, y \in X$. Show that $A \in \mathcal{B}(X)$.

11. Let X and Y be Hilbert spaces, and $A \in \mathcal{B}(X,Y)$. Show that the generalized inverse $A^\dagger : R(A) + R(A)^\perp \to X$ of A is continuous if and only if $R(A)$ is a closed subspace of Y (Recall the definition of $A^\dagger$ from Problem 24 in Chapter 3).

8

Dual Space Considerations

Suppose X is a normed linear space. We know that its dual X', the space of all bounded linear functionals on X, is a normed linear space, with the norm given by

$$\|f\| = \sup\{\,|f(x)| : x \in X,\ \|x\| = 1\},\qquad f \in X',$$

and X' is a Banach space with respect to this norm (Theorem 3.12). In Chapter 5, we discussed the question whether X' is nonzero or not. In fact, we proved that

$$\dim X' \geq \dim X.$$

We also know (*Do you?*) that if X is a finite dimensional space, then X' is linearly homeomorphic with X. In case X is a Hilbert space, then, by the Riesz representation theorem (Theorem 3.9), each $f \in X'$ is represented by a unique element $v_f \in X$ by the relation $f(x) = \langle x, v_f\rangle$, $x \in X$, and that the map $f \mapsto v_f$ is a conjugate linear isometry from X' onto X.

With X as any of the spaces $\ell^2(n)$, ℓ^2, $L^2[a,b]$, let us consider a new 'scalar product' $[\cdot,\cdot]_2$ on X (instead of its inner product), namely,

$$[x,y]_2 = \begin{cases} \sum_{j=1}^{n} x(j)y(j) & \text{if } x,y \in \ell^2(n),\\ \sum_{j=1}^{\infty} x(j)y(j) & \text{if } x,y \in \ell^2,\\ \int_a^b x(t)y(t)\,d\mu(t) & \text{if } x,y \in L^2[a,b]. \end{cases}$$

Then, using Theorem 3.9, it can be seen that the map $T : X \to X'$ defined by

$$(Ty)(x) = [x,y]_2,\qquad x,y \in X,$$

is a linear isometry from X onto X'.

This observation motivates us to look for linear isometries between the spaces $\ell^p(n)$, ℓ^p, $L^p[a,b]$ and their duals. For this, let $1 < p < \infty$, and let X_p be any of the spaces $\ell^p(n)$, ℓ^p, $L^p[a,b]$. Consider the map $[\cdot,\cdot]_{p,q} : X_p \times X_q \to \mathbb{K}$ defined by

$$[x,y]_{p,q} = \begin{cases} \sum_{j=1}^{n} x(j)y(j) & \text{if } x \in \ell^p(n),\ y \in \ell^q(n), \\ \sum_{j=1}^{\infty} x(j)y(j) & \text{if } x \in \ell^p,\ y \in \ell^q, \\ \int_a^b x(t)y(t)\,d\mu(t) & \text{if } x \in L^p[a,b],\ y \in L^q[a,b], \end{cases}$$

where q is the conjugate exponent of p. By Hölder's inequality, it follows that $[x,y]_{p,q}$ is well defined for all $x \in X_p$, $y \in X_q$, and it is linear in each of its variable. In fact, for each $y \in X_q$, the map $f_y : X_p \to \mathbb{K}$ defined by $f_y(x) = [x,y]_{p,q}$, $x \in X_p$, is a continuous linear functional on X_p and $\|f_y\| \le \|y\|_q$. The question is whether the correspondence $y \mapsto f_y$ is an isometry from X_q onto X_p'. We shall show that this is indeed true for all p with $1 \le p < \infty$.

In Section 8.1, we obtain representations of the duals of the spaces $\ell^p(n)$ for $1 \le p \le \infty$, ℓ^p for $1 \le p < \infty$, c_{00} with $\|\cdot\|_p$ for $1 \le p \le \infty$, c_0 and c with $\|\cdot\|_\infty$, $C[a,b]$ with $\|\cdot\|_\infty$, and $L^p[a,b]$ for $1 \le p < \infty$. We shall also make a revisit to separability, which helps us to infer why certain procedures do not work for the spaces ℓ^∞ and $L^\infty[a,b]$. In Section 8.2, we study certain properties of the spaces, namely reflexivity and weak convergence, using the duality consideration. The results based on these properties help us to infer more about the space and answer certain questions related to the Riesz lemma and best approximation.

8.1 Representation of Dual Spaces

8.1.1 Dual of $\ell^p(n)$

Recall from Example 2.1(viii) that the space $\ell^p(n)$ is the linear space $\mathbb{K}^n$ with the norm $\|\cdot\|_p$ defined by

$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^{n} |x(i)|^p\right)^{1/p} & \text{if } 1 \le p < \infty \\ \max\{|x(i)| : i = 1,\ldots,n\} & \text{if } p = \infty. \end{cases}$$

We show that the dual of $\ell^p(n)$ is linearly isometric with $\ell^q(n)$.

Theorem 8.1 *The map $T : (\ell^p(n))' \to \ell^q(n)$ defined by*

$$T(f) = (f(e_1), \dots, f(e_n)), \qquad f \in (\ell^p)',$$

is a surjective linear isometry, where $e_j \in \mathbb{K}^n$ is such that $e_j(i) = \delta_{ij}$ for $i, j = 1, \dots, n$.

Proof. Let $X = \ell^p(n)$, $1 \le p \le \infty$. First we note that the map $T : X' \to \ell^q(n)$ defined by

$$T(f) = (f(e_1), \dots, f(e_n)), \qquad f \in X',$$

is surjective. Indeed, for $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$, the function $f : X \to \mathbb{K}$ defined by

$$f(x) = \sum_{j=1}^{n} x(j)\beta_j, \qquad x \in X,$$

belongs to X' and $T(f) = (\beta_1, \dots, \beta_n)$.

It remains to show that $\|T(f)\|_q = \|f\|$ for every $f \in X'$. For this, let $f \in X'$. Since for every $x \in X$, $x = \sum_{i=1}^n x(i)e_i$ and $f(x) = \sum_{i=1}^n x(i)f(e_i)$, by Hölder's inequality, we have

$$|f(x)| \le \|x\|_p \|T(f)\|_q \quad \forall x \in X.$$

Therefore,

$$\|f\| \le \|T(f)\|_q \quad \forall f \in X'.$$

Also, since

$$|f(e_j)| \le \|f\| \, \|e_j\|_p = \|f\| \quad \forall j = 1, \dots, n,$$

we have $\|T(f)\|_\infty \le \|f\|$ for all $f \in X'$. In order to show that $\|T(f)\|_q \le \|f\|$ for $1 \le q < \infty$, we first observe that

$$\|T(f)\|_q^q = \sum_{j=1}^{n} |f(e_j)|^q = \sum_{j=1}^{n} \alpha_j f(e_j),$$

where

$$\alpha_j = \begin{cases} |f(e_j)|^q / f(e_j) & \text{if } f(e_j) \ne 0 \\ 0 & \text{if } f(e_j) = 0. \end{cases}$$

Thus, taking $x_0 = \sum_{j=1}^n \alpha_j e_j$, we have

$$\|T(f)\|_q^q = f(x_0) \le \|f\| \, \|x_0\|_p.$$

Note that $\|x_0\|_\infty \le 1$, and for $p \ne \infty$,

$$\|x_0\|_p^p = \sum_{j=1}^{n} |\alpha_j|^p = \sum_{j=1}^{n} |f(e_j)|^{pq-p} = \|T(f)\|_q^q.$$

Hence, for $p = \infty$,

$$\|T(f)\|_1 = |f(x_0)| \le \|f\|\, \|x_0\|_\infty \le \|f\|,$$

and for $1 < p < \infty$,

$$\|T(f)\|_q^q = |f(x_0)| \le \|f\|\, \|x_0\|_p = \|f\|\, \|T(f)\|_q^{q/p}.$$

Thus, $\|T(f)\|_q \le \|f\|$ for $1 \le p \le \infty$. This completes the proof. ∎

We recall from Chapter 2 (Remark 2.3) that if X is any linear space of dimension n, $E = \{u_1, \ldots, u_n\}$ is a basis of X, and $f_1, \ldots, f_n$ are the coordinate functionals on X associated with E, then

$$\|x\|_{E,p} = \begin{cases} \left(\sum_{i=1}^{n} |f_i(x)|^p\right)^{1/p} & \text{if } 1 \le p < \infty \\ \max\{|f_i(x)| : i = 1, \ldots, n\} & \text{if } p = \infty, \end{cases}$$

defines a norm on X for each $p \in [1, \infty]$. With this norm on X, following the arguments as in the above theorem, it can be seen that the map $T : X' \to \ell^q(n)$ defined by

$$T(f) = (f(u_1), \ldots, f(u_n)), \qquad f \in X'_{E,p},$$

is a surjective linear isometry.

8.1.2 Duals of Some Sequence Spaces

In this section, we shall consider the representation of the duals of the sequence spaces ℓ^p with $1 \le p < \infty$, c_{00} with $\|\cdot\|_p$ for $1 \le p \le \infty$, and c_0 and c with $\|\cdot\|_\infty$. In Section 8.1.4, we show that the dual of ℓ^∞ is not homeomorphic with any separable space.

We observe that if $x \in \ell^p$ and $y \in \ell^q$, where q is the conjugate exponent of p, then by Hölder's inequality, we have

$$\sum_{j=1}^{\infty} |x(j)y(j)| \le \|x\|_p \|y\|_q$$

so that the series $\sum_{j=1}^{\infty} x(j)y(j)$ converges. Also, recall that if $1 \leq p < \infty$, then for any $x \in \ell^p$,

$$\lim_{n\to\infty} \left\| x - \sum_{j=1}^{n} x(j)e_j \right\|_p = 0.$$

Here, and in what follows, $e_j \in \mathcal{F}(\mathbb{N}, \mathbb{K})$ is such that $e_j(i) = \delta_{ij}$ for $i, j \in \mathbb{N}$. Thus,

$$x = \sum_{j=1}^{\infty} x(j)e_j \quad \forall x \in \ell^p$$

for $1 \leq p < \infty$ which, in particular, shows that the set $\{e_1, e_2, \ldots\}$ is a *Schauder basis* of ℓ^p for $1 \leq p < \infty$.

Dual of ℓ^p, $1 \leq p < \infty$.

Theorem 8.2 *Let $1 \leq p \leq \infty$. For each $y \in \ell^q$, let $f_y : \ell^p \to \mathbb{K}$ be defined by*

$$f_y(x) = \sum_{j=1}^{\infty} x(j)y(j), \qquad x \in \ell^p.$$

Then $f_y \in (\ell^p)'$, and the map $T : \ell^q \to (\ell^p)'$ defined by

$$T(y) = f_y, \qquad y \in \ell^q,$$

is a linear isometry. If $1 \leq p < \infty$, then T is surjective. In fact, if $1 \leq p < \infty$, then for every $f \in (\ell^p)'$, $y := (f(e_1), f(e_2), \ldots) \in \ell^q$ and $T(y) = f$.

Proof. Let $y \in \ell^q$ and $f = f_y$ be defined by

$$f(x) = f_y(x) = \sum_{j=1}^{\infty} x(j)y(j), \qquad x \in \ell^p.$$

We have already observed that the above series is convergent, and

$$\sum_{j=1}^{\infty} |x(j)y(j)| \leq \|x\|_p \|y\|_q, \qquad x \in \ell^p.$$

In particular, we have

$$f \in (\ell^p)', \qquad \|f\| \leq \|y\|_q.$$

Next, we show that $\|y\|_q \le \|f\|$.

First we note that $f(e_j) = y(j)$ and $|y(j)| = |f(e_j)| \le \|f\|$ for all $j \in \mathbb{N}$ so that $\|y\|_\infty \le \|f\|$. Thus, for $p = 1$, we have $\|y\|_q \le \|f\|$. Now, let $1 < p \le \infty$, i.e., $1 \le q < \infty$. For each $n \in \mathbb{N}$, we have

$$\sum_{j=1}^{n} |y(j)|^q = \sum_{j=1}^{n} x_n(j)y(j) = \sum_{j=1}^{n} x_n(j)f(e_j),$$

where

$$x_n(j) = \begin{cases} |y(j)|^q/y(j) & \text{if } y(j) \ne 0 \text{ and } j \le n \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $x_n \in c_{00}$, and

$$\sum_{j=1}^{n} |y(j)|^q = f(x_n) \le \|f\| \, \|x_n\|_p.$$

Note that $\|x_n\|_\infty \le 1$ so that $\|y\|_1 \le \|f\|$. For $1 < p < \infty$,

$$\|x_n\|_p^p = \sum_{j=1}^{n} |x_n(j)|^p = \sum_{j=1}^{n} |y(j)|^q.$$

Therefore, for $1 < q < \infty$,

$$\sum_{j=1}^{n} |y(j)|^q = f(x_n) \le \|f\| \, \|x\|_p = \|f\| \left(\sum_{j=1}^{n} |y(j)|^q \right)^{1/p}$$

so that $\|y\|_q \le \|f\|$. Thus, we have shown that $T : \ell^q \to (\ell^p)'$ defined by

$$T(y) = f_y, \qquad y \in \ell^q,$$

is an isometry. Clearly, T is linear.

Next, we show that if $1 \le p < \infty$, then the above map T is surjective. So let $1 \le p < \infty$ and $f \in (\ell^p)'$. Then for every $x \in \ell^p$, we have

$$x = \lim_{n\to\infty} \sum_{j=1}^{n} x(j)e_j, \tag{8.1}$$

so that by continuity of f,

$$f(x) = \lim_{n\to\infty} \sum_{j=1}^{n} x(j)f(e_j) = \sum_{j=1}^{\infty} x(j)f(e_j).$$

Now, by taking $y = (f(e_1), f(e_2), \ldots)$, it is seen that

$$y \in \ell^q, \qquad \|y\|_q \le \|f\|.$$

Thus,

$$f(x) = \sum_{j=1}^{\infty} x(j) f(e_j) = f_y(x) \quad \forall\, x \in \ell^p,$$

showing that the map $y \mapsto f_y$ from ℓ^q into $(\ell^p)'$ is surjective. ∎

Corollary 8.3 *Let $1 < p < \infty$ and for each $j \in \mathbb{N}$, let $f_j(x) = x(j)$ for $x \in \ell^p$. Then, $f_j \in (\ell^p)'$ for all $j \in \mathbb{N}$, and $\{f_j : j \in \mathbb{N}\}$ is a Schauder basis of $(\ell^p)'$. In fact,*

$$f = \sum_{j=1}^{\infty} f(e_j) f_j \quad \forall\, f \in (\ell^p)'.$$

Proof. Clearly, $f_j \in (\ell^p)'$ for all $j \in \mathbb{N}$. Now let $f \in (\ell^p)'$. Then, by the above theorem,

$$f(x) = \sum_{j=1}^{\infty} x(j) y(j) \quad \forall\, x \in \ell^p,$$

where $y(j) = f(e_j)$ for all $j \in \mathbb{N}$. We observe that for every $x \in \ell^p$,

$$\Big| f(x) - \sum_{j=1}^{n} x(j) y(j) \Big| \le \sum_{j=n+1}^{\infty} |x(j) y(j)| \le \|x\|_p \Big(\sum_{j=n+1}^{\infty} |y(j)|^q \Big)^{1/q}.$$

Now, since $x(j) = f_j(x)$ and $y(j) = f(e_j)$ for all $j \in \mathbb{N}$, we have

$$f(x) - \sum_{j=1}^{n} x(j) y(j) = \Big(f - \sum_{j=1}^{n} f(e_j) f_j \Big)(x) \quad \forall\, x \in \ell^p.$$

Hence,

$$\Big\| f - \sum_{j=1}^{n} f(e_j) f_j \Big\|_p \le \Big(\sum_{j=n+1}^{\infty} |y(j)|^q \Big)^{1/q} \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, $f = \sum_{j=1}^{\infty} f(e_j) f_j$. ∎

Duals of $(c_{00}, \|\cdot\|_p)$ and $(c_0, \|\cdot\|_\infty)$.

We may observe that the proof of Theorem 8.2 holds good if we replace ℓ^p by $(c_{00}, \|\cdot\|_p)$ for any p with $1 \le p \le \infty$ or by $(c_0, \|\cdot\|_\infty)$. In the proof of the last part of Theorem 8.2, namely, the proof for the surjectivity of T, the condition $p \neq \infty$ is used only to have the relation (8.1). But we know (*Verify*) that if X is $(c_{00}, \|\cdot\|_p)$ or $(c_0, \|\cdot\|_\infty)$, then (8.1) does hold for every $x \in X$.

Thus, we obtain the following two theorems.

Theorem 8.4 *For $1 \le p \le \infty$, let X_p be the space c_{00} with $\|\cdot\|_p$. Then for every $y \in \ell^q$, $f_y : X_p \to \mathbb{K}$ defined by*

$$f_y(x) = \sum_{j=1}^{\infty} x(j)y(j), \qquad x \in X_p,$$

is an element of X_p', and the map $y \mapsto f_y$ is a linear isometry from ℓ^q onto X_p'.

Theorem 8.5 *Let X be the space c_0 with $\|\cdot\|_\infty$. Then for every $y \in \ell^1$, $f_y : X \to \mathbb{K}$ be defined by*

$$f_y(x) = \sum_{j=1}^{\infty} x(j)y(j), \qquad x \in X,$$

belongs to X', and the map $y \mapsto f_y$ is a linear isometry from ℓ^1 onto X'.

The following lemma shows that the dual of a dense subspace is linearly isometric with the dual of the space itself.

Lemma 8.6 *Let X_0 be a dense subspace of a normed linear space X and let $F : X' \to X_0'$ be defined by*

$$F(f) = f|_{X_0}, \qquad f \in X'.$$

Then F is a linear isometry from X' onto X_0'.

Proof. The result follows by invoking Theorem 3.18. ∎

Recall from Remark 2.8 that for $1 \le p < \infty$, $(c_{00}, \|\cdot\|_p)$ is dense in ℓ^p and $(c_{00}, \|\cdot\|_\infty)$ is dense in $(c_0, \|\cdot\|_\infty)$. Hence, using the above lemma, the facts that $(c_{00}, \|\cdot\|_p)'$ is linearly isometric with ℓ^q for $1 \le p < \infty$ and $(c_{00}, \|\cdot\|_\infty)'$ is linearly isometric with ℓ^1 follow also from Theorems 8.2 and 8.4, respectively.

Dual of $(c, \|\cdot\|_\infty)$.

Recall that c is the space of all convergent scalar sequences. Though the spaces c_{00} and c_0 are not dense in $(c, \|\cdot\|_\infty)$, we still see that its dual is ℓ^1. We may observe, taking $e_0 = (1, 1, \ldots) \in c$, that

$$c = \{u + \alpha e_0 : u \in c_0,\ \alpha \in \mathbb{K}\}.$$

Since every $u \in c_0$ can be written as $u = \sum_{j=1}^\infty u(j) e_j$, it follows that, for every $u \in c_0$, $\alpha \in \mathbb{K}$, $f \in c'$,

$$f(u + \alpha e_0) = \alpha f(e_0) + \sum_{j=1}^\infty u(j) f(e_j).$$

Here we used the fact that $(f(e_1), f(e_2), \ldots)$ belongs to ℓ^1, which is so, since the restriction of every $f \in c'$ to the space c_0 belongs to $(c_0)'$. From the above representation, it follows (*How?*) that

$$f(x) = \alpha\beta + \sum_{j=1}^\infty x(j) f(e_j) \quad \forall x \in c,$$

where

$$\alpha = \lim_{n\to\infty} x(n), \qquad \beta = f(e_0) - \sum_{j=1}^\infty f(e_j). \tag{8.2}$$

Theorem 8.7 *Let $X = c$ with $\|\cdot\|_\infty$. For every $f \in X'$, the sequence $(f(e_n))$ belongs to ℓ^1, and*

$$f(x) = \alpha\beta + \sum_{j=1}^\infty x(j) f(e_j) \quad \forall x \in X,$$

where α and β are as in (8.2). Moreover, the map

$$f \mapsto (\beta, f(e_1), f(e_2), \ldots), \qquad f \in X',$$

is a surjective linear isometry from X' onto ℓ^1.

Proof. For $f \in X'$, let $\beta_j = f(e_j)$ for all $j \in \mathbb{N}$, and let

$$\beta = f(e_0) - \sum_{j=1}^\infty \beta_j,$$

with $e_0(j) = 1$ for all $j \in \mathbb{N}$. We have already observed that $(\beta_1, \beta_2, \ldots) \in \ell^1$, and

$$f(x) = \alpha\beta + \sum_{j=1}^{\infty} x(j)\beta_j \quad \forall x \in X,$$

where $\alpha = \lim_{n\to\infty} x(n)$. It is also clear from this representation that

$$\|f\| \leq |\beta| + \sum_{j=1}^{\infty} |\beta_j|.$$

Next, we prove $\|f\| \geq |\beta| + \sum_{j=1}^{\infty} |\beta_j|$. For this, let

$$x_k = (\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha, \alpha, \ldots),$$

where

$$\alpha_j = \operatorname{sgn}(\beta_j), \quad j = 1, \ldots, k; \qquad \alpha = \operatorname{sgn}(\beta).$$

Then we have $\|x_k\|_\infty \leq 1$ for all $k \in \mathbb{N}$ and

$$f(x_k) = |\beta| + \sum_{j=1}^{k} |\beta_j| + \alpha \sum_{j=k+1}^{\infty} \beta_j.$$

Now, making use of the fact that $(\beta_1, \beta_2, \ldots) \in \ell^1$, we have

$$|\beta| + \sum_{j=1}^{k} |\beta_j| \leq |f(x_k)| + |\alpha| \sum_{j=k+1}^{\infty} |\beta_j| \leq \|f\| + |\alpha| \sum_{j=k+1}^{\infty} |\beta_j|$$

for all $k \in \mathbb{N}$. Hence, it follows, by letting $k \to \infty$,

$$|\beta| + \sum_{j=1}^{\infty} |\beta_j| \leq \|f\|.$$

Thus, we have proved that for every $f \in X'$,

$$y := (\beta, \beta_1, \beta_2, \ldots) \in \ell^1, \qquad f(x) = \alpha\beta + \sum_{j=1}^{\infty} x(j)\beta_j \quad \forall x \in c,$$

and $\|f\| = \|y\|_1$, where $\alpha = \lim_{n\to\infty} x(n)$.

Conversely, suppose $y := (\eta, \eta_1, \eta_2, \ldots) \in \ell^1$. Let $f : X \to \mathbb{K}$ be defined by

$$f(x) = \eta\alpha + \sum_{j=1}^{\infty} x(j)\eta_j, \qquad \alpha := \lim_{n\to\infty} x(n).$$

Then it is seen, by a similar argument as above, that $f \in X'$ and $\|f\| = \|y\|_1$.

Thus, we have obtained a linear isometry from X' onto ℓ^1. ∎

A question which most of the readers must have had in their minds is: what about a representation for the dual of ℓ^∞? Is it not linearly isometric with ℓ^1? In Section 8.1.4, we shall answer this question negatively. In fact, we show that $(\ell^\infty)'$ is not linearly homeomorphic with any separable space. A representation of $(\ell^\infty)'$ involves some ideas from general measure theory which we did not develop in this text. It can be shown that the dual of ℓ^∞ is linearly isometric with a space of measures on all subsets of $\mathbb{N}$ with *total variation* as norm. We urge the more inquisitive reader to look into Yosida ([36], Chapter IV, Section 9, Example 5).

8.1.3 Duals of $C[a, b]$ and $L^p[a, b]$

In order to find a representation for the dual of the space $C[a, b]$ with $\|\cdot\|_\infty$ we recall some definitions and results from real analysis (for details, see Royden [26] or de Barra [8]).

A function $\phi : [a, b] \to \mathbb{K}$ is said to be a **function of bounded variation** on $[a, b]$ if there exists $c_\phi > 0$ such that

$$\sum_{j=1}^{n} |\phi(t_j) - \phi(t_{j-1})| \le c_\phi$$

for all partition $\Pi : a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$. If ϕ is of bounded variation on $[a, b]$, then the quantity

$$V(\phi) := \sup_{\Pi} \sum_{j=1}^{n} |\phi(t_j) - \phi(t_{j-1})|$$

is called the **total variation** of ϕ. Here the supremum is taken over all partitions Π of $[a, b]$.

Let us denote the set of all functions of bounded variation on $[a, b]$ by $BV[a, b]$. It is easily seen that $BV[a, b]$ is a linear space with respect to addition and scalar multiplication defined pointwise.

It is known that every real-valued function of bounded variation can be written as the difference of two monotonically increasing functions. That is, if $\phi \in BV[a, b]$ is real-valued, then there exist monotonically increasing functions $\psi_1, \psi_2 : [a, b] \to \mathbb{R}$ such that

$$\phi(t) = \psi_1(t) - \psi_2(t) \quad \forall t \in [a, b].$$

Hence, using the definition of *Riemann-Stieltjes integral* (Rudin [27]) of real-valued functions with respect to monotonically increasing functions, we can define the Riemann-Stieltjes integral of $\mathbb{K}$-valued functions on $[a, b]$ with respect to functions of bounded variation as follows: Suppose $\phi \in BV[a, b]$ is real-valued and $\phi = \psi_1 - \psi_2$, where ψ_1, ψ_2 are real-valued monotonically increasing functions on $[a, b]$. Suppose $x : [a, b] \to \mathbb{R}$ is such that the Riemann-Stieltjes integrals

$$\int_a^b x \, d\psi_1, \qquad \int_a^b x \, d\psi_2$$

exist. Then the Riemann-Stieltjes integral of f with respect to ϕ is defined by

$$\int_a^b x \, d\phi := \int_a^b x \, d\psi_1 - \int_a^b x \, d\psi_2.$$

If x is complex-valued such that

$$\int_a^b \mathrm{Re}(x) d\phi, \qquad \int_a^b \mathrm{Im}\,(x) d\phi$$

exist, then we define

$$\int_a^b x \, d\phi := \int_a^b \mathrm{Re}(x) \, d\phi + i \int_a^b \mathrm{Im}\,(x) \, d\phi.$$

Now, suppose that $\phi \in BV[a, b]$ is complex-valued. Then it can be seen that $\phi_1 := \mathrm{Re}(\phi)$ and $\phi_2 := \mathrm{Im}\,(\phi)$ are real-valued functions of bounded variation on $[a, b]$. If x is $\mathbb{K}$-valued such that

$$\int_a^b x \, d\phi_1, \qquad \int_a^b x \, d\phi_2$$

exist, then we define

$$\int_a^b x\,d\phi := \int_a^b x\,d\phi_1 + i\int_a^b x\,d\phi_2.$$

It can be shown that, if $\phi \in BV[a,b]$ and $x \in C[a,b]$, then $\int_a^b x\,d\phi$ exists, and it is the unique number with the property that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every partition

$$\Pi : a = t_0 < t_1 < \ldots < t_n = b$$

and for every $\Delta = \{\tau_1, \ldots, \tau_n\}$ with $\tau_j \in [t_{j-1}, t_j]$, $j = 1, \ldots, n$,

$$\left|\sum_{j=1}^{n} x(\tau_j)[\phi(t_j) - \phi(t_{j-1})] - \int_a^b x\,d\phi\right| < \varepsilon$$

whenever $\max_{1\le j\le n}(t_j - t_{j-1}) < \delta$, and, in that case, we write

$$\lim_{n\to\infty} \sum_{j=1}^{n} x(\tau_j)[\phi(t_j) - \phi(t_{j-1})] = \int_a^b x\,d\phi.$$

We observe that if $x \in C[a,b]$ and $\phi \in BV[a,b]$, then for every partition $\Pi : a = t_0 < t_1 < \ldots < t_n = b$ of $[a,b]$,

$$\left|\sum_{j=1}^{n} x(\tau_j)[\phi(t_j) - \phi(t_{j-1})]\right| \le \|x\|_\infty V(\phi).$$

Hence, it follows that

$$\left|\int_a^b x\,d\phi\right| \le \|x\|_\infty V(\phi).$$

Thus, for each $\phi \in BV[a,b]$, $f_\phi : C[a,b] \to \mathbb{K}$ defined by

$$f_\phi(x) = \int_a^b x\,d\phi, \qquad x \in C[a,b],$$

is a continuous linear functional on $(C[a,b], \|\cdot\|_\infty)$.

Note that, for $\phi_1, \phi_2 \in BV[a,b]$, $f_{\phi_1} = f_{\phi_2}$ whenever $\phi_1 - \phi_2$ is a constant function. Two different functions in $BV[a,b]$ which vanish at a can also give the same functional. In order that the map $\phi \mapsto f_\phi$ is injective, we restrict it to a subspace $NBV[a,b]$ of $BV[a,b]$. The space $NBV[a,b]$ consists of all those $x \in BV[a,b]$ such that $x(a) = 0$ and x is continuous from right on (a,b). Elements of $NBV[a,b]$ are called **normalized functions of bounded variation** on $[a,b]$.

It can be shown (*Verify*) that $NBV[a,b]$ is a linear space and

$$\phi \mapsto V(\phi), \qquad \phi \in NBV[a,b],$$

defines a norm on $NBV[a,b]$. We observe that $NBV[a,b]$ is not a separable space. Indeed, $E = \{\chi_{[a,s)} : a < s < b\}$ is an uncountable subset of $NBV[a,b]$ such that $V(\phi_1 - \phi_2) = 2$ for every distinct $\phi_1, \phi_2 \in E$, and hence, by Proposition 2.37, $NBV[a,b]$ is not separable.

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

Now we obtain a representation for the dual of $(C[a,b], \|\cdot\|_\infty)$.

Theorem 8.8 *Let $C[a,b]$ be with $\|.\|_\infty$. For $\phi \in NBV[a,b]$, let*

$$f_\phi(x) = \int_a^b x\, d\phi, \qquad x \in C[a,b].$$

Then $f_\phi \in (C[a,b])'$, and the map $\phi \mapsto f_\phi$ is a linear isometry from $NBV[a,b]$ onto the dual of $C[a,b]$.

Proof. Let $\phi \in NBV[a,b]$ and $f_\phi(x) = \int_a^b x\, d\phi$, $x \in C[a,b]$. Then we have

$$|f_\phi(x)| \le \|x\|_\infty V(\phi) \quad \forall x \in C[a,b]$$

so that

$$f_\phi \in (C[a,b])', \qquad \|f_\phi\| \le V(\phi).$$

Now let $f \in (C[a,b])'$. Then, by the Hahn-Banach extension theorem (Theorem 5.1), there exists a continuous linear functional F on $\ell^\infty([a,b])$ such that

$$F(x) = f(x) \quad \forall x \in C[a,b], \qquad \|F\| = \|f\|.$$

Define

$$\phi(t) = F(u_t), \qquad a \le t \le b,$$

where

$$u_t = \begin{cases} 0 & \text{if } t = a \\ \chi_{(a,t]} & \text{if } a < t \le b. \end{cases}$$

Then

$$\phi \in BV[a,b], \qquad V(\phi) \le \|f\|, \qquad \phi(a) = 0.$$

To see this, consider a partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a,b]$ and note that

$$\sum_{j=1}^{n} |\phi(t_j) - \phi(t_{j-1})| = \sum_{j=1}^{n} \alpha_j[\phi(t_j) - \phi(t_{j-1})],$$

where $\alpha_j = \operatorname{sgn}[\phi(t_j) - \phi(t_{j-1})]$ for $j = 1, \ldots, n$. Thus,

$$\sum_{j=1}^{n} |\phi(t_j) - \phi(t_{j-1})| = \sum_{j=1}^{n} \alpha_j[F(u_{t_j}) - F(u_{t_{j-1}})] = F(u)$$

with $u = \sum_{j=1}^{n} \alpha_j[u_{t_j} - u_{t_{j-1}}]$. Since $\|u\|_\infty \le 1$, it follows that

$$\sum_{j=1}^{n} |\phi(t_j) - \phi(t_{j-1})| \le \|F\| = \|f\|.$$

This is true for all partitions of $[a,b]$. Therefore, $\phi \in BV[a,b]$ and $V(\phi) \le \|f\|$. Clearly, $\phi(a) = 0$.

Our next attempt is to show that $f(x) = \int_a^b x\,d\phi$ for every $x \in C[a,b]$ so that $f = f_\phi$ and $\|f\| = V(\phi)$. For this, let $x \in C[a,b]$. By the definition of the Riemann-Stieltjes integral, we have

$$\int_a^b x\,d\phi = \lim_{n\to\infty} \sum_{j=1}^{n} x(\tau_j)[\phi(\tau_j) - \phi(\tau_{j-1})],$$

where $a = \tau_0 < \tau_1 < \ldots < \tau_n = b$, and $\max\limits_{1\le j\le n} |\tau_j - \tau_{j-1}| \to 0$ as $n \to \infty$. Note that $x_n = \sum_{j=1}^{n} x(\tau_j)[u_{\tau_j} - u_{\tau_{j-1}}] \in \ell^\infty([a,b])$, and

$$\sum_{j=1}^{n} x(\tau_j)[\phi(\tau_j) - \phi(\tau_{j-1})] = F\left(\sum_{j=1}^{n} x(\tau_j)[u_{\tau_j} - u_{\tau_{j-1}}]\right) = F(x_n).$$

By uniform continuity of x,

$$\|x - x_n\|_\infty \to 0 \quad \text{as} \quad n \to \infty.$$

Therefore, the continuity of F implies

$$F(x_n) \to F(x) = f(x) \quad \text{as} \quad n \to \infty.$$

Thus,

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

Now, let

$$\tilde{\phi}(t) = \begin{cases} 0 & \text{if } t = a \\ \phi(t+) - \phi(a) & \text{if } a < t < b \\ \phi(b) - \phi(a) & \text{if } t = b. \end{cases}$$

Then it can be seen (cf. Limaye [20], Lemma 14.4) that $\tilde{\phi} \in NBV[a, b]$, and it is the unique element in $NBV[a, b]$ such that

$$\int_a^b x \, d\phi = \int_a^b x \, d\tilde{\phi} \quad \forall x \in C[a, b].$$

Moreover, $V(\tilde{\phi}) \leq V(\phi)$. Thus, $\|f\| = \|f_{\tilde{\phi}}\| \leq V(\tilde{\phi}) \leq V(\phi) = \|f\|$ completing the proof. ∎

Dual of $L^p[a, b]$, $1 \leq p < \infty$

Now we show that, for $1 \leq p < \infty$, the dual of $L^p[a, b]$ is linearly isometric with $L^q[a, b]$. In fact, we obtain a representation of each element of $(L^p[a, b])'$ by an element of $L^q[a, b]$.

Let us first recall, by Hölder's inequality, that if $x \in L^p[a, b]$, $y \in L^q[a, b]$, then

$$\int_a^b |xy| \, d\mu \leq \|x\|_p \|y\|_q.$$

Therefore, the integral $\int_a^b xy \, d\mu$ is well defined for every $x \in L^p[a, b]$, $y \in L^q[a, b]$.

Theorem 8.9 *Let $1 \leq p \leq \infty$. For $y \in L^q[a, b]$, let $f_y : L^p[a, b] \to \mathbb{K}$ be defined by*

$$f_y(x) := \int_a^b xy \, d\mu, \qquad x \in L^p[a, b].$$

Then $f_y \in (L^p[a, b])'$, and the map $y \mapsto f_y$ is a linear isometry from $L^q[a, b]$ into $(L^p[a, b])'$.

Proof. For $y \in L^q[a, b]$, let

$$f_y(x) = \int_a^b xy \, d\mu, \qquad x \in L^p[a, b].$$

Clearly, $f_y : L^p[a,b] \to \mathbb{K}$ is linear. Besides, by Hölder's inequality,

$$|f_y(x)| \leq \|x\|_p \|y\|_q \quad \forall x \in L^p[a,b]$$

so that $f_y \in (L^p[a,b])'$ and $\|f_y\| \leq \|y\|_q$. It also follows that the map $y \mapsto f_y$ is a linear operator from $L^q[a,b]$ into $(L^p[a,b])'$.

We claim that $\|y\|_q \leq \|f_y\|$. If $y = 0$, then this relation is obvious. Therefore, assume that $y \neq 0$, i.e., $\mu(\{t \in [a,b] : y(t) = 0\}) = 0$.

First consider the case $p = 1$. In this case, we have to show that $\|y\|_\infty \leq \|f_y\|$, i.e., to show that the set

$$E = \{t : |y(t)| > \|f_y\|\}$$

is of measure zero. Note that $E = \cup_{n=1}^{\infty} E_n$, where

$$E_n = \Big\{t : |y(t)| > \|f_y\| + \frac{1}{n}\Big\}.$$

Therefore, it is enough to prove that $\mu(E_n) = 0$ for all $n \in \mathbb{N}$. For this, let

$$x_n = \operatorname{sgn}(y)\chi_{E_n}, \qquad n \in \mathbb{N}.$$

Then,

$$f_y(x_n) = \int_a^b x_n y \, d\mu = \int_{E_n} |y| \, d\mu \geq \Big(\|f_y\| + \frac{1}{n}\Big)\mu(E_n).$$

Since $\|x_n\|_1 = \mu(E_n)$ for all $n \in \mathbb{N}$, we have

$$|f_y(x_n)| \leq \|f_y\| \, \|x_n\|_1 = \|f_y\|\mu(E_n).$$

Hence, for every $n \in \mathbb{N}$,

$$\|f_y\|\mu(E_n) \geq |f_y(x_n)| \geq \Big(\|f_y\| + \frac{1}{n}\Big)\mu(E_n),$$

showing that $\mu(E_n) = 0$ for all $n \in \mathbb{N}$.

Next, let $1 < p \leq \infty$, i.e., $1 \leq q < \infty$. In this case, taking $x = \operatorname{sgn}(y)|y|^{q-1}$, we have

$$\|y\|_q^q = \int_a^b |y|^q \, d\mu = \int_a^b xy \, d\mu.$$

Clearly, if $p = \infty$, then $q = 1$ and $x = \operatorname{sgn}(y)$, so that $x \in L^\infty[a,b]$, $\|x\|_\infty \leq 1$, and

$$\|y\|_q^q = \int_a^b xy \, d\mu = f_y(x) \leq \|f_y\| \, \|x\|_\infty \leq \|f_y\|.$$

If $1 < p < \infty$, then we have

$$\int_a^b |x|^p \, d\mu = \int_a^b |y|^q \, d\mu = \|y\|_q^q$$

so that $x \in L^p[a,b]$ and

$$\|y\|_q^q = \int_a^b xy \, d\mu = f_y(x) \leq \|f_y\| \, \|x\|_p = \|f_y\| \|y\|_q^{q/p}.$$

From this it follows that $\|y\|_q \leq \|f_y\|$.

Thus, we have proved that the map $y \mapsto f_y$ is a linear isometry from $L^q[a,b]$ into $(L^p[a,b])'$. ∎

Next, we show that if $p \neq \infty$, then the map $y \mapsto f_y$ in the above theorem is surjective. For this purpose, first we recall the definition of an *absolutely continuous function* and the statement of the *fundamental theorem of Lebesgue integration* (Royden [26]).

A function $v : [a,b] \to \mathbb{K}$ is said to be **absolutely continuous** if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every disjoint family of open intervals $\{(s_j, t_j) : j = 1, \ldots, n\}$ in $[a,b]$,

$$\sum_{j=1}^{n} (t_j - s_j) < \delta \implies \sum_{j=1}^{n} |v(t_j) - v(s_j)| < \varepsilon.$$

Theorem 8.10 (Fundamental theorem of Lebesgue integration) *A function $v : [a,b] \to \mathbb{K}$ is absolutely continuous if and only if there exists $x \in L^1[a,b]$ such that*

$$\int_a^t x \, d\mu = v(t) - v(a) \quad \forall t \in [a,b]$$

and, in that case, v is differentiable almost everywhere in $[a,b]$ and $v' = x$ almost everywhere.

We shall make use of the following result for which we give the proof as well.

Lemma 8.11 *Let $u_a = 0$, and for $a < t \leq b$, let $u_t = \chi_{(a,t]}$. Then, for every f in the dual of $L^p[a,b]$, $1 \leq p < \infty$, the function $v : [a,b] \to \mathbb{K}$ defined by*

$$v(t) = f(u_t), \qquad t \in [a,b],$$

is absolutely continuous.

Proof. Let $\{(t_j, s_j) : j = 1, \ldots, n\}$ be a family of disjoint open intervals in $[a, b]$ and let $f \in (L^p[a,b])'$, $1 \le p < \infty$. Then

$$\sum_{j=1}^{n} |v(s_j) - v(t_j)| = \sum_{j=1}^{n} \alpha_j [v(s_j) - v(t_j)] = \sum_{j=1}^{n} \alpha_j f(u_{s_j} - u_{t_j}),$$

where $\alpha_j = \operatorname{sgn}[v(s_j) - v(t_j)]$ for $j = 1, \ldots, n$. Thus, by taking $x = \sum_{j=1}^{n} \alpha_j (u_{s_j} - u_{t_j})$, we get

$$\sum_{j=1}^{n} |v(s_j) - v(t_j)| = f(x) \le \|f\| \, \|x\|_p.$$

Note that $|x(t)| = |u_{s_j}(t) - u_{t_j}(t)|$ whenever $t_j < t < s_j$, $j = 1, \ldots, n$. Hence, $\|x\|_p^p \le \sum_{j=1}^{n} (s_j - t_j)$ so that

$$\sum_{j=1}^{n} |v(s_j) - v(t_j)| \le \|f\| \left[\sum_{j=1}^{n} (s_j - t_j) \right]^{1/p},$$

and the result follows. ∎

Theorem 8.12 *If $1 \le p < \infty$, then the map $y \mapsto f_y$ from $L^q[a, b]$ into the dual of $L^p[a, b]$, defined in Theorem 8.9, is surjective.*

Proof. Let f be in the dual of $L^p[a, b]$, and let u_t and v be as in the above lemma, i.e., $u_a = 0$ and for $a < t \le b$, u_t is the characteristic function of the interval $(a, t]$, and $v(t) = f(u_t)$, $t \in [a, b]$. Then by the above lemma, v is absolutely continuous so that by Theorem 8.10, v is differentiable almost everywhere, v' is integrable and

$$v(t) = \int_a^t v'(s) \, d\mu(t).$$

Thus, taking $y = v'$, we have

$$f(u_t) = \int_a^t y \, d\mu = \int_a^b u_t y \, d\mu.$$

We show that

$$f(x) = \int_a^b xy \, d\mu \tag{8.3}$$

for every $x \in L^p[a,b]$. For this, first we observe that if x is the characteristic function of $(s,t]$, i.e., $x = u_t - u_s$ for some $s, t \in [a,b]$ with $s < t$, then

$$f(x) = f(u_t) - f(u_s) = \int_s^t y \, d\mu = \int_a^b xy \, d\mu.$$

This shows that (8.3) holds if x is of the form $\chi_{(s,t]}$. Since every open subset of $[a,b]$ is a countable disjoint union of sets of the form $(s,t]$, it follows, by usual measure theoretic arguments, that (8.3) holds if x is of the form χ_G for any open set G. Next, by using the regularity of the Lebesgue measure, we can conclude that (8.3) holds if x is of the form χ_E, where E is a measurable subset of $[a,b]$. From this, it also follows that (8.3) holds for any simple function x.

If $x \in L^p[a,b]$ and if $x \geq 0$, then we know from Proposition 2.8 that there exists an increasing sequence (x_n) of non-negative simple functions that converge pointwise to x. Hence, by monotone convergence theorem (Theorem 2.10), $\|x - x_n\|_p \to 0$ as $n \to \infty$, so that it also follows by the continuity of f that $f(x_n) \to f(x)$ as $n \to \infty$. Now writing a real-valued function $x \in L^p[a,b]$ as

$$x = x^+ - x^-$$

and a complex-valued function $x \in L^p[a,b]$ as

$$x = \mathrm{Re}\,(x) + i \,\mathrm{Im}\,(x),$$

it follows that (8.3) holds for every $x \in L^p[a,b]$. It remains to show that $y \in L^q[a,b]$.

First let $p = 1$, and let $t \in [a,b]$ be such that $v'(t)$ exists. Then we have

$$y(t) = v'(t) = \lim_{\delta \to 0} \frac{v(t+\delta) - v(t)}{\delta} = \lim_{\delta \to 0} \frac{f(u_{t+\delta} - u_t)}{\delta}.$$

But

$$|f(u_{t+\delta} - u_t)| \leq \|f\| \, \|u_{t+\delta} - u_t\|_1 \leq \|f\| \delta.$$

Since $v'(t)$ exists a.e. on $[a,b]$, it follows that $|y(t)| \leq \|f\|$ a.e. so that $\|y\|_\infty \leq \|f\|$.

Next, let $1 < p < \infty$. Then, with

$$D_n = \{t \in [a,b] : |y(t)| \leq n\},$$

we note that

$$\int_a^b |y|^q \, d\mu = \lim_{n\to\infty} \int_{D_n} |y|^q \, d\mu$$

since $D_n \subseteq D_{n+1}$ and $[a,b] = \cup_{n=1}^{\infty} D_n$. Now,

$$\int_{D_n} |y|^q \, d\mu = \int_a^b |y|^q \chi_{D_n} \, d\mu.$$

Taking

$$x = \operatorname{sgn}(y)|y|^{q-1}, \qquad x_n = x\chi_{D_n},$$

we have

$$\int_{D_n} |y|^q \, d\mu = \int_a^b |y|^q \chi_{D_n} \, d\mu = \int_a^b xy\chi_{D_n} \, d\mu = \int_a^b x_n y \, d\mu.$$

Since

$$\int_a^b |x_n|^p \, d\mu = \int_{D_n} |y|^q \, d\mu \leq n^q(b-a) < \infty,$$

it follows that

$$x_n \in L^p[a,b], \qquad \int_{D_n} |y|^q \, d\mu = \int_a^b x_n y \, d\mu = f(x_n) \quad \forall n \in \mathbb{N}.$$

From this, we have

$$\int_{D_n} |y|^q \, d\mu = |f(x_n)| \leq \|f\| \, \|x_n\|_p = \|f\| \left(\int_{D_n} |y|^q \, d\mu \right)^{1/p}$$

so that

$$\left(\int_{D_n} |y|^q \, d\mu \right)^{1/q} \leq \|f\| \quad \forall n \in \mathbb{N}.$$

Therefore,

$$\int_a^b |y|^q \, d\mu = \lim_{n\to\infty} \int_{D_n} |y|^q \, d\mu \leq \|f\|^q.$$

Consequently, $y \in L^q[a,b]$ and $\|y\|_q \leq \|f\|$. ∎

8.1.4 Separability Revisited

In the proofs of Theorems 8.2 and 8.12, for showing the surjectivity of the isometries, we used the fact that $p \neq \infty$. A question that naturally arises is whether it is possible to prove that $(\ell^\infty)'$ (respectively, $(L^\infty[a,b])'$) is linearly isometric or at least linearly homeomorphic with ℓ^1 (respectively, $L^1[a,b]$). We show that the answer is negative. To see this, we shall make use of the concept of separability.

Recall from Section 2.3 that a normed linear space is *separable* if it has a countable dense subset. In fact, since a (nonzero) normed linear space cannot have a finite dense subset, it follows that a normed linear space is separable if and only if it has a denumerable dense subset.

We may also recall from Example 2.10 that the spaces

(1) c_{00} and $\mathcal{P}[a,b]$ with respect to any norm,

(2) ℓ^p, $1 \leq p < \infty$,

(3) c_0 with $\|\cdot\|_\infty$,

(4) $C[a,b]$ with $\|\cdot\|_p$, $1 \leq p \leq \infty$, and

(5) $L^p[a,b]$, $1 \leq p < \infty$

are separable. Also, from Example 2.11, we know that the spaces $\ell^\infty(\mathbb{N})$, $\ell^\infty([a,b])$, $L^\infty[a,b]$ are not separable.

Theorem 8.13 *A normed linear space is separable if its dual is separable.*

Proof. Let X be a normed linear space such that its dual X' is separable. We show that X is separable.

The result is obvious if $X = \{0\}$. Suppose $X \neq \{0\}$. Since X' is separable, the set $S = \{f \in X' : \|f\| = 1\}$ is separable (by Proposition 2.34(i)). Let $\{f_1, f_2, \ldots\}$ be a dense subset of S. Since $\|f_n\| = 1$ for every $n \in \mathbb{N}$, there exists $x_n \in X$ such that

$$\|x_n\| = 1, \qquad |f_n(x_n)| > 1/2 \quad \forall n \in \mathbb{N}.$$

We show that the separable space $Y = \text{span}\{x_1, x_2, \ldots\}$ is dense in X. Suppose on the contrary, Y is not dense in X. Then, by Corollary 5.5 of Hahn-Banach theorem, there exists $f \in X'$ such that $\|f\| = 1$ and $f(y) = 0$ for all $y \in Y$. In particular, $f(x_n) = 0$ for all $n \in \mathbb{N}$, so that

$$\|f - f_n\| \geq |(f - f_n)(x_n)| = |f_n(x_n)| > 1/2$$

for every $n \in \mathbb{N}$. Since $f \in S$, the above relation leads to a contradiction to the fact that $\{f_1, f_2, \ldots\}$ is dense in S. ∎

The converse of the above theorem does not hold. For example, we have seen that the space ℓ^1 is separable and its dual is linearly isometric with ℓ^∞ which is not separable.

Since the continuous image of a separable space is separable (Proposition 2.34), the following result is an immediate consequence of Theorem 8.13.

Corollary 8.14 *The dual of a nonseparable normed linear space cannot be homeomorphic with a separable space.*

Recall that ℓ^∞, $\ell^\infty([a,b])$ and $L^\infty[a,b]$ are nonseparable spaces. Hence, as consequences of the above corollary, the duals of these spaces cannot be homeomorphic with any separable spaces. In particular, the dual of ℓ^∞ is not linearly isometric with ℓ^1, and the dual of $L^\infty[a,b]$ is not linearly isometric with $L^1[a,b]$.

8.2 Reflexivity and Weak Convergence

8.2.1 Reflexivity

Recall from Section 5.2 that a normed linear space X is said to be *reflexive* if the *canonical isometry* $J : X \to X''$ defined by

$$(Jx)(f) = f(x) \quad \forall x \in X, f \in X',$$

is surjective.

Since the dual of a normed linear space is a Banach space, it is obvious that every reflexive space is a Banach space. Thus, a normed linear space having a denumerable (Hamel) basis is not a reflexive space. As examples, the spaces c_{00} and $\mathcal{P}[a,b]$ are not reflexive spaces with respect to any norm on them.

In this section we prove some results on reflexivity by which we can infer whether some of the sequence spaces and function spaces are reflexive or not.

First, let us give a few examples of reflexive spaces.

EXAMPLE 8.1 (i) The spaces ℓ^p and $L^p[a,b]$ for $1 < p < \infty$ are reflexive:

First consider the case of $L^p[a,b]$. Let $\varphi \in (L^p[a,b])''$. We have to show that there exists $x \in L^p[a,b]$ such that $\varphi(f) = f(x)$ for all $f \in (L^p[a,b])'$. For this, let $f \in (L^p[a,b])'$. Recall from Theorem 8.12 that the map $F : L^q[a,b] \to (L^p[a,b])'$ defined by

$$F(y)(x) = \int_a^b xy \, d\mu, \qquad x \in L^p[a,b],\ y \in L^q[a,b],$$

is a surjective linear isometry. Hence, there is a $y \in L^q[a,b]$ such that $F(y) = f$. Thus,

$$\varphi(f) = \varphi(F(y)) = (\varphi \circ F)(y) \quad \forall y \in L^q[a,b].$$

Since $\varphi \circ F \in (L^q[a,b])'$, there exists $x \in L^p[a,b]$ such that $G(x) = \varphi \circ F$, where $G : L^p[a,b] \to (L^q[a,b])'$ is the surjective linear isometry defined by

$$G(x)(y) = \int_a^b yx \, d\mu, \qquad x \in L^p[a,b],\ y \in L^q[a,b].$$

Hence,

$$\varphi(f) = (\varphi \circ F)(y) = G(x)(y) = F(y)(x) = f(x) \quad \forall f \in (L^p[a,b])'.$$

The proof for the space ℓ^p with $1 < p < \infty$ follows in a similar manner by replacing $\int_a^b xy \, d\mu$ by $\sum_{j=1}^{\infty} x(j)y(j)$.

(ii) Every Hilbert space is reflexive: Let X be a Hilbert space. We have already observed in Section 3.3 (Remark 3.3) that X' is a Hilbert space with respect to the inner product

$$(f,g) \mapsto \langle f,g\rangle' := \langle v_g, v_f\rangle, \qquad (f,g) \in X' \times X',$$

where $v_f, v_g \in X$ are the element in X associated with $f, g \in X'$ according to the Riesz representation theorem.

Now, let $\varphi \in X''$. We show that there exists $x \in X$ such that $\varphi(f) = f(x)$ for all $f \in X'$. Note that Riesz representation theorem applied to the Hilbert space X' gives a unique element $g \in X'$ such that

$$\varphi(f) = \langle f,g\rangle' = \langle v_g, v_f\rangle \quad \forall f \in X'.$$

But $\langle v_g, v_f\rangle = f(v_g)$ for all $f \in X'$. Therefore, $\varphi(f) = f(v_g)$ for all $f \in X'$.

Exercise 8.1 Let X be a normed linear space. For $S \subseteq X'$, let

$$\widetilde{S} = \{x \in X : f(x) = 0 \quad \forall f \in S\}.$$

Show the following:

(i) If $\operatorname{span} S$ is dense in X', then $\widetilde{S} = \{0\}$.

(ii) If X is reflexive and $\widetilde{S} = \{0\}$, then $\operatorname{span} S$ is dense in X'. □

We have seen in Theorem 8.13 that a normed linear space is separable if its dual is separable, and its converse need not be true. Now we show that, if the space is reflexive, then the converse of Theorem 8.13 does hold.

Theorem 8.15 *Suppose X is reflexive. Then X is separable if and only if its dual X' is separable.*

Proof. Suppose X is a reflexive space. We have already mentioned above that X' separable implies X separable. Now assume that X is separable. Then, since it is reflexive, its second dual X'' which is linearly isometric with the separable space X is also separable. Hence, again by Theorem 8.13, the space X', the pre-dual of X'', is separable. ∎

This result helps us to show that certain Banach spaces are not reflexive.

EXAMPLE 8.2 The spaces ℓ^1, $L^1[a,b]$ and $C[a,b]$ with $\|\cdot\|_\infty$ are not reflexive, since they are separable, and their duals are linearly isometric with nonseparable spaces ℓ^∞, $L^\infty[a,b]$ and $NBV[a,b]$, respectively.

Now we consider reflexivity in more detail. In this regard we shall make use of the following result. First recall that if X and Y are normed linear spaces and $T \in \mathcal{B}(X,Y)$, then the *transpose* $T' : Y' \to X'$ of T is defined by

$$(T'f)(x) = f(Tx) \quad \forall f \in Y',\ x \in X,$$

and we have observed in Chapter 5 (Theorem 5.10) that $\|T'\| = \|T\|$.

Lemma 8.16 *Let X and Y be normed linear spaces. If $T : X \to Y$ is a surjective linear isometry, then $T' : Y' \to X'$ is a surjective linear isometry.*

Proof. Let $T : X \to Y$ be a surjective linear isometry. We have to show that $T' : Y' \to X'$ is surjective and $\|T'f\| = \|f\|$ for every $f \in Y'$. Since T is a bijective linear isometry, it is obvious that $T^{-1} : Y \to X$ is also a bijective linear isometry. In particular, $T^{-1} \in \mathcal{B}(Y, X)$. Hence, taking the transpose on both the sides of

$$TT^{-1} = I = T^{-1}T,$$

it follows that T' is bijective. Now let $f \in Y'$. Then using the fact that T is surjective and is an isometry, we have

$$\{y \in Y : \|y\| = 1\} = \{Tx : x \in X, \|x\| = 1\}.$$

Hence, for every $f \in Y'$,

$$\begin{aligned} \|T'f\| &= \sup\{|(T'f)(x)| : \|x\| = 1\} \\ &= \sup\{|f(Tx)| : \|x\| = 1\} \\ &= \sup\{|f(y)| : \|y\| = 1\} \\ &= \|f\|. \end{aligned}$$

This completes the proof. ∎

Theorem 8.17 (i) *Every closed subspace of a reflexive space is reflexive.*

(ii) *Every normed linear space which is linearly isometric with a reflexive space is reflexive.*

(iii) *A Banach space is reflexive if and only if its dual is reflexive.*

Proof. Let X be a normed linear space and $J : X \to X''$ be the canonical linear isometry.

(i) Suppose X is reflexive and X_0 is a closed subspace of X. We have to show that the canonical linear isometry $J_0 : X_0 \to X_0''$ is surjective. For this, let $F \in X_0''$, and define $\varphi : X' \to \mathbb{K}$ by

$$\varphi(f) = F(f|_{X_0}), \qquad f \in X'.$$

Then it follows that $\varphi \in X''$. Since X is reflexive, there exists $x \in X$ such that $Jx = \varphi$. We show that $x \in X_0$ and $J_0 x = F$.

If $x \notin X_0$, then, by Corollary 5.5 of the Hahn-Banach theorem, there exists $g \in X'$ such that

$$g(x) = \operatorname{dist}(x, X_0), \qquad g(u) = 0 \quad \forall u \in X_0$$

so that we get

$$0 \neq g(x) = (Jx)(g) = \varphi(g) = F(g|_{X_0}) = 0,$$

which is a contradiction. Thus, we have $x \in X_0$. Now for $f \in X_0'$, let $\tilde{f} \in X'$ be a Hahn-Banach extension of f (Theorem 3.18). Then we have

$$(J_0x)(f) = f(x) = \tilde{f}(x) = (Jx)(\tilde{f}) = \varphi(\tilde{f}) = F(\tilde{f}|_{X_0}) = F(f).$$

Thus, $J_0x = F$.

(ii) Suppose that X is a reflexive space, and Y is a normed linear space such that there is a surjective linear isometry $T : X \to Y$. Then, by Lemma 8.16, it is seen that $T'' : X'' \to Y''$, the transpose of T', is a surjective linear isometry. We show that the surjective linear isometry

$$T'' \circ J \circ T^{-1} : Y \to Y''$$

is, in fact, the canonical linear isometry. That is, to show that $\left(T'' \circ J \circ T^{-1}\right)(y)(f) = f(y)$ for all $y \in Y$, $f \in Y'$. So, let $y \in Y$ and $f \in Y'$. Then we have

$$\begin{aligned}
\left(T'' \circ J \circ T^{-1}\right)(y)(f) &= T''\left(JT^{-1}y\right)(f) \\
&= \left(JT^{-1}y\right)(T'f) \\
&= (T'f)((T^{-1}y) \\
&= f\left(TT^{-1}y\right) = f(y).
\end{aligned}$$

(iii) Let X be a reflexive space. Since J is a surjective linear isometry, by Lemma 8.16, the transpose $J' : X''' \to X'$ of J is also a surjective linear isometry, and hence, $\tilde{J} := (J')^{-1} : X' \to X'''$ is a surjective linear isometry as well. We show that $\tilde{J}$ is, in fact, the canonical linear isometry from X' onto X'''. For this, let $f \in X'$ and $\varphi \in X''$. Since X is reflexive, there exists $x \in X$ such that $\varphi(f) = (Jx)(f)$. Hence,

$$(\tilde{J}f)(\varphi) = (\tilde{J}f)(Jx) = J'[\tilde{J}f](x) = (J'\tilde{J}f)(x) = f(x) = \varphi(f)$$

showing that the surjective map $\tilde{J} : X' \to X'''$ is the canonical linear isometry. Thus, X' is reflexive.

Conversely, suppose that X is a Banach space such that X' is reflexive. Then by the first part, X'' is reflexive. Since X is a Banach space, $J(X)$, the image of X under the canonical linear isometry $J : X \to X''$, is a closed subspace of the reflexive space X''. Therefore, by (i), $J(X)$ is reflexive. Now, (ii) implies that $X = J^{-1}(J(X))$ is reflexive. ∎

Remark 8.1 Since every reflexive space has to be a Banach space, completeness assumption in Theorem 8.17 (iii) cannot be dropped. To illustrate this, we may recall that the dual of $X = c_{00}$, with $\|\cdot\|_p$ for $1 < p < \infty$, is linearly isometric with the reflexive space ℓ^q, whereas $(c_{00}, \|\cdot\|_p)$ is not reflexive.

Using Theorem 8.17, we give some more examples of non-reflexive Banach spaces.

EXAMPLE 8.3 (i) The space ℓ^∞ is not reflexive since its pre-dual ℓ^1 is not reflexive (cf. Theorem 8.17(iii)). This also follows from the fact that its closed subspace c_0 is not reflexive (cf. Theorem 8.17(i)), as its dual is linearly isometric with the non-reflexive space ℓ^1.

(ii) The space $\ell^\infty([a,b])$ is not reflexive, since its closed subspace $C[a,b]$ is not reflexive (cf. Theorem 8.17(i)).

(iii) The space $L^\infty[a,b]$ is not reflexive since its pre-dual $L^1[a,b]$ is not reflexive (cf. Theorem 8.17(iii)). Also, note that its closed subspace $C[a,b]$ is not reflexive (cf. Theorem 8.15).

(iv) The space $NBV[a,b]$ is not reflexive since its pre-dual $C[a,b]$ is not reflexive (cf. Theorem 8.17(iii)).

8.2.2 Weak Convergence

Suppose X is a finite dimensional normed linear space, say of dimension k, and $\{u_1, \ldots, u_k\}$ is a basis of X. Let $\{f_1, \ldots, f_k\}$ be the basis of X' consisting of coordinate functionals associated with the above basis of X. Then we know that

$$x = \sum_{j=1}^{k} f_j(x) u_j \quad \forall x \in X.$$

Therefore, a sequence (x_n) in X converges to $x \in X$ if and only if $(f_j(x_n))$ converges to $f_j(x)$ for every $j = 1, \ldots, k$. Also, since every $f \in X'$ is a finite linear combination of $f_1, \ldots, f_k$, we have

$$\|x_n - x\| \to 0 \iff f(x_n) \to f(x) \quad \forall f \in X'.$$

One may ask whether the above result is true for every normed linear space.

Clearly, if X is a normed linear space and if (x_n) in X converges to $x \in X$, then for every $f \in X'$,

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \, \|x_n - x\| \to 0 \quad \text{as} \quad n \to \infty.$$

But the converse need not hold. To see this, let X be an infinite dimensional Hilbert space and (x_n) be a sequence of orthonormal elements in X. Clearly, $\|x_n\| = 1$ for every $n \in \mathbb{N}$. But for every $f \in X'$, by the Riesz representation theorem, there is a unique $y \in X$ such that $f(x) = \langle x, y \rangle$ for every $x \in X$, so that using Bessel's inequality, it follows that

$$f(x_n) = \langle x_n, y \rangle \to 0 \quad \text{as} \quad n \to \infty.$$

As another instance, consider the sequence (e_n) in ℓ^p with $1 < p < \infty$. We know that for every $f \in (\ell^p)'$, there exists a unique $y \in \ell^q$ such that

$$f(x) = \sum_{j=1}^{\infty} x(j)y(j) \quad \forall x \in \ell^p.$$

Hence, $f(e_n) = y(n) \to 0$ for all $f \in (\ell^p)'$, but $\|e_n\|_p \not\to 0$.

Thus, we see that the requirement $f(x_n) \to f(x)$ for every $f \in X'$ on the sequence (x_n) is weaker than the requirement $\|x_n - x\| \to 0$. This observation motivates the following definition.

A sequence (x_n) in a normed linear space X is said to **converge weakly** to an element $x \in X$ if $f(x_n) \to f(x)$ as $n \to \infty$ for every $f \in X'$, and we write this fact as

$$x_n \to x \quad \text{weakly}.$$

If a sequence (x_n) in X converges weakly to an element in X, then we may say that (x_n) is a **weakly convergent sequence**.

Because of the above terminology, the usual convergence of a sequence (x_n) in a normed linear space, namely $\|x_n - x\| \to 0$, is sometimes called the **strong convergence**.

It is to be observed, as a consequence of the Riesz representation theorem, that if X is a Hilbert space, then a sequence (x_n) in X converges weakly to $x \in X$ if and only if

$$\langle x_n, y \rangle \to \langle x, y \rangle \quad \forall y \in X.$$

Exercise 8.2 (i) Let (x_n) in ℓ^p with $1 < p < \infty$ be defined by

$$x_n(j) = \begin{cases} 1/n^{1/p} & \text{if } 1 \leq j \leq n, \\ 0 & \text{if } j > n. \end{cases}$$

Show that (x_n) converges weakly to 0 but does not converge to 0.

(ii) Let $X = c_{00}$ with $\|\cdot\|_2$. Show that a sequence (x_n) in X converges weakly to $x \in X$ if and only if $\langle x_n, y\rangle \to \langle x, y\rangle$ for every $y \in X$. (Note that the space X is not a Hilbert space, and hence, we cannot use the Riesz representation theorem.) $\square$

Remark 8.2 The nomenclature "weak convergence" is also motivated by the following topological consideration.

Recall that a family $\mathcal{T}$ of subsets of a set Ω is called a **topology** on Ω if

(a) $\varnothing \in \mathcal{T}, \quad \Omega \in \mathcal{T}$,

(b) $\mathcal{T}_0 \subseteq \mathcal{T} \implies \cup_{S\in\mathcal{T}_0} S \in \mathcal{T}$,

(c) $S_1, S_2 \in \mathcal{T} \implies S_1 \cap S_2 \in \mathcal{T}$.

A set Ω, together with a topology $\mathcal{T}$, is called a **topological space.**

If $\mathcal{T}_1$ and $\mathcal{T}_2$ are topologies on a set Ω such that $\mathcal{T}_1 \subseteq \mathcal{T}_2$, then $\mathcal{T}_1$ is said to be **weaker than** $\mathcal{T}_2$; and $\mathcal{T}_2$ is said to be **stronger than** $\mathcal{T}_1$.

Given a family $\mathcal{S}$ of subsets of Ω, the intersection of all topologies on Ω each of which contains $\mathcal{S}$ is a topology on Ω, called the **topology generated by** $\mathcal{S}$, and we denote it by $\mathcal{T}_{\mathcal{S}}$.

Let Ω be a topological space with topology $\mathcal{T}$. Then a sequence (x_n) in Ω is said to **converge** to $x \in \Omega$ with respect to the topology $\mathcal{T}$, if for every $G \in \mathcal{T}$ with $x \in G$, there exists $N \in \mathbb{N}$ such that $x_n \in G$ for all $n \geq N$.

Let us list a few easily verifiable facts:

(1) If Ω is a metric space and $\mathcal{T}$ is the family of all open sets with respect to the metric, then $\mathcal{T}$ is a topology, and the convergence w.r.t. $\mathcal{T}$ is the convergence with respect to the metric on Ω.

(2) Suppose Ω is a set and $f : \Omega \to \mathbb{K}$ is a function. Consider the topology $\mathcal{T}_S$ generated by

$$\mathcal{S} = \{f^{-1}(V) : V \text{ open in } \mathbb{K}\}.$$

Then $\mathcal{T}_S$ is the weakest topology such that f is continuous, and the convergence of a sequence (x_n) to $x \in \Omega$ with respect to the topology $\mathcal{T}_S$ is the same as the convergence of $(f(x_n))$ to $f(x)$.

(3) Suppose $\mathcal{F}$ is a family of functions from a set Ω to $\mathbb{K}$. Consider the topology $\mathcal{T}_S$ generated by

$$\mathcal{S} = \{f^{-1}(V) : V \text{ open in } \mathbb{K} \text{ and } f \in \mathcal{F}\}.$$

Then $\mathcal{T}_S$ is the weakest topology such that every $f \in \mathcal{F}$ is continuous. This topology is called the **weak topology induced by** $\mathcal{F}$. The convergence of a sequence (x_n) to $x \in \Omega$ with respect to $\mathcal{T}_S$ is the same as the convergence of $(f(x_n))$ to $f(x)$ for every $f \in \mathcal{F}$.

By result (3) above, it is obvious that, a sequence (x_n) in a normed linear space X converges weakly to $x \in X$ if and only if (x_n) converges to x with respect to the weak topology on X induced by X'.

Schur's lemma

We have already seen that convergence of a sequence implies weak convergence, but the converse need not be true if the space is infinite dimensional. One may ask whether weak convergence and convergence are different in every infinite dimensional space. The answer is again in the negative. The following theorem illustrates this point.

Theorem 8.18 (Schur's lemma) *Every weakly convergent sequence in ℓ^1 is convergent.*

Proof. Suppose there exists a sequence (x_n) in ℓ^1 which converges weakly to $x \in \ell^1$, but $\|x_n - x\|_1 \not\to 0$. We shall arrive at a contradiction. Without loss of generality, we can assume that $x = 0$. Thus, $f(x_n) \to 0$ as $n \to \infty$ for all $f \in (\ell^1)'$, but there exists $\delta > 0$ such that $\|x_n\|_1 \geq \delta$ for infinitely many n. By passing to a subsequence, if necessary, we can assume that $\|x_n\|_1 \geq \delta$ for every n. Our idea is to find a subsequence (x_{n_k}) of (x_n) and an $f \in (\ell^1)'$ such that $f(x_{n_k}) \not\to 0$.

Let $\varepsilon > 0$ be given. By the assumption on (x_n), we know that $x_n(j) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$. Hence, there exists a positive integer $n_1 > 1$ such that $|x_n(1)| < \varepsilon$ for every $n \geq n_1$. In particular, $|x_{n_1}(1)| < \varepsilon$. Since $x_{n_1} \in \ell^1$, there exists an integer $m_1 > 1$ such that

$$\sum_{j=m_1+1}^{\infty} |x_{n_1}(j)| \leq \varepsilon.$$

Again, using the fact that $x_n(j) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$, there exists $n_2 > n_1$ such that $\sum_{j=1}^{m_1} |x_n(j)| < \varepsilon$ for every $n \geq n_2$. In particular,

$$\sum_{j=1}^{m_1} |x_{n_2}(j)| < \varepsilon.$$

Since $x_{n_2} \in \ell^1$, there exists $m_2 > m_1$ such that

$$\sum_{j=m_2+1}^{\infty} |x_{n_2}(j)| \leq \varepsilon.$$

Having obtained $m_1, \ldots, m_{k-1}$ and $n_1, \ldots, n_{k-1}$, let $n_k > n_{k-1}$ and $m_k > m_{k-1}$ be such that

$$\sum_{j=1}^{m_{k-1}} |x_{n_k}(j)| < \varepsilon, \qquad \sum_{j=m_k+1}^{\infty} |x_{n_k}(j)| \leq \varepsilon.$$

Now, let $y \in \ell^\infty$ with $\|y\|_\infty \leq 1$. Then, by Theorem 8.2, $f : \ell^1 \to \mathbb{K}$ defined by

$$f(x) = \sum_{j=1}^{\infty} x(j)y(j), \qquad x \in \ell^1$$

belongs to $(\ell^1)'$. We may choose y which satisfies

$$\sum_{j=m_{k-1}+1}^{m_k} x_{n_k}(j)y(j) = \sum_{j=m_{k-1}+1}^{m_k} |x_{n_k}(j)|, \qquad |y(j)| \leq 1 \quad \forall j \in \mathbb{N}.$$

This is accomplished by taking $y \in \ell^\infty$ such that $y(1) = 1$ and $y(j) = \operatorname{sgn}(x_{n_k}(j))$ whenever $m_{k-1} + 1 \leq j \leq m_k$, $k \in \mathbb{N}$, with $m_0 = 1$. For such y, we have

$$|f(x_{n_k}) - \|x_{n_k}\|_1| \leq 2 \sum_{j=1}^{m_{k-1}} |x_{n_k}(j)| + 2 \sum_{j=m_k+1}^{\infty} |x_{n_k}(j)| \leq 4\varepsilon.$$

From this, we have

$$|f(x_{n_k})| \geq \|x_{n_k}\|_1 - |f(x_{n_k}) - \|x_{n_k}\|_1| \geq \|x_{n_k}\|_1 - 4\varepsilon \geq \delta - 4\varepsilon.$$

Hence, by taking $\varepsilon < \delta/4$, we get $|f(x_{n_k})| \geq \delta - 4\varepsilon > 0$ for all $k \in \mathbb{N}$, which is a contradiction to the assumption that (x_{n_k}) converges weakly to 0. ∎

We now observe a property that a weakly convergent sequence shares with a convergent sequence.

Theorem 8.19 *Let X be a normed linear space and (x_n) converge weakly to $x \in X$. Then (x_n) is a bounded sequence and the weak limit is unique. Moreover,*

$$\|x\| \le \liminf_n \|x_n\|.$$

Proof. By the hypothesis, $f(x_n) \to f(x)$ as $n \to \infty$ for every $f \in X'$. In particular, the set $\{f(x_n) : n \in \mathbb{N}\}$ is bounded for every $f \in X'$. Hence, by a consequence of uniform boundedness principle (cf. Corollary 6.5), the sequence (x_n) is bounded. Also, since $f(x) = \lim_{n\to\infty} f(x_n)$ for every $f \in X'$, we have

$$|f(x)| = \lim_{n\to\infty} |f(x_n)| \le \|f\| \liminf_n \|x_n\| \quad \forall f \in X'$$

so that by a consequence of the Hahn-Banach theorem (cf. Theorem 5.9),

$$\|x\| \le \liminf_n \|x_n\|.$$

To see that the weak limit is unique, suppose that x and u in X are weak limits of the sequence (x_n), i.e.,

$$f(x_n) \to f(x), \qquad f(x_n) \to f(u) \quad \forall f \in X'.$$

Then we have $f(x) = f(u)$ for every $f \in X'$, so that by Corollary 5.6 of Hahn-Banach theorem, it follows that $x = u$. ∎

The definition of weak convergence involves all functionals in the dual space. Can we assert weak convergence of a sequence by knowing only a 'manageable' set of functionals? Of course, if we know a basis F of X', then

$$x_n \to x \text{ weakly} \iff f(x_n) \to f(x) \quad \forall f \in F.$$

But, we know that every basis of an infinite dimensional Banach space is uncountable, and its existence also cannot be asserted constructively. So, it is desirable to have a manageable subset of X' with which we can assert the weak convergence of a sequence. Let us first look at the spaces ℓ^p.

Let $f_1, f_2, \ldots$ be the *coordinate functionals* on ℓ^p, i.e., $f_j(x) = x(j)$, $x \in \ell^p$, $j \in \mathbb{N}$. Clearly, $f_j \in (\ell^p)'$ for every $j \in \mathbb{N}$. Therefore, if (x_n) is a sequence in ℓ^p which converges weakly to $x \in \ell^p$, then we have

$$x_n(j) \to x(j) \quad \text{as} \quad n \to \infty$$

for all $j \in \mathbb{N}$. It would be nice if the converse is also true. We show this for the case $1 < p < \infty$ under the assumption that (x_n) is a bounded sequence, and also show that the result is not true if $p \in \{1, \infty\}$.

First let us show the negative cases:
Case $p = 1$: Consider the sequence (e_n) in ℓ^1. Clearly, $e_n(j) \to 0$ for all $j \in \mathbb{N}$. Let $f : \ell^1 \to \mathbb{K}$ be defined by

$$f(x) = \sum_{j=1}^{\infty} x(j), \qquad x \in \ell^1.$$

Then we have $f \in (\ell^1)'$ and $f(e_n) = 1$ for every $n \in \mathbb{N}$. Thus, (e_n) does not converge weakly to 0.
Case $p = \infty$: Consider the sequence (x_n) in ℓ^∞ defined by

$$x_n(j) = \begin{cases} 1 & \text{if } j \le n, \\ 0 & \text{if } j > n \end{cases}$$

for every $n \in \mathbb{N}$. Taking $x = (1, 1, \ldots)$, we see that $x_n(j) \to x(j) = 1$ as $n \to \infty$ for all $j \in \mathbb{N}$. But (x_n) does not converge weakly to x. To see this, we observe that $x_n \in c_0$ for every $n \in \mathbb{N}$ and $x \in \ell^\infty \setminus c_0$, so that by Corollary 5.5 of the Hahn-Banach theorem, there exists $f \in (\ell^\infty)'$ such that $f(x) \neq 0$ but $f(x_n) = 0$ for every $n \in \mathbb{N}$.

Exercise 8.3 Let X be a normed linear space and $E \subseteq X$. Call E to be **weakly closed** if the weak limit of every weakly convergent sequence from E belongs to E. Show that every weakly closed subspace of a normed linear space is closed. □

Now the positive result.

Theorem 8.20 *Let $1 < p < \infty$, and let (x_n) be a bounded sequence in ℓ^p such that, for every $j \in \mathbb{N}$, $x_n(j) \to x(j)$ for some $x \in \ell^p$. Then (x_n) converges weakly to x.*

Proof. Without loss of generality, assume that $x = 0$. We have to show that $f(x_n) \to 0$ as $n \to \infty$ for all $f \in (\ell^p)'$. For this, let $f \in (\ell^p)'$ and $\varepsilon > 0$, and let $c > 0$ be such that $\|x_n\|_p \leq c$ for all $n \in \mathbb{N}$. By Theorem 8.2, we know that $y := (f(e_1), f(e_2), \ldots) \in \ell^q$ and

$$f(x) = \sum_{j=1}^{\infty} f(e_j)x(j) \quad \forall x \in \ell^p.$$

For every $m \in \mathbb{N}$, $n \in \mathbb{N}$, we have

$$|f(x_n)| \leq \sum_{j=1}^{m} |f(e_j)x_n(j)| + \sum_{j=m+1}^{\infty} |f(e_j)x_n(j)|,$$

so that by Hölder's inequality,

$$|f(x_n)| \leq \sum_{j=1}^{m} |f(e_j)x_n(j)| + \left(\sum_{j=m+1}^{\infty} |f(e_j)|^q \right)^{1/q} \|x_n\|_p. \tag{8.4}$$

Since $(f(e_1), f(e_2), \ldots) \in \ell^q$, it follows that there exists $k \in \mathbb{N}$ such that

$$\left(\sum_{j=k+1}^{\infty} |f(e_j)|^q \right)^{1/q} < \varepsilon.$$

Also, since $x_n(j) \to 0$ as $n \to \infty$ for each $j \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that

$$\left| \sum_{j=1}^{k} f(e_j)x_n(j) \right| < \varepsilon \quad \forall n \geq N.$$

Thus, taking $m = k$ in (8.4), we see that

$$|f(x_n)| < (c+1)\varepsilon \quad \forall n \geq N.$$

This shows the $f(x_n) \to 0$ as $n \to \infty$. ∎

Since the span of the coordinate functionals $f_1, f_2, \ldots$ is dense in $(\ell^p)'$ for $1 < p < \infty$ (Corollary 8.3), the above theorem is a particular case of the following.

Theorem 8.21 *Let $E \subseteq X'$ be such that* span E *is dense in X', (x_n) be a bounded sequence in X, and $x \in X$. If $f(x_n) \to f(x)$ for every $f \in E$, then (x_n) converges weakly to x.*

Proof. Let $D = \operatorname{span} E$. Then, by hypothesis, $f(x_n) \to f(x)$ as $n \to \infty$ for every $f \in D$. For $n \in \mathbb{N}$, let $\varphi_n = J(x_n)$ and $\varphi = J(x)$, where $J : X \to X''$ is the canonical linear isometry. Then it follows that (φ_n) is a uniformly bounded sequence in $X'' = \mathcal{B}(X', \mathbb{K})$ which converges pointwise to φ on the dense subset D of X'. Therefore, by Theorem 3.11,

$$f(x_n) = \varphi_n(f) \to \varphi(f) = f(x) \quad \forall f \in X'. \quad \blacksquare$$

Now, we state results analogous to Theorem 8.20 for the spaces $C[a,b]$ and $L^p[a,b]$. We omit their proofs, as they require some more measure theoretic arguments which we have not developed. For proofs, the reader may refer Limaye [20].

(1) A sequence (x_n) in the space $C[a,b]$ with $\|\cdot\|_\infty$ converges weakly to $x \in C[a,b]$ if and only if (x_n) is bounded and

$$\lim_{n\to\infty} x_n(t) = x(t) \quad \forall t \in [a,b].$$

(2) A sequence (x_n) in $L^1[a,b]$ converges weakly to $x \in L^1[a,b]$ if and only if (x_n) is bounded and

$$\lim_{n\to\infty} \int_E x_n \, d\mu = \int_E x \, d\mu$$

for every measurable subset E of $[a,b]$.

(3) A sequence (x_n) in $L^p[a,b]$, $1 < p < \infty$, converges weakly to $x \in L^p[a,b]$ if and only if (x_n) is bounded and

$$\lim_{n\to\infty} \int_{[c,d]} x_n \, d\mu = \int_{[c,d]} x \, d\mu \quad \forall [c,d] \subseteq [a,b].$$

It is to be borne in mind that for a sequence (x_n) in X, convergence of $(f(x_n))$ for every $f \in X'$ does not imply that (x_n) converges weakly. It can happen that $(f(x_n))$ converges to some scalar which is not $f(x)$ for any $x \in X$. Look at the following example.

EXAMPLE 8.4 Let $X = c_{00}$ with $\|\cdot\|_p$, $1 < p < \infty$, and (x_n) in c_{00} be defined by

$$x_n(j) = \begin{cases} 1/j & \text{if } j \le n \\ 0 & \text{if } j > n, \end{cases}$$

for all $n \in \mathbb{N}$. If $f_1, f_2, \ldots$ are the coordinate functionals on X, i.e., $f_j(x) = x(j)$, $x \in X$, $j \in \mathbb{N}$, then for every $j \in \mathbb{N}$, we have

$$f_j(x_n) = x_n(j) \to \frac{1}{j} \quad \text{as} \quad n \to \infty.$$

Thus, defining $\varphi_n = J(x_n)$ for every $n \in \mathbb{N}$, where $J : X \to X''$ is the canonical linear isometry, we have the convergence of the sequence $(\varphi_n(f_j))$ for every $j \in \mathbb{N}$. Since X' is linearly isometric with $(\ell^p)'$, the sequence $(\|\varphi_n\|)$ is bounded, and span $\{f_j : j \in \mathbb{N}\}$ is dense in $(\ell^p)'$ (Theorem 8.4 and Corollary 8.3), it follows from Theorem 3.11 that $(\varphi_n(f))$ converges for every $f \in X'$. In fact, the map $\varphi : X' \to \mathbb{K}$ defined by

$$\varphi(f) = \lim_{n\to\infty} \varphi_n(f), \qquad f \in X',$$

belongs to X''. Thus, $f(x_n) \to \varphi(f)$ for all $f \in X'$. But, there is no $x \in X = c_{00}$ such that $f(x_n) \to f(x)$ for every $f \in X'$. For, if there is $x \in X$ such that $f(x_n) \to f(x)$ for every $f \in X'$, then it would imply that

$$x_n(j) = f_j(x_n) \to f_j(x) = x(j) \quad \forall j \in \mathbb{N}$$

and, consequently, $x(j) = 1/j$ for every $j \in \mathbb{N}$. This is a contradiction, since $x \in c_{00}$.

Looking at the above example, one may anticipate that such a situation does not arise if the space is reflexive. The result is, in fact, true as the following theorem shows. Its proof is using the arguments involved in the above example, and hence, we leave it as an exercise.

Theorem 8.22 *Suppose X is a reflexive space and (x_n) is a sequence in X such that $(f(x_n))$ converges for every $f \in X'$. Then there exists $x \in X$ such that (x_n) converges to x weakly.*

Of course, a sequence (f_n) in X' converges weakly to an element $f \in X'$ if and only if for every $\varphi \in X''$, $\varphi(f_n) \to \varphi(f)$ as $n \to \infty$. Using the canonical isometry $J : X \to X''$, it can be seen that if (f_n) converges weakly to f in X', then $(f_n(x))$ converges to $f(x)$ for every $x \in X$. In case X is a reflexive space, then we have the converse as well.

Theorem 8.23 *If X is a reflexive space, then a sequence (f_n) in X' is weakly convergent to $f \in X'$ if and only if $f_n(x) \to f(x)$ for every $x \in X$.*

Proof. The proof is left as an exercise for the reader. ∎

If X is not reflexive, then pointwise convergence of (f_n) to f does not imply its weak convergence in X'. To see this, consider the following example.

EXAMPLE 8.5 Let $X = c_{00}$ with $\|\cdot\|_\infty$. For each $n \in \mathbb{N}$, let

$$u_n(j) = \begin{cases} 1/n & \text{if } j \leq n, \\ 0 & \text{if } j > n \end{cases}$$

and let $f_n : X \to \mathbb{K}$ be defined by

$$f_n(x) = \sum_{j=1}^{\infty} u_n(j)x(j), \qquad x \in X.$$

It is seen that $f_n \in X'$ for all $n \in \mathbb{N}$. Also, note that for $x \in X$, if $k_x \in \mathbb{N}$ is such that $x(j) = 0$ for all $j > k_x$, then $|f_n(x)| \leq \|x\|_\infty k_x/n$ for all $x \in c_{00}$, so that we have $\lim_{n\to\infty} f_n(x) = 0$, that is, (f_n) converges pointwise to $f = 0$. We show that there is a $\varphi \in X''$ such that $\varphi(f_n) \not\to \varphi(f)$. For this, recall from Theorem 8.4 that every $g \in X'$ is represented uniquely by an element $v_g \in \ell^1$, in the sense that

$$g(x) = \sum_{j=1}^{\infty} x(j)v_g(j) \quad \forall x \in X.$$

Note that the element in ℓ^1 which represents f_n is u_n. Now, let $\varphi : X' \to \mathbb{K}$ be defined by

$$\varphi(g) = \sum_{j=1}^{\infty} v_g(j), \qquad g \in X'.$$

Clearly, $\varphi \in X''$, and $\varphi(f_n) = \sum_{j=1}^{\infty} u_n(j) = 1$ for all $n \in \mathbb{N}$. Thus, we have proved that $f_n(x) \to f(x) = 0$ for all $x \in X$, but for the above $\varphi \in X''$, $\varphi(f_n) \not\to \varphi(f)$, so that (f_n) does not converge weakly to f.

The above example and the discussion preceding Theorem 8.23 show that the pointwise convergence of a sequence in X' is weaker than its weak convergence. Pointwise convergence of a sequence in X' is also called **weak* convergence** (weak-star convergence). One may also observe that the weak* convergence in X' is the convergence with respect to the topology on X' induced by the set $J(X)$, where J is the canonical isometry from X to X''.

Eberlein-Shmulyan theorem

We know by Theorem 2.39 that if every bounded sequence in a normed linear space has a convergent subsequence, then the space is finite dimensional. One may ask whether every bounded sequence in a normed linear space has a weak convergent subsequence. The answer is in the negative. To see this, consider the following situation:

Let $X = \ell^1$. Then the sequence (e_n) is bounded and has no weak convergent subsequence. For, if there is a weak convergent subsequence, say (x_n), then, by Schur's lemma (Theorem 8.18), (x_n) must converge to some $x \in \ell^1$. Also, since $x_n(j) \to 0$ as $n \to \infty$ for every $j \in \mathbb{N}$, we must have $x(j) = 0$ for every $j \in \mathbb{N}$, i.e., $x = 0$. It is a contradiction since $\|x_n\| = 1$ for all $n \in \mathbb{N}$.

The Eberlein-Shmulyan theorem says that the answer to the question raised above is affirmative if the space X is reflexive. First, we prove this result with an additional assumption that the dual space is separable. In fact, we prove a slightly more general result.

Proposition 8.24 *Suppose X' is separable and (x_n) is a bounded sequence in X. Then there is a subsequence $(\tilde{x}_n)$ of (x_n) such that $(f(\tilde{x}_n))$ converges for every $f \in X'$.*

If, in addition, X is reflexive, then (x_n) has a weakly convergent subsequence.

Proof. Suppose X' is separable and (x_n) is a bounded sequence in X. Let $D := \{f_1, f_2, \ldots\}$ be a denumerable dense subset of X'. We know that every bounded sequence in $\mathbb{K}$ has a convergent subsequence. Therefore,

$$(f_1(x_n)) \text{ has a convergent subsequence, say } (f_1(x_{n,1})),$$
$$(f_2(x_{n,1})) \text{ has a convergent subsequence, say } (f_2(x_{n,2})),$$
$$(f_3(x_{n,2})) \text{ has a convergent subsequence, say } (f_3(x_{n,3})),$$

and so on. Note that for each $j \in \mathbb{N}$, $(x_{n,j+1})$ is a subsequence of $(x_{n,j})$. Hence, it follows that the sequence $(f_j(x_{n,n}))$ is a subsequence of the convergent sequence $(f_j(x_{n,j}))$, so that $(f_j(x_{n,n}))$ converges. Now, let $\varphi_n = J(x_{n,n})$, $n \in \mathbb{N}$, where $J : X \to X''$ is the canonical linear isometry. Then (φ_n) is a bounded sequence in $X'' = \mathcal{B}(X', \mathbb{K})$ such that the sequence $(\varphi_n(f))$ converges for every $f \in D$. Therefore, using denseness of the set D, it follows from Theorem 3.11 that there exists $\varphi \in X''$ such that $\varphi_n(f) \to \varphi(f)$ for every $f \in X'$. Thus,

$$f(x_{n,n}) \to \varphi(f) \quad \forall\, f \in X'.$$

If, in addition, X is reflexive, then there exists $x \in X$ such that $\varphi = J(x)$. Hence,

$$f(x_{n,n}) = \varphi_n(f) \to \varphi(f) = (Jx)(f) = f(x) \quad \forall\, f \in X'.$$

This shows that $(x_{n,n})$ converges weakly to x. ∎

Remark 8.3 Recall that neither the space ℓ^1 is reflexive, nor its dual is separable. So the above proposition cannot be applied to ℓ^1. Whereas, it can be applied to spaces ℓ^p and $L^p[a,b]$ for $1 < p < \infty$. Also, the first part of the above theorem can be applied to the space $(c_{00}, \|\cdot\|_p)$ with $1 < p \leq \infty$, and to the spaces c_0, c.

Now we prove the last part of the above proposition by dropping the assumption that the dual space is separable.

Theorem 8.25 (Eberlein-Shmulyan) *Every bounded sequence in a reflexive space has a weakly convergent subsequence.*

Proof. Let (x_n) be a bounded sequence in a reflexive space X, and let

$$X_0 := \mathrm{cl}\,(\mathrm{span}\,\{x_n : n \in \mathbb{N}\}).$$

Since $\mathrm{span}\,\{x_n : n \in \mathbb{N}\}$ is separable, X_0 is also separable. Moreover, by Theorems 8.17 and 8.15, X_0 is reflexive and X_0' is separable. Hence, by Theorem 8.24, there is a subsequence $(\tilde{x}_n)$ of (x_n) and $x \in X_0$ such that

$$f(\tilde{x}_n) \to f(x) \quad \forall\, f \in X_0'.$$

Now, for any $g \in X'$, taking $f = g_{|X_0} \in X_0'$, we have

$$g(\tilde{x}_n) = f(\tilde{x}_n) \to f(x) = g(x).$$

Thus, $(\tilde{x}_n)$ converges weakly to x. ∎

The Eberlein-Shmulyan theorem in its full generality contains the converse of the above result as well (cf. [11], [29]). Thus, if a space X is not reflexive, then there always exists a bounded sequence having no weakly convergent subsequence.

Now we give an important application of the above theorem.

8.2.3 Best Approximation in Reflexive Spaces

Recall from Theorems 2.43 and 2.44 that if X_0 is a finite dimensional subspace of a normed linear space X or a complete subspace of an inner product space X, then for every $x \in X$, there exists $x_0 \in X_0$ such that

$$\|x - x_0\| = \operatorname{dist}(x, X_0) = \inf\{\|x - v\| : v \in X_0\}.$$

As a consequence, the Riesz lemma (Theorem 2.40) holds for $r = 1$ as well (cf. Theorem 2.42). Also, if X is an infinite dimensional space, then for $f \in X'$, there need not exist $x \in X$ such that $\|x\| = 1$ and $|f(x)| = \|f\|$ (*Why?*). Now we show that these results hold for every reflexive space X.

Theorem 8.26 *Suppose X is a reflexive space. Then we have the following:*

(i) *If Y is a closed subspace of X, then for every $x \in X$, there exists $y \in Y$ such that*

$$\|x - y\| = \operatorname{dist}(x, Y).$$

(ii) *If Y is a proper closed subspace of X, then there exists $x_0 \in X$ such that*

$$\|x_0\| = 1, \qquad \operatorname{dist}(x_0, Y) = 1.$$

(iii) *For every $f \in X'$, there exists $x \in X$ such that*

$$\|x\| = 1, \qquad |f(x)| = \|f\|.$$

Proof. (i) Let Y be a closed subspace of X and $x \in X$. If $x \in Y$, then we may take $y = x$. If $x \notin Y$, consider the closed set

$$E = \{x - v : v \in Y\}.$$

Then we have

$$\operatorname{dist}(x, Y) = \inf\{\|u\| : u \in E\}.$$

Let $d = \operatorname{dist}(x, Y)$ and let (u_n) in E be such that $\|u_n\| \to d$. By Theorem 8.25, (u_n) has a weakly convergent subsequence, say $(\tilde{u}_n)$. Suppose $(\tilde{u}_n)$ converges weakly to $u \in X$. Then, by the last part of Theorem 8.19,

$$\|u\| \leq \liminf_n \|\tilde{u}_n\|.$$

Since $\|\tilde{u}_n\| \to d$, we have $\|u\| \leq d$. Therefore, if we show that $u \in E$, then it follows that there is $v \in Y$ such that $u = x - v$ and

$$d \leq \|u\| = \|x - v\| \leq d.$$

Assume for a moment $u \notin E$, i.e., $x - u \notin Y$. Since $\tilde{u}_n \in E$ for all $n \in \mathbb{N}$, by a consequence of the Hahn-Banach theorem (Corollary 5.5), there exists $f \in X'$ such that

$$f(x - u) \neq 0, \qquad f(x - \tilde{u}_n) = 0 \quad \forall n \in N.$$

This is a contradiction since $f(x - \tilde{u}_n) \to f(x - u)$ as $n \to \infty$.

(ii) Let Y be a proper closed subspace of X and $x \in X \setminus Y$. By (i), there exists $y \in Y$ such that

$$\|x - y\| = \operatorname{dist}(x, Y) = \operatorname{dist}(x - y, Y).$$

Taking $x_0 = (x-y)/\|x-y\|$, it follows that $\|x_0\| = 1, \operatorname{dist}(x_0, Y) = 1$.

(iii) Let $0 \neq f \in X'$. By (ii), there exists $x_0 \in X$ such that $\|x_0\| = 1$ and $\operatorname{dist}(x_0, N(f)) = 1$. Therefore, by Theorem 3.8,

$$\|f\| = \frac{|f(x_0)|}{\operatorname{dist}\,(x_0, N(f))} = |f(x_0)|. \quad \blacksquare$$

Now, the following theorem is a consequence of Theorems 8.26(i) and 2.46.

Theorem 8.27 *Suppose X is a reflexive space which is strictly convex as well, and Y is a closed subspace of X. Then for every $x \in X$, there exists a unique $y \in Y$ such that*

$$\|x - y\| = \operatorname{dist}(x, Y).$$

Remark 8.4 Recall that every inner product space is strictly convex (Theorem 2.16), but not necessarily reflexive; for example, c_{00} with $\|\cdot\|_2$. The spaces ℓ^p, $L^p[a, b]$ for $1 < p < \infty$, and all Hilbert spaces are reflexive as well as strictly convex.

PROBLEMS

1. For $1 \leq p < \infty$, reformulate the proof of the last assertion in Theorem 8.2 by showing that $(f(e_1), f(e_2), \ldots) \in \ell^q$ for every $f \in (\ell^p)'$, and the map $f \mapsto (f(e_1), f(e_2), \ldots)$ is a linear isometry from $(\ell^p)'$ onto ℓ^q.

2. Give details of the proof of Lemma 8.6.

3. Show that for every f in the dual of $(c_0, \|\cdot\|_\infty)$, the sequence $(f(e_n))$ belongs to ℓ^1, $f(x) = \sum_{j=1}^\infty x(j)f(e_j)$ for all $x \in c_0$, and the map $f \mapsto (f(e_1), f(e_2), \ldots)$ is a linear isometry from c_0' onto ℓ^1.

4. If f is in the dual of $(c, \|\cdot\|_\infty)$, then show, without relying on a representation of c_{00}', that $(f(e_1), f(e_2), \ldots) \in \ell^1$.

5. For $(\eta, \eta_1, \eta_2, \ldots) \in \ell^1$, let $f : c \to \mathbb{K}$ be defined by

$$f(x) = \eta\alpha + \sum_{j=1}^{\infty} x(j)\eta_j, \qquad x \in c,$$

where $\alpha := \lim_{n\to\infty} x(n)$. Show that f belongs to the dual of $(c, \|\cdot\|_\infty)$ and $\|f\| = |\eta_0| + \sum_{j=1}^\infty |\eta_j|$.

6. Show that $\|\phi\| := |\phi(a)| + V(\phi)$ defines a norm on $BV[a,b]$. (Here, $V(\phi)$ denotes the total variation of $\phi \in BV[a,b]$.)

7. Show that $NBV[a,b]$, the space of all normalized functions of bounded variation on [a,b] with the norm $\phi \mapsto V(\phi)$ is a Banach space.

8. Supply the missing details of the proof for the fact that (8.3) holds for all $x \in L^p[a,b]$.

9. Let X be a Hilbert space and (x_n) be a sequence in X. Show that if (x_n) converges weakly to $x \in X$ and $\|x_n\| \to \|x\|$, then $\|x_n - x\| \to 0$.

10. Let X be a normed linear space and $S = \{f \in X' : \|f\| = 1\}$. Let (x_n) be a sequence in X. Show that

(a) (x_n) converges weakly to $x \in X$ if and only if $f(x_n) \to f(x)$ for every $f \in S$, and

(b) (x_n) converges to $x \in X$ if and only if $f(x_n) \to f(x)$ uniformly for $f \in S$.

11. Let (x_n) be a sequence in a normed linear space X, and let $x \in X$. Show that if $\operatorname{Re} f(x_n) \to \operatorname{Re} f(x)$ for every $f \in X'$, then (x_n) converges weakly to x.

12. Let X be a normed linear space. Show that the weak topology on X induced by X' is the same as the topology induced by the norm on X if and only if X is finite dimensional.

13. Let X be a normed linear space. Show that a subspace X_0 is closed if and only if it is closed with respect to the weak topology on X induced by X'.

[*Hint*: X_0 closed in X implies that $X_0 = \cap\{N(f) : f \in X_0^a\}$, where $X_0^a = \{f \in X' : f(x) = 0 \; \forall x \in X_0\}$.]

14. Let X be a normed linear space and (x_n) be a sequence in X. Show that, if (x_n) converges weakly to $x \in X$, then there exists a sequence (y_n) in $\operatorname{span}\{x_1, x_2, \ldots\}$ such that (y_n) converges to x.

15. Let f_n and f in $(\ell^1)'$ be defined by

$$f_n(x) = \sum_{j=1}^{n} x(j), \qquad f(x) = \sum_{j=1}^{\infty} x(j),$$

respectively, for $x \in \ell^1$. Show that (f_n) does not converge weakly to f in $(\ell^1)'$.

[*Hint*: Let $u_n, u \in \ell^\infty$ be such that $T(u_n) = f_n$, $T(u) = f$, where $T : \ell^\infty \to (\ell^1)'$ is as in Theorem 8.2. Show that there exists $g \in (\ell^\infty)'$ such that $g(u_n) = 0$ for every $n \in \mathbb{N}$, but $g(u) \neq 0$. Consider $\varphi \in (\ell^1)''$ such that $T'\varphi = \varphi$, and see that $\varphi(f_n) = 0$ for every $n \in \mathbb{N}$, but $\varphi(f) \neq 0$.]

16. Let (x_n) be in X which converges weakly to $x \in X$. Show that

(a) $x \in \operatorname{cl}(\operatorname{span}\{x_n : n \in \mathbb{N}\})$, and

(b) if $A \in \mathcal{B}(X, Y)$, then (Ax_n) converges weakly to Ax in Y.

9

Compact Operators

Recall from Chapter 3 that, if $A : X \to Y$ is a bounded operator between normed linear spaces X and Y, then $\mathrm{cl}\{Ax : \|x\| \leq 1\}$ is a closed and bounded subset of Y. Since every closed and bounded subset of a finite dimensional normed linear space is compact, it follows that, if $A : X \to Y$ is a bounded operator of finite rank, then the above set is compact as well. This is not necessarily true if $\operatorname{rank} A = \infty$. For example, suppose A is the identity operator on an infinite dimensional normed linear space X. Then the above set is the same as the closed unit ball of X which, by Theorem 2.39, is not compact.

A linear operator $A : X \to Y$ between normed linear spaces X and Y is said to be a **compact operator** if the set $\mathrm{cl}\{Ax : \|x\| \leq 1\}$ is compact in Y.

The following results are easy to verify:

(1) Every bounded operator of finite rank is a compact operator.

(2) The identity operator on a normed linear space is a compact operator if and only if the space is of finite dimension.

(3) If X_0 is a subspace of a normed linear space X, then the inclusion operator $I_0 : X_0 \to X$, i.e., $I_0 x = x$ for all $x \in X_0$, is a compact operator if and only if X_0 is finite dimensional.

(4) If $P : X \to X$ is a projection operator, then it is compact if and only if $\operatorname{rank} P < \infty$, because $P|_{R(P)} : R(P) \to X$ is an inclusion operator.

Compact operators arise naturally in applications as integral operators and limits of finite rank operators. We shall discuss these examples in this chapter and in subsequent chapters. Also, see Problems 3 – 6 at the end of this chapter.

9.1 Some Characterizations

The characterizations for compact operators given in the following theorem and some of the properties of the set of compact operators given in the next section will help us identify compact operators, and check whether an operator is compact or not.

We shall make use of the fact that a subset E of a metric space Ω is compact if and only if every sequence in E has a subsequence which converges in E. Using this, it can be easily seen that, if $A : X \to Y$ is a linear operator between normed linear spaces X and Y, then for any $r > 0$,

$$\text{cl}\,\{Ax : \|x\| \le r\} \text{ compact } \iff \text{cl}\,\{Ax : \|x\| \le 1\} \text{ compact},$$

$$\text{cl}\,\{Ax : \|x\| < r\} \text{ compact } \iff \text{cl}\,\{Ax : \|x\| < 1\} \text{ compact}.$$

Theorem 9.1 *Let X and Y be normed linear spaces and $A : X \to Y$ be a linear operator. Then the following are equivalent:*

(i) *A is a compact operator.*

(ii) $\text{cl}\,\{Ax : \|x\| < 1\}$ *is compact in Y.*

(iii) *For every bounded subset E of X, $\text{cl}\,A(E)$ is compact in Y.*

(iv) *For every bounded sequence (x_n) in X, the sequence (Ax_n) has a convergent subsequence in Y.*

Proof. Clearly, (iii) implies (i) and (ii). Now assume that (i) holds, i.e., $\text{cl}\,\{Ax : \|x\| \le 1\}$ is compact. Let E be a bounded subset of X. Then, we know that there exists $r > 0$ such that

$$E \subseteq \{x \in X : \|x\| < r\}.$$

Now, from the relations

$$\text{cl}\,[A(E)] \subseteq \text{cl}\,\{Ax : x \in X,\ \|x\| < r\} \subseteq \text{cl}\,\{Ax : x \in X,\ \|x\| \le r\},$$

and the fact that a closed subset of a compact set is compact, it follows that (i) implies (ii) and (iii), and (ii) implies (iii).

Now we prove the equivalence of (iii) and (iv). Assume that (iii) holds, and let (x_n) be a bounded sequence in X. Suppose that $\|x_n\| \le c$ for every $n \in \mathbb{N}$, and let $E = \{x \in X : \|x\| \le c\}$. Then (Ax_n) is a sequence in the compact set $\text{cl}\,[A(E)]$, so that it has a convergent subsequence. Thus, (iv) holds. Conversely, assume that (iv) holds, and let E be a bounded subset of X. To show that

$\mathrm{cl}\,[A(E)]$ is compact, it is enough to prove that every sequence in it has a convergent subsequence. To show this, suppose that (y_n) is a sequence in $\mathrm{cl}\,[A(E)]$. Then there exists (x_n) in E such that $\|y_n - Ax_n\| \leq 1/n$ for all $n \in \mathbb{N}$. Now by the hypothesis (iv), (Ax_n) has a convergent subsequence, say (Ax_{n_j}). Then it follows that (y_{n_j}) is a convergent subsequence of (y_n). ∎

We already know that the identity operator on an infinite dimensional normed linear space is not a compact operator. Now, using one of the characterizations given in the above theorem, we give some more examples of bounded operators which are not compact.

EXAMPLE 9.1 Let $1 \leq p \leq \infty$. (i) Let $A : \ell^p \to \ell^p$ be the *right shift operator* on ℓ^p, i.e., A is defined by

$$(Ax)(i) = \begin{cases} 0 & \text{if } i = 1, \\ x(i-1) & \text{if } i > 1 \end{cases}$$

Since

$$Ae_n = e_{n+1}, \qquad \|e_n - e_m\|_p = \begin{cases} 2^{1/p} & \text{if } 1 \leq p < \infty, \\ 1 & \text{if } p = \infty, \end{cases}$$

for all $n, m \in \mathbb{N}$, $n \neq m$, it follows that, corresponding to the bounded sequence (e_n), the sequence (Ae_n) does not have a convergent subsequence. Hence, by Theorem 9.1(iv), the operator A is not compact.

(ii) Using similar arguments as in (i) above, it can be seen that the *left shift operator* on ℓ^p is also not a compact operator on ℓ^p for any p with $1 \leq p \leq \infty$.

(iii) Let (λ_n) be a sequence of scalars such that $\lambda_n \to \lambda \neq 0$, and let $A : \ell^p \to \ell^p$ be the *diagonal operator* associated with this sequence, i.e., A is defined by

$$(Ax)(i) = \lambda_i x(i), \quad i \in \mathbb{N};\ x \in \ell^p.$$

Then, we know that $A \in \mathcal{B}(\ell^p)$, and $\|A\| = \sup_n |\lambda_n|$ (see Exercise 3.5). Since

$$Ae_n = \lambda_n e_n, \quad \|e_n - e_m\|_p = c_p := \begin{cases} 2^{1/p} & \text{if } 1 \leq p < \infty, \\ 1 & \text{if } p = \infty, \end{cases} \quad \forall n, m \in \mathbb{N},$$

for all $n, m \in \mathbb{N}$, it follows that

$$\begin{aligned}\|Ae_n - Ae_m\|_p &= \|\lambda_n e_n - \lambda_m e_m\|_p \\ &\geq \|\lambda_n(e_n - e_m)\|_p - \|(\lambda_n - \lambda_m)e_m\|_p \\ &= c_p|\lambda_n| - |\lambda_n - \lambda_m|.\end{aligned}$$

Since $\lambda_n \to \lambda \neq 0$, it follows that, (Ae_n) does not have any convergent subsequence, and hence, A is not a compact operator.

We shall show in Example 9.2(i) that if $\lambda_n \to 0$, then the operator A in the above example is compact.

9.2 Space of Compact Operators

We shall denote the set of all compact operators from X to Y by $\mathcal{K}(X,Y)$, and we write $\mathcal{K}(X)$ for $\mathcal{K}(X,X)$. Clearly, $\mathcal{K}(X,Y)$ is a subset of $\mathcal{B}(X,Y)$.

Theorem 9.2 *Let X, Y, Z be normed linear spaces. Then the following results hold:*

(i) *$\mathcal{K}(X,Y)$ is a subspace of $\mathcal{B}(X,Y)$.*

(ii) *If $A \in \mathcal{B}(X,Y)$ and $B \in \mathcal{B}(Y,Z)$ and if one of them is compact, then $BA \in \mathcal{K}(X,Z)$.*

Proof. For showing an operator to be compact, we shall use the characterization given in Theorem 9.1(iv).

(i) It is enough to prove that for every $A_1, A_2 \in \mathcal{K}(X,Y)$ and $\alpha \in \mathbb{K}$, the operator $A_1 + \alpha A_2$ belongs to $\mathcal{K}(X,Y)$. For this, let $A_1, A_2 \in \mathcal{K}(X,Y)$, $\alpha \in \mathbb{K}$, and let (x_n) be a bounded sequence in X. We show that the sequence $(A_1x_n + \alpha A_2x_n)$ has a convergent subsequence in Y.

Since A_1 is compact, there is a subsequence $(\tilde{x}_n)$ of (x_n) such that $(A_1\tilde{x}_n)$ converges in Y. Also, since $(\tilde{x}_n)$ is bounded and A_2 is compact, there is a subsequence $(\hat{x}_n)$ of $(\tilde{x}_n)$ such that $(A_2\hat{x}_n)$ converges. Hence the sequence $(A_1\hat{x}_n + \alpha A_2\hat{x}_n)$ converges.

(ii) Let $A \in \mathcal{B}(X,Y)$ and $B \in \mathcal{B}(Y,Z)$, and one of them compact. For showing $BA \in \mathcal{K}(X,Z)$, it is enough to show that for every bounded sequence (x_n) in X, the sequence (BAx_n) has a convergent subsequence. So, let (x_n) be a bounded sequence in X.

First assume that $A \in \mathcal{K}(X,Y)$. Then there is a subsequence $(\tilde{x}_n)$ of (x_n) such that $(A\tilde{x}_n)$ converges in Y. By continuity of B, $(BA\tilde{x}_n)$ converges in Z.

Next, assume that $B \in \mathcal{K}(Y,Z)$. Since $A \in \mathcal{B}(X,Y)$, the sequence (Ax_n) is bounded in Y so that using compactness of B, (BAx_n) has a convergent subsequence in Z. ∎

If the space Y in the above theorem is a Banach space, then the set $\mathcal{K}(X,Y)$ is a closed subspace of $\mathcal{B}(X,Y)$, as the following theorem shows.

Theorem 9.3 *Let X be a normed linear space, Y be a Banach space and $A \in \mathcal{B}(X,Y)$. If (A_n) is a sequence in $\mathcal{K}(X,Y)$ such that $\|A_n - A\| \to 0$ as $n \to \infty$, then $A \in \mathcal{K}(X,Y)$.*

Proof. Let (A_n) be in $\mathcal{K}(X,Y)$ such that $\|A_n - A\| \to 0$. In order to show that $A \in \mathcal{K}(X,Y)$, by Theorem 9.1 (iv), it is enough to show that for every bounded sequence (x_n) in X, the sequence (Ax_n) has a convergent subsequence in Y.

Since each A_k is compact and (x_n) is bounded, (x_n) has a subsequence $(x_n^{(k)})$ such that $(A_k x_n^{(k)})$ converges. Without loss of generality, we may assume that $(x_n^{(k+1)})$ is a subsequence of $(x_n^{(k)})$. Note that by construction, $(x_{k+n}^{(k+n)})$ is a subsequence of $(x_{k+n}^{(k)})$. Therefore, by taking $u_n = x_n^{(n)}$, $n \in \mathbb{N}$, the sequence $(A_k u_n)$ converges. Now, let $\varepsilon > 0$ be given. Since $\|A_n - A\| \to 0$, there exists $k \in \mathbb{N}$ such that $\|A - A_k\| < \varepsilon$. Since $(A_k u_n)$ converges, there exists $N \in \mathbb{N}$ such that

$$\|A_k u_n - A_k u_m\| < \varepsilon \quad \forall n, m \geq N.$$

Therefore, for all $m, n \in \mathbb{N}$,

$$\begin{aligned} \|Au_n - Au_m\| &\leq \|(A - A_k)u_n\| + \|A_k u_n - A_k u_m\| + \|(A_k - A)u_m\| \\ &\leq c\varepsilon + \varepsilon + c\varepsilon \\ &= (2c + 1)\varepsilon. \end{aligned}$$

Thus, (Au_n) is a Cauchy subsequence of (Ax_n). This competes the proof. ∎

As a particular case of the above theorem, we can say that if X is a Banach space, and if (A_n) is a sequence of finite rank operators in $\mathcal{B}(X)$ such that $\|A - A_n\| \to 0$ as $n \to \infty$ for some $A \in \mathcal{B}(X)$, then A is a compact operator. An obvious question is whether every compact operator A on X is the limit of some sequence of finite rank operators. This question has been answered negatively by Enflo [12]. However, we shall show that if X is a Hilbert space, then the answer to the above question is in the affirmative (cf. Remark 13.1).

Remark 9.1 By Theorems 9.2 and 9.3, we can say that $\mathcal{K}(X)$ is an *ideal* of the *normed algebra* $\mathcal{B}(X)$; and if X is a Banach space, then $\mathcal{K}(X)$ is a closed ideal of the *Banach algebra* $\mathcal{B}(X)$.

By a **normed algebra** we mean a normed linear space $\mathcal{X}$, together with an additional structure of *multiplication*, say $(x, y) \mapsto xy$, such that the following are satisfied for all $x, y, z \in \mathcal{X}$:

(a) $(xy)z = x(yz)$,

(b) $(x + y)z = xz + yz$,

(c) $\alpha(xy) = (\alpha x)y = x(\alpha y)$,

(d) $\|xy\| \leq \|x\|\,\|y\|$.

A normed algebra, which is also a Banach space, is called a **Banach algebra.**

A subspace $\mathcal{I}$ of a normed algebra $\mathcal{X}$ is called an **ideal** of $\mathcal{X}$ if $xy, yx \in \mathcal{I}$ for all $x \in \mathcal{I}$, $y \in \mathcal{X}$.

It is easy to see that the space $C[a, b]$ with $\|\cdot\|_\infty$ is a Banach algebra with respect to pointwise multiplication, and for any subset Δ of $[a, b]$, the set $\mathcal{I} := \{x \in C[a, b] : x(t) = 0 \;\forall\, t \in \Delta\}$ is a closed ideal of $C[a, b]$.

In this text, we do not discuss any result related to normed algebra and ideals. For such a discussion with direct bearing on operator theory, one may refer Sunder [31].

We now give a few examples of compact operators of infinite rank.

EXAMPLE 9.2 (i) Let (λ_n) be a bounded sequence of scalars. Let $A : \ell^p \to \ell^p$ be defined by

$$(Ax)(i) = \lambda_i\, x(i), \qquad x \in \ell^p,\ i \in \mathbb{N}.$$

We have seen in Example 9.1(iii) that if (λ_n) converges to a nonzero scalar, then A is not a compact operator. Now, suppose that $\lambda_n \to 0$ as $n \to \infty$. In this case, we show that A is a compact operator.

For each $n \in \mathbb{N}$, let

$$(A_n x)(i) = \begin{cases} \lambda_i x(i) & \text{if } 1 \leq i \leq n \\ 0 & \text{if } i > n. \end{cases}$$

Clearly, $A_n : \ell^p \to \ell^p$ is a bounded operator of finite rank. In particular, each A_n is a compact operator. It also follows that

$$\|(A - A_n)x\|_p \leq (\sup_{i>n} |\lambda_i|) \|x\|_p \quad \forall x \in \ell^p, \ n \in \mathbb{N}.$$

Hence, by the assumption on (λ_n), we have

$$\|A - A_n\| \leq \sup_{i>n} |\lambda_i| \to 0 \quad \text{as} \quad n \to \infty$$

so that by Theorem 9.3, A is a compact operator. Since $Ae_n = \lambda_n e_n$ for all $n \in \mathbb{N}$, A is of infinite rank whenever $\lambda_n \neq 0$ for infinitely many n.

In (ii) – (v) below, we consider an operator on sequence spaces induced by an infinite matrix (a_{ij}) of scalars as in Example 3.3, i.e., on a suitable sequence space X, we define the operator A by

$$(Ax)(i) = \sum_{j=1}^{\infty} a_{ij} x(j), \qquad x \in X; \ i \in \mathbb{N}.$$

Of course, in order that the operator makes sense, the series on the right-hand side of the above equality must converge.

(ii) Let (a_{ij}) be an infinite matrix of scalars such that

$$\alpha := \sup_j \sum_{i=1}^{\infty} |a_{ij}| < \infty.$$

We know from Example 3.3(vi) that

$$A \in \mathcal{B}(\ell^1), \qquad \|A\| = \alpha.$$

We show that, if $\alpha_j := \sum_{i=1}^{\infty} |a_{ij}| \to 0$ as $j \to \infty$, then $A \in \mathcal{K}(\ell^1)$.

For each $n \in \mathbb{N}$, define

$$(A_n x)(i) = \sum_{j=1}^{n} a_{ij} x(j), \qquad x \in \ell^1; \ i \in \mathbb{N}.$$

Then $A_n \in \mathcal{B}(\ell^1)$ and is of finite rank. In fact, it is seen that for every $x \in \ell^1$,

$$A_n x = \sum_{j=1}^{n} x(j) v_j \quad \text{with} \quad v_j = \sum_{i=1}^{\infty} a_{ij} e_i;\ j = 1, \ldots, n.$$

Also, we have (*Verify*)

$$\|A_n\| \leq \sup_{1 \leq j \leq n} \alpha_j, \qquad \|A - A_n\| \leq \sup_{j > n} \alpha_j.$$

Therefore, if $\alpha_j \to 0$ as $j \to \infty$, then $\|A - A_n\| \to 0$. Thus, by Theorem 9.3, $A \in \mathcal{K}(\ell^1)$.

(iii) Let (a_{ij}) be an infinite matrix of scalars such that

$$\beta := \sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty.$$

Again, from Example 3.3(vii), we know that

$$A \in \mathcal{B}(\ell^\infty), \qquad \|A\| \leq \beta.$$

We show that, if $\beta_i := \sum_{j=1}^{\infty} |a_{ij}| \to 0$ as $i \to \infty$, then $A \in \mathcal{K}(\ell^\infty)$.

Now, for each $n \in \mathbb{N}$, define

$$(A_n x)(i) = \begin{cases} (Ax)(i) & \text{if } 1 \leq i \leq n \\ 0 & \text{if } i > n. \end{cases}$$

Then, $A_n \in \mathcal{B}(\ell^\infty)$, and is of finite rank. In this case, we have

$$A_n x = \sum_{i=1}^{n} (Ax)(i) e_i, \qquad x \in \ell^\infty,$$

$$\|A_n\| \leq \sup_{1 \leq i \leq n} \beta_i, \qquad \|A - A_n\| \leq \sup_{i > n} \beta_i.$$

Therefore, by Theorem 9.3, if $\beta_i \to 0$ as $i \to \infty$, then $A \in \mathcal{K}(\ell^\infty)$.

(iv) Let $1 < p < \infty$. In this case, the assumptions

$$\alpha := \sup_j \sum_{i=1}^{\infty} |a_{ij}| < \infty, \qquad \beta := \sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty$$

on the matrix (a_{ij}) imply (see Example 3.4(i)) that

$$A \in \mathcal{B}(\ell^p), \qquad \|A\| \le \alpha^{1/p}\beta^{1/q}.$$

For $x \in \ell^p$ and $n, i \in \mathbb{N}$, let

$$(A_n x)(i) = \sum_{j=1}^{n} a_{ij}x(j), \qquad (\tilde{A}_n x)(i) = \begin{cases} (Ax)(i) & \text{if } 1 \le i \le n \\ 0 & \text{if } i > n. \end{cases}$$

Then, $A_n, \tilde{A}_n \in \mathcal{B}(\ell^p)$ and are of finite rank. It can be seen (*Verify*) that

$$\|A - A_n\| \le \sup_{i\in\mathbb{N}} \beta_i^{1/q} \sup_{j>n} \alpha_j^{1/p},$$

$$\|A - \tilde{A}_n\| \le \sup_{i>n} \beta_i^{1/q} \sup_{j\in\mathbb{N}} \alpha_j^{1/p},$$

where α_j and β_i are as in (ii) and (iii) above. Therefore, if

$$\alpha_n \to 0 \quad \text{or} \quad \beta_n \to 0 \quad \text{as} \quad n \to \infty,$$

then, by Theorem 9.3, $A \in \mathcal{K}(\ell^p)$.

(v) Let $1 < p < \infty$ and $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|^q < \infty$. Then we know from Example 3.4(i) that

$$A \in \mathcal{B}(\ell^p, \ell^q), \qquad \|A\|^q \le \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|^q.$$

For each n, define A_n as in (iii) above. Then A_n is of finite rank, $A_n \in \mathcal{B}(\ell^p, \ell^q)$, and

$$\|A - A_n\|^q \le \sum_{i=n+1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|^q \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, by Theorem 9.3, $A \in \mathcal{K}(\ell^p, \ell^q)$.

If the matrix (a_{ij}) is diagonal, i.e., $a_{ij} = 0$ for $i \ne j$, then the quantities α_n, β_n in (ii) – (iv) coincide with $\lambda_n := a_{nn}$. In this special case, convergence of (λ_n) to zero ensures the compactness of the corresponding operator A. Thus, in the above, (i) is a particular case of (ii) – (iv). Note that the (i) is not a special case of (v).

Next, we consider examples of operators between function spaces.

(vi) For $k(\cdot,\cdot) \in C([a,b] \times [a,b])$, consider the integral operator K defined by

$$(Kx)(s) = \int_a^b k(s,t)x(t)\,d\mu(t), \qquad x \in L^1[a,b].$$

Note that for every $x \in L^1[a,b]$, $s, \tau \in [a,b]$,

$$\begin{aligned} |(Kx)(s)| &\leq \int_a^b |k(s,t)||x(t)|\,d\mu(t) \\ &\leq \sup_{s,t\in[a,b]} |k(s,t)|\|x\|_1, \\ |(Kx)(s) - (Kx)(\tau)| &\leq \int_a^b |k(s,t) - k(\tau,t)||x(t)|\,d\mu(t) \\ &\leq \sup_{t\in[a,b]} |k(s,t) - k(\tau,t)|\|x\|_1. \end{aligned}$$

Hence by continuity of $k(\cdot,\cdot)$, it follows that $Kx \in C[a,b]$ and

$$\|Kx\|_\infty \leq \sup_{s,t\in[a,b]} |k(s,t)|\|x\|_1.$$

We now show that if X, Y are any of the spaces $C[a,b]$, $L^p[a,b]$, $1 \leq p \leq \infty$, then $K : X \to Y$ is a compact operator. For this purpose, first we prove the following theorem.

Theorem 9.4 *For every bounded sequence (x_n) in $L^1[a,b]$, the sequence (Kx_n) has a convergent subsequence in $(C[a,b], \|\cdot\|_\infty)$.*

Proof. Let (x_n) be a bounded sequence in $L^1[a,b]$. By the observations preceding the statement of the theorem, it follows that the sequence (Kx_n) is bounded and equicontinuous in $C[a,b]$. Therefore, by Arzela-Ascoli's theorem (Theorem 6.7), it has a convergent subsequence in $C[a,b]$. ∎

EXAMPLE 9.2 (*cont.*)(vi) Let $1 \leq p \leq \infty$, and let q be the conjugate exponent of p. Since

$$\|x\|_1 \leq (b-a)^{1/q}\|x\|_p \quad \forall x \in L^p[a,b],\ p \neq 1,$$

$$\|x\|_p \leq (b-a)^{1/p}\|x\|_\infty \quad \forall x \in C[a,b],\ p \neq \infty,$$

it follows that

$$C[a,b] \subseteq L^p[a,b] \subseteq L^1[a,b],$$

and a bounded sequence in $L^p[a,b]$ remains bounded in $L^1[a,b]$. Also, if a sequence in $C[a,b]$ converges w.r.t. $\|\cdot\|_\infty$, then it converges w.r.t. $\|.\|_p$ as well, for every $p \in [1,\infty]$. Therefore, by the above theorem, and the characterization of a compact operator given in Theorem 9.1, $K : X \to Y$ is a compact operator for any $X, Y \in \{C[a,b],\ L^p[a,b]\}$.

(vii) Let p and q be conjugate exponents with $1 < p < \infty$. Let $k(\cdot,\cdot) \in L^q([a,b] \times [a,b])$, that is, for each $s \in [a,b]$, the function $t \mapsto k(s,t)$ belongs to $L^q[a,b]$, and the function

$$s \mapsto \int_a^b |k(s,t)|^q \, d\mu(t), \qquad s \in [a,b],$$

is integrable so that

$$\|k\|_q := \left[\int_a^b \int_a^b |k(s,t)|^q \, d\mu(t) \, d\mu(s)\right]^{1/q} < \infty.$$

Consider the integral operator K, defined by

$$(Kx)(s) = \int_a^b k(s,t)x(t) \, d\mu(t), \qquad x \in L^p[a,b].$$

Then, for $x \in L^p[a,b]$, we have

$$|(Kx)(s)| \leq \int_a^b |k(s,t)||x(t)| \, d\mu(t) \leq \left(\int_a^b |k(s,t)|^q \, d\mu(t)\right)^{1/q} \|x\|_p,$$

$$\left(\int_a^b |(Kx)(s)|^q \, d\mu(s)\right)^{1/q} \leq \|k\|_q \|x\|_p.$$

Thus, $K \in \mathcal{B}(L^p, L^q)$ and $\|K\| \leq \|k\|_q$.

Now, suppose that $(k_n(\cdot,\cdot))$ is a sequence in $C([a,b] \times [a,b])$ such that

$$\|k - k_n\|_q^q := \int_a^b \int_a^b |k(s,t) - k_n(s,t)|^q \, d\mu(t) \, d\mu(s) \to 0.$$

It is known that such sequence always exists (cf. Rudin [28]). Let K_n be the integral operator with kernel $k_n(\cdot,\cdot)$. Then, by the results in (v) above, K_n is a compact operator for each n. Moreover, we have

$$\|K - K_n\| \leq \|k - k_n\|_q \to 0,$$

so that by Theorem 9.3, $K \in \mathcal{K}(L^p, L^q)$.

9.3 Further Properties

Here is an important property of compact operators.

Theorem 9.5 *Let X and Y be normed linear spaces and $A : X \to Y$ be a linear operator. If $A \in \mathcal{K}(X,Y)$, then for every sequence (x_n) in X,*

$$x_n \to x \quad \text{weakly} \implies Ax_n \to Ax.$$

For the proof of the above theorem, we make use of the following easily verifiable result.

Lemma 9.6 *Let (Ω, d) be a metric space, (x_n) be a sequence in Ω and $x \in \Omega$. Then (x_n) converges to x in Ω if and only if every subsequence of (x_n) has a subsequence which converges to x.*

Proof of Theorem 9.5. Let (x_n) be a sequence in X which converges weakly to $x \in X$. By the above lemma, it is enough to show that every subsequence of (Ax_n) has a subsequence which converges to Ax.

Let $(\tilde{y}_n)$ be a subsequence of (Ax_n). Then $(\tilde{y}_n) = (A\tilde{x}_n)$ for some subsequence $(\tilde{x}_n)$ of (x_n). Since $(\tilde{x}_n)$ also converges weakly, by Theorem 8.19, it is bounded so that by compactness of A, $(\tilde{y}_n)$ has a convergent subsequence, i.e., there exists a subsequence $(\hat{x}_n)$ of $(\tilde{x}_n)$ such that $(A\hat{x}_n)$ converges to some $v \in Y$. It is enough to show that $v = Ax$. By Corollary 5.6 of the Hahn-Banach theorem (Theorem 5.1, it amounts to show that $f(v) = f(Ax)$ for every $f \in Y'$. Note that

$$f(v) = \lim_{n\to\infty} f(A\hat{x}_n) = \lim_{n\to\infty} (fA)(\hat{x}_n) \quad \forall f \in Y'.$$

Since $(\hat{x}_n)$ converges weakly to x, and since $fA \in X'$ for every $f \in Y'$, we have

$$\lim_{n\to\infty} (fA)(\hat{x}_n) = (fA)(x) = f(Ax).$$

Thus, $f(v) = f(Ax)$ for every $f \in Y'$. This completes the proof. ∎

Exercise 9.1 Using Theorem 9.5, show that the shift operators

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots), \qquad B : (\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots)$$

are not compact on ℓ^p for any p with $1 < p < \infty$. □

What about the converse of Theorem 9.5? Well, the converse is not true. For example, if $X = \ell^1$, then by Schur's lemma (Theorem 8.18), every weakly convergent sequence is convergent. Thus, every bounded operator on ℓ^1 has the property that it maps every weakly convergent sequence onto a convergent sequence. Obviously, every bounded operator on ℓ^1 is not compact. But, if the space X is reflexive, then the converse of Theorem 9.5 does hold.

Theorem 9.7 *Suppose X is a reflexive space, and $A : X \to Y$ is a linear operator such that for (x_n) in X,*

$$x_n \to x \quad \text{weakly} \implies Ax_n \to Ax.$$

Then $A \in \mathcal{K}(X, Y)$.

Proof. It is enough to show that, if (u_n) is a bounded sequence in X, then (Au_n) has a convergent subsequence. So, let (u_n) be a bounded sequence in X. By the Eberlein-Shmulyan theorem (Theorem 8.25), (u_n) has a weakly convergent subsequence, say $(\tilde{u}_n)$. Then, by the hypothesis, $(A\tilde{u}_n)$ converges. ∎

A linear operator $A : X \to Y$ is said to be a **completely continuous operator** if for every sequence (x_n) in X,

$$x_n \to x \quad \text{weakly} \implies Ax_n \to Ax.$$

Thus, every compact operator is completely continuous, but a completely continuous operator need not be compact.

The following result is important for applications to compact operator equations.

Theorem 9.8 *Let X and Y be normed linear spaces and $A : X \to Y$ be an injective compact operator. Then $A^{-1} : R(A) \to X$ is continuous if and only if* rank $A < \infty$.

Proof. Clearly, if rank $A < \infty$, then $A^{-1} : R(A) \to X$ is continuous.

Next, suppose that $A^{-1} : R(A) \to X$ is continuous. Since A is compact, it follows from Theorem 9.2 (ii) that the inclusion operator from $R(A)$ into Y, namely, $AA^{-1} : R(A) \to Y$, is a compact operator. Hence, $R(A)$ is finite dimensional. ∎

Corollary 9.9 *Let X and Y be normed linear spaces and $A : X \to Y$ be an infinite rank linear operator which is bounded below. Then A is not a compact operator.*

Proof. Since A is bounded below, by Proposition 3.2, A is injective and $A^{-1} : R(A) \to X$ is continuous, so that by Theorem 9.8, A cannot be a compact operator. ∎

By Theorem 9.8, if $A : X \to Y$ is an injective compact operator of infinite rank, then the solution (whenever it exists) of the operator equation $Ax = y$ is not stable, that is, for every $y \in R(A)$, there exists a sequence (y_n) in $R(A)$ such that $y_n \to y$ but $A^{-1}y_n \not\to A^{-1}y$.

Exercise 9.2 Show that the operators considered in (i) of Example 9.1 are bounded below, and are of infinite rank. □

We know that if $A : X \to Y$ is a bounded operator and X_0 is a subspace of X, then the restriction of A to X_0 is a bounded operator. Also, recall from Lemma 7.8 that if X_0 is a closed subspace of X such that $X_0 \subseteq N(A)$, then the quotient operator $\tilde{A} : X/X_0 \to Y$ defined by

$$\tilde{A}(x + X_0) = Ax, \quad x \in X,$$

is a bounded operator. Do such results hold if boundedness is replaced by compactness? It is easy to see that if $A \in \mathcal{K}(X, Y)$, then $A|_{X_0} \in \mathcal{K}(X_0, Y)$ for every subspace of X. The result holds for quotient operators as well.

Theorem 9.10 *Let X, Y be normed linear spaces and $X_0 \subseteq N(A)$ be a closed subspace of X. If $A : X \to Y$ is a compact operator, then $\tilde{A} : X/X_0 \to Y$ is also a compact operator.*

Proof. Suppose $A : X \to Y$ is a compact operator. Let $(x_n + X_0)$ be a bounded sequence in X/X_0, that is, there exists $c > 0$ such that

$$\|x_n + X_0\| := \inf \{\|x_n + u\| : u \in X_0\} \le c \quad \forall n \in \mathbb{N}.$$

Then it follows that there exists a sequence (u_n) in X_0 such that

$$\|x_n + u_n\| \le 2c \quad \forall n \in \mathbb{N}.$$

Now, by compactness of A, the sequence $(A(x_n + u_n))$ has a convergent subsequence. Since $A(x_n + u_n) = \tilde{A}(x_n + X_0)$ for every $n \in N$, it follows that $\tilde{A}$ is a compact operator. ∎

A particular case of the following theorem shows that if X and Y are Banach spaces and $A : X \to Y$ is a compact operator of infinite rank, then there exists $y \in Y$ such that the operator equation $Ax = y$ has no solution.

Theorem 9.11 *Let X and Y be Banach spaces, and $A : X \to Y$ be a compact operator with $R(A)$ closed in Y. Then* $\operatorname{rank} A < \infty$.

Proof. By Theorem 9.10, the operator $\tilde{A} : X/N(A) \to Y$ is a compact operator. Since $X/N(A)$ is a Banach space, $\tilde{A}$ is injective and $R(\tilde{A}) = R(A)$ is a Banach space, as a consequence of bounded inverse theorem (Theorem 7.4), the inverse of $\tilde{A}$ is continuous. Therefore, by Proposition 9.8, $R(A) = R(\tilde{A})$ is finite dimensional. ∎

Exercise 9.3 Using the above theorem, show that the following operators are not compact:

(a) The identity operator on an infinite dimensional normed linear space.

(b) The operator $(\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots)$ on ℓ^p.

(c) The operator $(\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots)$ on ℓ^p. □

Compactness of the transpose

Theorem 9.12 *Let X and Y be normed linear spaces. Then we have the following:*

(i) *If $A \in \mathcal{K}(X,Y)$, then $A' \in \mathcal{K}(Y', X')$.*

(ii) *Converse of* (i) *holds if Y is a Banach space.*

Proof. (i) Suppose $A \in \mathcal{K}(X,Y)$ and (f_n) is a bounded sequence in Y'. Since X' is a Banach space, it is enough to show that $(A'f_n)$ has a Cauchy subsequence. Note that for every $n, m \in \mathbb{N}$,

$$\begin{aligned} \|A'f_n - A'f_m\| &= \sup_{\|x\|\le 1} |A'(f_n - f_m)(x)| \\ &= \sup_{\|x\|\le 1} |(f_n - f_m)(Ax)| \\ &= \sup_{v\in\Omega} |f_n(v) - f_m(v)|, \end{aligned}$$

where $\Omega = \operatorname{cl}\{Ax : \|x\| \le 1\}$. Thus, it is enough to show that (f_n) has a subsequence which converges uniformly on Ω. For this purpose, first we observe that

$$|f_n(v_1) - f_n(v_2)| \le \|f_n\|\|v_1 - v_2\| \quad \forall v_1, v_2 \in \Omega$$

so that (g_n) with $g_n = f_n|_\Omega$, $n \in \mathbb{N}$, is a bounded equicontinuous sequence in $C(\Omega)$. Hence, by Arzela-Ascoli's theorem (Theorem 6.7), (g_n) has a convergent subsequence, say (g_{n_k}), so that the subsequence (f_{n_k}) of (f_n) converges uniformly on Ω.

(ii) Suppose Y is a Banach space and $A' : Y' \to X'$ is a compact operator. Let (x_n) be a bounded sequence in X. Since Y is a Banach space, it is enough to show that (Ax_n) has a Cauchy subsequence. By (i), $A'' : X'' \to Y''$ is a compact operator. Denoting $\phi_n = Jx_n$, $n \in \mathbb{N}$, where $J : X \to X''$ is the canonical linear isometry, it follows that (ϕ_n) is a bounded sequence in X''. Therefore, by compactness of A'', the sequence $(A''\phi_n)$ has a convergent subsequence, say $(A''\phi_{n_k})$. Therefore, using Corollary 5.9 of the Hahn-Banach theorem, we get

$$\begin{aligned} \|Ax_{n_k} - Ax_{n_m}\| &= \sup\{|f(Ax_{n_k} - Ax_{n_m})| : f \in Y', \|f\| = 1\} \\ &= \|A''\phi_{n_k} - A''\phi_{n_m}\| \end{aligned}$$

for all $k, m \in \mathbb{N}$. Thus, (Ax_{n_k}) is a Cauchy subsequence of (Ax_n). ∎

PROBLEMS

1. If X is an infinite dimensional normed linear space and $K \in \mathcal{K}(X)$, then show that, if λ is a nonzero scalar, then $\lambda I - K$ is not a compact operator . Deduce that the operator

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto \left(\alpha_1 + \alpha_2,\ \alpha_2 + \frac{\alpha_3}{2},\ \alpha_3 + \frac{\alpha_4}{3},\ \ldots\right)$$

is not a compact operator on ℓ^p, $1 \le p \le \infty$.

2. For $1 \le p \le \infty$, let q be the conjugate exponent of p. Let (a_{ij}) be an infinite matrix with $a_{ij} \in \mathbb{K}$, $i, j \in \mathbb{N}$. Show, in each of the following cases, that $(Ax)(i) = \sum_{j=1}^\infty a_{ij}x(j)$, $x \in \ell^p$, $i \in \mathbb{N}$, is well defined and $A : \ell^p \to \ell^r$ is a compact operator.

(a) $1 \le p \le \infty$, $1 \le r \le \infty$ and $\sum_{j=1}^\infty |a_{ij}| \to 0$ as $i \to \infty$.

(b) $1 \le p \le \infty$, $1 \le r < \infty$ and $\sum_{i=1}^\infty \left(\sum_{j=1}^\infty |a_{ij}|\right)^r < \infty$.

(c) $1 < p \le \infty$, $1 \le r \le \infty$ and $\sum_{j=1}^\infty |a_{ij}|^q \to 0$ as $i \to \infty$.

(d) $1 < p \le \infty$, $1 \le r < \infty$ and $\sum_{i=1}^\infty \left(\sum_{j=1}^\infty |a_{ij}|^q\right)^{r/q} < \infty$.

3. Show that the range of a compact operator is separable.

[*Hint*: Use the facts (to be proved) that totally bounded sets are separable, and the range of a compact operator is a countable union of totally bounded sets.]

The next three exercises are useful in numerical solution of integral equations (cf. Anselone [3] and Atkinson [4]).

4. Let X be a Banach space and $K \in \mathcal{K}(X)$. Let (P_n) be a sequence of finite rank projections on X such that $\|P_n x - x\| \to 0$ for every $x \in X$. Then prove that $\|K - P_n K\| \to 0$ as $n \to \infty$.

If $K_n \in \{P_n K, KP_n, P_n K P_n\}$, then show that

$$\|(K - K_n)K\| \to 0, \qquad \|(K - K_n)K_n\| \to 0.$$

5. Let $X = C[a,b]$ with $\|\cdot\|_\infty$ and $K : X \to X$ be defined by

$$(Kx)(s) = \int_a^b k(s,t)x(t)\,d\mu(t), \qquad x \in X,$$

where $k(\cdot,\cdot) \in C([a,b] \times [a,b])$. Let (K_n) be the Nyström approximation of K corresponding to a convergent quadrature formula with nodes $t_{1,n}, \ldots, t_{n,n}$ in $[a,b]$ and weights $w_{1,n}, \ldots, w_{n,n}$ in $\mathbb{K}$, i.e.,

$$(K_n x)(s) = \sum_{j=1}^{n} k(t_{j,n}, t)x(t_{j,n})w_{j,n}, \qquad x \in X,\ n \in \mathbb{N},$$

where nodes and weights are such that

$$\sum_{j=1}^{n} x(t_{j,n})w_{j,n} \to \int_a^b x(t)\,dt \quad \text{as} \quad n \to \infty \quad \text{for every } x \in C[a,b].$$

Then show that

(a) $\|Kx - K_n x\| \to 0$ for every $x \in X$,

(b) $\|(K - K_n)K\| \to 0$ and $\|(K - K_n)K_n\| \to 0$.

[*Hint*: (a): Show that for each $u \in C[a,b]$, $((K_n u(s))$ converges to $(Ku)(s)$, $\{K_n u : n \in \mathbb{N}\}$ is equicontinuous, and then conclude the result. (b): Use the result in (a) and the fact that $\{Ku : \|u\|_\infty \le 1\}$ and $\{K_n u : \|u\|_\infty \le 1,\ n \in \mathbb{N}\}$ are equicontinuous.]

6. Let X be a Banach space. A family $\{A_\alpha : \alpha \in \Lambda\}$ in $\mathcal{L}(X)$ such that the closure of $\cup_{\alpha\in\Lambda}\{A_\alpha x : x \in X,\ \|x\| \le 1\}$ is compact in X is called a *collectively compact set* of operators.

Suppose (A_n) is a sequence in $\mathcal{L}(X)$ such that $(A_n x)$ converges for every $x \in X$ and $\{A_n : n \in \mathbb{N}\}$ is a collectively compact set. Let $A : X \to X$ be defined by $Ax = \lim_{n\to\infty} A_n x$, $x \in X$. Prove the following:

(a) A is a compact operator.

(b) $\|(A - A_n)A\| \to 0$ and $\|(A - A_n)A_n\| \to 0$ as $n \to \infty$.

(c) $\{K_n : n \in \mathbb{N}\}$ in Problems 4 and 5 above are collectively compact sets.

7. Suppose X is a Hilbert space and $A \in \mathcal{B}(X)$. Show that $A \in \mathcal{K}(X)$ if and only if for every sequence (x_n) in X,

$$\langle x_n, u\rangle \to \langle x, u\rangle \quad \forall u \in X \implies Ax_n \to Ax.$$

8. Give an alternative proof for Theorem 9.11 using open mapping theorem (Theorem 7.9) and Theorem 2.39.

9. Let $1 \le p < \infty$. Is the set

$$\left\{x = (\alpha_1 \alpha_2, \ldots) \in \ell^p : \sum_{j=1}^{\infty} j^p |\alpha_j|^p < \infty\right\}$$

a closed subspace of ℓ^p?

10. Suppose X and Y are infinite dimensional normed linear spaces. Show that if $A : X \to Y$ is a surjective linear operator, then $A \notin \mathcal{K}(X, Y)$.

10

Spectral Results for Banach Space Operators

10.1 Eigenvalues and Approximate Eigenvalues

We recall from Section 1.2.6 that if $A : X \to X$ is a linear operator on a linear space X, then a scalar $\lambda \in \mathbb{K}$ is called an **eigenvalue** of A if there exists a nonzero $x \in X$ such that

$$Ax = \lambda x,$$

and in that case x is called an **eigenvector** of A corresponding to the eigenvalue λ. The set of all eigenvalues of A is called the **eigenspectrum** of A. Eigenspectrum is also known as **point spectrum.** We denote the eigenspectrum of A by $\sigma_{\text{eig}}(A)$. Thus,

$$\sigma_{\text{eig}}(A) = \{\lambda \in \mathbb{K} : \exists\, x \neq 0 \text{ such that } Ax = \lambda x\}.$$

Note that

$$\lambda \in \sigma_{\text{eig}}(A) \iff N(A - \lambda I) \neq \{0\},$$

and nonzero elements of $N(A - \lambda I)$ are the eigenvectors of A corresponding to the eigenvalue λ. The subspace $N(A - \lambda I)$ is called the **eigenspace** of A corresponding to the eigenvalue λ.

We know that a linear operator need not have eigenvalues at all. For example, recall from Example 1.8(ii), (v) that the linear operator $A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$A((\alpha_1, \alpha_2)) = (\alpha_2, -\alpha_1), \qquad (\alpha_1, \alpha_2) \in \mathbb{R}^2,$$

and the operator $A : \mathcal{P} \to \mathcal{P}$ defined by

$$(Ax)(t) = tx(t), \qquad x \in \mathcal{P},$$

have no eigenvalues. However, every linear operator on a finite dimensional linear space over $\mathbb{C}$ has at least one eigenvalue (Theorem 1.14).

EXAMPLE 10.1 In the following, X is any of the sequence spaces c_{00}, c_0, c, ℓ^p.

(i) Let (λ_n) be a bounded sequence of scalars. Let $A : X \to X$ be the *diagonal operator* defined by

$$(Ax)(j) = \lambda_j\, x(j), \qquad x \in X;\ j \in \mathbb{N}.$$

Then it is easily seen that, for $\lambda \in \mathbb{K}$, the equation $Ax = \lambda x$ is satisfied for a nonzero $x \in X$ if and only if $\lambda = \lambda_j$ for some $j \in \mathbb{N}$. Hence,

$$\sigma_{\text{eig}}(A) = \{\lambda_1, \lambda_2, \ldots\}.$$

In fact, for $n \in \mathbb{N}$, $e_n \in X$ defined by $e_n(j) = \delta_{ij}$ is an eigenvector of A corresponding to the eigenvalue λ_n.

(ii) Let $A : X \to X$ be the *right shift operator*, i.e.,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots).$$

Let $\lambda \in \mathbb{K}$. Then we see that the equation $Ax = \lambda x$ is satisfied for some $x = (\alpha_1, \alpha_2, \ldots) \in X$ if and only if

$$0 = \lambda\alpha_1, \qquad \alpha_j = \lambda\alpha_{j+1} \quad \forall j \in \mathbb{N}.$$

This is possible only if $\alpha_j = 0$ for all $j \in \mathbb{N}$. Thus,

$$\sigma_{\text{eig}}(A) = \varnothing.$$

(iii) Let $A : X \to X$ be the *left shift operator*, i.e.,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots).$$

We see that, for $x = (\alpha_1, \alpha_2, \ldots) \in X$ and $\lambda \in \mathbb{K}$,

$$Ax = \lambda x \iff \alpha_{n+1} = \lambda^n \alpha_1.$$

From this we can infer the following: Clearly, $\lambda = 0$ is an eigenvalue of A with a corresponding eigenvector e_1. Now suppose that $\lambda \neq 0$. If λ is an eigenvalue, then a corresponding eigenvector is of the form

$x = \alpha_1(1, \lambda, \lambda^2, \lambda^3, \ldots)$ for some nonzero α_1. Note that if $\alpha_1 \neq 0$ and $\lambda \neq 0$, then $x = \alpha_1(1, \lambda, \lambda^2, \lambda^3, \ldots)$ does not belong to c_{00}. Thus, if $X = c_{00}$, then

$$\sigma_{\text{eig}}(A) = \{0\}.$$

Next, consider the cases of $X = c_0$ or $X = \ell^p$ for $1 \leq p < \infty$. In these cases we see that $(1, \lambda, \lambda^2, \lambda^3, \ldots) \in X$ if and only if $|\lambda| < 1$, so that

$$\sigma_{\text{eig}}(A) = \{\lambda : |\lambda| < 1\}.$$

For the case of $X = c$, we see that $(1, \lambda, \lambda^2, \lambda^3, \ldots) \in X$ if and only if either $|\lambda| < 1$ or $\lambda = 1$. Thus, in this case,

$$\sigma_{\text{eig}}(A) = \{\lambda : |\lambda| < 1\} \cup \{1\}.$$

If $X = \ell^\infty$, then $(1, \lambda, \lambda^2, \lambda^3, \ldots) \in X$ if and only if $|\lambda| \leq 1$. Thus, in this case,

$$\sigma_{\text{eig}}(A) = \{\lambda : |\lambda| \leq 1\}.$$

We can see as in Example 9.2(i) that if $\lambda_n \to 0$ as $n \to \infty$, and if we endow the spaces c_{00} and ℓ^p with $\|\cdot\|_p$ for $1 \leq p \leq \infty$, and the spaces c_0 and c with $\|\cdot\|_\infty$, then the operator A in Example 10.1(i) above is a compact operator. Also, we have seen that if (λ_n) converges to a nonzero scalar, then the operator is not compact. This is, in fact, a common feature of all compact operators as the following theorem shows.

Theorem 10.1 *Let X be a normed linear space and $A : X \to X$ be a compact operator. Then $\sigma_{\text{eig}}(A)$ is a countable subset of $\mathbb{K}$, and zero is the only possible limit point of $\sigma_{\text{eig}}(A)$.*

Proof. We observe that

$$\sigma_{\text{eig}}(A) \setminus \{0\} = \bigcup_{n=1}^{\infty} \{\lambda \in \sigma_{\text{eig}}(A) : |\lambda| \geq 1/n\}.$$

Hence, it is enough to show that

$$E_r := \{\lambda \in \sigma_{\text{eig}}(A) : |\lambda| \geq r\}$$

is a finite set for each $r > 0$.

Suppose there is an $r > 0$ such that E_r is an infinite set. Let (λ_n) be a sequence of distinct elements in E_r, i.e., (λ_n) is a sequence of distinct eigenvalues of A such that $|\lambda_n| \geq r$. For $n \in \mathbb{N}$, let x_n be an eigenvector of A corresponding to the eigenvalue λ_n, and let

$$X_n = \operatorname{span}\{x_1, \ldots, x_n\}, \qquad n \in \mathbb{N}.$$

Then each X_n is a proper closed subspace of X_{n+1} so that by the Riesz lemma (Theorem 2.40), there exists a sequence (u_n) in X such that $u_n \in X_n$, $\|u_n\| = 1$ for all $n \in \mathbb{N}$, and

$$\operatorname{dist}(u_n, X_m) \geq \frac{1}{2} \quad \forall m < n.$$

Therefore, for every $m, n \in \mathbb{N}$ with $m < n$, we have

$$\begin{aligned} \|Au_n - Au_m\| &= \|(A - \lambda_n I)u_n - (A - \lambda_m I)u_m + \lambda_n u_n - \lambda_m u_m\| \\ &= \|\lambda_n u_n - [\lambda_m u_m + (A - \lambda_m I)u_m - (A - \lambda_n I)u_n]\|. \end{aligned}$$

Note that $u_m \in X_m \subseteq X_{n-1}$ and (*Verify*)

$$(A - \lambda_n I)u_n \in X_{n-1}, \quad (A - \lambda_m I)u_m \in X_{m-1} \subseteq X_{n-1}.$$

Hence we have

$$\|Au_n - Au_m\| \geq |\lambda_n| \operatorname{dist}(u_n, X_{n-1}) \geq \frac{|\lambda_n|}{2} \geq \frac{r}{2}.$$

Thus, (Au_n) has no convergent subsequence, contradicting the fact that A is a compact operator. ∎

Let X be a normed linear space and $A : X \to X$ be a linear operator. Suppose that λ is not an eigenvalue of A. Then we can say that for $y \in X$, the operator equation

$$Ax - \lambda x = y$$

can have atmost one solution. In examples which occur in applications, one would like to know, not only that the above equation has a unique solution, but also that the solution depends continuously on the 'data' y. In other words, one would like to know that the inverse operator

$$(A - \lambda I)^{-1} : R(A - \lambda I) \to X$$

is continuous. This is equivalent to saying that the operator $A - \lambda I$ is bounded below, i.e., there exists $c > 0$ such that

$$\|Ax - \lambda x\| \geq c\|x\| \quad \forall x \in X.$$

Motivated by this requirement, we generalize the concept of eigenspectrum.

Let X be a normed linear space and $A : X \to X$ be a linear operator. A scalar λ is called an **approximate eigenvalue** of A if $A - \lambda I$ is not bounded below. The set of all approximate eigenvalues of A is called the **approximate eigenspectrum** of A, and we denote it by $\sigma_{\rm app}(A)$.

Thus, the approximate eigenspectrum of A is

$$\sigma_{\rm app}(A) = \{\lambda \in \mathbb{K} : A - \lambda I \text{ not bounded below}\},$$

and by Proposition 3.2, $\lambda \notin \sigma_{\rm app}(A)$ if and only if $A - \lambda I$ is injective and $(A - \lambda I)^{-1} : R(A - \lambda I) \to X$ is continuous.

The following characterization of $\sigma_{\rm app}(A)$ best explains the suitability of the terminology *approximate eigenvalue*.

Theorem 10.2 *Let X be a normed linear space, $A : X \to X$ be a linear operator and $\lambda \in \mathbb{K}$. Then $\lambda \in \sigma_{\rm app}(A)$ if and only if there exists a sequence (x_n) in X such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$, and*

$$\|Ax_n - \lambda x_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

Proof. Clearly, if $\lambda \notin \sigma_{\rm app}(A)$, i.e., if there exists $c > 0$ such that $\|Ax - \lambda x\| \geq c\|x\|$ for every $x \in X$, then there cannot exist (x_n) in X such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\|Ax_n - \lambda x_n\| \to 0$ as $n \to \infty$.

Conversely, suppose $\lambda \in \sigma_{\rm app}(A)$, i.e., there does not exist $c > 0$ such that $\|Ax - \lambda x\| \geq c\|x\|$ for all $x \in X$. Then, for every $n \in \mathbb{N}$, there exists $u_n \in X$ such that

$$\|Au_n - \lambda u_n\| < \frac{1}{n}\|u_n\| \quad \forall n \in \mathbb{N}.$$

Clearly, $u_n \neq 0$ for all $n \in \mathbb{N}$. Taking $x_n = u_n/\|u_n\|$, $n \in \mathbb{N}$, we have

$$\|x_n\| = 1 \quad \forall n \in \mathbb{N} \quad \text{and} \quad \|Ax_n - \lambda x_n\| < \frac{1}{n} \to 0.$$

This completes the proof. ∎

Theorem 10.3 *Let X be a normed linear space and $A : X \to X$ be a linear operator. Then,*

$$\sigma_{\rm eig}(A) \subseteq \sigma_{\rm app}(A).$$

If the space X is finite dimensional, then

$$\sigma_{\text{eig}}(A) = \sigma_{\text{app}}(A).$$

Proof. Clearly, $\lambda \notin \sigma_{\text{app}}(A)$ implies $A - \lambda I$ is injective so that $\lambda \notin \sigma_{\text{eig}}(A)$. Thus, $\sigma_{\text{eig}}(A) \subseteq \sigma_{\text{app}}(A)$. Now, assume that X is a finite dimensional space. If $\lambda \notin \sigma_{\text{eig}}(A)$, then $A - \lambda I$ is injective so that using the finite dimensionality of X, it follows that $A - \lambda I$ is surjective as well, and hence the operator $(A - \lambda I)^{-1}$ is continuous. Consequently, $A - \lambda I$ is bounded below, i.e., $\lambda \notin \sigma_{\text{app}}(A)$. Thus, if X is finite dimensional, then $\sigma_{\text{eig}}(A) = \sigma_{\text{app}}(A)$. ∎

Strict inclusion in $\sigma_{\text{eig}}(A) \subsetneq \sigma_{\text{app}}(A)$ can occur if the space X is infinite dimensional, as the following example shows.

EXAMPLE 10.2 Let X be any of the sequence spaces c_{00}, c_0, c, ℓ^p with any norm satisfying $\|e_n\| = 1$ for all $n \in \mathbb{N}$. Let $A : X \to X$ be defined by

$$(Ax)(j) = \lambda_j x(j) \quad \forall x \in X;\ j \in \mathbb{N},$$

where (λ_n) is a bounded sequence of scalars. We have already observed in Example 10.1(i) that

$$\sigma_{\text{eig}}(A) = \{\lambda_1, \lambda_2, \ldots\}.$$

Next, suppose that $\lambda_n \to \lambda$ as $n \to \infty$. Then we have

$$\|Ae_n - \lambda e_n\| = |\lambda_n - \lambda|\, \|e_n\| = |\lambda_n - \lambda| \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, we can conclude that $\lambda \in \sigma_{\text{app}}(A)$. Note that, if $\lambda \neq \lambda_n$ for every $n \in \mathbb{N}$, then $\lambda \notin \sigma_{\text{eig}}(A)$.

We see as in Example 9.2(i) that if $\lambda = 0$ in the above example, then A is a compact operator. Also, A is of infinite rank. The property that $0 \in \sigma_{\text{app}}(A)$ is true for all such operators. In fact, we can say something more.

Theorem 10.4 *Let X be a normed linear space and $A : X \to X$ be a compact operator. Then the following results hold:*

(i) $\sigma_{\text{app}}(A) \setminus \{0\} = \sigma_{\text{eig}}(A) \setminus \{0\}$.

(ii) *If A is a finite rank operator, then $\sigma_{\text{app}}(A) = \sigma_{\text{eig}}(A)$.*

(iii) *If X is infinite dimensional, then $0 \in \sigma_{\text{app}}(A)$.*

(iv) 0 *is the only possible limit point of* $\sigma_{\text{app}}(A)$.

Proof. (i) We have already observed that $\sigma_{\mathrm{eig}}(A) \subseteq \sigma_{\mathrm{app}}(A)$. Now, suppose $0 \neq \lambda \in \sigma_{\mathrm{app}}(A)$. We show that $\lambda \in \sigma_{\mathrm{eig}}(A)$.

Let (x_n) be a sequence in X be such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\|Ax_n - \lambda x_n\| \to 0$ as $n \to \infty$. Since A is a compact operator, there exists a subsequence $(\tilde{x}_n)$ of (x_n) and $y \in X$ such that $A\tilde{x}_n \to y$. Hence,

$$\lambda \tilde{x}_n = A\tilde{x}_n - (A\tilde{x}_n - \lambda \tilde{x}_n) \to y.$$

Then it follows that $\|y\| = |\lambda|$ and

$$y = \lim_{n\to\infty} A\tilde{x}_n = A(y/\lambda)$$

so that $Ay = \lambda y$, showing that $\lambda \in \sigma_{\mathrm{eig}}(A)$.

(ii) Suppose A is a finite rank operator. Having proved (i), it is enough to show that $0 \in \sigma_{\mathrm{app}}(A)$ implies $0 \in \sigma_{\mathrm{eig}}(A)$. Suppose $0 \notin \sigma_{\mathrm{eig}}(A)$. Then A is injective so that by the hypothesis that A is of finite rank, X is finite dimensional. Therefore, $\sigma_{\mathrm{app}}(A) = \sigma_{\mathrm{eig}}(A)$, and, consequently, $0 \notin \sigma_{\mathrm{app}}(A)$.

(iii) Let X be infinite dimensional. Suppose $0 \notin \sigma_{\mathrm{app}}(A)$, i.e., A is bounded below. We show that every bounded sequence in X has a Cauchy subsequence so that X would be of finite dimension, contradicting the assumption.

Let (x_n) be a bounded sequence in X. Since A is compact, there is a subsequence $(\tilde{x}_n)$ of (x_n) such that $(A\tilde{x}_n)$ converges. Since A is bounded below, it follows that $(\tilde{x}_n)$ is a Cauchy subsequence of (x_n).

(iv) Proof of this part follows from (i) and Theorem 10.1. ∎

An alternative argument for the proof of (iii). Suppose X is an infinite dimensional space. Assume for the moment that $0 \notin \sigma_{\mathrm{app}}(A)$. Then A is bounded below, injective, and its inverse $A^{-1} : R(A) \to X$ is continuous. Hence, by Theorem 9.8, A is of finite rank, and by the injectivity of A, X is a finite dimensional space. This is contradiction to the assumption that X is infinite dimensional. ∎

Part (iii) of the above Theorem can be used to infer that an operator defined on an infinite dimensional space is not compact. To illustrate this point, let us look at the following examples.

EXAMPLE 10.3 Let $X = \ell^p$ with $1 \le p \le \infty$. Let A be the *right shift operator* on X,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots),$$

or the diagonal operator on X,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (\lambda_1 \alpha_1, \lambda_2 \alpha_2, \ldots)$$

associated with a sequence (λ_n) of nonzero scalars which converges to a nonzero scalar. We showed in Example 9.1 (i) and (iii), that A is not a compact operator. Also, we know that A is bounded below (cf. Exercise 9.2). Hence $0 \notin \sigma_{\mathrm{app}}(A)$. Thus, the fact that A is not compact also follows from Theorem 10.4(iii).

We know that the range of an infinite rank compact operator on a Banach space is not closed. Does Theorem 10.4 (iii) hold for every bounded operator with non-closed range as well? The answer is in the affirmative if X is a Banach space, as the following theorem shows.

Theorem 10.5 *Let X be a Banach space and $A \in \mathcal{B}(X)$. If $R(A)$ is not closed in X, then $0 \in \sigma_{\mathrm{app}}(A)$.*

Proof. The proof follows from Proposition 3.3. ∎

An alternative argument. If $0 \notin \sigma_{\mathrm{app}}(A)$, then by Theorem 10.2, we know that $A^{-1} : R(A) \to X$ is continuous as well as a closed operator. Hence, by Theorem 3.17, $R(A)$ is closed. ∎

Now we prove a topological property of $\sigma_{\mathrm{app}}(A)$.

Theorem 10.6 *Let X be a normed linear space and $A \in \mathcal{B}(X)$. Then $\sigma_{\mathrm{app}}(A)$ is a closed subset of $\mathbb{K}$.*

Proof. Let (λ_n) be a sequence in $\sigma_{\mathrm{app}}(A)$ such that $\lambda_n \to \lambda$ for some $\lambda \in \mathbb{K}$. Suppose $\lambda \notin \sigma_{\mathrm{app}}(A)$. Let $c > 0$ be such that

$$\|Ax - \lambda x\| \ge c\|x\| \quad \forall x \in X.$$

Observe that, for every $x \in X$, $n \in \mathbb{N}$,

$$\begin{aligned} \|Ax - \lambda_n x\| &= \|(Ax - \lambda x) - (\lambda_n - \lambda)x\| \\ &\ge \|Ax - \lambda x\| - |\lambda_n - \lambda| \|x\| \\ &\ge (c - |\lambda_n - \lambda|)\|x\|. \end{aligned}$$

Thus, for all large enough n, $A - \lambda_n I$ is bounded below. More precisely, let $N \in \mathbb{N}$ be such that $|\lambda_n - \lambda| \le c/2$ for all $n \ge N$. Then we have

$$\|Ax - \lambda_n x\| \ge \frac{c}{2}\|x\| \quad \forall x \in X, \forall n \ge N,$$

which shows that $\lambda_n \notin \sigma_{\mathrm{app}}(A)$ for all $n \ge N$. Thus, we arrive at a contradiction. ∎

The above result, in particular, shows that if (λ_n) is a sequence of eigenvalues of $A \in \mathcal{B}(X)$ such that $\lambda_n \to \lambda$, then λ is an approximate eigenvalue. One may ask whether every approximate eigenvalue arises in this manner. The answer is, in general, negative, as the following example shows.

EXAMPLE 10.4 Let $X = \ell^1$ and A be the *right shift operator* on ℓ^1. Then we know (Example 10.1(ii)) that $\sigma_{\mathrm{eig}}(A) = \varnothing$. We show that $\sigma_{\mathrm{app}}(A) \neq \varnothing$.

Let (x_n) in ℓ^1 be defined by

$$x_n(j) = \begin{cases} 1/n & \text{if } j \le n, \\ 0 & \text{if } j > n. \end{cases}$$

Then we see that $\|x_n\| = 1$ for all $n \in \mathbb{N}$, and $\|Ax_n - x_n\|_1 = 2/n \to 0$ as $n \to \infty$ so that $1 \in \sigma_{\mathrm{app}}(A)$.

Before closing this section we give a few examples of operators describing eigenspectrum and approximate eigenspectrum completely.

EXAMPLE 10.5 (i) Let $X = \ell^p$, $1 \le p \le \infty$ and let $A : X \to X$ be the diagonal operator, i.e.,

$$(Ax)(i) = \lambda_i x(i), \qquad x \in X, i \in \mathbb{N},$$

where (λ_n) is a bounded sequence of scalars. We have already seen in Example 10.1(i) that

$$\sigma_{\mathrm{eig}}(A) = \{\lambda_1, \lambda_2, \ldots\}.$$

Let $S = \mathrm{cl}\,\{\lambda_n : n \in \mathbb{N}\}$. Then, by Theorem 10.6, $S \subseteq \sigma_{\mathrm{app}}(A)$. We now show that $S \supseteq \sigma_{\mathrm{app}}(A)$. For this let $\lambda \notin S$, and let $\delta > 0$ be such that $|\lambda - \mu| \ge \delta$ for every $\mu \in S$. Then it is seen that

$$\|Ax - \lambda x\|_p \ge \delta \|x\|_p \quad \forall x \in \ell^p$$

so that $\lambda \notin \sigma_{\text{app}}(A)$. Thus, we have proved that $\sigma_{\text{app}}(A) = S$.

(ii) Let $X = \ell^p$, $1 \le p \le \infty$, and let A be the right shift operator on ℓ^p. We have already seen in Example 10.1(ii) that

$$\sigma_{\text{eig}}(A) = \varnothing.$$

We show that

$$\sigma_{\text{app}}(A) = \{\lambda \in \mathbb{K} : |\lambda| = 1\}.$$

Since $\|Ax\|_p = \|x\|_p$ for every $x \in X$, it follows that, for every $\lambda \in \mathbb{K}$,

$$\|Ax - \lambda x\|_p \ge \big|\|Ax\|_p - \|\lambda x\|_p\big| = |1 - |\lambda|\,|\,\|x\|_p.$$

Thus, $A - \lambda I$ is bounded below for every $\lambda \in \mathbb{K}$ with $|\lambda| \ne 1$, and hence

$$\sigma_{\text{app}}(A) \subseteq \{\lambda \in \mathbb{K} : |\lambda| = 1\}.$$

To see the other way inclusion, first let $1 \le p < \infty$ and let $\lambda \in \mathbb{K}$ be such that $|\lambda| = 1$. Consider the sequence (x_n) in X defined by

$$x_n(j) = \begin{cases} 1/n^{1/p}\lambda^{j-1} & \text{if } j \le n, \\ 0 & \text{if } j > n. \end{cases}$$

Then it is seen that $\|x_n\|_p = 1$ for every $n \in \mathbb{N}$, and

$$\|Ax_n - \lambda x_n\|_p^p = 2/n \to 0 \quad \text{as} \quad n \to \infty.$$

Hence $\lambda \in \sigma_{\text{app}}(A)$. Thus, we have proved that for $1 \le p < \infty$,

$$\{\lambda \in \mathbb{K} : |\lambda| = 1\} = \sigma_{\text{app}}(A).$$

The case of $p = \infty$ will be considered in Example 10.7(ii).

(iii) Let $X = \ell^p$, $1 \le p \le \infty$, and A be the *left shift operator* on ℓ^p, i.e.,

$$(Ax)(i) = x(i+1), \qquad x \in X, \quad i \in \mathbb{N},$$

We have seen in Example 10.1(iii) that

$$\sigma_{\text{eig}}(A) = \begin{cases} \{\lambda : |\lambda| < 1\} & \text{if } 1 \le p < \infty \\ \{\lambda : |\lambda| \le 1\} & \text{if } p = \infty. \end{cases}$$

Since $\sigma_{\text{app}}(A)$ is a closed set (Theorem 10.2 (ii)), we have

$$\{\lambda : |\lambda| \le 1\} \subseteq \sigma_{\text{app}}(A).$$

Now, since the relation $\|Ax\|_p \leq \|x\|_p$ holds for all $x \in \ell^p$, we have

$$\|Ax - \lambda x\|_p \geq \|\lambda x\|_p - \|Ax\|_p \geq (|\lambda| - 1)\|x\|_p$$

for every $\lambda \in \mathbb{K}$ so that $A - \lambda I$ is bounded below for every $\lambda \in \mathbb{K}$ with $|\lambda| > 1$. Hence,

$$\sigma_{\text{app}}(A) \subseteq \{\lambda \in \mathbb{K} : |\lambda| \leq 1\}.$$

Thus, we have proved that

$$\sigma_{\text{app}}(A) = \{\lambda \in \mathbb{K} : |\lambda| \leq 1\}.$$

(iv) Let $X = C[a, b]$ with $\|\cdot\|_\infty$, and $A : X \to X$ be defined by

$$(Ax)(t) = tx(t) \quad \forall x \in X;\ t \in [a, b].$$

Clearly, $A \in \mathcal{B}(X)$. We show that

$$\sigma_{\text{eig}}(A) = \varnothing, \qquad \sigma_{\text{app}}(A) = [a, b].$$

If $\lambda \in \mathbb{K}$ and $x \in X$ such that $Ax = \lambda x$, then we have $(t - \lambda)x(t) = 0$ for all $t \in [a, b]$. Then we have $x = 0$; otherwise $x(t) \neq 0$ for t in an interval $J \subseteq [a, b]$, leading to $\lambda = t$ for every $t \in J$. This is absurd. Thus, we have proved that $\sigma_{\text{eig}}(A) = \varnothing$.

Now, suppose that $\lambda \notin [a, b]$ and let $d_\lambda = \inf\{|t - \lambda| : t \in [a, b]\}$. Then, for every $x \in X$, we have

$$|(Ax - \lambda x)(t)| = |(t - \lambda)x(t)| \geq d_\lambda |x(t)| \quad \forall t \in [a, b]$$

so that

$$\|Ax - \lambda x\|_\infty \geq d_\lambda \|x\|_\infty \quad \forall x \in X.$$

Hence $\lambda \notin \sigma_{\text{app}}(A)$. Thus,

$$\sigma_{\text{app}}(A) \subseteq [a, b].$$

Now we show that $[a, b] \subseteq \sigma_{\text{app}}(A)$. Since $\sigma_{\text{app}}(A)$ is a closed set, it is enough to show that $(a, b) \subseteq \sigma_{\text{app}}(A)$. Let $\lambda \in (a, b)$, and let $N \in \mathbb{N}$ be such that $(\lambda - 1/n, \lambda + 1/n) \subset (a, b)$ for every $n \geq N$. For $t \in [a, b]$ and $n \geq N$, let

$$x_n(t) = \begin{cases} n(t - \lambda + 1/n) & \text{if } \lambda - 1/n < t \leq \lambda, \\ -n(t - \lambda - 1/n) & \text{if } \lambda \leq t < \lambda + 1/n, \\ 0 & \text{if } |\lambda - t| \geq 1/n. \end{cases}$$

Then it follows that $x_n \in X$, $\|x_n\|_\infty = 1$ for all $n \in \mathbb{N}$ and

$$\|Ax_n - \lambda x_n\|_\infty = \sup_{|\lambda - t| \le 1/n} |t - \lambda|\, |x_n(t)| \le \frac{1}{n} \to 0.$$

Thus, $\lambda \in \sigma_{\mathrm{app}}(A)$.

10.2 Spectrum and Resolvent Set

Let X be a normed linear space and $A : X \to X$ be a linear operator. Recall that a scalar λ is not an approximate eigenvalue of A if and only if for every $y \in R(A - \lambda I)$, there exists a unique $x \in X$ such that

$$Ax - \lambda x = y,$$

and the map $y \mapsto x$ is continuous. Thus, if x and y are as above, and if (y_n) is a sequence in $R(A - \lambda I)$ such that $y_n \to y$, and (x_n) in X satisfy $Ax_n - \lambda x_n = y_n$, then $x_n \to x$.

In applications, one would like to have the above situation not only for every $y \in R(A - \lambda I)$, but also for every $y \in Y$. Motivated by this requirement, we generalize the concept of approximate eigenspectrum to what is known as the *spectrum* of A.

We define the **resolvent set** of A to be the set

$$\rho(A) = \{\lambda \in \mathbb{K} : A - \lambda I \text{ is bijective and } (A - \lambda I)^{-1} \in \mathcal{B}(X)\}.$$

The complement of $\rho(A)$ in $\mathbb{K}$ is called the **spectrum** of A and is denoted by $\sigma(A)$. Thus, $\lambda \in \sigma(A)$ if and only if either $A - \lambda I$ is not bijective or else $(A - \lambda I)^{-1} \notin \mathcal{B}(X)$. The elements of the spectrum are called the **spectral values** of A.

We observe that, for $A \in \mathcal{B}(X)$,

$$0 \in \rho(A) \iff \exists B \in \mathcal{B}(X) \quad \text{such that} \quad AB = I = BA,$$

and, in that case, $B = A^{-1}$. If $0 \in \rho(A)$, then we say that A is **invertible** in $\mathcal{B}(X)$. We note that if $A, B \in \mathcal{B}(X)$ are invertible, then AB is invertible, and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Exercise 10.1 Give an example of a Banach space X, operators $A, B \in \mathcal{B}(X)$ such that AB invertible but at least one of A and B is not invertible. Also, show that if X is finite dimensional, then for every $A, B \in \mathcal{B}(X)$, AB invertible implies both A and B are invertible. □

In view of Proposition 3.2, if $\lambda \in \rho(A)$, then $A - \lambda I$ is bounded below. Hence, every approximate eigenvalue is a spectral value, i.e.,

$$\sigma_{\text{app}}(A) \subseteq \sigma(A).$$

Clearly, if X is a finite dimensional space, then

$$\sigma_{\text{eig}}(A) = \sigma_{\text{app}}(A) = \sigma(A).$$

We have seen examples of infinite rank operators A for which $\sigma_{\text{eig}}(A) \neq \sigma_{\text{app}}(A)$. The following example shows that strict inclusion is possible in $\sigma_{\text{app}}(A) \subseteq \sigma(A)$ as well.

EXAMPLE 10.6 Let $X = \ell^p$, $1 \leq p \leq \infty$, and A be the right shift operator on X. We have seen in Example 10.3 that $0 \notin \sigma_{\text{app}}(A)$. But $0 \in \sigma(A)$, since A is not onto. In fact, $e_1 \notin R(A)$.

A characterization of the spectrum

Theorem 10.7 *Suppose X is a Banach space, $A \in \mathcal{B}(X)$ and $\lambda \in \mathbb{K}$. Then $\lambda \in \sigma(A)$ if and only if either $\lambda \in \sigma_{\text{app}}(A)$ or $R(A - \lambda I)$ is not dense in X.*

Proof. Clearly, if $\lambda \in \sigma_{\text{app}}(A)$ or $R(A - \lambda I)$ is not dense in X, then $\lambda \in \sigma(A)$.

Conversely, suppose that $\lambda \in \sigma(A)$. If $\lambda \notin \sigma_{\text{app}}(A)$, then by Propositions 3.2 and 3.3, the operator $A - \lambda I$ is injective, and its inverse $(A - \lambda I)^{-1} : R(A - \lambda I) \to X$ is continuous, and $R(A - \lambda I)$ is closed. Hence, $R(A - \lambda I)$ is not dense in X; otherwise, $A - \lambda I$ would become bijective and $(A - \lambda I)^{-1} \in \mathcal{B}(X)$, which is a contradiction to the assumption that $\lambda \in \sigma(A)$. ∎

For $A \in \mathcal{B}(X)$, the set

$$\sigma_{\text{com}}(A) := \{\lambda \in \mathbb{K} : R(A - \lambda I) \text{ not dense in } X\}$$

is called the **compression spectrum** of A. In view of the above theorem, if X is a Banach space, then we have

$$\sigma(A) = \sigma_{\text{app}}(A) \cup \sigma_{\text{com}}(A).$$

If X is a Banach space, then we know, by bounded inverse theorem (Theorem 7.4), that the continuity of the inverse of an operator is a consequence of its bijectivity. Thus, if X is a Banach space and $A \in \mathcal{B}(X)$, then, in the definition of $\rho(A)$, the requirement that $(A - \lambda I)^{-1}$ is continuous is redundant. Therefore,

$$\begin{aligned}
\rho(A) &= \{\lambda \in \mathbb{K} : A - \lambda I \text{ is bijective}\}, \\
\sigma(A) &= \{\lambda \in \mathbb{K} : A - \lambda I \text{ is either not one-one or not onto}\}.
\end{aligned}$$

Now we prove some topological properties of the spectrum, again for the case when X is a Banach space. First, let us prove the following theorem.

Theorem 10.8 *Let X be a Banach space and $A \in \mathcal{B}(X)$. If $\|A\| < 1$, then $I - A$ is invertible and*

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

Proof. Suppose $\|A\| < 1$. In order to show that $I - A$ is invertible, it is enough to show that $I - A$ is bijective. Note that for every $x \in X$,

$$\|(I - A)x\| \geq (1 - \|A\|)\|x\| \quad \forall x \in X \tag{10.1}$$

so that $I - A$ is bounded below. Therefore, $I - A$ is injective and $R(I - A)$ is closed in X. If $R(I - A)$ is not the whole of X, then by a consequence (Corollary 5.5) of the Hahn-Banach Theorem, there exists a nonzero $f \in X'$ such that $f(y) = 0$ for all $y \in \overline{R(I - A)}$. In particular, $f(x - Ax) = 0$ for all $x \in X$, that is,

$$f(x) = f(Ax) \quad \forall x \in X.$$

Thus,

$$|f(x)| \leq \|f\| \, \|A\| \, \|x\| \quad \forall x \in X$$

so that

$$\|f\| \leq \|A\| \, \|f\|.$$

This is a contradiction since $\|f\| \neq 0$ and $\|A\| < 1$. Hence, $R(I - A) = X$. Now, by Proposition 3.2, relation (10.1) implies that

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

This completes the proof. ∎

Corollary 10.9 *Let X be a Banach space and $A \in \mathcal{B}(X)$.*

(i) *If $\|A^n\| < 1$ for some $n \in \mathbb{N}$, then $I - A$ is invertible, and*

$$\|(I-A)^{-1}\| \leq \frac{\|I + A + A^2 + \cdots + A^{n-1}\|}{1 - \|A^n\|}.$$

(ii) *If $\|A\| < 1$, then*

$$\|(I-A)^{-1} - (I + A + A^2 + \cdots + A^{n-1})\| \leq \frac{\|A^n\|}{1 - \|A\|}$$

for every $n \in \mathbb{N}$, and the series $\sum_{k=0}^{\infty} A^k$ converges with

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

Proof. Note that for every $n \in \mathbb{N}$,

$$\begin{aligned} I - A^n &= (I-A)(I + A + A^2 + \cdots + A^{n-1}) \\ &= (I + A + A^2 + \cdots + A^{n-1})(I-A). \end{aligned}$$

(i) Suppose that $\|A^n\| < 1$ for some $n \in \mathbb{N}$. Then, by Theorem 10.8, $I - A^n$ is invertible in $\mathcal{B}(X)$. In particular, $I - A^n$ is bijective. Hence, from the above identity, it follows that $I - A$ is also bijective, and hence it is invertible in $\mathcal{B}(X)$. It also follows from the above identity that

$$(I-A)^{-1} = (I - A^n)^{-1}(I + A + A^2 + \cdots + A^{n-1}). \tag{10.2}$$

Hence, again by Theorem 10.8,

$$\begin{aligned} \|(I-A)^{-1}\| &\leq \|(I-A^n)^{-1}\| \, \|I + A + A^2 + \cdots + A^{n-1}\| \\ &\leq \frac{\|I + A + A^2 + \cdots + A^{n-1}\|}{1 - \|A^n\|}. \end{aligned}$$

(ii) Now, suppose that $\|A\| < 1$. From relation (10.2), we have

$$(I-A)^{-1}(I - A^n) = \sum_{k=0}^{n-1} A^k.$$

Hence,

$$(I-A)^{-1} - \sum_{k=0}^{n-1} A^k = (I-A)^{-1}A^n.$$

Since $\|A\| < 1$, by Theorem 10.8, we have $\|(I-A)^{-1}\| \leq \frac{1}{1-\|A\|}$. Therefore,

$$\begin{aligned}\|(I-A)^{-1} - \sum_{k=0}^{n-1} A^k\| &\leq \|(I-A)^{-1}\|\|A^n\| \\ &\leq \frac{\|A^n\|}{1-\|A\|}.\end{aligned}$$

Again, since Since $\|A\| < 1$, we have $\|A^n\| \leq \|A\|^n \to 0$ as $n \to \infty$ so that it follows from the above inequality that the series $\sum_{k=0}^{\infty} A^k$ converges, and the limit is the operator $(I-A)^{-1}$. This completes the proof. ∎

Theorem 10.10 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then the following results hold.*

(i) *If $\lambda \in \mathbb{K}$ such that $|\lambda| > \|A\|$, then $\lambda \in \rho(A)$ and*

$$\|(A-\lambda I)^{-1}\| \leq \frac{1}{|\lambda| - \|A\|}.$$

In particular,

$$\sigma(A) \subseteq \{\lambda : |\lambda| \leq \|A\|\}$$

and

$$\|(A-\lambda I)^{-1}\| \to 0 \quad as \quad |\lambda| \to \infty.$$

(ii) *For every $\lambda \in \rho(A)$,*

$$\left\{\mu \in \mathbb{K} : |\mu - \lambda| < \frac{1}{\|(A-\lambda I)^{-1}\|}\right\} \subseteq \rho(A).$$

In particular, $\rho(A)$ is an open subset of $\mathbb{K}$.

Proof. (i) Let $\lambda \in \mathbb{K}$ satisfies $\|A\| < |\lambda|$. Then, by Theorem 10.8, $I - A/\lambda$ is invertible in $\mathcal{B}(X)$ and

$$\|(I - A/\lambda)^{-1}\| \leq \frac{1}{1 - \|A/\lambda\|}.$$

Thus, $\lambda \in \rho(A)$ and

$$\|(A-\lambda I)^{-1}\| \leq \frac{1}{|\lambda| - \|A\|}.$$

From this, it also follows that

$$\sigma(A) \subseteq \{\lambda \in \mathbb{K} : |\lambda| \leq A\|\},$$

$$\|(A - \lambda I)^{-1}\| \to 0 \quad \text{as} \quad |\lambda| \to \infty.$$

(ii) Let $\lambda \in \rho(A)$. Then for every $\mu \in \mathbb{K}$, we have

$$A - \mu I = [I - (\mu - \lambda)(A - \lambda I)^{-1}](A - \lambda I).$$

Hence, by Theorem 10.8, if μ is such that $|\mu - \lambda| < 1/\|(A - \lambda I)^{-1}\|$, then $I - (\mu - \lambda)(A - \lambda I)^{-1}$ is invertible in $\mathcal{B}(X)$, and consequently, from the above identity, $A - \mu I$ is also invertible in $\mathcal{B}(X)$. Thus,

$$\left\{\mu : |\mu - \lambda| < \frac{1}{\|(A - \lambda I)^{-1}\|}\right\} \subseteq \rho(A).$$

In particular, $\rho(A)$ is an open subset of $\mathbb{K}$. ∎

Theorem 10.11 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then, $\sigma(A)$ is a compact subset of $\mathbb{K}$.*

Proof. By the particular cases in (i) and (ii) of Theorem 10.10, we see that $\sigma(A)$ is a closed and bounded subset of $\mathbb{K}$ so that it is compact in $\mathbb{K}$. ∎

Theorem 10.12 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then every boundary point of $\sigma(A)$ is an approximate eigenvalue of A.*

Proof. Let λ be a boundary point of $\sigma(A)$. Then there exists a sequence (μ_n) in $\rho(A)$ such that $\mu_n \to \lambda$ as $n \to \infty$. Suppose $\lambda \notin \sigma_{\text{app}}(A)$, and let $c > 0$ be such that

$$\|Ax - \lambda x\| \geq c\|x\| \quad \forall\, x \in X.$$

We shall arrive at a contradiction once we show that $\lambda \in \rho(A)$, because we already know, by Theorem 10.11, that $\lambda \in \sigma(A)$.

We note that

$$A - \lambda I = [I - (\lambda - \mu_n)(A - \mu_n)^{-1}](A - \mu_n I) \tag{10.3}$$

for all $n \in \mathbb{N}$, and

$$\begin{aligned} \|Ax - \mu_n x\| &= \|(Ax - \lambda x) - (\mu_n - \lambda)x\| \\ &\geq \|Ax - \lambda x\| - |\mu_n - \lambda|\|x\| \\ &\geq (c - |\mu_n - \lambda|)\|x\| \end{aligned}$$

for all $n \in \mathbb{N}$ and for all $x \in X$. Since $\mu_n \to \lambda$, there exists N such that $|\mu_n - \lambda| < c$ for all $n \geq N$. Hence, by Proposition 3.2,

$$\|(A - \mu_n I)^{-1}\| \leq \frac{1}{c - |\mu_n - \lambda|} \quad \forall n \geq N$$

so that

$$\|(\lambda - \mu_n)(A - \mu_n I)^{-1}\| \leq \frac{|\lambda - \mu_n|}{c - |\mu_n - \lambda|} \quad \forall n \geq N.$$

From this, we see that

$$\|(\lambda - \mu_n)(A - \mu_n I)^{-1}\| < 1 \quad \text{whenever} \quad |\lambda - \mu_n| < \frac{c}{2}.$$

Suppose $N_1 \in \mathbb{N}$ be such that $|\lambda - \mu_n| < c/2$ for all $n \geq N_1$. Thus, in view of Theorem 10.8, it follows from (10.3) that $\lambda \in \rho(A)$. This completes the proof. ∎

Before studying further, let us find the spectrum of some familiar operators.

EXAMPLE 10.7 (i) Let $X = \ell^p$, $1 \leq p \leq \infty$, and let $A : \ell^p \to \ell^p$ be the diagonal operator defined by

$$(Ax)(i) = \lambda_i x(i), \qquad x \in X,\ i \in \mathbb{N},$$

where (λ_n) is a bounded sequence of scalars. We have already seen in Example 10.5(i) that

$$\sigma_{\text{eig}}(A) = \{\lambda_n : n \in \mathbb{N}\}, \qquad \sigma_{\text{app}}(A) = S := \text{cl}\,\{\lambda_n : n \in \mathbb{N}\}.$$

We show that $\sigma(A) = S$. It is enough to show that, if $\lambda \notin S$, then $\lambda \notin \sigma(A)$. Now, suppose $\lambda \notin S$ and let $\delta > 0$ be such that $|\lambda - \mu| \geq \delta$ for every $\mu \in S$. Since $\lambda \notin S = \sigma_{\text{app}}(A)$, $A - \lambda I$ is injective. To see that $A - \lambda I$ is surjective, let $y = (\beta_1, \beta_2, \ldots) \in X$. Then we see that

$$x = (\alpha_1, \alpha_2, \ldots) \quad \text{with} \quad \alpha_j = -\frac{\beta_j}{\lambda - \lambda_j}, \ j \in \mathbb{N},$$

is a solution of the equation $Ax - \lambda x = y$. In fact, $\|x\|_p \leq \|y\|_p/\delta$.

(ii) Let $X = \ell^p$, $1 \leq p \leq \infty$, and let A be the *right shift operator* on X, i.e.,

$$A : x = (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots) \quad \forall x \in X.$$

We have already seen in Example 10.5(ii) that

$$\sigma_{\rm eig}(A) = \varnothing, \qquad \sigma_{\rm app}(A) \subseteq \{\lambda \in \mathbb{K} : |\lambda| = 1\},$$

and for $1 \le p < \infty$,

$$\sigma_{\rm app}(A) = \{\lambda \in \mathbb{K} : |\lambda| = 1\}.$$

Now, we show that

$$\sigma(A) = \{\lambda : |\lambda| \le 1\}.$$

Once we show this, it follows from Theorem 10.12 that

$$\sigma_{\rm app}(A) = \{\lambda \in \mathbb{K} : |\lambda| = 1\}$$

for $p = \infty$ as well.

Note that $e_1 \notin R(A)$, so that $0 \in \sigma(A)$. If $\lambda \neq 0$, then we see that $x := (\alpha_1, \alpha_2, \ldots)$ satisfies the equation $Ax - \lambda x = e_1$ if and only if

$$\alpha_n = -\frac{1}{\lambda^n} \quad \forall n \in \mathbb{N}.$$

Therefore, if $|\lambda| < 1$, then $e_1 \notin R(A - \lambda I)$, and hence

$$\{\lambda : |\lambda| < 1\} \subseteq \sigma(A).$$

Also, since $\|A\| = 1$, by Theorem 10.10(i),

$$\sigma(A) \subseteq \{\lambda : |\lambda| \le 1\}.$$

Thus,

$$\{\lambda : |\lambda| < 1\} \subseteq \sigma(A) \subseteq \{\lambda : |\lambda| \le 1\}.$$

Hence, using the fact that $\sigma(A)$ is closed (Theorem 10.11), it follows that

$$\sigma(A) = \{\lambda : |\lambda| \le 1\}.$$

(iii) Let $X = \ell^p$, $1 \le p < \infty$, and let A be the *left shift operator* on X, i.e.,

$$A : x = (\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots), \quad \forall x \in X.$$

We have already seen in Example 10.5(iii) that

$$\sigma_{\rm app}(A) = \{\lambda \in \mathbb{K} : |\lambda| \le 1\}.$$

Hence, the equality

$$\sigma(A) = \{\lambda \in \mathbb{K} : |\lambda| \leq 1\}$$

follows since $\sigma(A) \subseteq \{\lambda \in \mathbb{K} : |\lambda| \leq \|A\|\}$ (cf. Theorem 10.10(i)) and $\|A\| = 1$.

(iv) Let $X = C[a, b]$, with $\|\cdot\|_\infty$, and let $A : X \to X$ be defined by

$$(Ax)(t) = tx(t) \quad \forall x \in X;\ t \in [a, b].$$

We have seen in Example 10.5(iv) that

$$\sigma_{\text{eig}}(A) = \varnothing, \qquad \sigma_{\text{app}}(A) = [a, b].$$

In particular, $[a, b] \subseteq \sigma(A)$. This is also seen by noting that if $y_0(t) = 1$ for all $t \in [a, b]$, then there is no $x \in X$ such that $Ax - \lambda x = y_0$ for any $\lambda \in [a, b]$, since $(Ax - \lambda x)(\lambda) = 0$, whereas $y_0(\lambda) = 1$. Thus, $A - \lambda I$ is not onto for any $\lambda \in [a, b]$.

Now we show that $\sigma(A) \subseteq [a, b]$. For this, suppose $\lambda \notin [a, b]$. We have to show that $\lambda \notin \sigma(A)$. We know that $A - \lambda I$ is injective. Since for every $y \in X$, the function x defined by $x(t) = y(t)/(t - \lambda)$ belongs to X, and satisfies $Ax - \lambda x = y$, $A - \lambda I$ is surjective as well. Therefore, $\lambda \notin \sigma(A)$.

10.2.1 Spectral Radius

Let X be a Banach space and $A \in \mathcal{B}(X)$. Then by Theorem 10.11, we know that the spectrum of A is a compact subset of $\mathbb{K}$. In particular,

$$r_\sigma(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\} < \infty.$$

The quantity $r_\sigma(A)$ is called **the spectral radius of** A.

By Theorem 10.10 (i), we have

$$r_\sigma(A) \leq \|A\|.$$

By convention, if an operator has no spectral values, then we take the spectral radius to be zero.

One may ask the questions: Is the estimate $r_\sigma(A) \leq \|A\|$ sharp? Can strict inequality occur?

The answers to both the above questions are in the affirmative. To see the first case, consider A to be the identity operator on X. Then we have

$$\sigma(A) = \{1\}, \qquad r_\sigma(A) = 1 = \|A\|.$$

For the second question, consider $X = \mathbb{K}^2$ with any norm, and

$$A : (\alpha_1, \alpha_2) \to (\alpha_2, 0).$$

Then we see that

$$\sigma(A) = \{0\}, \quad r_\sigma(A) = 0, \qquad \|A\| \neq 0.$$

Another example of this category is the operator

$$A : (\alpha_1, \alpha_2) \mapsto (\alpha_2, -\alpha_1)$$

on $\mathbb{R}^2$ with any norm on $\mathbb{R}^2$. In this case,

$$\sigma(A) = \sigma_{\text{eig}}(A) = \varnothing, \qquad r_\sigma(A) = 0, \qquad \|A\| \neq 0.$$

The following theorem provides, in some cases, a better estimate for the spectral radius.

Theorem 10.13 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then*

(i) $\{\lambda^k : \lambda \in \sigma(A)\} \subseteq \sigma(A^k) \quad \forall k \in \mathbb{N}$,

(ii) $r_\sigma(A) \leq \inf\{\|A^k\|^{1/k} : k \in \mathbb{N}\}$.

Proof. Let $k \in \mathbb{N}$ and $\lambda \in \mathbb{K}$. Then we have

$$\begin{aligned} A^k - \lambda^k I &= (A - \lambda I)(A^{k-1} + \lambda A^{k-2} + \cdots + \lambda^{k-2} A + \lambda^{k-1} I) \\ &= (A^{k-1} + \lambda A^{k-2} + \cdots + \lambda^{k-2} A + \lambda^{k-1} I)(A - \lambda I). \end{aligned}$$

From this it follows that if $A^k - \lambda^k I$ is bijective, then $A - \lambda I$ is bijective. Thus,

$$\lambda \in \sigma(A) \implies \lambda^k \in \sigma((A)^k)$$

proving (i). From (i), we obtain

$$|\lambda|^k \leq r_\sigma(A^k) \leq \|A^k\| \quad \forall \lambda \in \sigma(A)$$

for all $k \in \mathbb{N}$ showing that

$$r_\sigma(A) \leq \inf\{\|A^k\|^{1/k} : k \in \mathbb{N}\}.$$

Thus, the proof of (ii) is complete. ∎

The following example shows that the estimate for $r_\sigma(A)$ given in the above theorem is better than the estimate $r_\sigma(A) \le \|A\|$.

EXAMPLE 10.8 Let $X = \mathbb{K}^2$ with any norm, and $A : X \to X$ be defined by

$$Ax = (\alpha_2, 0), \qquad x = (\alpha_1, \alpha_2) \in X.$$

Then, we have

$$\sigma(A) = \{0\}, \qquad r_\sigma(A) = 0, \qquad \|A\| \ne 0.$$

Note that $A^2 = 0$, so that $\inf\{\|A^k\|^{1/k} : k \in \mathbb{N}\} = 0$.

In the following example, we discuss a case wherein strict inclusion and strict inequality can occur in Theorem 10.13.

EXAMPLE 10.9 Let $X = \mathbb{R}^3$ with norm $\|\cdot\|_p$ and $A : X \to X$ be defined by

$$Ax = (\alpha_2, -\alpha_1, 0), \qquad x = (\alpha_1, \alpha_2, \alpha_3) \in X.$$

It is seen that 0 is the only spectral value of A so that $r_\sigma(A) = 0$, whereas

$$A^{4k+1} = A, \qquad A^{4k+2} = -B, \qquad A^{4k+3} = -A, \qquad A^{4k+4} = B,$$

where $B : (\alpha_1, \alpha_2, \alpha_3) \mapsto (\alpha_1, \alpha_2, 0)$. Thus, we have $\|A\| = 1 = \|B\|$,

$$\sigma(A^{4k+1}) = \{0\} = \sigma(A^{4k+3}),$$

$$\sigma(A^{4k+2}) = \sigma(-B) = \{0, -1\}, \qquad \sigma(A^{4k+4}) = \sigma(B) = \{0, 1\},$$

$$\inf \|A^k\|^{1/k} = 1.$$

We shall show that if the scalar field is the set of complex numbers, then in Theorem 10.13 we do have equalities in place of inclusion and inequality. First we prove a result more general than Theorem 10.13(i).

10.2.2 Spectral Mapping Theorem

Recall from Section 1.2.4 that if $p(t)$ is a polynomial with coefficients in $\mathbb{K}$, say $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ and if $A : X \to X$ is a linear operator, then the operator $p(A) : X \to X$ is defined by

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n.$$

Theorem 10.14 *Let X be a Banach space, $p(t)$ be a polynomial with coefficients in $\mathbb{K}$ and $A \in \mathcal{B}(X)$. Then*

$$\{p(\lambda) : \lambda \in \sigma(A)\} \subseteq \sigma(p(A)).$$

If $\mathbb{K} = \mathbb{C}$, then

$$\{p(\lambda) : \lambda \in \sigma(A)\} = \sigma(p(A)).$$

Proof.

Proof. Let $p(t)$ be a polynomial with coefficients in $\mathbb{K}$ and $A \in \mathcal{B}(X)$, and let $\lambda \in \mathbb{K}$. Since $p(t) - p(\lambda)$ vanishes at λ,

$$p(t) - p(\lambda) = (t - \lambda)q(t) = q(t)(t - \lambda)$$

for some polynomial $q(t)$. Therefore, it is seen that

$$p(A) - p(\lambda)I = (A - \lambda I)q(A) = q(A)(A - \lambda I).$$

From this it follows that if $p(A) - p(\lambda)I$ is bijective, then $A - \lambda I$ is bijective. Thus,

$$\{p(\lambda) : \lambda \in \sigma(A)\} \subseteq \sigma(p(A)).$$

Now, let $\mathbb{K} = \mathbb{C}$ and $\mu \in \sigma(p(A))$. Then there exist $c, t_1, \ldots, t_n$ in $\mathbb{C}$ such that

$$p(t) - \mu = c(t - t_1)(t - t_2)\ldots(t - t_n)$$

so that

$$p(A) - \mu I = c(A - t_1 I)(A - t_2 I)\ldots(A - t_n I).$$

Since $\mu \in \sigma(p(A))$, at least one of $(A - t_1 I), (A - t_2 I), \ldots, (A - t_n I)$ is not bijective, i.e., there exists $j \in \{1, \ldots, n\}$ such that $t_j \in \sigma(A)$. Then, since $p(t_j) - \mu = 0$, we have

$$\mu = p(t_j) \in \{p(\lambda) : \lambda \in \sigma(A)\}.$$

Thus, we also have

$$\sigma(p(A)) \subseteq \{p(\lambda) : \lambda \in \sigma(A)\}.$$

This completes the proof. ∎

The second part of the above theorem is called the **spectral mapping theorem.**

We have seen in Example 10.9 that if

$$A : (\alpha_1, \alpha_2, \alpha_3) \mapsto (\alpha_2, -\alpha_1, 0)$$

on $X = \mathbb{R}^3$, then

$$\sigma(A) = \{0\}, \qquad \sigma(A^{4k+2}) = \{0, -1\}, \qquad \sigma(A^{4k+4}) = \{0, 1\}.$$

This example shows that the conclusion in spectral mapping theorem, i.e., the second part of Theorem 10.14, need not hold if $\mathbb{K} = \mathbb{R}$.

One may ask: For what kind of operators on real Banach spaces do we have equality in the above theorem? We shall show later that if A is a *self-adjoint operator* on a Hilbert space, then equality does hold even for $\mathbb{K} = \mathbb{R}$.

Spectral mapping theorem prompts us to inquire into the nature of the spectrum of the operator A^{-1} whenever it exists as a bounded operator.

Theorem 10.15 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then, for $\mu \in \rho(A)$,*

$$\begin{aligned} \sigma\left((A - \mu I)^{-1}\right) &= \left\{(\lambda - \mu)^{-1} : \lambda \in \sigma(A)\right\}, \\ r_\sigma\left((A - \mu I)^{-1}\right) &= \frac{1}{\operatorname{dist}(\mu, \sigma(A))}. \end{aligned}$$

Proof. For $\mu \in \rho(A)$ and $\lambda \in \mathbb{K}$ with $\lambda \neq \mu$, we note the relation

$$(A - \mu I)^{-1} - (\lambda - \mu)^{-1} I = -(\lambda - \mu)^{-1}(A - \lambda I)(A - \mu I)^{-1}.$$

From this, we get

$$\lambda \in \sigma(A) \iff (\lambda - \mu)^{-1} \in \sigma((A - \mu I)^{-1}),$$

so that the results follow. ∎

10.2.3 More Results Based on Resolvent

We have seen that the spectrum of an operator can be empty if the scalar field is $\mathbb{R}$. We show that if X is a Banach space over the field of complex numbers and $A \in \mathcal{B}(X)$, then the spectrum of A is nonempty. First we establish a preparatory result. For $\lambda \in \rho(A)$, let

$$R(\lambda, A) = (A - \lambda I)^{-1}.$$

The map $\lambda \mapsto R(\lambda, A)$, $\lambda \in \rho(A)$, is called the **resolvent** of A, and the operator $R(\lambda, A)$ is called the **resolvent operator** of A at $\lambda \in \rho(A)$. Once the operator A is fixed, we shall denote the resolvent operator $R(\lambda, A)$ by $R(\lambda)$.

The identity

$$R(\lambda) - R(\mu) = (\lambda - \mu)R(\lambda)R(\mu) \quad \forall\, \lambda, \mu \in \rho(A),$$

called the **resolvent identity**, can be verified easily.

Theorem 10.16 *Let X be a Banach space and $A \in \mathcal{B}(X)$. Then we have the following:*

(i) *The map $\lambda \mapsto R(\lambda)$ is continuous on $\rho(A)$.*

(ii) *For every $\lambda \in \rho(A)$,*

$$\lim_{\mu \to \lambda} \frac{R(\mu) - R(\lambda)}{\mu - \lambda} = [R(\lambda)]^2.$$

(iii) *For every $f \in (\mathcal{B}(X))'$, the map $\phi_f : \rho(A) \to \mathbb{C}$ defined by*

$$\phi_f(\lambda) = f(R(\lambda)), \qquad \lambda \in \rho(A),$$

is differentiable on $\rho(A)$.

Proof. (i) Let $\lambda_0 \in \rho(A)$ and $\lambda \in \mathbb{K}$. Then from the relation

$$\begin{aligned} A - \lambda I &= (A - \lambda_0 I) - (\lambda - \lambda_0)I \\ &= [I - (\lambda - \lambda_0)R(\lambda_0)](A - \lambda_0 I) \end{aligned}$$

and from Theorem 10.10(i), it follows that if $|\lambda - \lambda_0| < 1/\|R(\lambda_0)\|$, then $\lambda \in \rho(A)$ and

$$R(\lambda) = R(\lambda_0)[I - (\lambda - \lambda_0)R(\lambda_0)]^{-1}.$$

From this, again by Theorem 10.10(i), we have

$$\|R(\lambda)\| \le \frac{\|R(\lambda_0)\|}{1 - |\lambda - \lambda_0|\, \|R(\lambda_0)\|}.$$

Hence from the resolvent identity

$$R(\lambda) - R(\lambda_0) = (\lambda - \lambda_0)R(\lambda)R(\lambda_0),$$

we have

$$\|R(\lambda) - R(\lambda_0)\| \le \frac{|\lambda - \lambda_0|\,\|R(\lambda_0)\|}{1 - |\lambda - \lambda_0|\,\|R(\lambda_0)\|}.$$

Thus, $\|R(\lambda) - R(\lambda_0)\| \to 0$ as $\lambda \to \lambda_0$, showing that $\lambda \mapsto R(\lambda)$ is continuous on $\rho(A)$.

(ii) From the resolvent identity

$$R(\mu) - R(\lambda) = (\mu - \lambda)R(\mu)R(\lambda), \qquad \mu, \lambda \in \rho(A),$$

we have

$$\frac{R(\mu) - R(\lambda)}{\mu - \lambda} = R(\mu)R(\lambda) \quad \text{whenever} \quad \mu \ne \lambda.$$

Hence, by continuity of the map $\mu \mapsto R(\mu)$, proved in (i), we have

$$\lim_{\mu \to \lambda} \frac{R(\mu) - R(\lambda)}{\mu - \lambda} = [R(\lambda)]^2.$$

(iii) Let $f \in (\mathcal{B}(X))'$, and let $\lambda_0, \lambda \in \rho(A)$ with $\lambda \ne \lambda_0$. Then from the resolvent identity

$$R(\lambda) - R(\lambda_0) = (\lambda - \lambda_0)R(\lambda)R(\lambda_0)$$

and the result in (i), we have

$$\lim_{\lambda \to \lambda_0} \frac{\phi_f(\lambda) - \phi_f(\lambda_0)}{\lambda - \lambda_0} = \lim_{\lambda \to \lambda_0} f\left(R(\lambda)R(\lambda_0)\right) = f\left([R(\lambda_0)]^2\right).$$

Thus, ϕ_f is differentiable on $\rho(A)$. ∎

By the above theorem, we know that, if X is a Banach space over $\mathbb{C}$ and $A \in \mathcal{B}(X)$, then the function

$$\lambda \mapsto f(R(\lambda, A)), \qquad \lambda \in \mathbb{K},$$

is analytic in $\rho(A)$. In fact, we can define the concept of differentiability and analyticity of a Banach space valued function in the same manner as we do for complex-valued functions.

Let X be a Banach space over $\mathbb{C}$, and Ω be an open subset of $\mathbb{C}$. Then a function $F : \Omega \to X$ is said to be **differentiable** at $z_0 \in \Omega$, if

$$\lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0}$$

exists, and in that case, the above limit is called the derivative of F at z_0, denoted by $F'(z_0)$. If the function $F : \Omega \to X$ is differentiable at every point in Ω, then F is said to be **analytic** on Ω.

We may observe, by Theorem 10.16(ii) that, if $\mathbb{K} = \mathbb{C}$, then the resolvent map $R(\cdot) : \rho(A) \to \mathcal{B}(X)$ of $A \in \mathcal{B}(X)$ is analytic on $\rho(A)$, and its derivative at $\lambda \in \rho(A)$ is $[R(\lambda)]^2$.

Exercise 10.2 Let X be a Banach space over $\mathbb{C}$, $A \in \mathcal{B}(X)$, and $\zeta \in \rho(A)$. Prove that, if $z \in \mathbb{C}$ is such that $|z - \zeta| < 1/\|R(\zeta)\|$, then $z \in \rho(A)$ and

$$R(z) = R(\zeta) \sum_{k=0}^{\infty} [R(\zeta)]^k (z - \zeta)^k. \quad \square$$

Theorem 10.17 *Let X be a Banach space over $\mathbb{C}$ and $A \in \mathcal{B}(X)$. Then*

(i) **(Gelfand-Mazur theorem)** $\sigma(A) \neq \varnothing$,

(ii) **(Spectral radius formula)** $\lim_{k\to\infty} \|A^k\|^{1/k}$ *exists and*

$$r_\sigma(A) = \lim_{k\to\infty} \|A^k\|^{1/k}.$$

Proof. For $f \in (\mathcal{B}(X))'$, let

$$\phi_f(z) := f(R(z)), \qquad z \in \rho(A).$$

We have already observed in Theorem 10.16(iii) that the map

$$z \mapsto \phi_f(z), \qquad z \in \rho(A)$$

is an analytic function for every $f \in (\mathcal{B}(X))'$. Also, since $\|R(z)\| \to 0$ as $|z| \to \infty$ (cf. Theorem 10.10 (i)), we have

$$\phi_f(z) \to 0 \quad \text{as} \quad |z| \to \infty.$$

Thus, ϕ_f is a bounded analytic function on $\rho(A)$.

(i) Now suppose that $\sigma(A) = \varnothing$. Then ϕ_f is a bounded entire function. Therefore, by Liouville's theorem in complex analysis (cf. Ahlfors [2]), ϕ_f is a constant function. Since $\phi_f(z) \to 0$ as $|z| \to \infty$, it follows that $\phi_f(z) = 0$ for every $z \in \mathbb{C}$. Let $z \in \mathbb{C}$. Then $f(R(z)) = 0$ for every $f \in (\mathcal{B}(X))'$. Hence, by a consequence of the Hahn-Banach theorem (see Exercise 5.3), $R(z) = 0$. This is not possible. Hence, we can conclude that $\sigma(A)$ cannot be empty.

(ii) By Theorem 10.13,

$$r_\sigma(A) \le \inf\{\|A^k\|^{1/k} : k \in \mathbb{N}\}$$

so that

$$r_\sigma(A) \le \liminf\{\|A^k\|^{1/k} : k \in \mathbb{N}\}.$$

Hence it is enough to show that

$$\limsup\{\|A^k\|^{1/k} : k \in \mathbb{N}\} \le r_\sigma(A).$$

Let $z \in \mathbb{C}$ with $|z| > \|A\|$. Then by Corollary 10.9(ii), the series $\sum_{j=0}^\infty A^j/z^j$ converges in $\mathcal{B}(X)$, and

$$R(z) = -\frac{1}{z}\left(I - \frac{A}{z}\right)^{-1} = -\frac{1}{z}\sum_{j=0}^\infty \frac{A^j}{z^j}.$$

Hence, for $f \in (\mathcal{B}(X))'$, the function ϕ_f, which is analytic in $\rho(A)$, has the Laurent series expansion

$$\phi_f(z) = -\frac{1}{z}\sum_{j=0}^\infty f\left(\frac{A^j}{z^j}\right) \quad \text{for} \quad |z| > \|A\|.$$

Recall from Theorem 10.16 that ϕ_f is analytic on $\rho(A)$. In particular, it is analytic on

$$\{z \in \mathbb{C} : |z| > r_\sigma(A)\}.$$

Hence, the Laurent series expansion of ϕ_f is valid for $|z| > r_\sigma(A)$. Therefore, it also follows that for every z with $|z| > r_\sigma(A)$, the sequence $(f(A^n)/z^n)$ is bounded. This is true for all $f \in (\mathcal{B}(X))'$, so that by Corollary 6.5 of the uniform boundedness principle, the sequence (A^n/z^n) is also bounded.

Let $z \in \mathbb{C}$ be such that $|z| > r_\sigma(A)$, and let $c > 0$ such that $\|A^n/z^n\| \le c$ for all $j \in \mathbb{N}$. Then we have

$$\|A^n\|^{1/n} \le c^{1/n}|z| \quad \forall n \in \mathbb{N}$$

so that

$$\limsup_n \|A^n\|^{1/n} \le |z|.$$

Since this is true for all z such that $|z| > r_\sigma(A)$, we have

$$\limsup_j \|A^j\|^{1/j} \le r_\sigma(A).$$

This completes the proof. ∎

Exercise 10.3 Let X be a Banach space over $\mathbb{C}$, $A \in \mathcal{B}(X)$, and $\zeta \in \rho(A)$. Let $\delta_\zeta = 1/\lim_{k\to\infty} \|R(\zeta)^k\|^{1/k}$. Prove that

$$R(z) = R(\zeta)\sum_{k=0}^{\infty} [R(\zeta)]^k (z-\zeta)^k$$

for every $z \in \Delta_\zeta := \{\lambda \in \mathbb{K} : |\lambda - \zeta| < \delta_\zeta\}$. □

We have already seen examples where the conclusions of the above theorem do not hold if the scalar field is the set of real numbers. Next, one may ask for what type of operators A we have the relation $r_\sigma(A) = \|A\|$. Certainly, not for all operators, even if $\mathbb{K} = \mathbb{C}$. For example, consider $X = \mathbb{K}^2$ with any norm, and

$$A : (\alpha_1, \alpha_2) \mapsto (\alpha_2, 0), \qquad (\alpha_1, \alpha_2) \in X.$$

Then we see that $\sigma(A) = \{0\}$ so that

$$r_\sigma(A) = 0 \quad \text{but} \quad \|A\| \ne 0.$$

We shall see in Chapter 12 (Theorems 12.7 and 12.8) that there are classes of operators, namely, the *self-adjoint operators* on Hilbert spaces and *normal operators* on complex Hilbert spaces, for which we do have the relation $r_\sigma(A) = \|A\|$.

10.3 Integration of Operator-Valued Functions

In this section we define the concept of integration of operator-valued functions of a complex variable over rectifiable curves, and use this concept to define *spectral projection* associated with a part of the spectrum of an operator.

Recall that a *curve* in $\mathbb{C}$ is a continuous function $\gamma : [0,1] \to \mathbb{C}$. By a *simple closed curve*, we mean a curve γ such that $\gamma(t_1) \ne \gamma(t_2)$ for every distinct $t_1, t_2 \in [0,1]$, except that $\gamma(0) = \gamma(1)$. A curve

$\gamma : [0,1] \to \mathbb{C}$ is said to be a *rectifiable curve* if there exists $\kappa > 0$ such that

$$\sum_{j=1}^{n} |\gamma(t_j) - \gamma(t_{j-1})| \le \kappa$$

for every partition $0 = t_0 < t_1 < \ldots < t_n = 1$. We denote the range of γ by Γ, and say that γ is a parametrization of Γ. We shall call Γ also a rectifiable curve.

Let X be a Banach space. Suppose Γ is a rectifiable curve in $\mathbb{C}$ and $F : \Gamma \to \mathcal{B}(X)$ is a continuous function. Let $\gamma : [0,1] \to \Gamma$ be a parametrization of Γ. We define the integral of F along Γ as in the case of complex-valued function.

Let $\Pi = \{t_j : j = 0, 1, \ldots, k\}$ be a partition of $[0,1]$, i.e.,

$$\Pi : 0 = t_0 < t_1 < \ldots < t_k = 1,$$

and let $\Delta := \{s_1, \ldots, s_k\} \subset [0,1]$ be the set of *tags* on Π, i.e.,

$$t_{j-1} \le s_j \le t_j, \qquad j \in \{1, \ldots, k\}.$$

Let us denote the *mesh* of the partition by $m(\Pi)$, i.e.,

$$m(\Pi) = \max\{t_j - t_{j-1} : j = 1, \ldots, k\}.$$

Then consider the *Riemann-Stieltjes sum* of F associated with the with Π and Δ:

$$\mathcal{S}(\Pi, \Delta, F) = \sum_{j=1}^{k} [\gamma(t_j) - \gamma(t_{j-1})] F(\gamma(s_j)).$$

Then the following theorem can be proved using the uniform continuity of the function $t \to F(\gamma(t))$, $t \in [0,1]$, see Limaye [21] for a proof. In Section 11.1.3, we shall give a proof of this result when X is a Hilbert space.

Theorem 10.18 *There exists a unique operator $\Phi \in \mathcal{B}(X)$ with the property that for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\|\mathcal{S}(\Pi, \Delta, F) - \Phi\| < \varepsilon \quad \textit{whenever} \quad m(\Pi) < \delta.$$

The operator Φ in the above theorem is called the **integral of F along Γ**, and is denoted by

$$\int_\Gamma F(z)\, dz.$$

Thus, if (Π_n) is a sequence of partitions of $[0,1]$ such that $m(\Pi_n) \to 0$ as $n \to \infty$, and if $F : \Gamma \to \mathcal{B}(X)$ is a continuous function defined on a rectifiable curve Γ, then

$$S(\Pi_n, \Delta_n, F) \to \int_\Gamma F(z)\, dz,$$

where Δ_n is a set of tags on Π_n for each $n \in \mathbb{N}$. If $A \in \mathcal{B}(X)$, $f \in X'$ and $x \in X$, then it can be seen (*Verify*), using the continuity of A and f, that

$$A \int_\Gamma F(z)\, dz = \int_\Gamma AF(z)\, dz,$$

$$f\left(\left[\int_\Gamma F(z)\, dz\right] x\right) = \int_\Gamma f(F(z)x)\, dz.$$

This result, in particular, shows that Cauchy's theorem is valid in the context of operator-valued functions as well. In fact, if Φ_1 and Φ_2 are integrals of F over curves Γ_1 and Γ_2, respectively, then by Corollary 5.6,

$$f(\Phi_1 x) = f(\Phi_2 x) \quad \forall\, (x, f) \in X \times X' \iff \Phi_1 = \Phi_2.$$

10.3.1 Spectral Projections

Let X be a Banach space and $A \in \mathcal{B}(X)$. Now we consider integration of the resolvent operator

$$z \mapsto R(z) := (A - zI)^{-1}, \qquad z \in \rho(A),$$

over a simple closed rectifiable curve $\Gamma \subseteq \rho(A)$. We shall denote by Δ_Γ, the domain enclosed by Γ.

Theorem 10.19 *Let $A \in \mathcal{B}(X)$ and let Γ be a simple closed rectifiable curve in $\rho(A)$. Then*

$$P = -\frac{1}{2\pi i} \int_\Gamma R(z)\, dz$$

is a projection operator in $\mathcal{B}(X)$. Moreover,

$$P = 0 \iff \Delta_\Gamma \subseteq \rho(A),$$

$$P = I \iff \sigma(A) \subseteq \Delta_\Gamma.$$

Proof. By the definition of integral, it is clear that P is a bounded operator on X. To see that it is a projection, first we observe, following the arguments in complex analysis, that for any simple closed rectifiable curve $\tilde{\Gamma} \subseteq \rho(A)$ lying inside Γ having no spectral values in the domain lying between Γ and $\tilde{\Gamma}$, we have

$$P = -\frac{1}{2\pi i}\int_{\tilde{\Gamma}} R(\zeta)\,d\zeta.$$

Hence, using the resolvent identity $R(z) - R(\zeta) = (z-\zeta)R(z)R(\zeta)$, we have

$$\begin{aligned}
P^2 &= \Big(-\frac{1}{2\pi i}\int_\Gamma R(z)\,dz\Big)\Big(-\frac{1}{2\pi i}\int_{\tilde{\Gamma}} R(\zeta)\,d\zeta\Big)\\
&= \Big(-\frac{1}{2\pi i}\Big)^2 \int_\Gamma\int_{\tilde{\Gamma}} R(z)R(\zeta)\,dz\,d\zeta\\
&= \Big(-\frac{1}{2\pi i}\Big)^2 \int_\Gamma\int_{\tilde{\Gamma}} \frac{R(z)-R(\zeta)}{z-\zeta}\,dz\,d\zeta\\
&= \Big(-\frac{1}{2\pi i}\Big)^2 \Big\{\int_\Gamma\int_{\tilde{\Gamma}} \frac{R(z)}{z-\zeta}\,dz\,d\zeta - \int_\Gamma\int_{\tilde{\Gamma}} \frac{R(\zeta)}{z-\zeta}\,dz\,d\zeta\Big\}
\end{aligned}$$

Since

$$\frac{1}{2\pi i}\int_{\tilde{\Gamma}} \frac{1}{z-\zeta}\,d\zeta = 0 \quad \forall z\in\Gamma, \qquad \frac{1}{2\pi i}\int_\Gamma \frac{1}{z-\zeta}\,dz = 1 \quad \forall \zeta\in\tilde{\Gamma},$$

we have

$$\int_\Gamma\int_{\tilde{\Gamma}} \frac{R(z)}{z-\zeta}\,dz\,d\zeta = \int_\Gamma R(z)\left(\int_{\tilde{\Gamma}} \frac{1}{z-\zeta}\,d\zeta\right)\,d\zeta = 0,$$

$$\int_\Gamma\int_{\tilde{\Gamma}} \frac{R(\zeta)}{z-\zeta}\,dz\,d\zeta = \int_{\tilde{\Gamma}} R(\zeta)\left(\int_\Gamma \frac{1}{z-\zeta}\,dz\right)\,d\zeta = 2\pi i.$$

Hence, it follows that $P^2 = P$. Thus, P is a projection.

For $\zeta \notin \Gamma$, let

$$B_\zeta = \frac{1}{2\pi i}\int_\Gamma \frac{R(z)}{z-\zeta}\,dz.$$

Then, using the relation $AR(z) = I + zR(z)$, we have

$$\begin{aligned}(A-\zeta I)B_\zeta &= \frac{1}{2\pi i}\int_\Gamma \frac{(A-\zeta I)R(z)}{z-\zeta}\,dz \\ &= \frac{1}{2\pi i}\int_\Gamma \frac{I+(z-\zeta)R(z)}{z-\zeta}\,dz \\ &= \frac{1}{2\pi i}\int_\Gamma \frac{I}{z-\zeta}\,dz + \frac{1}{2\pi i}\int_\Gamma R(z)\,dz \\ &= \begin{cases} I-P & \text{if } \zeta \in \Delta_\Gamma \\ -P & \text{if } \zeta \notin \Delta_\Gamma. \end{cases}\end{aligned}$$

From this it follows that

$$P = 0 \implies \Delta_\Gamma \subseteq \rho(A), \qquad P = I \implies \sigma(A) \subseteq \Delta_\Gamma.$$

To obtain the reverse implications, suppose first that $\Delta_\Gamma \subseteq \rho(A)$. Then the resolvent $R(\cdot)$ is analytic on and inside Γ, so that by Cauchy's theorem, $P = 0$.

Next, suppose that $\sigma(A) \subseteq \Delta_\Gamma$. Then, again by Cauchy's theorem, we have

$$P = -\frac{1}{2\pi i}\int_{\Gamma_r} R(z)\,dz,$$

where Γ_r is a circle with centre 0 and radius $r > \|A\|$. Then, using the relations

$$\frac{1}{2\pi i}\int_{\Gamma_r} \frac{I}{z}\,dz = I, \qquad AR(z) = I + zR(z),$$

we have

$$\begin{aligned} I-P &= \frac{1}{2\pi i}\int_{\Gamma_r}\left[\frac{1}{z} + R(z)\right]dz \\ &= \frac{1}{2\pi i}\int_{\Gamma_r}\frac{1}{z}[I+zR(z)]\,dz \\ &= \frac{1}{2\pi i}\int_{\Gamma_r}\frac{1}{z}AR(z)\,dz.\end{aligned}$$

Hence, for every $(x, f) \in X \times X'$,

$$f(x - Px) = \frac{1}{2\pi i}\int_{\Gamma_r} f\Big(\frac{1}{z}AR(z)x\Big)\,dz.$$

Since (by Theorem 10.10)

$$\left\|\frac{1}{z}R(z)\right\| \leq \frac{1}{|z|(|z|-\|A\|)} = \frac{1}{r(r-\|A\|)} \quad \forall z \in \Gamma_r,$$

we have

$$\|f(x-Px)\| = \frac{\|A\|}{r-\|A\|}\|x\|\,\|f\|$$

for every $(x,f) \in X \times X'$. Since this is true for every $r > \|A\|$, by letting $r \to \infty$, we have

$$f(x-Px) = 0 \quad \forall (x,f) \in X \times X'$$

From this, by a consequence (Corollary 5.6) of the Hahn-Banach theorem, we obtain $P = I$. ∎

Note that the projection operator P in the above theorem is nonzero if and only if $\sigma(A) \cap \Delta_\Gamma \neq \varnothing$. This projection P is called the **spectral projection** associated with the part of the spectrum of A lying inside Γ. The range of a spectral projection is called a **spectral subspace**. The notion of a spectral projection has been extensively used in eigenvalue problems and perturbation theory, eigenvalue problems and their numerical approximations (cf. Kato [17], Chatelin [10], Limaye [21], and Ahues, Largillier and Limaye [1]).

Exercise 10.4 Let Γ and P be as in the above theorem. Prove the following:

(i) If f is a polynomial, then $f(A)P = -\frac{1}{2\pi i}\int_\Gamma f(z)R(z)\,dz$.

(ii) $R(P)$ is invariant under A, i.e., $Ax \in R(P)$ for all $x \in R(P)$.

(iii) $\sigma(A|_{R(P)}) = \sigma(A) \cap \Delta_\Gamma$.

(iv) If rank $P = m < \infty$, then Γ encloses a finite number of eigenvalues of A, say $\lambda_1, \ldots, \lambda_k$, and the total algebraic multiplicities of these eigenvalues, as eigenvalues of $A|_{R(P)}$, is m.

Here, by *algebraic multiplicity* of an eigenvalue λ of a linear operator $B : Y \to Y$ on a finite dimensional linear space Y we mean the dimension of the subspace $\cup_{k=1}^{\infty}\{u \in Y : (B-\lambda I)^k x = 0\}$. □

10.4 Riesz-Schauder Theory

In this section we derive results which are special to compact operators on a Banach space.

We have already seen in Theorem 10.4 that if $A : X \to X$ is a compact operator on a normed linear space, then every nonzero approximate eigenvalue of A is an eigenvalue, i.e.,

$$\sigma_{\rm app}(A) \setminus \{0\} = \sigma_{\rm eig}(A) \setminus \{0\}.$$

Now, we show that for a compact operator A on a Banach space, we, in fact, have

$$\sigma(A) \setminus \{0\} = \sigma_{\rm eig}(A) \setminus \{0\}.$$

First we prove some preliminary results.

Lemma 10.20 *Let X be a normed linear space, $A : X \to X$ be a compact operator and $0 \neq \lambda \in \mathbb{K}$. Suppose X_0 is a closed subspace of X and the restriction of $A - \lambda I$ to X_0 is injective. Then,*

(i) *$\exists\, c > 0$ such that $\|Ax - \lambda x\| \geq c\|x\|$ for all $x \in X_0$, and*

(ii) *$\{Ax - \lambda x : x \in X_0\}$ is a closed subspace of X.*

Proof. (i) Suppose the conclusion does not hold. Then there exists a sequence (x_n) in X_0 such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\|Ax_n - \lambda x_n\| \to 0$ as $n \to \infty$. Using the compactness of A, there exists a subsequence $(\tilde{x}_n)$ of (x_n) and $y \in Y$ such that $A\tilde{x}_n \to y$. Hence, we have

$$\lambda\tilde{x}_n = A\tilde{x}_n - (A\tilde{x}_n - \lambda\tilde{x}_n) \to y.$$

Since X_0 is a closed subspace of X and $\|\lambda\tilde{x}_n\| = |\lambda| \neq 0$ for all $n \in \mathbb{N}$, it follows that $y \in X_0$ and $y \neq 0$. Now using the continuity of A, we have

$$\|(A - \lambda I)y\| = \lim_{n\to\infty} \|(A - \lambda I)(\lambda\tilde{x}_n)\| = |\lambda| \lim_{n\to\infty} \|A\tilde{x}_n - \lambda\tilde{x}_n\| = 0.$$

This contradicts the fact that $(A - \lambda I)|_{X_0}$ is injective.

(ii) Now suppose that $y \in {\rm cl}\,\{Ax - \lambda x : x \in X_0\}$ and (u_n) in X_0 is such that $Au_n - \lambda u_n \to y$. We show that $y = Ax - \lambda x$ for some $x \in X_0$. By (i) it follows that (u_n) is a bounded sequence, and hence by compactness of A, there exists a subsequence $(\tilde{u}_n)$ of (u_n) and $v \in X$ such that $A\tilde{u}_n \to v$. Hence, we have

$$\lambda\tilde{u}_n = A\tilde{u}_n - (A\tilde{u}_n - \lambda\tilde{u}_n) \to v - y.$$

Since X_0 is a closed subspace and A is continuous, we have $v-y \in X_0$ and

$$(A - \lambda I)(\lambda \tilde{u}_n) \to (A - \lambda I)(v - y).$$

But, since $y = \lim_{n\to\infty} (A - \lambda I)\tilde{u}_n$, it follows that $(A - \lambda I)(v - y) = \lambda y$, i.e., $Ax - \lambda x = y$ with $x = (v - y)/\lambda \in X_0$. ∎

Remark 10.1 We note the following:

(a) Theorem 10.4 (i) is obtained from Lemma 10.20(i) by taking $X_0 = X$.

(b) If X is a Banach space, then Lemma 10.20(ii) follows from Lemma 10.20(i), without using the compactness of the operator.

(c) From Lemma 10.20(i), it also follows that if X is a Hilbert space and $A : X \to X$ is a compact operator, then for every nonzero scalar λ, there exists $c > 0$ such that

$$\|Ax - \lambda x\| \geq c\|x\| \quad \forall x \in N(A - \lambda I)^{\perp}.$$

Proposition 10.21 *Let X be a normed linear space, $A : X \to X$ be a compact operator, and λ be a nonzero scalar. Then*

(i) *$N(A - \lambda I)$ is finite dimensional, and*

(ii) *$R(A - \lambda I)$ is closed.*

Proof. (i) It is enough to show that every bounded sequence in $N(A - \lambda I)$ has a convergent subsequence. Let (x_n) be a bounded sequence in $N(A - \lambda I)$. Since A is compact, there exists a subsequence $(\tilde{x}_n)$ of (x_n) and $y \in X$ such that $A\tilde{x}_n \to y$. Thus, $\lambda\tilde{x}_n = A\tilde{x}_n \to y$ and $y \in N(A - \lambda I)$, showing that $(\tilde{x}_n)$ is a convergent subsequence of (x_n).

(ii) By (i), and using Corollary 5.8, there exists a closed subspace Y of X such that X is the direct sum of $N(A - \lambda I)$ and Y. Then it follows that $(A - \lambda I)|_Y : Y \to X$ is injective and $R(A - \lambda I) = R((A - \lambda I)|_Y)$. Hence, by Lemma 10.20 (ii), $R(A - \lambda I)$ is closed in X. ∎

The following corollary is immediate from the first part of the above proposition.

Corollary 10.22 *Let X be a normed linear space, $A : X \to X$ be a compact operator, and λ be a nonzero scalar. If λ is an eigenvalue* *of A, and for $y \in X$, if x_0 is a (particular) solution of the equation $Ax - \lambda x = y$, then the set of all solutions is*

$$\left\{x_0 + \sum_{j=1}^{k} \alpha_j u_j : (\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k\right\},$$

where $\{u_1, \ldots, u_k\}$ is a basis of $N(A - \lambda I)$.

Theorem 10.23 *Let X be a normed linear space and $A : X \to X$ be a compact operator. Then every nonzero spectral value of A is an eigenvalue of A.*

Proof. Suppose λ is a nonzero scalar which is not an eigenvalue of A. We show that $\lambda \notin \sigma(A)$. By Theorem 10.4(i), we know that $\lambda \notin \sigma_{\rm app}(A)$. Hence it is enough to show that $R(A - \lambda I) = X$.

First observe from Proposition 10.21(ii) that $R(A - \lambda I)$ is closed in X. Let $Y_0 = X$, and for $n \in \mathbb{N}$, let

$$Y_n = R((A - \lambda I)^n).$$

Note that

$$Y_n \supseteq Y_{n+1} \quad \forall n \in \mathbb{N}.$$

Since $(A - \lambda I)^n$ can be written in the form $B - \mu I$ with B a compact operator and μ a nonzero scalar, it again follows from Proposition 10.21(ii), that Y_n is closed for every $n \in \mathbb{N}$. We first show that there exists $k \in \{0, 1, 2, \ldots\}$ such that

$$Y_k = Y_{k+1},$$

and use this fact to show that $Y_0 = Y_1$.

Suppose there is no k such that $Y_k = Y_{k+1}$. Then Y_{n+1} is a proper closed subspace of Y_n for every $n \in \{0, 1, 2, \ldots\}$ so that by Riesz lemma (Theorem 2.40), there exists a sequence (x_n) in X such that

$$x_n \in Y_n, \qquad \|x_n\| = 1, \qquad \operatorname{dist}(x_n, Y_{n+1}) \geq \frac{1}{2}.$$

Note that for $m > n$,

$$\begin{aligned} \|Ax_n - Ax_m\| &= \|(A - \lambda I)x_n - (A - \lambda I)x_m + \lambda(x_n - x_m)\| \\ &= \|\lambda x_n - [\lambda x_m - (A - \lambda I)x_n + (A - \lambda I)x_m]\|. \end{aligned}$$

Since $x_m \in Y_m \subseteq Y_{n+1}$, and

$$(A - \lambda I)x_n \in Y_{n+1}, \qquad (A - \lambda I)x_m \in Y_{m+1} \subseteq Y_{n+1},$$

it follows that

$$\|Ax_n - Ax_m\| \geq |\lambda| \operatorname{dist}(x_n, Y_{n+1}) \geq \frac{|\lambda|}{2}.$$

Thus, (x_n) is a bounded sequence such that (Ax_n) has no convergent subsequence, contradicting the compactness of A. Thus, there exists $k \in \{0, 1, 2, \ldots\}$ such that $Y_k = Y_{k+1}$. (If $k = 0$, then the proof is over, for then we have $X = Y_0 = Y_1 = R(A - \lambda I)$.)

We now show that, if $Y_n = Y_{n+1}$ for some $n \in \mathbb{N}$, then $Y_{n-1} = Y_n$. Once this is proved, then (since we already know that there exists $k \in \{0, 1, 2, \ldots\}$ such that $Y_k = Y_{k+1}$) by induction it follows that $Y_0 = Y_1 = \ldots = Y_n$. In particular, $X = Y_0 = Y_1 = R(A - \lambda I)$.

So, let $Y_n = Y_{n+1}$ for some $n \in \mathbb{N}$, and let $x \in Y_{n-1}$. Let $u \in X$ be such that $x = (A - \lambda I)^{n-1}u$. Then we have

$$(A - \lambda I)x = (A - \lambda I)^n u \in Y_n = Y_{n+1}.$$

Therefore, there exists $v \in X$ such that $(A - \lambda I)x = (A - \lambda I)^{n+1}v$. Now, injectivity of $A - \lambda I$ implies $x = (A - \lambda I)^n v \in Y_n$. Thus, $Y_{n-1} = Y_n$.

This completes the proof. ∎

The following result, known as a **Fredholm alternative**, is a consequence of the above theorem.

Corollary 10.24 *Let X be a normed linear space, $A : X \to X$ be a compact operator, and λ be a nonzero scalar. Then one and only one of the following statements holds:*

(i) *The equation $Ax - \lambda x = 0$ has a nonzero solution.*

(ii) *For every $y \in X$, there exists a unique $x \in X$ such that $Ax - \lambda x = y$.*

In view of Theorem 10.1 and Theorem 10.23, we have the following result.

Theorem 10.25 *Let X be a normed linear space and $A : X \to X$ be a compact operator. Then the only possible limit point of $\sigma(A)$ is zero. In particular, $\sigma(A)$ is countable.*

PROBLEMS

1. Let $X = C[0,1]$ and $A : X \to X$ be defined by

$$(Ax)(t) = \int_0^t x(s)\,ds, \qquad x \in X,\ 0 \le t \le 1.$$

Show that $\sigma_{\rm eig}(A) = \varnothing$ and $\sigma_{\rm app}(A) = \{0\}$.

2. Let X be a normed linear space and $A : X \to X$ be a linear operator. Suppose $\lambda \in \mathbb{K}$ is such that $A - \lambda I$ is injective. Show that $(A - \lambda I)^{-1} : R(A - \lambda I) \to X$ is continuous if and only if λ is not an approximate eigenvalue.

3. Give an example of an approximate eigenvalue which is neither an eigenvalue nor a limit of a sequence of eigenvalues.

4. Let X be a Banach space and $A \in \mathcal{B}(X)$. Show that the series $\sum_{n=1}^{\infty} A^n/n!$ converges.

The above series is usually denoted by $\exp(A)$.

5. Give an example of a bijective operator A on a normed linear space X such that $0 \in \sigma(A)$.

6. Let Δ be a compact subset of $\mathbb{K}$. Give an example of a Banach space X and an operator $A \in \mathcal{B}(X)$ such that $\sigma(A) = \Delta = \sigma_{\rm app}(A)$.

7. Suppose X is a normed linear space (not necessarily a Banach space), and $A \in \mathcal{B}(X)$. If $\lambda \in \mathbb{K}$ is such that $|\lambda| > \|A\|$, then show that

(a) $A - \lambda I$ is bounded below,

(b) $R(A - \lambda I)$ dense in X, and

(c) $(A - \lambda I)^{-1} : R(A - \lambda I) \to X$ is continuous.

8. Let X be a Banach space and $A \in \mathcal{B}(X)$. Then show that

(a) $\lambda \in \sigma_{\rm eig}(A)$ if and only if there exists a nonzero $B \in \mathcal{B}(X)$ such that $(A - \lambda I)B = 0$,

(b) $\lambda \in \sigma_{\rm app}(A)$ if and only if $A' - \lambda I$ is not surjective,

(c) $\lambda \in \sigma_{\rm com}(A)$ if and only if there exists $B \neq 0$ in $\mathcal{B}(X)$ such that $B(A - \lambda I) = 0$.

9. Let X be a Banach space and $A \in \mathcal{B}(X)$. Consider the following subsets of $\sigma(A)$:

$$\sigma_r(A) := \sigma_{\rm com}(A) \setminus \sigma_{\rm eig}(A), \qquad \sigma_c(A) := \sigma(A) \setminus (\sigma_{\rm com}(A) \cup \sigma_{\rm eig}(A)).$$

Show that $\sigma(A)$ is the disjoint union of $\sigma_{\rm eig}$, $\sigma_r(A)$ and $\sigma_c(A)$.

The sets $\sigma_r(A)$ and $\sigma_c(A)$ are, called the *residual spectrum* and *continuous spectrum* of A, respectively.

10. Let X be a Banach space and $A \in \mathcal{B}(X)$. Show that for every $\mu \in \rho(A)$,

$$\|(A - \mu I)^{-1}\| \geq \frac{1}{\operatorname{dist}(\mu, \sigma(A))}.$$

11. Let X be a Banach space and $A \in \mathcal{B}(X)$. If $\lambda \in \rho(A)$, and if $B \in \mathcal{B}(X)$ satisfies the inequality $\|A - B\| < 1/\|(A - \lambda I)^{-1}\|$, then show that $\lambda \in \rho(B)$.

Also, show that if Δ is a compact subset of $\mathbb{K}$ contained in $\rho(A)$, then there exists $\varepsilon > 0$ such that $\Delta \subseteq \rho(B)$ for every $B \in \mathcal{B}(X)$ with $\|A - B\| < \varepsilon$.

12. Let X be a normed linear space and $A : X \to X$ be a compact operator. For a nonzero scalar λ and $k \in \mathbb{N}$, let $N_k := N\left((A - \lambda I)^k\right)$ and $R_k := R\left((A - \lambda I)^k\right)$. Prove that there exist non-negative integers m and n such that

$$N_m = N_{m+j}, \qquad R_n = R_{n+j} \quad \forall j \in \mathbb{N}.$$

Also, prove that, if $\ell := \min\{m : N_m = N_{m+j} \ \forall j \in \mathbb{N}\}$ and $\kappa := \min\{n : R_n = R_{n+j} \ \forall j \in \mathbb{N}\}$, then $\ell = \kappa$, and

$$X = R\left((A - \lambda I)^r\right) \oplus N\left((A - \lambda I)^r\right), \qquad r = \kappa = \ell.$$

11

Operators on Hilbert Spaces

In this chapter we deal with some special properties of operators on inner product spaces and Hilbert spaces.

We may recall that if X is an inner product space, then for every $x \in X$,

$$\|x\| = \sup\{|\langle x, u\rangle| : u \in X, \|u\| \leq 1\}$$

In particular, $x = 0$ if and only if $\langle x, u\rangle = 0$ for all $u \in X$. Hence, it also follows that if $A : X \to Y$ is a linear operator between inner product spaces X and Y, then

$$A = 0 \iff \langle Ax, y\rangle = 0 \quad \forall x \in X,\ y \in Y,$$

and if $A \in \mathcal{B}(X, Y)$, then (*Verify*)

$$\|A\| = \sup\{|\langle Ax, y\rangle| : x \in X,\ y \in Y,\ \|x\| \leq 1,\ \|y\| \leq 1\}.$$

NOTATION: We shall denote the inner product on any space by the same notation $\langle \cdot, \cdot \rangle$.

11.1 Adjoint of an Operator

We introduce the notion of *adjoint* of an operator as a generalization of the concept of *conjugate transpose* of a matrix.

Let (a_{ij}) be an $m \times n$ matrix of scalars. Let $A : \mathbb{K}^n \to \mathbb{K}^m$ be the linear operator associated with the above matrix. That is, A is defined by

$$(Ax)(i) = \sum_{j=1}^{n} a_{ij} x(j), \qquad i = 1, \ldots, m;\ x \in \mathbb{K}^n.$$

Let $B : \mathbb{K}^m \to \mathbb{K}^n$ be the linear operator associated with the conjugate transpose of (a_{ij}), i.e.,

$$(By)(i) = \sum_{j=1}^{m} \overline{a_{ji}}\, y(j), \qquad i = 1, \ldots, n; \;\; y \in \mathbb{K}^m.$$

Then, it is easily seen that

$$\langle Ax, y\rangle_{\mathbb{K}^m} = \langle x, By\rangle_{\mathbb{K}^n} \quad \forall x \in \mathbb{K}^n, \; y \in \mathbb{K}^m.$$

Here, $\langle \cdot, \cdot \rangle_{\mathbb{K}^m}$ and $\langle \cdot, \cdot \rangle_{\mathbb{K}^n}$ denote the standard inner products on $\mathbb{K}^m$ and $\mathbb{K}^n$, respectively.

Since every linear operator from $\mathbb{K}^n$ to $\mathbb{K}^m$ is induced by an $m \times n$ matrix, it can be seen easily that, corresponding to every linear operator $A : \mathbb{K}^n \to \mathbb{K}^m$, there is a linear operator $B : \mathbb{K}^m \to \mathbb{K}^n$ such that

$$\langle Ax, y\rangle_{\mathbb{K}^m} = \langle x, By\rangle_{\mathbb{K}^n} \quad \forall x \in \mathbb{K}^n, \; y \in \mathbb{K}^m.$$

Since $\mathbb{K}^n$ is a prototype of any n-dimensional inner product space, it is also not difficult to show (*Verify*) that, if X and Y are finite dimensional inner product spaces, and if $A : X \to Y$ is a linear operator, then there is a linear operator $B : Y \to X$ such that

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X, \; y \in Y.$$

Now, let X and Y be any (not necessarily finite dimensional) inner product spaces and $A : X \to Y$ be a linear operator. Then A is said to have an **adjoint** if there exists a linear operator $B : Y \to X$ such that

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X, \; y \in Y.$$

Such an operator B is called an **adjoint** of A.

We may observe that, if an adjoint exists, then it is unique:

Suppose B_1, B_2 are adjoints of a linear operator $A : X \to Y$. Then we have

$$\langle x, B_1 y\rangle = \langle Ax, y\rangle = \langle x, B_2 y\rangle \quad \forall x \in X, \; y \in Y$$

so that $\langle x, (B_1 - B_2)y\rangle = 0$ for all $x \in X$, $y \in Y$. Consequently, $B_1 = B_2$.

The adjoint of a linear operator A, if it exists, is denoted by A^*.

Note that, if a linear operator $A : X \to Y$ has the adjoint A^*, then

$$\langle A^*y, x\rangle = \overline{\langle x, A^*y\rangle} = \overline{\langle Ax, y\rangle} = \langle y, Ax\rangle \quad \forall x \in X,\ y \in Y,$$

and, consequently,

$$(A^*)^* = A.$$

Exercise 11.1 Let $\mathcal{U}$ and $\mathcal{V}$ be orthonormal bases of finite dimensional inner product spaces X and Y, respectively, and let $A : X \to Y$ be a linear operator. Show that if (a_{ij}) is the matrix representation of A with respect to the bases $\mathcal{U}$ and $\mathcal{V}$, then the matrix representation of the adjoint of A with respect to $\mathcal{V}$ and $\mathcal{U}$ is $(\overline{a_{ji}})$. □

We have already mentioned that every linear operator between finite dimensional inner product spaces has an adjoint. What about the situation in which X, Y are infinite dimensional spaces? Well, in this case, the adjoint may not exist. Here is an example to that effect.

EXAMPLE 11.1 Let $X = c_{00}$ with ℓ^2-inner product, i.e.,

$$\langle x, y\rangle = \sum_{j=1}^{\infty} x(j)\,\overline{y(j)}, \quad x, y \in X.$$

Consider the linear operator $A : X \to X$ defined by

$$Ax = \left(\sum_{j=1}^{\infty} \frac{x(j)}{j}\right) e_1, \quad x \in X.$$

Note that $A \in \mathcal{B}(X)$, and

$$\langle Ae_n, e_1\rangle = \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Now if there exists an operator $B : X \to X$ such that

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X,\ y \in Y,$$

then, as a particular case of the above relation, we will have

$$\frac{1}{n} = \langle Ae_n, e_1\rangle = \langle e_n, Be_1\rangle = \overline{(Be_1)(n)} \quad \forall n \in \mathbb{N},$$

which is not possible since Be_1 must belong to c_{00}. Thus, in this example, there exists no linear operator $B : X \to X$ such that $\langle Ax, y\rangle = \langle x, By\rangle$ for all $x, y \in X$.

However, if we consider the same operator as an operator on ℓ^2, then the adjoint does exist. Indeed, for every $x, y \in \ell^2$,

$$\langle Ax, y\rangle = \sum_{j=1}^{\infty} \frac{x(j)}{j} \langle e_1, y\rangle = \sum_{j=1}^{\infty} \frac{x(j)}{j}\, \overline{y(1)} = \langle x, By\rangle,$$

where $B : \ell^2 \to \ell^2$ is defined by

$$(A^*y)(j) = \frac{y(1)}{j}, \qquad j \in \mathbb{N};\ y \in \ell^2.$$

We now give a few more examples of linear operators between infinite dimensional inner product spaces having adjoints.

EXAMPLE 11.2 Let X and Y be $C[a,b]$ or $L^2[a,b]$ with $\|\cdot\|_2$.

(i) For $u \in C[a,b]$, consider the linear operator $A : X \to X$ defined by

$$(Ax)(t) = u(t)x(t), \qquad x \in X,\ t \in [a,b].$$

We note that for all $x, y \in X$,

$$\langle Ax, y\rangle = \int_a^b u(t)x(t)\overline{y(t)}\, d\mu(t).$$

From this, we see that A^* exists and is given by

$$(A^*y)(t) = \overline{u(t)}\, y(t) \quad \forall y \in X,\ t \in [a,b]$$

which is the adjoint of A.

(ii) Suppose $k(\cdot,\cdot) \in C([a,b] \times [a,b])$. Then we know (Example 9.2(vi)) that for every $x \in L^2[a,b]$ the function Ax defined by

$$(Ax)(s) = \int_a^b k(s,t)x(t)d\mu(t), \qquad x \in L^2[a,b];\ s \in [a,b],$$

belongs to $C[a,b]$, and $A : X \to Y$ is a compact operator. Note that, for all $x, y \in L^2[a,b]$,

$$\begin{aligned}
\langle Ax, y\rangle &= \int_a^b (Ax)(s)\overline{y(s)}\, d\mu(s) \\
&= \int_a^b \left(\int_a^b k(s,t)x(t)d\mu(t)\right)\overline{y(s)}\, d\mu(s) \\
&= \int_a^b x(t)\overline{\left(\int_a^b \overline{k(s,t)}y(s)d\mu(s)\right)}\, d\mu(t).
\end{aligned}$$

Thus, A^* exists and is given by

$$(A^*y)(s) = \int_a^b \overline{k(t,s)}y(t)d\mu(t) \quad \forall y \in Y,\ s \in [a,b].$$

Exercise 11.2 Let $X = L^2[a,b]$, and for $\phi \in L^\infty[a,b]$, let $Ax := \phi x$, $x \in X$. Show that the operator $B : X \to X$ defined by $Bx = \bar{\phi}\, x$, $x \in X$, is the adjoint of A. $\square$

We shall soon show that if X is a Hilbert space, then every bounded linear operator $A : X \to Y$ has the adjoint.

First we observe the following.

Theorem 11.1 *Let X, Y be inner product spaces, and $A : X \to Y$ be a linear operator such that the adjoint A^* exists. If $A \in \mathcal{B}(X,Y)$, then $A^* \in \mathcal{B}(Y,X)$, and*

$$\|A^*\| = \|A\|, \qquad \|A^*A\| = \|A\|^2.$$

Proof. Suppose $A \in \mathcal{B}(X,Y)$. For every $y \in Y$,

$$\begin{aligned} \|A^*y\| &= \sup\{|\langle A^*y, u\rangle| : u \in X,\ \|u\| \le 1\} \\ &= \sup\{|\langle y, Au\rangle| : u \in X,\ \|u\| \le 1\}. \end{aligned}$$

Hence, by continuity of A, we have

$$A^* \in \mathcal{B}(Y,X), \qquad \|A^*\| \le \|A\|.$$

Since $(A^*)^* = A$, it follows from the above that

$$\|A\| = \|(A^*)^*\| \le \|A^*\|.$$

Thus, we have proved that $\|A^*\| = \|A\|$.

Also, for every $x \in X$,

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle \le \|A^*Ax\|\,\|x\| \le \|A\|^2\|x\|^2.$$

From this we obtain $\|A\|^2 = \|A^*A\|$. ∎

Now we state the theorem that we promised.

Existence of the adjoint for linear operators

Theorem 11.2 *Let X be a Hilbert space and Y be an inner product space. Then every $A \in \mathcal{B}(X,Y)$ has the adjoint.*

Proof. Let $A \in \mathcal{B}(X,Y)$. We have to show that there exists a linear operator $B : Y \to X$ such that

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X,\ y \in Y.$$

For this purpose, for each $y \in Y$, consider the map $f_y : X \to \mathbb{K}$ defined by $f_y(x) = \langle Ax, y\rangle$, $x \in X$. Clearly, $f_y \in X'$. Hence, by the Riesz representation theorem (Theorem 3.9), there exists a unique $u_y \in X$ such that $f_y(x) = \langle x, u_y\rangle$ for all $x \in X$, i.e.,

$$\langle Ax, y\rangle = \langle x, u_y\rangle \quad \forall x \in X.$$

It is easily seen that the map $B : Y \to X$ defined by

$$By = u_y, \qquad y \in Y$$

is a linear operator, and it satisfies

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X,\ y \in Y.$$

Thus, the adjoint A^* exists and is given by $A^*y = u_y$ for all $y \in Y$. ∎

Remark 11.1 In the above theorem, although we used completeness of the space X for the existence of an adjoint of an operator, it is not a necessary condition. Recall that in Example 11.2, the space $C[a,b]$ with $\|\cdot\|_2$ is not a Hilbert space.

We have already observed that $(A^*)^* = A$ for every $A \in \mathcal{B}(X,Y)$. In the following theorem, let us list a few more properties of the adjoint. The proofs will follow easily from the definition.

Theorem 11.3 *Let X, Y, Z be Hilbert spaces.*

(i) *For $A, A_1, A_2 \in \mathcal{B}(X,Y)$ and $\alpha \in \mathbb{K}$,*

$$(A_1 + A_2)^* = A_1^* + A_2^*, \qquad (\alpha A)^* = \bar{\alpha}A^*.$$

(ii) *For $A_1 \in \mathcal{B}(X,Y)$ and $A_2 \in \mathcal{B}(Y,Z)$,*

$$(A_2A_1)^* = A_1^*A_2^*.$$

By part (i) in the above theorem and the fact that $(A^*)^* = A$ and $\|A^*\| = \|A\|$ for all $A \in \mathcal{B}(X,Y)$, it is clear that the map $A \mapsto A^*$ is a conjugate linear isometry from $\mathcal{B}(X,Y)$ onto $\mathcal{B}(Y,X)$.

Let us find the adjoint of a few bounded linear operators.

EXAMPLE 11.3 (i) Let X and Y be infinite dimensional separable Hilbert spaces, $\mathcal{U} = \{u_1, u_2, \ldots\}$ and $\mathcal{V} = \{v_1, v_2, \ldots\}$ be orthonormal bases of X and Y, respectively. Let (λ_n) be a bounded sequence of scalars. For $x \in X$, define

$$Ax = \sum_n \lambda_n \langle x, u_n\rangle v_n, \qquad x \in X.$$

Clearly (*How?*), $A \in \mathcal{B}(X, Y)$. It is easily seen that the adjoint of A is given by

$$A^* y = \sum_n \bar{\lambda}_n \langle y, v_n\rangle u_n, \qquad y \in Y.$$

The following example includes the previous one as a particular case.

(ii) Let X, Y, $\mathcal{U} = \{u_1, u_2, \ldots\}$, $\mathcal{V} = \{v_1, v_2, \ldots\}$ be as in (i). Let (a_{ij}) be an infinite matrix of scalars, and let

$$\alpha := \sum_{i,j} |a_{ij}|^2, \qquad \beta := \sup_i \sum_j |a_{ij}|, \qquad \gamma := \sup_j \sum_i |a_{ij}|.$$

Suppose that $\min\{\alpha, \beta\gamma\} < \infty$. Then, as in Example 3.4(i), we see that

$$Ax := \sum_i \Big(\sum_j a_{ij} \langle x, u_j\rangle \Big) v_i, \qquad x \in X, \tag{11.1}$$

belongs to Y for all $x \in X$, and the map $A : X \to Y$ is a bounded linear operator (*Verify*). It can be easily seen that A^* is given by

$$A^* y = \sum_i \Big(\sum_j \overline{a_{ji}} \langle y, v_j\rangle \Big) u_i, \qquad y \in Y. \tag{11.2}$$

Note that

$$\langle Au_j, v_i\rangle = a_{ij}, \qquad \langle A^* v_j, u_i\rangle = \langle v_j, Au_i\rangle = \overline{a_{ji}} \quad \forall i, j \in \mathbb{N}.$$

In the above example, the operator A is defined by a given matrix (a_{ij}) which satisfies certain conditions. The following example shows that every bounded linear operator $A : X \to Y$ between separable Hilbert spaces X and Y can be represented by a matrix in the sense of (11.1), and then its adjoint is given by (11.2).

(iii) Let $X, Y, \mathcal{U} = \{u_1, u_2, \ldots\}$, $\mathcal{V} = \{v_1, v_2, \ldots\}$ be as in (i), and let $A \in \mathcal{B}(X, Y)$. Using the Fourier expansion on X (Theorem 4.9) and the continuity of the operator A, it follows that

$$x = \sum_j \langle x, u_j \rangle u_j, \qquad Ax = \sum_j \langle x, u_j \rangle A u_j \quad \forall x \in X.$$

Now, using the Fourier expansion on Y and the continuity of the inner product, we have

$$Ax = \sum_i \langle Ax, v_i \rangle v_i = \sum_i \Big(\sum_j \langle Au_j, v_i \rangle \langle x, u_j \rangle \Big) v_i, \qquad x \in X.$$

Similarly, for every $y \in Y$,

$$A^* y = \sum_i \Big(\sum_j \langle A^* v_j, u_i \rangle \langle y, v_j \rangle \Big) u_i.$$

Since

$$\langle A^* v_j, u_i \rangle = \langle v_j, Au_i \rangle = \overline{\langle Au_i, v_j \rangle},$$

we have

$$A^* y = \sum_i \Big(\sum_j \overline{\langle Au_i, v_j \rangle} \langle y, v_j \rangle \Big) u_i, \qquad y \in Y.$$

In the above situation, we say that the matrix (a_{ij}) defined by $a_{ij} = \langle Au_j, v_i \rangle$, $i, j \in \{1, 2, \ldots\}$ is the matrix representation of A with respect to the orthonormal bases $\mathcal{U}$ and $\mathcal{V}$, and we may write this matrix as $[A]_{\mathcal{U},\mathcal{V}}$. Thus, from the above observation, we have $[A^*]_{\mathcal{V},\mathcal{U}} = (b_{ij})$ with $b_{ij} = \overline{a_{ij}}$.

As particular cases of the above example, let us look at the adjoints of the right shift and left shift operators: Let $X = \ell^2$ and A be the right shift operator, i.e.,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots).$$

In this case, we have $\langle Ae_j, e_i \rangle = \delta_{j+1,i}$ for all $i, j \in \mathbb{N}$, and

$$Ax = \sum_i x(i) e_{i+1}, \qquad A^* x = \sum_i x(i+1) e_i \quad \forall x \in \ell^2.$$

Note that A^* is the left shift operator $(\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots)$.

Since $(A^*)^* = A$, it follows at once that the adjoint of the left shift operator is the right shift operator.

(iv) Let $X = L^2[a,b]$ and for $k(\cdot,\cdot) \in L^2([a,b] \times [a,b])$, define

$$(Ax)(s) = \int_a^b k(s,t)x(t)\,d\mu(t), \qquad x \in X,\ s \in [a,b].$$

Then we know from Example 9.2(vii) that $A : X \to X$ is a compact operator. As in Example 11.2(ii), we see that, for all $x, y \in X$,

$$\langle Ax, y\rangle = \int_a^b x(t)\overline{\left(\int_a^b \overline{k(s,t)}y(s)\,d\mu(s)\right)}\,d\mu(t).$$

Thus the adjoint of A is given by

$$(A^*y)(s) = \int_a^b \overline{k(t,s)}y(t)\,d\mu(t), \qquad y \in X,\ s \in [a,b].$$

In the following proposition, we list a few results concerning ranges and null spaces of A and its adjoint which will be useful for deriving spectral results for Hilbert space operators.

Proposition 11.4 *Let X and Y be Hilbert spaces, and $A \in \mathcal{B}(X,Y)$. Then we have the following:*

(i) $N(A) = R(A^*)^\perp$.

(ii) $N(A^*) = R(A)^\perp$.

(iii) $N(A)^\perp = \overline{R(A^*)}$.

(iv) $N(A^*)^\perp = \overline{R(A)}$.

(v) $N(A^*A) = N(A)$.

Proof. We note that for $x \in X$,

$$\begin{aligned} x \in R(A^*)^\perp &\iff \langle x, A^*y\rangle = 0 \quad \forall y \in Y\\ &\iff \langle Ax, y\rangle = 0 \quad \forall y \in Y\\ &\iff Ax = 0\\ &\iff x \in N(A). \end{aligned}$$

Thus we have proved (i). The result in (ii) is obtained by replacing A in (i) by A^*. The results in (iii) and (iv) follow from (i) and (ii) by observing that for every subset S of a Hilbert space, $(S^\perp)^\perp = \overline{\operatorname{span} S}$ (*Verify*). For the proof of (v), first we observe that $N(A) \subseteq N(A^*A)$. The reverse inclusion is a consequence of the relation $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle$ for all $x \in X$. ∎

11.1.1 Compactness of the Adjoint Operator

Let X and Y be Hilbert spaces and $A \in \mathcal{B}(X,Y)$. Let $T : X' \to X$ and $S : Y' \to Y$ be the maps that relate the elements of the duals of the spaces X and Y to the elements of the spaces via the Riesz representation theorem, i.e., for $f \in X'$ and $g \in Y'$, $Tf \in X$ and $Sg \in Y$ are such that

$$f(x) = \langle x, Tf \rangle, \qquad g(y) = \langle y, Sg \rangle \quad \forall x \in X,\ y \in Y.$$

Let $A' : Y' \to X'$ be the *transpose* of A, i.e.,

$$(A'g)(x) = g(Ax) \quad \forall g \in Y',\ x \in X.$$

Then it can be seen (*Verify*) that

$$A^* = TA'S^{-1}, \quad \text{or equivalently,} \quad A' = T^{-1}A^*S.$$

Since T, S and their inverses are bounded linear operators, the following result will follow from Theorems 9.2 and 9.12.

Theorem 11.5 *Let $A \in \mathcal{B}(X,Y)$. Then*

$$A \in \mathcal{K}(X,Y) \iff A^* \in \mathcal{K}(Y,X).$$

An alternate proof for Theorem 11.5. Let $A \in \mathcal{B}(X,Y)$ be a compact operator. By Theorem 9.1(iv), it is enough to show that for every bounded sequence (y_n) in Y, the sequence (A^*y_n) has a convergent sequence. So let (y_n) be a bounded sequence in Y. Since $A \in \mathcal{K}(X,Y)$ and $A^* \in \mathcal{B}(Y,X)$, by Theorem 9.2(ii), it follows that $AA^* \in \mathcal{K}(Y)$. Hence, (AA^*y_n) has a convergent subsequence, say $(AA^*y_{n_k})$. Then we have

$$\begin{aligned} \|A^*y_{n_k} - A^*y_{n_m}\|^2 &= \langle A^*y_{n_k} - A^*y_{n_m}, A^*y_{n_k} - A^*y_{n_m} \rangle \\ &= \langle AA^*y_{n_k} - AA^*y_{n_m}, y_{n_k} - y_{n_m} \rangle \\ &\leq \|AA^*y_{n_k} - AA^*y_{n_m}\| \, \|y_{n_k} - y_{n_m}\| \\ &\leq 2c\,\|AA^*y_{n_k} - AA^*y_{n_m}\|, \end{aligned}$$

where $c > 0$ is such that $\|y_n\| \leq c$ for all $n \in \mathbb{N}$. The above relation shows that $(A^*y_{n_k})$ is a Cauchy sequence in X so that it converges.

The converse part is a consequence of the fact that $A^{**} = A$. ∎

Exercise 11.3 Let (P_n) be a sequence of orthogonal projections on a Hilbert space X and $K : X \to X$ be a compact operator. If $P_n x \to x$ as $n \to \infty$ for every $x \in X$, then show that $\|K(I - P_n)\| \to 0$ as $n \to \infty$. □

11.1.2 Sesquilinear Functionals

In Section 2.1.6, we have defined a sesquilinear functional on a linear space. More generally, if X and Y are linear spaces, then a function $\psi : X \times Y \to \mathbb{K}$ is called a **sesquilinear functional** if for every $x, y \in X$ and $u, v \in Y$, and $\alpha, \beta \in \mathbb{K}$, we have

$$\begin{aligned} \psi(\alpha x + \beta y, u) &= \alpha\psi(x, u) + \beta\psi(y, u), \\ \psi(x, \alpha u + \beta v) &= \bar{\alpha}\psi(x, u) + \bar{\beta}\psi(x, v), \end{aligned}$$

An immediate example of a sesquilinear functional is the following: Suppose X and Y are inner product spaces and $A : X \to Y$ is a linear operator. Then the function $\psi_A : X \times Y \to \mathbb{K}$ defined by

$$\psi_A(x, y) = \langle Ax, y\rangle, \quad (x, y) \in X \times Y,$$

is a sesquilinear functional.

Now, suppose that X and Y are normed linear spaces. Then a sesquilinear functional $\psi : X \times Y \to \mathbb{K}$ is said to be **bounded** if there exists $c > 0$ such that

$$|\psi(x, y)| \le c\|x\|\,\|y\| \quad \forall\, (x, y) \in X \times Y.$$

Thus, if X and Y are inner product spaces and $A \in \mathcal{B}(X, Y)$, then the function $\psi_A : X \times Y \to \mathbb{K}$ defined by

$$\psi_A(x, y) = \langle Ax, y\rangle, \quad (x, y) \in X \times Y,$$

is a bounded sesquilinear functional. The question is whether every bounded sesquilinear functional $\psi : X \times Y \to \mathbb{K}$ with X, Y inner product spaces arise in this manner. We answer this question affirmatively when X and Y are Hilbert spaces.

Theorem 11.6 *Let X and Y be Hilbert spaces and $\psi : X \times Y \to \mathbb{K}$ be a bounded sesquilinear functional. Then there exists a unique bounded linear operator $A : X \to Y$ such that*

$$\psi(x, y) = \langle Ax, y\rangle$$

for all $(x, y) \in X \times Y$.

Proof. Let $c > 0$ be such that

$$|\psi(x, y)| \le c\|x\|\,\|y\| \quad \forall\, (x, y) \in X \times Y.$$

For each $y \in Y$, consider the function $\psi_y : X \to \mathbb{K}$ defined by

$$\psi_y(x) = \psi(x, y), \quad x \in X.$$

Then we see that $\psi_y \in X'$ so that by the Riesz representation theorem 3.9), there exists a unique $u_y \in X$ such that

$$\psi_y(x) = \langle x, u_y \rangle \quad \forall x \in X.$$

Let us denote the map $y \mapsto u_y$ by B so that

$$\psi(x, y) = \psi_y(x) = \langle x, By \rangle \quad \forall (x, y) \in X \times Y.$$

It is easily seen that $B : Y \to X$ is a linear operator. Also, we have

$$|\langle Bx, y \rangle| = |\psi(x, y)| \le c\|x\| \, \|y\| \quad \forall (x, y) \in X \times Y,$$

so that $B \in \mathcal{B}(Y, X)$ and $\|B\| \le c$ (*How?*). The operator $A = B^* : X \to Y$ satisfies the requirements of the theorem. ∎

Exercise 11.4 Derive Theorem 11.2 from Theorem 11.6. □

Next, we consider an important result connecting the ranges of an operator and its adjoint.

Theorem 11.7 (Closed range theorem) *Let X and Y be Hilbert spaces and $A \in \mathcal{B}(X, Y)$. Then $R(A)$ is closed if and only if $R(A^*)$ is closed.*

Proof. Suppose $Y_0 := R(A)$ is closed. Let $A_0 : X \to Y_0$ be defined by $A_0 x = Ax$ for all $x \in X$. Clearly, A_0 is a bounded linear operator from X onto Y_0. We show that $R(A_0^*)$ is closed and $R(A_0^*) = R(A^*)$.

Since $A_0 : X \to Y_0$ is a surjective bounded linear operator, by a consequence of open mapping theorem (Corollary 7.11), there exists $c > 0$ such that for every $y \in Y_0$, there exists $x \in X$ satisfying $A_0 x = y$ and $\|x\| \le c\,\|y\|$.

Now, let $c > 0$ be as above, $y \in Y_0$, and let $x \in X$ be such that $A_0 x = y$ and $\|x\| \le c\,\|y\|$. Then, for every $u \in Y_0$, we have

$$|\langle y, u \rangle| = |\langle A_0 x, u \rangle| = |\langle x, A_0^* u \rangle| \le \|x\| \, \|A_0^* u\| \le c\,\|y\| \, \|A_0^* u\|.$$

This is true for all $y \in Y_0$. Hence,

$$\|u\| \le c\,\|A_0^* u\| \quad \forall u \in Y_0$$

so that A_0^* is bounded below and ,consequently, by Proposition 3.3, $R(A_0^*)$ is closed.

It remains to be shown that $R(A^*) = R(A_0^*)$. For this, observe that

$$\langle A_0 x, y\rangle = \langle Ax, y\rangle = \langle x, A^* y\rangle \quad \forall x \in X, y \in Y_0.$$

Hence, it follows that

$$A_0^* y = A^* y \quad \forall y \in Y_0.$$

Since $Y_0 = R(A)$ is closed in Y, by projection theorem, we have $Y = R(A) + R(A)^\perp$, and by Proposition 11.4, $R(A)^\perp = N(A^*)$. Hence we have

$$\begin{aligned} R(A_0^*) &= \{A_0^* y : y \in R(A)\} = \{A^* y : y \in R(A)\} \\ &= \{A^* y : y \in Y\} = R(A^*). \end{aligned}$$

The converse part follows from the first part by observing the relation $A = (A^*)^*$. ∎

11.1.3 Integration of Operator-Valued Functions Revisited

Using Theorem 11.6, we prove Theorem 10.18 in the setting of Hilbert spaces.

Theorem 11.8 *Let X and Y be Hilbert spaces over $\mathbb{C}$, and Γ be a rectifiable curve in $\mathbb{C}$. Let $F : \Gamma \to \mathcal{B}(X, Y)$ be a continuous function. Then there exists a unique operator $\Phi \in \mathcal{B}(X, Y)$ such that*

$$\langle \Phi x, y\rangle = \int_\Gamma \langle F(z)x, y\rangle\, dz \quad \forall (x, y) \in X \times Y.$$

Moreover,

$$\Phi = \int_\Gamma F(z)\, dz,$$

where the integral is in the sense of Riemann–Stieltjes, defined in Section 10.3.

Proof. We first note that for each $(x, y) \in X \times Y$, the function $f_{x,y} : \Gamma \to \mathbb{C}$ defined by

$$f_{x,y}(z) = \langle F(z)x, y\rangle, \qquad z \in \Gamma,$$

is a continuous function on Γ, and hence the integral

$$\int_\Gamma \langle F(z)x, y\rangle\, dz$$

is well defined. It is easily seen that the function $\psi : X \times Y \to \mathbb{C}$ defined by

$$\psi(x,y) = \int_\Gamma \langle F(z)x, y\rangle\, dz, \qquad (x,y) \in X \times Y,$$

is a bounded sesquilinear functional. Hence, by Theorem 11.6, there exists a unique operator $\Phi \in \mathcal{B}(X,Y)$ such that

$$\langle \Phi x, y\rangle = \int_\Gamma \langle F(z)x, y\rangle\, dz \quad \forall\, (x,y) \in X \times Y.$$

Consider a parametrization $\gamma : [0,1] \to \mathbb{C}$ of Γ. Let

$$\Pi_n = \{t_j : j = 1, \ldots, n\}$$

be a partition of $[0,1]$. Let

$$\Phi_n := \sum_{j=1}^{n} [\gamma(t_j) - \gamma(t_{j-1}]F(\gamma(s_j)),$$

the Riemann–Stieltjes sum for F associated with Π_n and the set $\Delta_n = \{s_j : j = 1, \ldots, n\}$ of tags, i.e., $t_{j-1} \le s_j \le t_j$, $j = 1, \ldots, n$. If the sequence (Π_n) is such that $m(\Pi_n) \to 0$ as $n \to \infty$, then we see that

$$\begin{aligned}
\langle \Phi_n x, y\rangle &= \sum_{j=1}^{n} [\gamma(t_j) - \gamma(t_{j-1}]\langle F(\gamma(s_j))x, y\rangle \\
&\to \int_\Gamma \langle F(z)x, y\rangle\, dz \\
&= \langle \Phi x, y\rangle
\end{aligned}$$

as $n \to \infty$, for every $(x,y) \in X \times Y$. To complete the proof, it is sufficient to prove that $\|\Phi - \Phi_n\| \to 0$ as $n \to \infty$. This is done as follows: For each $j = 1, \ldots, n$, let Γ_j be the part of Γ joining t_{j-1} to t_j. Then, it follows that

$$\langle (\Phi - \Phi_n)\, x, y\rangle = \sum_{j=1}^{n} \int_{\Gamma_j} \langle (F(z) - F(\gamma(s_j))\, x, y\rangle\, dz.$$

Hence, using uniform continuity of $z \mapsto F(z)$, it can be seen (*give details*) that for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|\langle(\Phi - \Phi_n)x, y\rangle| \leq \ell(\Gamma)\varepsilon\|x\| \, \|y\| \quad \forall (x, y) \in X \times Y,$$

where $\ell(\Gamma) > 0$ is the length of Γ. Hence, we get

$$\|\Phi - \Phi_n\| \leq \ell(\Gamma)\varepsilon \quad \forall n \geq N.$$

Thus we have shown that $\|\Phi - \Phi_n\| \to 0$ as $n \to \infty$. ∎

11.1.4 Adjoint of an Unbounded Operator

We may observe that for proving Theorem 11.2, the continuity of the operator A is used only in establishing the continuity of the functional $f_y : x \mapsto \langle Ax, y\rangle$, $x \in X$, for each $y \in Y$. One may wonder whether we can define the adjoint of an unbounded linear operator A by restricting the domain of the adjoint (to be defined) to be those $y \in Y$ for which the function $f_y : x \mapsto \langle Ax, y\rangle$, $x \in X$, is continuous. That is exactly what we do for unbounded operators. It is important to define the adjoint of an unbounded operator, as most of the differential operators are unbounded operators. Let us see how we can do this.

Let X be a Hilbert space, Y be an inner product space and $A : X_0 \to Y$ be a linear operator defined on a *dense* subspace X_0 of X. For $y \in Y$, let $f_y : X_0 \to \mathbb{K}$ be defined by

$$f_y(x) = \langle Ax, y\rangle, \qquad x \in X_0.$$

Clearly, for every $y \in Y$, $f_y : X_0 \to \mathbb{K}$ is a linear functional on X, and is continuous if and only if there exists $c_y > 0$ such that

$$|\langle Ax, y\rangle| \leq c_y\|x\| \quad \forall x \in X_0.$$

Let

$$Y_0 = \{y \in Y : f_y \text{ continuous}\}.$$

It is seen that Y_0 is a subspace of Y. By Theorem 3.18, we know that, for every $y \in Y_0$, there exists a unique continuous linear functional $\tilde{f}_y : X \to \mathbb{K}$ which is an extension of f_y, and by the Riesz representation theorem (Theorem 3.9), there exists a unique $u_y \in X$ such that

$$\tilde{f}_y(x) = \langle x, u_y\rangle \quad \forall x \in X.$$

In particular, we have

$$\langle Ax, y\rangle = f_y(x) = \tilde{f}_y(x) = \langle x, u_y\rangle \quad \forall x \in X_0,\ y \in Y_0.$$

It is easily seen that the map $B : Y_0 \to X$ defined by $By = u_y$, $y \in Y_0$, is a linear operator, and it satisfies the relation

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X_0,\ y \in Y_0.$$

Using the fact that X_0 is dense in X, it can be seen that there is only one such operator B. Thus we have proved the following theorem.

Theorem 11.9 *Let X be a Hilbert space, Y be an inner product space and $A : X_0 \to Y$ be a linear operator defined on a dense subspace X_0 of X. Let $Y_0 = \{y \in Y : x \mapsto \langle Ax, y\rangle$ continuous on $X_0\}$. Then there exists a unique linear operator $B : Y_0 \to X$ such that*

$$\langle Ax, y\rangle = \langle x, By\rangle \quad \forall x \in X_0,\ y \in Y_0.$$

The operator $B : Y_0 \to X$ in the above theorem is also called the **adjoint** of the operator $A : X_0 \to Y$, and is also denoted by A^*.

In the study of unbounded linear operators, the domain of a linear operator A is usually denoted by $D(A)$.

Now, we give a few examples of unbounded operators and their adjoints.

EXAMPLE 11.4 (i). Let $X = \ell^2$, $X_0 = c_{00}$, and $A : X_0 \to X$ be defined by

$$Ax = \left(\sum_{j=1}^{\infty} x(j)\right) e_1, \qquad x \in X_0.$$

Clearly, X_0 is a dense subspace of X, and for $x \in c_{00}$, $y \in \ell^2$,

$$\langle Ax, y\rangle = \sum_{j=1}^{\infty} x(j)\, \overline{y(1)}.$$

Hence, for $y \in X$, the linear functional $f_y : X_0 \to \mathbb{K}$ defined by

$$f_y(x) = \langle Ax, y\rangle, \qquad x \in X_0,$$

is continuous if and only if $y(1) = 0$ (*Verify*). Thus, it follows that

$$D(A^*) = \{y \in \ell^2 : y(1) = 0\}, \qquad A^*y = 0 \quad \forall y \in D(A^*).$$

In this example, we may observe that A is neither a continuous nor a closed operator, and $D(A^*)$ is a closed proper subspace of X.

(ii) Let $X = \ell^2$ and (λ_n) be a sequence of positive scalars such that $\lambda_n \to 0$ as $n \to \infty$. Let

$$X_0 = \left\{ x \in \ell^2 : \sum_{j=1}^{\infty} \frac{|x(j)|^2}{\lambda_j^2} < \infty \right\},$$

and $A : X_0 \to X$ be defined by

$$(Ax)(j) = \frac{x(j)}{\lambda_j}, \qquad x \in X; \ j \in \mathbb{N}.$$

We have seen in Example 3.12 that A is a closed operator which is not continuous. It can be easily seen that $D(A^*) = X_0$ and $A^*x = Ax$ for all $x \in X_0$.

(iii) Let $X = L^2[0,1]$ and $X_0 = \{x \in C^1[0,1] : x(0) = x(1)\}$. Let $A : X_0 \to X$ be defined by

$$Ax = x', \qquad x \in X_0.$$

Then it can be seen that

$$A^* = -A.$$

(iv) Let $X = L^2[0,1]$ be over $\mathbb{C}$, and let X_0 be as in (iii) above. Let $A : X_0 \to X$ be defined by

$$Ax = i\,x', \qquad x \in X_0.$$

Then it can be seen that

$$A^* = A.$$

Exercise 11.5 Let X and Y be Hilbert spaces. Prove the following:

(i) If $A : X_0 \subseteq X \to Y$ is a densely defined operator, then the adjoint operator $A^* : Y_0 \subseteq Y \to X$ is a closed operator.

(ii) If $A : X_0 \subseteq X \to Y$ is a closed densely defined operator, then its adjoint $A^* : Y_0 \subseteq Y \to X$ is a closed densely defined operator, and in that case $D(A^{**}) = D(A)$ and $A = A^{**}$. Here, A^{**} is the adjoint of A^*. $\square$

11.2 Self-Adjoint, Normal and Unitary Operators

In this section we consider a few important classes of operators which are defined using the adjoint of an operator, and which have many interesting spectral properties.

We know from Theorem 4.10 that if X is a separable Hilbert space and if E is an orthonormal basis of X, then E is countable. We may write such countable set E as $\{u_j : j \in \Lambda\}$, where $\Lambda = \mathbb{N}$ if E is denumerable, and $\Lambda = \{1, \ldots, k\}$ if E is a finite set consisting of k elements.

Let X be a Hilbert space and $A \in \mathcal{B}(X)$. Then A is said to be a

(a) **self-adjoint operator** if $A^* = A$,

(b) **normal operator** if $A^*A = AA^*$,

(c) **unitary operator** if $A^*A = I = AA^*$.

We observe that for every $A \in \mathcal{B}(X)$, the operators A^*A and AA^* are self-adjoint operators. Also, the identity operator is self-adjoint, and self-adjoint operators and unitary operators are normal. But a normal operator need not be unitary or self-adjoint, as the following example shows.

EXAMPLE 11.5 Let $X = \mathbb{C}^2$, and for $\lambda_1, \lambda_2 \in \mathbb{C}$, let

$$A : (\alpha_1, \alpha_2) \mapsto (\lambda_1\alpha_1, \lambda_2\alpha_2), \qquad (\alpha_1, \alpha_2) \in \mathbb{C}^2.$$

Then it is seen that

(a) A is a normal operator,

(b) A is self-adjoint if and only if $\lambda_1, \lambda_2 \in \mathbb{R}$, and

(c) A is unitary if and only if $|\lambda_1| = 1 = |\lambda_2|$.

EXAMPLE 11.6 Let X be a separable Hilbert space, and let $\{u_j : j \in \Lambda\}$ be an orthonormal basis of X.

(i) Let $\{\lambda_j : j \in \Lambda\}$ be a bounded set of scalars, and let

$$Ax = \sum_j \lambda_j \langle x, u_j\rangle u_j, \qquad x \in X.$$

We know from Example 11.3(i) that $A \in \mathcal{B}(X)$, and its adjoint is given by

$$A^*x = \sum_j \bar{\lambda}_j \langle x, u_j \rangle u_j, \qquad x \in X$$

and we see that

(a) A is a normal operator,

(b) A is self-adjoint if and only if $\lambda_j \in \mathbb{R}$ for all $j \in \Lambda$, and

(c) A is unitary if and only if $|\lambda_j| = 1$ for all $j \in \Lambda$.

(ii) Let (a_{ij}) be a matrix of scalars such that

$$Ax = \sum_i \Big(\sum_j a_{ij} \langle x, u_j \rangle \Big) v_i, \qquad x \in X$$

defines a bounded linear operator on X. We have seen in Example 11.3(ii) that

$$A^*y = \sum_i \Big(\sum_j \overline{a_{ji}} \langle y, v_j \rangle \Big) u_i, \qquad y \in Y.$$

Thus, that A is self-adjoint if and only if $a_{ij} = \overline{a_{ji}}$.

(iii) Let $X = L^2[a,b]$, $\phi \in L^\infty[a,b]$ and A be defined by

$$Ax = \phi x, \qquad x \in X.$$

We have seen that $A \in \mathcal{B}(X)$, and $A^*y = \bar{\phi} y$ for every $y \in X$ so that

$$A^*Ax = |\phi|^2 x \quad \forall x \in X.$$

Hence,

(a) A is normal,

(b) A is self-adjoint if and only if ϕ is real-valued, and

(c) A is unitary if and only if $|\phi(t)| = 1$ a.e. on $[a,b]$.

Recall that if $A : X \to X$ is a bounded linear operator, then

$$\|A\| = \sup \{ |\langle Ax, y \rangle| : x, y \in X, \ \|x\| = 1 = \|y\| \}.$$

The following theorem shows that if A is a self-adjoint operator, then the set on which supremüm is taken can be substantially small.

Theorem 11.10 *Let X be a Hilbert space and $A \in \mathcal{B}(X)$ be a self-adjoint operator. Then*

$$\|A\| = \sup\{|\langle Ax, x\rangle| : x \in X,\ \|x\| = 1\}.$$

In particular, $A = 0$ if and only if $\langle Ax, x\rangle = 0$ for all $x \in X$.

Proof. Let

$$r = \sup\{|\langle Ax, x\rangle| : x \in X, \|x\| = 1\}.$$

Clearly, $r \leq \|A\|$. We show that $\|Ax\| \leq r$ for every $x \in X$ with $\|x\| = 1$. Observe that for every $x, y \in X$,

$$\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle = 4\,\mathrm{Re}\,\langle Ax, y\rangle$$

so that if $\|x\| = 1 = \|y\|$, then

$$\begin{aligned} 4\,\mathrm{Re}\langle Ax, y\rangle &\leq |\langle A(x+y), x+y\rangle| + |\langle A(x-y), x-y\rangle| \\ &\leq r\left(\|x+y\|^2 + \|x-y\|^2\right) \\ &\leq 2r\left(\|x\|^2 + \|y\|^2\right) \\ &= 4r. \end{aligned}$$

Let $x \in X$ be such that $\|x\| = 1$. If $Ax = 0$, then obviously, $\|Ax\| \leq r$. If $Ax \neq 0$, then taking $y = Ax/\|Ax\|$, we have $\|y\| = 1$ and

$$\|Ax\| = \langle Ax, y\rangle = \mathrm{Re}\langle Ax, y\rangle$$

so that we have $\|Ax\| \leq r$. Thus, we have proved that $\|A\| \leq r$. The last part of the theorem is, now, obvious. ∎

The above result shows that if A and B are self-adjoint operators on a Hilbert space X, then

$$A = B \iff \langle Ax, x\rangle = \langle Bx, x\rangle \quad \forall x \in X.$$

11.2.1 Numerical Range and Numerical Radius

For $A \in \mathcal{B}(X)$, the set

$$w(A) := \{\langle Ax, x\rangle : x \in X, \|x\| = 1\}$$

is called the **numerical range** of A, and the quantity

$$r_w(A) := \sup\{|\lambda| : \lambda \in w(A)\}$$

is called the **numerical radius** of A.

The elements of $w(A)$ are also called **Rayleigh quotients.**

It is obvious from the definition that if $A \in \mathcal{B}(X)$ is a self-adjoint operator, then $\langle Ax, x\rangle \in \mathbb{R}$ for every $x \in X$. This observation, together with Theorem 11.10, yields the following result.

Theorem 11.11 *If $A \in \mathcal{B}(X)$ is a self-adjoint operator, then*

$$w(A) \subseteq \mathbb{R}, \qquad r_w(A) = \|A\|.$$

In fact, if the scalar field is $\mathbb{C}$, then the condition $w(A) \subseteq \mathbb{R}$ is also sufficient for A to be a self-adjoint operator, as the following result shows.

Theorem 11.12 *Let X be a Hilbert space over $\mathbb{C}$ and $A \in \mathcal{B}(X)$ be such that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in X$. Then A is self-adjoint.*

Proof. Let $x, y \in X$. We show that $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in X$. By hypothesis, we have

$$\langle A(x+y), x+y\rangle = \langle x+y, A(x+y)\rangle.$$

From this we get

$$\langle Ax, y\rangle + \langle Ay, x\rangle = \langle x, Ay\rangle + \langle y, Ax\rangle.$$

Hence

$$\operatorname{Im}\langle Ax, y\rangle = \operatorname{Im}\langle x, Ay\rangle.$$

Again, by hypothesis, taking iy in place of y, we have

$$\langle A(x+iy), x+iy\rangle = \langle x+iy, A(x+iy)\rangle$$

so that

$$-i\langle Ax, y\rangle + i\langle Ay, x\rangle = -i\langle x, Ay\rangle + i\langle y, Ax\rangle.$$

Therefore,

$$\operatorname{Re}\langle Ax, y\rangle = \operatorname{Re}\langle x, Ay\rangle.$$

Thus,

$$\langle Ax, y\rangle = \langle x, Ay\rangle \quad \forall x, y \in X,$$

showing that A is a self-adjoint operator. ∎

It is easily seen (*Verify*) that the above theorem does not hold if the scalar field is $\mathbb{R}$.

A linear operator $A : X \to X$ is said to be a **positive operator** if

$$w(A) \subseteq [0, \infty),$$

i.e., if $\langle Ax, x\rangle \geq 0$ for all $x \in X$.

If $A \in \mathcal{B}(X)$ is a positive operator, then we write $A \geq 0$, and if $A_1, A_2 \in \mathcal{B}(X)$ are such that $A_1 - A_2 \geq 0$, then we write $A_1 \geq A_2$ or $A_2 \leq A_1$.

Exercise 11.6 Let $A \in \mathcal{B}(X)$ be a positive operator.

(i) Prove that, if $\mathbb{K} = \mathbb{C}$, then A is a self-adjoint operator.

(ii) Show that the conclusion in (i) need not hold if $\mathbb{K} = \mathbb{R}$. □

11.2.2 Some Characterizations

Theorem 11.13 *Let X be a Hilbert space and $A \in \mathcal{B}(X)$. Then*

(i) *A is normal if and only if $\|Ax\| = \|A^*x\|$ for every $x \in X$,*

(ii) *A is unitary if and only if A is surjective and $\|Ax\| = \|x\|$ for every $x \in X$.*

In particular, if A is a unitary operator, then $\|A\| = 1$.

Proof. We observe that for every $x \in X$,

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle,$$

$$\|A^*x\|^2 = \langle A^*x, A^*x\rangle = \langle AA^*x, x\rangle.$$

Since A^*A and AA^* are self-adjoint operators, it follows from the above relations and Theorem 11.10 that

$$A \quad \text{is normal} \iff \|A^*x\| = \|Ax\| \quad \forall x \in X.$$

Now, suppose that A is unitary. Then it is surjective, and

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = \langle x, x\rangle = \|x\|^2$$

for all $x \in X$. Conversely, suppose that A is surjective and $\|Ax\| = \|x\|$ for every $x \in X$, i.e.,

$$\langle (A^*A - I)x, x\rangle = 0 \quad \forall x \in X.$$

Again, using Theorem 11.10 and the fact that $A^*A - I$ is a self-adjoint operator, we see that $A^*A = I$. This also implies that A is injective. Thus, A is invertible in $\mathcal{B}(X)$ and $A^* = A^{-1}$. Thus, $A^*A = I = AA^*$. ∎

One of the ways to extend some of the results available for self-adjoint operators to all bounded operators is by observing the following:

Let X be a complex Hilbert space and $A \in \mathcal{B}(X)$. Then, we have

$$A = B + iC \quad \text{with} \quad B = \frac{A + A^*}{2}, \quad C = \frac{A - A^*}{2i}.$$

We observe that B and C are self-adjoint operators. In addition,

(a) A is self-adjoint if and only if $C = O$,

(b) A is normal if and only if $BC = CB$, and

(c) A is unitary if and only if $BC = CB$ and $B^2 + C^2 = I$.

The above representation of $A = B + iC$, with B and C self-adjoint, is akin to the representation of a complex number z as $z = x + iy$ with $x, y \in \mathbb{R}$. Note that

$x = (z + \bar{z})/2$ and $y = (z - \bar{z})/2i$,

z is real if and only if $y = 0$, and

z is on the unit circle if and only if $x^2 + y^2 = 1$.

Because of this analogy, the operators

$$B = \frac{A + A^*}{2}, \qquad C = \frac{A - A^*}{2i}$$

are often called the *real* and *imaginary parts* of A.

We now consider a subclass of the class of compact operators, namely, the class of *Hilbert-Schmidt operators.*

11.3 Hilbert-Schmidt Operators

Let X be an infinite dimensional separable Hilbert space, and let $\{u_1, u_2, \ldots\}$ be an orthonormal basis of X. A linear operator A on X is said to be a **Hilbert-Schmidt operator** if

$$\sum_{i,j} |\langle Au_j, u_i \rangle|^2 < \infty.$$

The following theorem shows that the above definition is independent of the choice of the orthonormal basis.

Theorem 11.14 *Let X be a separable Hilbert space and A be a linear operator on X. Suppose $\{u_1, u_2, \ldots\}$ and $\{v_1, v_2, \ldots\}$ are orthonormal bases of X. Then*

$$\sum_{i,j} |\langle Au_j, u_i\rangle|^2 < \infty \iff \sum_{i,j} |\langle Av_j, v_i\rangle|^2 < \infty.$$

Proof. By Parseval's formula (Theorem 4.9 (iii)), we have

$$\sum_i |\langle Av_j, v_i\rangle|^2 = \|Av_j\|^2 = \sum_i |\langle Av_j, u_i\rangle|^2 = \sum_i |\langle v_j, A^*u_i\rangle|^2.$$

Therefore,

$$\sum_{i,j} |\langle Av_j, v_i\rangle|^2 = \sum_i \|A^*u_i\|^2 = \sum_{i,j} |\langle u_j, A^*u_i\rangle|^2 = \sum_{i,j} |\langle Au_j, u_i\rangle|^2.$$

Thus the result follows. ∎

We may observe (*Verify*) that if $A : X \to X$ is a Hilbert-Schmidt operator, then

$$\|Ax\|^2 \le \Big(\sum_{i,j} |\langle Au_j, u_i\rangle|^2\Big)\|x\|^2$$

so that

$$A \in \mathcal{B}(X), \qquad \|A\| \le \Big(\sum_{i,j} |\langle Au_j, u_i\rangle|^2\Big)^{1/2}$$

In fact, we have the following result.

Theorem 11.15 *Let X be a separable Hilbert space. Then every Hilbert-Schmidt operator on X is a compact operator.*

Proof. Suppose $A : X \to X$ is a Hilbert-Schmidt operator, i.e.,

$$\sum_{i,j} |\langle Au_j, u_i\rangle|^2 < \infty,$$

where $\{u_1, u_2, \ldots\}$ is an orthonormal basis of X. Consider the operator

$$\tilde{A} = TAT^{-1},$$

where $T : X \to \ell^2$ is defined by

$$Tx = (\langle x, u_1 \rangle, \langle x, u_2 \rangle, \ldots), \qquad x \in X.$$

Recall from Theorem 4.12 that T is a surjective linear isometry between X and ℓ^2. It is seen that

$$(\tilde{A}v)(i) = \sum_j \langle Au_j, u_i \rangle v(j), \qquad v \in \ell^2;\ i \in \mathbb{N}.$$

Since $\sum_{i,j} |\langle Au_j, u_i \rangle|^2 < \infty$, it follows from Example 9.2(v) that $\tilde{A}$ is a compact operator. Hence, by Theorem 9.2(ii), $A = T^{-1}\tilde{A}T$ is a compact operator. ∎

What about the converse of the above theorem? Well, not every compact operator is a Hilbert-Schmidt operator. Here is an example.

EXAMPLE 11.7 Let $X = \ell^2$ and

$$Ax = \sum_{n=1}^{\infty} \frac{x(n)}{\sqrt{n}} e_n, \qquad x \in \ell^2.$$

Then, by Example 9.2(i), A is a compact operator on ℓ^2, but

$$\sum_{i,j} |\langle Ae_j, e_i \rangle|^2 \not< \infty$$

so that A is not a Hilbert-Schmidt operator.

How often do Hilbert-Schmidt operators appear in applications, and how big is the class of Hilbert-Schmidt operators? Before attempting to answer this question, we prove a result concerning orthonormal basis of $L^2([a,b] \times [a,b])$. We may recall that the inner product in $L^2([a,b] \times [a,b])$ is given by

$$\langle x, y \rangle := \int_a^b \int_a^b x(s,t)\overline{y(s,t)}\, d\mu(s)\, d\mu(t), \qquad x, y \in L^2([a,b] \times [a,b]).$$

Proposition 11.16 *Suppose $\{u_1, u_2, \ldots\}$ is an orthonormal basis of $L^2[a,b]$ and*

$$\varphi_{ij}(s,t) = u_i(s)\overline{u_j(t)}$$

for $(s,t) \in [a,b] \times [a,b]$, $(i,j) \in \mathbb{N} \times \mathbb{N}$. Then

$$\{\varphi_{ij} : (i,j) \in \mathbb{N} \times \mathbb{N}\}$$

is an orthonormal basis of $L^2([a,b] \times [a,b])$.

Proof. It is seen that for $i, j, k, l \in \mathbb{N}$,

$$\langle \varphi_{ij}, \varphi_{kl} \rangle = \langle u_i, u_k \rangle \langle u_j, u_l \rangle$$

so that $\{\varphi_{ij} : i, j \in \mathbb{N}\}$ is an orthonormal set in $L^2([a, b] \times [a, b])$.

Now, suppose that $f \in L^2([a, b] \times [a, b])$ be such that $\langle f, \varphi_{ij} \rangle = 0$ for all $i, j \in \mathbb{N}$. Corresponding to this f, let $K : L^2[a, b] \to L^2[a, b]$ be defined by

$$(Kx)(s) = \int_a^b f(s, t)x(t)\, d\mu(t), \qquad x \in L^2[a, b];\ s \in [a, b].$$

Then we have

$$\begin{aligned} \langle f, \varphi_{ij} \rangle &= \int_a^b \int_a^b f(s, t)\overline{u_i(s)}u_j(t)d\mu(s)\, d\mu(t) \\ &= \int_a^b \left(\int_a^b f(s, t)u_j(t)\, d\mu(t) \right) \overline{u_i(s)}d\mu(s) \\ &= \langle Ku_j, u_i \rangle \end{aligned}$$

for all $i, j \in \mathbb{N}$. Since $\{u_1, u_2, \ldots\}$ is an orthonormal basis of $L^2[a, b]$, and since $\langle f, \varphi_{ij} \rangle = 0$ for all i,j, it follows that

$$Kx = 0 \quad \forall x \in L^2[a, b].$$

Consequently, $f = 0$. Thus, by Theorem 4.2, $\{\varphi_{ij} : i, j \in \mathbb{N}\}$ is an orthonormal basis for $L^2([a, b] \times [a, b])$. ∎

The next result shows that the class of Hilbert-Schmidt operators on $L^2[a, b]$ is precisely the set of all Fredholm integral operators with kernels belonging to $L^2([a, b] \times [a, b])$.

Theorem 11.17 *For $k(\cdot, \cdot) \in L^2([a, b] \times [a, b])$, let*

$$(Ax)(s) = \int_a^b k(s, t)x(t)\, d\mu(t), \quad x \in L^2[a, b].$$

Then $A : L^2[a, b] \to L^2[a, b]$ is a Hilbert-Schmidt operator on $L^2[a, b]$. Conversely, every Hilbert-Schmidt operator A on $L^2[a, b]$ is precisely of the above form for some $k(\cdot, \cdot) \in L^2([a, b] \times [a, b])$.

Proof. Let $\{u_1, u_2, \ldots\}$ be an orthonormal basis of $L^2[a,b]$. Then, using the fact that $\{\varphi_{ij} : (i,j) \in \mathbb{N} \times \mathbb{N}\}$, with φ_{ij} defined as in Proposition 11.16, is an orthonormal basis for $L^2([a,b] \times [a,b])$, it follows that

$$k = \sum_{i,j} \langle k, \varphi_{ij} \rangle \varphi_{ij}, \qquad \sum_{i,j} |\langle k, \varphi_{ij} \rangle|^2 = \|k\|^2.$$

Also, we see that

$$\langle k, \varphi_{ij} \rangle = \langle Au_j, u_i \rangle \quad \forall\, (i,j) \in \mathbb{N} \times \mathbb{N}$$

so that

$$\sum_{i,j} |\langle Au_j, u_i \rangle|^2 = \sum_{i,j} |\langle k, \varphi_{ij} \rangle|^2 = \|k\|_2^2.$$

Therefore, A is a Hilbert-Schmidt operator.

Conversely, suppose that A is a Hilbert-Schmidt operator on $L^2[a,b]$. Define

$$k = \sum_{i,j} \langle Au_j, u_i \rangle \varphi_{ij}.$$

Then

$$\|k\|_2^2 = \sum_{i,j} |\langle Au_j, u_i \rangle|^2 < \infty$$

so that $k \in L^2([a,b] \times [a,b])$. Also,

$$\langle k, \varphi_{ij} \rangle = \langle Au_j, u_i \rangle \quad \forall\, (i,j) \in \mathbb{N} \times \mathbb{N}.$$

Next, define

$$(Bx)(s) = \int_a^b k(s,t)x(t)\, d\mu(t), \qquad x \in L^2[a,b].$$

Then it follows that

$$\langle Bu_j, u_i \rangle = \langle k, \varphi_{ij} \rangle = \langle Au_j, u_i \rangle \quad \forall\, (i,j) \in \mathbb{N} \times \mathbb{N}$$

showing that $B = A$. ∎

Exercise 11.7 Let $\{u_1, u_2, \ldots\}$ be an orthonormal basis of a separable Hilbert space X and $HS(X)$ be the set of all Hilbert-Schmidt operators on X. For $A, B \in HS(X)$, define

$$\langle A, B \rangle_{HS} = \sum_j \langle Au_j, Bu_j \rangle.$$

Prove the following:

(i) $HS(X)$ with $\langle \cdot, \cdot \rangle_{HS}$ is an inner product space.

(ii) $\langle \cdot, \cdot \rangle_{HS}$ is independent of the choice of the orthonormal basis.

(iii) $HS(X)$ is a Hilbert space with respect to $\langle \cdot, \cdot \rangle_{HS}$. $\square$

PROBLEMS

In the following problems, X and Y are Hilbert spaces.

1. Suppose $\dim X < \infty$ and $A : X \to Y$ is a linear operator. If $\{u_1, \ldots, u_n\}$ is an orthonormal basis of X, then show that

$$\|A\| \leq \Big(\sum_{j=1}^{n} \|Au_j\|^2\Big)^{1/2}.$$

2. Show that a sesquilinear functional $\psi : X \times X \to \mathbb{K}$ is bounded if and only if it is continuous.

3. Let X be a complex Hilbert space and $\psi : X \times X \to \mathbb{K}$ be a sesquilinear functional. Let $q : X \to \mathbb{C}$ be the associated *quadratic form*, i.e., $q(x) = \psi(x, x)$ for all $x \in X$. Verify the following:

(a) $4\,\psi(x, y) = q(x + y) - q(x - y) + iq(x + iy) - iq(x - iy)$.

(b) ψ is *symmetric*, i.e., $\psi(x, y) = \overline{\psi(y, x)}$, if and only if q is real-valued.

4. Let $A \in \mathcal{B}(X)$ and M be a subspace of X. Show that M is invariant under A if and only if $M^\perp$ is invariant under A^*.

5. Let $A \in \mathcal{B}(X, Y)$. Show that $R(A)$ is dense in Y if and only if A^* is injective.

6. Show that for $A \in \mathcal{B}(X, Y)$, if $R(A) = Y$, then A^* is bounded below.

[*Hint:* One may use the closed range theorem (Theorem 11.7) and the open mapping theorem (Theorem 7.9).]

7. Suppose $A \in \mathcal{B}(X, Y)$ such that the equation $A^* u = v$ is solvable for every $v \in X$. Show that the equation $Ax = y$ is uniquely solvable for every $y \in R(A)$, and the solution is stable, in the sense that if

(y_n) in $R(A)$ is such that $y_n \to y$ as $n \to \infty$ for some $y \in R(A)$, then $x_n \to x$ as $n \to \infty$, where $Ax = y$ and $Ax_n = y_n$ for all $n \in \mathbb{N}$.

8. Let $P \in \mathcal{B}(X)$ be a projection operator. Show that P is self-adjoint if and only if it is orthogonal, i.e., $R(P) \perp N(P)$, and in that case, P is a positive operator.

9. Let $A \in \mathcal{B}(X)$ be a self-adjoint operator and $P \in \mathcal{B}(X)$ be an orthogonal projection. Show that $PA|_{R(P)} : R(P) \to R(P)$ is a self-adjoint operator on $R(P)$.

10. Let $A \in \mathcal{B}(X)$ be a self-adjoint operator. Show that

$$\|A^n\| = \|A\|^n \quad \forall n \in \mathbb{N}.$$

11. Let $A \in \mathcal{B}(X)$ be a normal operator. Show that

$$(A^*A)^n = (A^*)^n A^n \quad \forall n \in \mathbb{N}.$$

Deduce that

$$\|A^n\| = \|A\|^n \quad \forall n \in \mathbb{N}.$$

[*Hint*: Use the self-adjointness of A^*A and the previous problem.]

12. For $A \in \mathcal{B}(X)$, let $\exp(A) := \sum_{n=1}^{\infty} A^n/n!$. Show that, if A is a self-adjoint operator, then $\exp(iA)$ is a unitary operator.

13. Give an example of a Hilbert space X and $A \in \mathcal{B}(X)$ such that $A^*A = I$, but $AA^* \neq I$. Can this happen if $\dim X < \infty$?

14. Let X be a separable Hilbert space and $\{u_1, u_2, \ldots\}$ be an orthonormal basis of X. Show that a linear operator A on X is a Hilbert-Schmidt operator if and only if $\sum_j \|Au_j\|^2 < \infty$.

12

Spectral Results for Hilbert Space Operators

In this chapter we derive certain spectral results which are special to operators on Hilbert spaces.

Recall from Chapter 10 that if X is a Banach space and A is a bounded linear operator on X, then then the eigenspectrum, approximate eigenspectrum, and the spectrum of A are the sets

$$\begin{aligned}
\sigma_{\rm eig}(A) &= \{\lambda \in \mathbb{K} : A - \lambda I \text{ not injective}\}, \\
\sigma_{\rm app}(A) &= \{\lambda \in \mathbb{K} : A - \lambda I \text{ not bounded below}\}, \\
\sigma(A) &= \{\lambda \in \mathbb{K} : A - \lambda I \text{ not bijective}\},
\end{aligned}$$

respectively.

Throughout this chapter, we assume that X is a Hilbert space.

12.1 Some Properties of the Spectrum

Theorem 12.1 *Let $A \in \mathcal{B}(X)$. Then*

$$\sigma(A) = \sigma_{\rm app}(A) \cup \{\lambda : \bar{\lambda} \in \sigma_{\rm eig}(A^*)\}.$$

Proof. Let $\lambda \in \mathbb{K}$. We recall from Theorem 10.7 that $\lambda \in \sigma(A)$ if and only if either $\lambda \in \sigma_{\rm app}(A)$ or $R(A - \lambda I)$ is not dense in X. Also, by Proposition 11.4(iv), $R(A - \lambda I)$ is dense in X if and only if $N(A^* - \bar{\lambda} I) = \{0\}$, i.e., if and only if $\bar{\lambda} \notin \sigma_{\rm eig}(A^*)$. Thus the proof is complete. ∎

Next, we show that if $\mathbb{K} = \mathbb{C}$, then the spectrum of A^* is the *mirror image* of the spectrum of $\sigma(A)$ with respect to the real line.

Theorem 12.2 *Let $A \in \mathcal{B}(X)$. Then*

(i) $\lambda \in \sigma(A) \iff \bar{\lambda} \in \sigma(A^*)$,

(ii) $[(A - zI)^{-1}]^* = (A^* - \bar{z}I)^{-1} \quad \forall z \in \rho(A)$.

Proof. We observe that for $z \in \mathbb{K}$ and $B \in \mathcal{B}(X)$,

$$(A - zI)B = I = B(A - zI) \iff B^*(A^* - \bar{z}I) = (A^* - \bar{z}I)B^*.$$

Hence, it follows that

$$z \in \rho(A) \iff \bar{z} \in \rho(A^*)$$

so that the result in (i) holds. From the above observation, it also follows that if $z \in \rho(A)$ and if $B = (A - zI)^{-1}$, then $B^* = (A^* - \bar{z}I)^{-1}$. Thus we obtain the result in (ii) as well. ∎

In view of the above theorem, one may ask whether we have the relation

$$\lambda \in \sigma_{\rm eig}(A) \iff \bar{\lambda} \in \sigma_{\rm eig}(A^*). \tag{12.1}$$

The above relation holds if the space X is finite dimensional (*Verify*), but not necessarily if X is infinite dimensional.

EXAMPLE 12.1 Let $X = \ell^2$ and let A be the left shift operator, i.e.,

$$A : (\alpha_1, \alpha_2, \ldots) \mapsto (\alpha_2, \alpha_3, \ldots).$$

Then we know that A^* is the right shift operator, i.e.,

$$A^* : (\alpha_1, \alpha_2, \ldots) \mapsto (0, \alpha_1, \alpha_2, \ldots),$$

$$\{\lambda : |\lambda| < 1\} \subseteq \sigma_{\rm eig}(A), \qquad \sigma_{\rm eig}(A^*) = \varnothing.$$

12.1.1 Results for Normal and Self-Adjoint Operators

It can be easily seen that if $A \in \mathcal{B}(X)$ is a self-adjoint operator, then

$$\sigma_{\rm eig}(A) \subseteq \mathbb{R}.$$

Now, we show that (12.1) does hold if A is a normal operator. In fact, we have the following result.

Theorem 12.3 *Suppose $A \in \mathcal{B}(X)$ is a normal operator. Then, for $x \in X$ and $\lambda \in \mathbb{K}$,*

$$Ax = \lambda x \iff A^*x = \bar{\lambda}x.$$

Moreover, if $x_1, x_2 \in X$ and $\lambda_1, \lambda_2 \in \mathbb{K}$ are such that $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$ and $\lambda_1 \neq \lambda_2$, then $\langle x_1, x_2 \rangle = 0$. In particular, eigenvectors associated with distinct eigenvalues of a normal operator are orthogonal.

Proof. Let $A \in \mathcal{B}(X)$ be a normal operator and $\lambda \in \mathbb{K}$. Then we see that $A - \lambda I$ is also a normal operator, so that by Theorem 11.13(i), for every $x \in X$,

$$\|(A - \lambda I)x\| = \|(A^* - \bar{\lambda}I)x\|.$$

In particular,

$$Ax = \lambda x \iff A^*x = \bar{\lambda}x.$$

Next, suppose that $x_1, x_2 \in X$ and $\lambda_1, \lambda_2 \in \mathbb{K}$ are such that

$$Ax_1 = \lambda_1 x_1, \quad Ax_2 = \lambda_2 x_2.$$

By what we have just proved, we have $A^*x_2 = \bar{\lambda}_2 x_2$. Hence,

$$\langle \lambda_1 x_1, x_2 \rangle = \langle Ax_1, x_2 \rangle = \langle x_1, A^*x_2 \rangle = \langle x_1, \bar{\lambda}_2 x_2 \rangle = \langle \lambda_2 x_1, x_2 \rangle.$$

Thus,

$$(\lambda_1 - \lambda_2)\langle x_1, x_2 \rangle = 0.$$

Therefore, if $\lambda_1 \neq \lambda_2$ then $\langle x_1, x_2 \rangle = 0$. ∎

As a particular case of the above theorem we have

$$\lambda \in \sigma_{\text{eig}}(A) \iff \bar{\lambda} \in \sigma_{\text{eig}}(A^*)$$

for every normal operator $A \in \mathcal{B}(X)$.

As a consequence of Theorems 12.1 and 12.3, we obtain the following result.

Corollary 12.4 *If $A \in \mathcal{B}(X)$ is a normal operator, then*

$$\sigma(A) = \sigma_{\text{app}}(A).$$

Remark 12.1 We observe that if A is a normal as well as a compact operator, then the conclusion of Theorem 10.23, namely, every nonzero spectral value of A is an eigenvalue, is also a consequence of Theorem 10.4(i) and Corollary 12.4.

We have already observed that if $A \in \mathcal{B}(X)$ is a self-adjoint operator, then $\sigma_{\text{eig}}(A) \subseteq \mathbb{R}$. Now, we shall show that $\sigma(A) \subseteq \mathbb{R}$. First we shall connect the concepts of spectrum and numerical range of an operator in a nice way. Recall from Chapter 11 that if $A \in \mathcal{B}(X)$, then the *numerical range* of A is defined by

$$w(A) = \{\langle Ax, x\rangle : x \in X, \ \|x\| = 1\}$$

and the *numerical radius* of A is the number

$$r_w(A) = \sup\{|\lambda| : \lambda \in w(A)\}.$$

Recall also from Theorem 11.11 that if A is a self-adjoint operator, then

$$w(A) \subseteq \mathbb{R}.$$

We observe that if λ is an eigenvalue of A with a corresponding eigenvector x, then with $u = x/\|x\|$, we have $\lambda = \langle Au, u\rangle \in w(A)$. Thus,

$$\sigma_{\text{eig}}(A) \subseteq w(A).$$

The following theorem shows how spectrum is related to the numerical range.

Theorem 12.5 *If $A \in \mathcal{B}(X)$, then*

$$\sigma(A) \subseteq \overline{w(A)}, \qquad r_\sigma(A) \leq r_w(A).$$

In particular, if A is a self-adjoint operator, then

$$\sigma(A) \subseteq \overline{w(A)} \subseteq \mathbb{R}.$$

Proof. In view of Theorem 12.1, it is enough to prove that

$$\sigma_{\text{app}}(A) \cup \{\lambda : \bar{\lambda} \in \sigma_{\text{eig}}(A^*)\} \subseteq \overline{w(A)}.$$

Let $\lambda \in \mathbb{K}$ be such that $\bar{\lambda} \in \sigma_{\text{eig}}(A^*)$, and let $x \in X$ be such that $\|x\| = 1$ and $A^*x = \bar{\lambda}x$. Then we have

$$\lambda = \langle x, A^*x\rangle = \langle Ax, x\rangle \in w(A).$$

Now, let $\lambda \in \sigma_{\text{app}}(A)$, and let (x_n) be in X such that $\|x_n\| = 1$ for every $n \in \mathbb{N}$ and $\|Ax_n - \lambda x_n\| \to 0$ as $n \to \infty$. Then we have

$$|\langle Ax_n, x_n\rangle - \lambda| = |\langle Ax_n - \lambda x_n, x_n\rangle| \leq \|Ax_n - \lambda x_n\| \to 0$$

as $n \to \infty$. Thus, $\lambda \in \overline{w(A)}$.

The last part of the theorem is a consequence of the fact that, for self-adjoint A, $w(A) \subseteq \mathbb{R}$ (Theorem 11.11). ∎

The following result specifies a region in $\mathbb{K}$ in which spectra of unitary operators lie.

Theorem 12.6 *If $A \in \mathcal{B}(X)$ is a unitary operator, then*

$$\sigma(A) \subseteq \{\lambda : |\lambda| = 1\}.$$

Proof. Let $A \in \mathcal{B}(X)$ be a unitary operator. We have already seen in Theorem 11.13 that the norm of a unitary operator is 1. Therefore, by Theorem 10.10(i),

$$\sigma(A) \subseteq \{\lambda : |\lambda| \leq 1\}.$$

By Theorem 10.15, it also follows that

$$\sigma(A^*) = \sigma(A^{-1}) \subseteq \left\{\frac{1}{\lambda} : \lambda \in \sigma(A)\right\} \subseteq \{\lambda : |\lambda| \geq 1\}.$$

Hence, by Theorem 12.2(i),

$$\sigma(A) \subseteq \{\lambda : |\lambda| \geq 1\}.$$

Thus,

$$\sigma(A) \subseteq \{\lambda : |\lambda| = 1\}.$$

This completes the proof. ∎

We now furnish examples to illustrate that (i) the spectrum need not be a subset of the numerical range and (ii) the spectrum can be a proper subset of the numerical range.

EXAMPLE 12.2 (i) Let $X = \ell^2$, and $A : \ell^2 \to \ell^2$ be defined by

$$(Ax)(i) = \frac{x(j)}{j}, \qquad x \in \ell^2;\ i \in \mathbb{N}.$$

We have seen in Examples 10.5(i) and 10.7(i) that

$$\sigma_{\rm eig} = \{1,\ 1/2,\ 1/3, \ldots\}, \qquad \sigma(A) = \sigma_{\rm app} = \{0,\ 1,\ 1/2,\ 1/3, \ldots\}.$$

But $0 \notin w(A)$, because, for $x \in X$, $\langle Ax, x\rangle = 0$ if and only if $x = 0$. Thus, in this example, $\sigma(A) \not\subseteq w(A)$.

(ii) Let $X = \mathbb{K}^2$ and $A : (\alpha_1, \alpha_2) \mapsto (0, \alpha_2)$. Then it is seen that

$$Ae_1 = 0, \qquad Ae_2 = e_2, \qquad \langle Ae_1, e_1\rangle = 0, \qquad \langle Ae_2, e_2\rangle = 1$$

so that

$$\{0, 1\} = \sigma_{\rm eig}(A) = \sigma(A).$$

In this case, $\sigma(A)$ is a proper subset of the numerical range. To see this, note that $u := (e_1 + e_2)/\sqrt{2}$ satisfies $\|u\| = 1$ and

$$\langle Au, u\rangle = \frac{1}{2}\langle A(e_1 + e_2),\, e_1 + e_2\rangle = \frac{1}{2}\langle e_2,\, e_1 + e_2\rangle = \frac{1}{2} \notin \sigma(A).$$

In fact, in this example, we have (*Verify*)

$$w(A) = \big\{|\alpha_2|^2 : |\alpha_1|^2 + |\alpha_2|^2 = 1\big\} = [0, 1].$$

Remark 12.2 In Example 12.2(ii), the operator A is self-adjoint, and the numerical range of A is the interval with end points as the smallest and the largest spectral values of A. This is, in fact, a particular case of the so-called *Toeplitz-Hausdorff theorem* which states that the numerical range of every bounded linear operator on a Hilbert space is a convex subset of the scalar field.

In the following example, we describe the spectrum of a multiplication operator on $L^2[a, b]$.

EXAMPLE 12.3 (i) Let $X := L^2[a, b]$ and $\phi \in L^\infty[a, b]$. We know that $\phi x \in X$ whenever $x \in L^2[a, b]$ and the map $A : X \to X$ defined by

$$Ax = \phi x, \qquad x \in X,$$

is a normal operator (Example 11.6(iii)). Hence, $\sigma(A) = \sigma_{\rm app}(A)$. We show that $\sigma_{\rm app}(A)$ is the *essential range* of ϕ, i.e.,

$$\sigma_{\rm app}(A) = \{\lambda \in \mathbb{K} : \forall \varepsilon > 0,\ \mu(\{t : |\phi(t) - \lambda| < \varepsilon\}) > 0\},$$

and

$$\sigma_{\rm eig}(A) = \{\lambda \in \mathbb{K} : \mu(\{t : \phi(t) = \lambda\}) > 0\}.$$

Suppose λ is not in the essential range of ϕ, i.e., $\lambda \in \mathbb{K}$ is such that there exists $\varepsilon > 0$ with $\mu(\{t : |\phi(t) - \lambda| < \varepsilon\}) = 0$. Then, $|\phi(t) - \lambda| \geq \varepsilon$, a.e., so that for every $x \in X$, we have

$$\|Ax - \lambda x\|_2^2 = \int_a^b |\phi(t) - \lambda|^2 |x(t)|^2 \, d\mu \geq \varepsilon^2 \int_a^b |x(t)|^2 \, d\mu \geq \varepsilon^2 \|x\|_2^2.$$

Hence, $A - \lambda I$ is bounded below, and therefore, $\lambda \notin \sigma_{\text{app}}(A)$. Thus, the proof of

$$\sigma_{\text{app}}(A) \subseteq \{\lambda \in \mathbb{K} : \forall \varepsilon > 0, \ \mu(\{t : |\phi(t) - \lambda| < \varepsilon\}) > 0\}.$$

Now, suppose that λ is in the essential range of u. Then for every $n \in \mathbb{N}$, the set

$$S_n = \{t : |\phi(t) - \lambda| < 1/n\}$$

is of positive measure. For $n \in \mathbb{N}$, let x_n be the characteristic function of S_n. Then we have $\|x_n\|_2^2 = \mu(S_n) \neq 0$, and

$$\|Ax_n - \lambda x_n\|_2^2 = \int_{S_n} |\phi(t) - \lambda|^2 \, d\mu \leq \frac{1}{n^2} \mu(S_n) = \frac{1}{n^2} \|x_n\|_2^2.$$

Taking $u_n = x_n / \|x_n\|_2$, it follows that $\|u_n\|_2 = 1$ for every $n \in \mathbb{N}$, and

$$\|Au_n - \lambda u_n\| \leq \frac{1}{n} \to 0 \quad \text{as} \quad n \to \infty,$$

showing that $\lambda \in \sigma_{\text{app}}(A)$. Thus, the proof of

$$\sigma_{\text{app}}(A) = \{\lambda \in \mathbb{K} : \forall \varepsilon > 0, \ \mu(\{t : |\phi(t) - \lambda| < \varepsilon\}) > 0\}$$

is complete.

In order to show that $\sigma_{\text{eig}}(A) = \{\lambda \in \mathbb{K} : \mu(\{t : \phi(t) = \lambda\}) > 0\}$, first let $\lambda \in \sigma_{\text{eig}}(A)$, and let x be a corresponding eigenvector. Then $(\phi(t) - \lambda)x(t) = 0$ for almost all $t \in [a, b]$. Let $E = \{t \in [a, b] : x(t) \neq 0\}$. Then $\mu(E) > 0$, and from the relation $Ax = \lambda x$, i.e., $(\phi(t) - \lambda)x(t) = 0$, a.e., it follows that $\phi(t) = \lambda$ for every $t \in E$. In particular,

$$E \subseteq \{t \in [a, b] : \phi(t) = \lambda\}, \qquad \mu(\{t \in [a, b] : \phi(t) = \lambda\}) > 0.$$

Thus,

$$\sigma_{\text{eig}}(A) \subseteq \{\lambda \in \mathbb{K} : \mu(\{t : \phi(t) = \lambda\}) > 0\}.$$

To show the reverse incluđion, suppose that $\lambda \in \mathbb{K}$ is such that $\mu(\{t \in [a,b] : \phi(t) = \lambda\}) > 0$. Taking $\Delta = \{t \in [a,b] : \phi(t) = \lambda\}$ and x_0 as the characteristic function of Δ, we have

$$\|Ax_0 - \lambda x_0\|_2^2 = \int_a^b |\phi(t) - \lambda|^2 |x_0(t)|^2 \, d\mu = \int_\Delta |\phi(t) - \lambda|^2 \, d\mu = 0,$$

showing that

$$\sigma_{\text{eig}}(A) \supseteq \{\lambda \in \mathbb{K} : \mu(\{t : \phi(t) = \lambda\}) > 0\}.$$

ii) Suppose $\phi \in C[a,b]$, and let A be the operator as in the above example. Then the essential range of ϕ is the range of ϕ, and for $\lambda \in \mathbb{K}$, if $\mu(\{t : \phi(t) = \lambda\}) > 0$, then the set $\phi^{-1}(\{\lambda\})$ contains an open interval. Hence, in this case, we have

$$\sigma_{\text{app}}(A) = \{\lambda \in \mathbb{K} : \phi(t) = \lambda \text{ for some } t \in [a,b]\},$$

$$\sigma_{\text{eig}}(A) = \{\lambda \in [a,b] : \phi^{-1}(\{\lambda\}) \text{ contains an open interval}\}.$$

By Theorems 11.11 and 12.5, we know that if $A \in \mathcal{B}(X)$ is a self-adjoint operator, then

$$r_\sigma(A) \leq r_w(A) = \|A\|.$$

The following example shows that if A is not a self–adjoint operator, then strict inequality can occur in place of inequality and equality.

EXAMPLE 12.4 Let $X = \mathbb{R}^2$ and $A : (\alpha_1, \alpha_2) \mapsto (0, \alpha_1)$. Clearly,

$$\sigma(A) = \sigma_{\text{eig}}(A) = \{0\}$$

so that $r_\sigma(A) = 0$. But, it is seen (*Verify*) that

$$r_w(A) = 1/2, \quad \|A\| = 1.$$

But, for normal operators A on Hilbert spaces over $\mathbb{C}$, the three quantities $r_\sigma(A)$, $r_w(A)$ and $\|A\|$ are the same, as the following theorem shows.

Theorem 12.7 *Let $\mathbb{K} = \mathbb{C}$ and $A \in \mathcal{B}(X)$ be a normal operator. Then,*

$$r_\sigma(A) = \|A\|.$$

Proof. By know, by Theorem 11.1, that $\|A\|^2 = \|A^*A\|$ for every $A \in \mathcal{B}(X)$. Suppose A is normal. Then by Theorem 11.13,

$$\|A^*Ax\| = \|A^*(Ax)\| = \|A(Ax)\| = \|A^2x\|$$

for every $x \in X$ so that

$$\|A^*A\| = \|A^2\|.$$

Thus we have

$$\|A^2\| = \|A\|^2.$$

Again, using the fact that A^k is a normal operator for every positive integer k, we get

$$\|A^{2^k}\| = \|A\|^{2^k}.$$

Hence,

$$r_\sigma(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \|A^{2^k}\|^{1/2^k} = \|A\|.$$

This completes the proof. ∎

Remark 12.3 (a) The conclusion of the above theorem need not hold if the scalar field is $\mathbb{R}$. For example, we know that the nonzero linear operator $A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $A((\alpha_1, \alpha_2)) = (\alpha_2, -\alpha_1)$ has empty spectrum. Clearly, A is a normal operator.

(b) By the above theorem, if A is a normal operator on a complex Hilbert space X, then $\sigma(A) \neq \varnothing$. Indeed, if $A = 0$, then $\sigma(A) = \{0\}$, and if $A \neq 0$, then $r_\sigma(A) = \|A\| \neq 0$ so that A has a nonzero spectral value. In fact, using compactness of $\sigma(A)$, there exists a spectral value λ such that $|\lambda| = \|A\|$. Thus, for normal operators, we now have a simpler proof for the Gelfand-Mazur theorem (Theorem 10.17(i)).

12.2 More Results on the Spectra of Self-Adjoint Operators

We have seen in Theorem 12.7 that if A is a normal operator on a complex Hilbert space, then $r_\sigma(A) = \|A\|$. Now we prove that if A is a self-adjoint operator, then the above conclusion is true for a real Hilbert space as well.

First we recall from Theorem 12.5 that if $A \in \mathcal{B}(X)$ is a self-adjoint operator, then

$$\sigma(A) \subseteq \overline{w(A)} \subseteq \mathbb{R}.$$

In fact, for a self-adjoint operator A,

$$\sigma(A) \subseteq \overline{w(A)} \subseteq [\alpha_A, \beta_A],$$

where

$$\alpha_A = \inf w(A), \qquad \beta_A = \sup w(A).$$

Also, we recall from Chapter 2, as a special case of the generalized Schwarz inequality (Theorem 2.21), that if $B \in \mathcal{B}(X)$ is self-adjoint and positive, i.e.,

$$B^* = B, \qquad w(B) \subseteq [0, \infty)$$

then

$$|\langle Bx, y\rangle|^2 \leq \langle Bx, x\rangle\langle By, y\rangle \quad \forall x, y \in X. \tag{12.2}$$

Theorem 12.8 *If $A \in \mathcal{B}(X)$ is a self-adjoint operator, then we have the following:*

(i) $\{\alpha_A, \beta_A\} \subseteq \sigma(A) \subseteq [\alpha_A, \beta_A]$.

(ii) $r_\sigma(A) = \|A\| = r_w(A) = \max\{|\alpha_A|, |\beta_A|\}$.

Proof. Let $A \in \mathcal{B}(X)$ be a self-adjoint operator. We have already observed that

$$\sigma(A) \subseteq \overline{w(A)} \subseteq [\alpha_A, \beta_A].$$

Also, by Theorem 11.11, we have

$$r_w(A) = \|A\|.$$

Thus we have

$$r_\sigma(A) \leq \|A\| = r_w(A) = \max\{|\alpha_A|, |\beta_A|\}.$$

Hence the results in (i) and (ii) follow once we prove that

$$\{\alpha_A, \beta_A\} \subseteq \sigma(A).$$

Since A is a self-adjoint operator, by Corollary 12.4, $\sigma(A) = \sigma_{\text{app}}(A)$. Hence, it is enough to prove that $\{\alpha_A, \beta_A\} \subseteq \sigma_{\text{app}}(A)$.

By the definition of α_A, there exists (x_n) in X such that

$$\|x_n\| = 1 \quad \forall n \in \mathbb{N}, \qquad \langle Ax_n, x_n\rangle \to \alpha_A \quad \text{as} \quad n \to \infty.$$

Hence,

$$\langle (A - \alpha_A I)x_n, x_n\rangle \to 0 \quad \text{as} \quad n \to \infty.$$

We show that

$$\|(A - \alpha_A I)x_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

Now taking $B = A - \alpha_A I$, we see that $\langle Bx, x\rangle \geq 0$ for every $x \in X$ and $B^* = B$. Therefore, by (12.2), we have

$$|\langle Bx, y\rangle|^2 \leq \langle Bx, x\rangle\langle By, y\rangle \quad \forall x, y \in X.$$

Therefore, taking $x = x_n$ and $y = Bx_n$ in the above inequality, we obtain

$$|\langle Bx_n, Bx_n\rangle|^2 \leq \langle Bx_n, x_n\rangle\langle B^2x_n, Bx_n\rangle \quad \forall n \in \mathbb{N}.$$

Hence,

$$\|Bx_n\|^4 \leq \langle Bx_n, x_n\rangle \|B\| \, \|Bx_n\|^2 \quad \forall n \in \mathbb{N}.$$

Thus,

$$\|(A - \alpha_A I)x_n\|^2 = \|Bx_n\|^2 \leq \langle Bx_n, x_n\rangle \|B\| \to 0 \quad \text{as} \quad n \to \infty.$$

Now, let (u_n) in X be such that

$$\|u_n\| = 1 \quad \forall n, \qquad \langle Au_n, u_n\rangle \to \beta_A \quad \text{as} \quad n \to \infty.$$

Proceeding as above, and taking $B = \beta_A I - A$ and (u_n) in place of (x_n) above, it is seen that

$$\|(\beta_A I - A)u_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, we have shown that $\alpha_A, \beta_A \in \sigma_{\text{app}}(A)$. ∎

Remark 12.4 The above theorem, in particular, shows that if A is a self-adjoint operator, then the Gelfand-Mazur theorem (Theorem 10.17(i)) holds for a real Hilbert space as well.

Corollary 12.9 *For $A \in \mathcal{B}(X)$,*

$$\|A\|^2 = r_\sigma(A^*A).$$

Proof. Since A^*A is a self-adjoint operator, by the above theorem, we have $r_\sigma(A^*A) = \|A^*A\|$. We have seen in Theorem 11.13 that $\|A^*A\| = \|A\|^2$. ∎

The following corollary is another important consequence of the above theorem, together with Theorem 10.1 (also, see the remark following Corollary 12.4) and the fact that every eigenvalue of a self-adjoint operator is real.

Corollary 12.10 *If A is a compact self-adjoint operator on a Hilbert space X, then either $\|A\|$ or $-\|A\|$ is an eigenvalue of A.*

12.2.1 An Approximation Result

The following theorem indicates how we can approximate the biggest (respectively the smallest) approximate eigenvalue of a self-adjoint operator using the biggest (respectively the smallest) eigenvalues of a sequence of finite rank operators. Before stating the result, we may observe that if X_0 is a finite dimensional subspace of a Hilbert space X, and $A \in \mathcal{B}(X)$ is a self-adjoint operator, then

$$\alpha_0 := \inf\{\langle Au, u\rangle : u \in X_o, \|u\| = 1\},$$

$$\beta_0 := \sup\{\langle Au, u\rangle : u \in X_0, \|u\| = 1\}$$

are the smallest and biggest eigenvalues of the self-adjoint operator

$$A_0 = P_0 A|_{X_0} : X_0 \to X_0,$$

where $P_0 : X \to X$ is the orthogonal projection onto X_0.

Theorem 12.11 *Let $A \in \mathcal{B}(X)$ be a self-adjoint operator and let (X_n) be a sequence of finite dimensional subspaces of X such that $X_n \subseteq X_{n+1}$ for every $n \in \mathbb{N}$ and $\cup_{n=1}^\infty X_n$ is dense in X. Let*

$$\alpha_n = \inf\{\langle Au, u\rangle : u \in X_n, \|u\| = 1\},$$

$$\beta_n = \sup\{\langle Au, u\rangle : u \in X_n, \|u\| = 1\}.$$

Then

$$\alpha_A \le \alpha_{n+1} \le \alpha_n \le \beta_n \le \beta_{n+1} \le \beta_A$$

for all $n \in \mathbb{N}$, and

$$\alpha_n \to \alpha_A, \quad \beta_n \to \beta_A \quad as \quad n \to \infty.$$

Proof. The inequalities

$$\alpha_A \le \alpha_{n+1} \le \alpha_n \le \beta_n \le \beta_{n+1} \le \beta_A$$

are obvious from the fact that $X_n \subseteq X_{n+1}$ for every $n \in \mathbb{N}$. Also, since (α_n) is a decreasing sequence which is bounded below by α_A and (β_n) is an increasing sequence which is bounded above by β_A, there exists α and β such that

$$\lim_n \alpha_n = \alpha \ge \alpha_A, \qquad \lim_n \beta_n = \beta \le \beta_A.$$

Clearly,

$$\alpha \le \alpha_n, \qquad \beta \ge \beta_n \quad \forall n \in \mathbb{N}.$$

We show that $\alpha = \alpha_A$ and $\beta = \beta_A$.

If $\alpha > \alpha_A$, then by the definition of α_A, there exists $x \in X$ such that

$$\|x\| = 1, \qquad \langle Ax, x\rangle < \alpha.$$

By the denseness of $Y := \cup_n X_n$ in X, there exists (x_n) in Y such that $x_n \to x$ as $n \to \infty$. Since $\|x\| = 1$, we can assume, without loss of generality, that $x_n \ne 0$ for all n. Let $u_n = x_n/\|x_n\|$ for all $n \in \mathbb{N}$. Then, we have $\|u_n\| = 1$ and $u_n \to x$ as $n \to \infty$. Now, since $u_n \in Y$, there exists $i_n \in \mathbb{N}$ such that $u_n \in X_{i_n}$. Hence, we get

$$\alpha_{i_n} \le \langle Au_n, u_n\rangle \to \langle Ax, x\rangle \quad \text{as} \quad n \to \infty.$$

Thus, there is $N \in \mathbb{N}$ such that

$$\alpha_{i_N} < \alpha \le \alpha_{i_N},$$

which is a contradiction. Similarly, if $\beta < \beta_A$, then we will arrive at a contradiction.

This completes the proof. ∎

Remark 12.5 (a) If X is a separable Hilbert space and $\{u_1, u_2, \ldots\}$ is an orthonormal basis of X, then $X_n = \operatorname{span}\{u_1, \ldots, u_n\}$ satisfies the requirements of the above theorem.

(b) Let X_n be as in the above theorem and $P_n : X \to X$ be the orthogonal projection onto X_n. Let

$$A_n := P_n A|_{X_n}, \qquad n \in \mathbb{N}.$$

Then we see that

$$\alpha_n = \inf w(A_n), \qquad \beta_n = \sup w(A_n)$$

and α_n and β_n are, respectively, the smallest and the largest eigenvalues of A. Thus, by the above theorem, the largest and smallest eigenvalues of A_n 'approximates' β_A and α_A, respectively. But, α_A and β_A need not be eigenvalues of A (*recall an appropriate example*). Under what situation can one assert that α_A (or β_A) is an eigenvalue of A? As a consequence of a theorem that we shall prove in the next section (Theorem 12.15(ii)), we can say that if α_A (respectively, β_A) is an isolated point of the spectrum of A, then it is an eigenvalue.

12.2.2 Spectral Mapping Theorem Revisited

Recall that the spectral mapping theorem (Theorem 10.14) was proved for bounded linear operators on complex Banach spaces. In this section we shall consider the same theorem for self-adjoint operators for real Hilbert spaces as well.

Theorem 12.12 *Let $A \in \mathcal{B}(X)$ be a self-adjoint operator on a Hilbert space X and $p(t)$ be a polynomial with coefficients in $\mathbb{K}$. Then,*

$$\sigma(p(A)) = \{p(\lambda) : \lambda \in \sigma(A)\}.$$

Proof. For the case $\mathbb{K} = \mathbb{C}$, the result is nothing but Theorem 10.14. Now we prove the theorem for the case $\mathbb{K} = \mathbb{R}$. We adopt the procedure as in the proof of Theorem 10.14, but use the fact that complex (which are not real) roots of a polynomial with real coefficients occur in conjugate pairs.

Let $p(t)$ be a polynomial with real coefficients. By the first part of Theorem 10.14, we have

$$\{p(\lambda) : \lambda \in \sigma(A)\} \subseteq \sigma(p(A)).$$

Hence it is enough to prove that

$$\sigma(p(A)) \subseteq \{p(\lambda) : \lambda \in \sigma(A)\}.$$

For this, let $\mu \in \sigma(p(A))$. Let $t_1, \ldots, t_k$ be the real roots and let $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_m, \bar{\lambda}_m$ be the complex roots of $p(t) - \mu$. Let $\lambda_j = \alpha_j + i\beta_j$ with $\alpha_j, \beta_j \in \mathbb{R}$ and $j = 1, \ldots, m$. Then,

$$\begin{aligned} p(t) - \mu &= c(t - t_1) \ldots (t - t_k)(t - \lambda_1)(t - \bar{\lambda}_1) \ldots (t - \lambda_m)(t - \bar{\lambda}_m) \\ &= c(t - t_1) \ldots (t - t_k)[(t - \alpha_1)^2 + \beta_1^2] \ldots [(t - \alpha_m)^2 + \beta_m^2] \end{aligned}$$

for some real number c. Hence,

$$\begin{aligned} p(A) - \mu I &= c(A - t_1 I) \ldots (A - t_k I) \\ &\quad \times [(A - \alpha_1 I)^2 + \beta_1^2 I] \ldots [(A - \alpha_m I)^2 + \beta_m^2 I]. \end{aligned}$$

Now, since $\mu \in \sigma(p(A))$, one of the factors on the right-hand side of the above equation is not bijective. Since $\sigma((A - \alpha_i I)^2) \subseteq [0, \infty)$ and $\beta_i \neq 0$ for each $i \in \{1, \ldots, m\}$, the operator $(A - \alpha_i I)^2 + \beta_i^2 I$ is bijective for every $i \in \{1, \ldots, m\}$. Hence, there exists $j \in \{1, \ldots, k\}$ such that $A - t_j I$ is not bijective so that $t_j \in \sigma(A)$. Note that $p(t_j) = \mu$. Thus we have shown that μ is of the form $p(\lambda)$ for some $\lambda \in \sigma(A)$. In case $p(t) - \mu$ does not have any non-real complex roots, then expressions of the forms $(t - \alpha_i)^2 + \beta_i^2$ and $(A - \alpha_i I)^2 + \beta_i^2 I$ do not appear in the above analysis. ∎

We may observe that if $A \in \mathcal{B}(X)$ is a self-adjoint operator and if $p(t)$ is a polynomial, then $p(A)$ is a normal operator. Hence, the following corollary is a consequence of Theorems 12.12 and 12.7.

Corollary 12.13 *Let $A \in \mathcal{B}(X)$ be a self-adjoint operator. Then for every polynomial $p(t)$,*

$$\|p(A)\| = \sup\{|p(\lambda)| : \lambda \in \sigma(A)\},$$

Let $\mathcal{P}$ be the set of all polynomials with coefficients in $\mathbb{K}$. For a self-adjoint operator $A \in \mathcal{B}(X)$, consider the map $\Psi_0 : \mathcal{P} \to \mathcal{B}(X)$ defined by

$$\Psi_0(p) = p(A), \quad p \in \mathcal{P}.$$

Then it is easy to see that for every $p, q \in \mathcal{P}$ and $\alpha, \beta \in \mathbb{K}$,

$$\begin{aligned} \Psi_0(\alpha p + \beta q) &= \alpha \Psi_0(p) + \beta \Psi_0(q) && \text{(linear)} \\ \Psi_0(pq) &= \Psi_0(p)\Psi_0(q) && \text{(multiplicative)} \\ \Psi_0(\bar{p}) &= [\Psi_0(p)]^* && \text{(involutive).} \end{aligned}$$

Moreover, if p is real-valued, then $\Psi_0(p)$ is self-adjoint.

By Corollary 12.13, we have

$$\|\Psi_0(p)\| \leq \sup\{|p(\lambda)| : \alpha_A \leq \lambda \leq \beta_A\} \quad \forall p \in \mathcal{P}[\alpha_A, \beta_A],$$

where α_A and β_A are the infimum and supremum, respectively, of the numerical range of A, so that Ψ_0 is a bounded linear operator on the normed linear space $\mathcal{P}[\alpha_A, \beta_A]$ with respect to $\|\cdot\|_\infty$.

Theorem 12.14 *Let $A \in \mathcal{B}(X)$ be a self-adjoint operator and let Ψ_0 be as above. Then there exists a function $\Psi : C[\alpha_A, \beta_A] \to \mathcal{B}(X)$ which is an extension of $\Psi_0 : \mathcal{P} \to \mathcal{B}(X)$ and satisfies the following:*

(i) $\Psi(\alpha f + \beta g) = \alpha\Psi(f) + \beta\Psi(g)$ (linear)

(ii) $\Psi(fg) = \Psi(f)\Psi(g)$ (multiplicative)

(iii) $\Psi(\bar{f}) = [\Psi(f)]^*$ (involutive)

Moreover,

(iv) *$\Psi(f)$ is a normal operator for every $f \in C[\alpha_A, \beta_A]$,*

(v) *$\|\Psi(f)\| \le \|f\|_\infty$, and*

(vi) *$\Psi(f)$ is self-adjoint if f is real-valued.*

Proof. Recall that $\Psi_0 : \mathcal{P}[\alpha_A, \beta_A] \to \mathcal{B}(X)$ is a bounded linear operator. Since $\mathcal{P}[\alpha_A, \beta_A]$ is dense in $C[\alpha_A, \beta_A]$ with respect to the norm $\|\cdot\|_\infty$, and since $\mathcal{B}(X)$ is a Banach space, by Theorem 3.18, the bounded operator Ψ_0 has a unique extension to $C[\alpha_A, \beta_A]$, say $\Psi : C[\alpha_A, \beta_A] \to \mathcal{B}(X)$. In fact, for $f \in C[\alpha_A, \beta_A]$, Ψ is defined by

$$\Psi(f) = \lim_{n\to\infty} \Psi_0(p_n),$$

where (p_n) is a sequence in $\mathcal{P}[\alpha_A, \beta_A]$ such that $\|f - p_n\|_\infty \to 0$ as $n \to \infty$. Now the properties (i)–(vi) follow easily using the properties of Ψ_0. ∎

For self-adjoint $A \in \mathcal{B}(X)$ and $f \in C[\alpha_A, \beta_A]$, the operator $\Psi(f)$ given by the above theorem is usually denoted by $f(A)$. Thus, for example, if $f(t) = |t|$, then the corresponding operator $f(A)$ is denoted by $|A|$, and if A is a positive operator, then the operator corresponding to the function $f(t) = t^\nu$ for $\nu > 0$ is denoted by A^ν.

12.2.3 A Result Using Spectral Projection

For normal operators on complex Hilbert spaces, the following result is worth noticing.

Theorem 12.15 *Let $\mathbb{K} = \mathbb{C}$ and $A \in \mathcal{B}(X)$ be a normal operator. Then we have the following:*

(i) *For every $z \in \rho(A)$, $(A - zI)^{-1}$ is a normal operator, and*

$$\|(A - zI)^{-1}\| = r_\sigma\left((A - zI)^{-1}\right) = \frac{1}{\operatorname{dist}(z, \sigma(A))}.$$

(ii) *Every isolated spectral value of A is an eigenvalue.*

Proof. (i) Let $z \in \rho(A)$. Then it is easily seen that $(A - zI)^{-1}$ is a normal operator, and hence, the result follows by Theorem 12.7.

(ii) Let λ be an isolated spectral value of A, and let $r > 0$ be such that

$$\sigma(A) \cap \{z \in \sigma(A) : |z - \lambda| \le r\} = \{\lambda\},$$

and let C_r be the circle with centre λ and radius r. Then by Theorem 10.19, the operator

$$P = -\frac{1}{2\pi i} \int_{C_r} (A - zI)^{-1}\, dz$$

is nonzero. Also using the relation

$$A(A-zI)^{-1} = I + z(A-zI)^{-1} = I + (z-\lambda)(A-zI)^{-1} + \lambda(A-zI)^{-1},$$

and the fact that $\int_{C_r} dz = 0$, we get

$$AP = \lambda P - \frac{1}{2\pi i} \int_{C_r} (z - \lambda)(A - zI)^{-1}\, dz$$

so that for every $x, y \in X$,

$$\langle (AP - \lambda P)x, y \rangle = -\frac{1}{2\pi i} \int_{C_r} \langle (z - \lambda)(A - zI)^{-1} x, y \rangle\, dz.$$

Suppose r is such that $2r < \operatorname{dist}(\lambda, \sigma(A) \setminus \{\lambda\})$. Then using the result in (i), we observe that for every $x, y \in X$ and $z \in C_r$,

$$\begin{aligned} |\langle (z - \lambda)(A - zI)^{-1} x, y \rangle| &\le r\|x\|\, \|y\| \|(A - zI)^{-1}\| \\ &\le \frac{r\|x\|\, \|y\|}{\operatorname{dist}(z, \sigma(A))} = \|x\|\, \|y\|. \end{aligned}$$

Therefore,

$$|\langle (AP - \lambda P)x, y \rangle| \le r\|x\|\, \|y\| \quad \forall x, y \in X$$

so that

$$\|AP - \lambda P\| \le r.$$

This is true for all r such that $0 < 2r < \operatorname{dist}(\lambda, \sigma(A) \setminus \{\lambda\})$. Hence, $AP = \lambda P$. Since P is a nonzero projection operator, it follows that λ is an eigenvalue of A. ∎

PROBLEMS

Unless stated otherwise, X denotes a Hilbert space and A denotes a bounded linear operator from X into itself.

1. Show that, if A is a normal operator on X, then

$$X = \overline{R(A)} \oplus N(A).$$

2. If $\lambda \in \mathbb{K}$ is such that $\operatorname{dist}(\lambda, w(A)) > 0$, then show that $\lambda \in \rho(A)$ and

$$\|(A - \lambda I)^{-1}\| \leq \frac{1}{\operatorname{dist}(\lambda, w(A))}.$$

3. If there exists $\gamma > 0$ such that $|\langle Ax, x\rangle|^2 \geq \gamma\|x\|^2$ for every $x \in X$, then show that A is bijective and $\|A^{-1}\| \leq 1/\gamma$.

4. Let Δ be a compact subset of $\mathbb{K}$. Give an example of a normal operator A on ℓ^2 such that $\sigma(A) = \Delta$.

5. Give an example of a unitary operator A on ℓ^2 such that $\sigma(A)$ is the set $\{\lambda \in \mathbb{K} : |\lambda| = 1\}$.

6. If A is a self-adjoint operator, then show that $\sigma\left((A - \alpha I)^{2n}\right) \subseteq [0, \infty)$ for every $\alpha \in \mathbb{R}$, $n \in \mathbb{N}$.

7. Let X be a complex Hilbert space and A be a self-adjoint operator. Show that A is a positive operator if and only if $\sigma(A) \subseteq [0, \infty)$.

8. Let $A \in \mathcal{B}(X)$ be a compact operator and (P_n) be a sequence of finite rank orthogonal projections such that $\|P_n x - x\| \to 0$ as $n \to \infty$. Let $T_n = P_n A^* A|_{R(P_n)}$. Show that, for each $n \in \mathbb{N}$, T_n is a positive compact self-adjoint operator on $R(P_n)$. Show also that, if λ_n is the largest eigenvalue of T_n, $n \in \mathbb{N}$, then (λ_n) converges and the limit is the largest eigenvalue of A^*A.

9. Show that the operator $A : L^2[a,b] \to L^2[a,b]$ defined by $(Ax)(t) = t^2 x(t)$, $t \in [a,b]$, has no eigenvalues.

10. Write detailed proofs for Corollaries 12.13 and 12.10.

11. Verify (i)–(vi) in Theorem 12.14.

12. Show that if $A \in \mathcal{B}(X)$ is a normal operator and $\lambda \in \rho(A)$, then $(A - \lambda I)^{-1}$ is a normal operator.

13

Spectral Representations

At the end of a basic linear algebra course one learns that if X is a finite dimensional inner product space and $A : X \to X$ is a linear operator satisfying

$$\langle Ax, y\rangle = \langle x, Ay\rangle \quad \forall\, x, y \in X,$$

that is, A is a self-adjoint operator, then there exists an orthonormal basis $\mathcal{U} = \{u_1, \ldots, u_n\}$ for X such that the matrix representation of A with respect to $\mathcal{U}$ is a diagonal matrix. Equivalently, there exist $\mu_1, \ldots, \mu_n$ in $\mathbb{K}$ and orthonormal vectors $u_1, \ldots, u_n$ in X such that

$$Ax = \sum_{j=1}^{n} \mu_j \langle x, u_j\rangle u_j, \qquad x \in X. \tag{13.1}$$

Note that

$$Au_j = \mu_j u_j, \qquad j = 1, \ldots, n,$$

so that $\mu_1, \ldots, \mu_n$ are eigenvalues of A with corresponding eigenvectors $u_1, \ldots, u_n$. It is easily seen that $\lambda \in \mathbb{K}$ is an eigenvalue of A if and only if $\lambda = \mu_j$ for some $j \in \{1, \ldots, n\}$.

Suppose $\lambda_1, \ldots, \lambda_k$ are the distinct scalars among $\mu_1, \ldots, \mu_n$, with λ_i repeated m_i times for $i = 1, \ldots, k$. Then it can be seen (*Verify*) that

$$m_i := \dim N(A - \lambda_i I), \qquad i = 1, \ldots, k,$$

and $n = m_1 + \cdots + m_k$. Rearranging $u_1, \ldots, u_n$ as

$$u_1^{(1)}, \ldots, u_{m_1}^{(1)}, u_1^{(2)}, \ldots, u_{m_2}^{(2)}, \ldots, u_1^{(k)}, \ldots, u_{m_k}^{(k)}$$

such that for each $i = 1, \ldots, k$, the elements $u_1^{(i)}, \ldots, u_{m_i}^{(i)}$ are those eigenvectors associated with the eigenvalue λ_i, (13.1) takes the form

$$Ax = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \lambda_i \langle x, u_j^{(i)}\rangle u_j^{(i)}, \qquad x \in X.$$

Writing

$$P_i x = \sum_{j=1}^{m_i} \langle x, u_j^{(i)} \rangle u_j^{(i)}, \qquad x \in X,$$

we have the representation of A as

$$A = \sum_{i=1}^{k} \lambda_i P_i. \tag{13.2}$$

We may observe (*Verify*) that, for each $i \in \{1, \ldots, k\}$, $P_i : X \to X$ is an orthogonal projection onto $N(A - \lambda_i I)$. In addition,

$$P_1 + \cdots + P_k = I.$$

Is a representation similar to (13.2) possible for a self-adjoint operator on an infinite dimensional Hilbert space?

An affirmative answer to this question is too much to expect, because, we know that a self-adjoint operator need not have eigenvalues at all, and the spectrum need not be even a countable set. For example, recall (cf. Example 12.3) that if $A : L^2[a, b] \to L^2[a, b]$ is defined by

$$(Ax)(t) = tx(t), \qquad x \in L^2[a, b],$$

then

$$\sigma_{\text{eig}}(A) = \varnothing; \qquad \sigma(A) = [a, b].$$

However, if A is a nonzero compact self-adjoint operator, then it has a nonzero eigenvalue, and all its nonzero spectral values are eigenvalues (Corollary 12.10 and Remark 12.1). In this case, we obtain a spectral representation such as the above, with finite sum replaced by infinite sum. We shall do this in the first section. In fact, our procedure includes the case of finite dimensional X as well.

What about for a general compact operator?

Again, a non-self-adjoint compact operator need not have eigenvalues at all. For example, if $A : L^2[a, b] \to L^2[a, b]$ is defined by

$$(Ax)(t) = \int_a^t x(s)\, ds, \qquad x \in L^2[a, b],$$

then it can be seen (*Verify*) that A is a compact operator and

$$\sigma_{\text{eig}}(A) = \varnothing, \qquad \sigma(A) = \{0\}.$$

However, if A is a compact operator on X, then we do obtain a representation of A involving the eigenvalues of the compact self-adjoint operator A^*A. Section 13.2 deals with this situation.

The situation with self-adjoint, possibly non-compact, operator is also not too pessimistic. In this case, we have a representation, involving an integral and a family of orthogonal projections. We consider this situation in Section 13.3.

Before trying to obtain the above-mentioned representations, we first observe the following easily verifiable fact.

Lemma 13.1 *Let X be a Hilbert space, Y be a clsoed subspace of X and $A \in \mathcal{B}(X)$. If Y is invariant under A, then $Y^\perp$ is invariant under A^*. In particular, if A is a self-adjoint operator, then a subspace Y of X is invariant under A if and only if $Y^\perp$ is invariant under A.*

Recall that a subspace X_0 of X is *invariant* under $A \in \mathcal{B}(X)$ if $Ax \in X_0$ for every $x \in X_0$.

In this chapter, we assume that the space X under consideration is a Hilbert space.

13.1 Spectral Representation of Compact Self-Adjoint Operators

We recall the following results from Chapters 10 and 12:

(1) Suppose $A : X \to X$ is a compact operator. Then every nonzero spectral value of A is an eigenvalue, and the corresponding eigenspace is finite dimensional (Theorem 10.23 and Proposition 10.21(i)). Moreover, the spectrum $\sigma(A)$ of A is countable, and zero is the only possible limit point of $\sigma(A)$ (cf. Theorem 10.1).

(2) If $A : X \to X$ is a normal operator and if λ and μ are distinct eigenvalues of A, then $\langle x, y\rangle = 0$ for all $x \in N(A-\lambda I)$, $y \in N(A-\mu I)$ (Theorem 12.3).

(3) If $A : X \to X$ is a nonzero compact self-adjoint operator, then A has a nonzero eigenvalue and $\sigma(A) \setminus \{0\} = \sigma_{\text{eig}}(A) \setminus \{0\} \subseteq \mathbb{R}$ (cf. Corollary 12.10 and Theorem 12.5).

In this section, we assume that $A : X \to X$ is a compact self-adjoint operator. Suppose $\{\lambda_i : i \in \Delta\}$ is the set of all nonzero eigenvalues of A, where $\Delta = \mathbb{N}$ if $\sigma(A)$ is infinite, and $\Delta = \{1, \ldots, k\}$ for some $k \in \mathbb{N}$ if $\sigma(A)$ is finite.

Theorem 13.2 *Let A be a compact self-adjoint operator on X and let $\{\lambda_i : i \in \Delta\}$ be the set of all nonzero eigenvalues of A. For each $i \in \Delta$, let $\{u_1^{(i)}, \dots, u_{m_i}^{(i)}\}$ be an orthonormal basis of $N(A - \lambda_i I)$. Then*

$$Ax = \sum_{i\in\Delta} \sum_{j=1}^{m_i} \lambda_i \langle x, u_j^{(i)} \rangle u_j^{(i)} \quad \forall x \in X$$

and $\cup_{i\in\Delta}\{u_1^{(i)}, \dots, u_{m_i}^{(i)}\}$ is an orthonormal basis of $N(A)^\perp$. In particular,

$$A = \sum_{i\in\Delta} \lambda_i P_i,$$

where P_i is the orthogonal projection onto $N(A - \lambda_i I)$. If A is of infinite rank, then we also have

$$\left\| A - \sum_{j=1}^{n} \lambda_i P_i \right\| \le \max_{i>n} |\lambda_i| \to 0 \quad \text{as } n \to \infty.$$

Proof. Let X_0 be the closure of the span of $\cup_{i\in\Delta}\{u_1^{(i)}, \dots, u_{m_i}^{(i)}\}$. Then by Lemma 13.1, X_0 and $X_0^\perp$ are invariant under A, i.e.,

$$A(X_0) \subseteq X_0, \quad A(X_0^\perp) \subseteq X_0^\perp.$$

By projection theorem (Theorem 2.47), every $x \in X$ can be written as

$$x = y + z \quad \text{with} \quad y \in X_0, \ z \in X_0^\perp.$$

Since $\cup_{i\in\Delta}\{u_1^{(i)}, \dots, u_{m_i}^{(i)}\}$ is an orthonormal basis of X_0 (*Why?*), by Fourier expansion theorem (Theorem 4.9(ii)),

$$y = \sum_{i\in\Delta} \sum_{j=1}^{m_i} \langle y, u_j^{(i)} \rangle u_j^{(i)}.$$

Hence, using the fact that $\langle x, u \rangle = \langle y, u \rangle$ for every $u \in X_0$, we have

$$\begin{aligned} Ax &= Ay + Az \\ &= \sum_{i\in\Delta} \sum_{j=1}^{m_i} \langle y, u_j^{(i)} \rangle A u_j^{(i)} + Az \\ &= \sum_{i\in\Delta} \lambda_i \sum_{j=1}^{m_i} \langle x, u_j^{(i)} \rangle u_j^{(i)} + Az. \end{aligned}$$

Thus, once we show that $Az = 0$, we obtain the required representation for Ax. For this, let B be the restriction of A to the space $X_0^\perp$. Note that (*Verify*)

$$B : X_0^\perp \to X_0^\perp$$

is a compact self-adjoint operator. We claim that $B = 0$.

If $B \neq 0$, then, by Corollary 12.10, B has a nonzero eigenvalue, say $\lambda \neq 0$. Let $x \in X_0^\perp$ be a corresponding eigenvector of B. Then we have

$$\lambda x = Bx = Ax$$

so that λ is an eigenvalue of A as well. This implies that there exists $i \in \Delta$ such that $\lambda = \lambda_i$ and $x \in N(A - \lambda_i I) \subseteq X_0$, a contradiction to the fact that $0 \neq x \in X_0^\perp$. Thus, $B = 0$. In particular, $Az = 0$ so that for every $x \in X$,

$$Ax = \sum_{i\in\Delta} \lambda_i \sum_{j=1}^{m_i} \langle x, u_j^{(i)} \rangle u_j^{(i)}.$$

Now, let $x \in N(A)^\perp$, and $\langle x, u_j^{(i)} \rangle = 0$ for all $j = 1, \ldots, m_i$; $i \in \Delta$. Then from the above representation of Ax, we have $Ax = 0$ so that $x \in N(A)$. This is possible only if $x = 0$. Thus, by Theorem 4.2, $\cup_{i\in\Delta}\{u_1^{(i)}, \ldots, u_{m_i}^{(i)}\}$ is an orthonormal basis of $N(A)^\perp$.

Also, we have

$$Ax = \sum_{i\in\Delta} \lambda_i P_i x, \qquad x \in X,$$

where $P_i : X \to X$ for each $i \in \Delta$ is the orthogonal projection onto $N(A - \lambda_i I)$, i.e., P_i is given by

$$P_i x = \sum_{j=1}^{m_i} \langle x, u_j^{(i)} \rangle u_j^{(i)}, \qquad x \in X.$$

Since $R(P_i) \cap R(P_k) = \{0\}$ for every $i \neq k$, it follows that A is of finite rank if and only if there are only a finite number of nonzero eigenvalues for A.

If A is of finite rank with nonzero eigenvalues $\lambda_1, \ldots, \lambda_m$, then from the above representation of A, it is obvious that

$$A = \sum_{i=1}^{m} \lambda_i P_i.$$

Now, suppose that A is of infinite rank. Then the nonzero eigenvalues of A constitute a denumerable set $\{\lambda_i : i \in \mathbb{N}\}$ so that $\Delta = \mathbb{N}$. In this case, by Theorem 10.1, $\lambda_n \to 0$ as $n \to \infty$. For $n \in \mathbb{N}$, let

$$A_n x = \sum_{j=1}^{n} \sum_{j=1}^{m_i} \lambda_i \langle x, u_j^{(i)} \rangle u_j^{(i)}, \qquad x \in X.$$

Then we have

$$\|Ax - A_n x\|^2 = \sum_{i>n} \sum_{j=1}^{m_i} |\lambda_i|^2 |\langle x, u_j^{(i)} \rangle|^2 \leq \max_{i>n} |\lambda_i|^2 \|x\|^2$$

for all $x \in X$. Thus,

$$\|A - A_n\| \leq \max_{i>n} |\lambda_i| \to 0 \quad \text{as} \quad n \to \infty$$

so that

$$A = \sum_i \lambda_i P_i.$$

This completes the proof. ▮

Remark 13.1 The above theorem, in particular, shows that every compact self-adjoint operator A is the (norm) limit of a sequence of finite rank bounded operators. Can we have this conclusion for every compact (not necessarily self-adjoint) operator A? The answer is in the affirmative if the scalar field is the set of complex numbers. This follows from the fact that if $\mathbb{K} = \mathbb{C}$, then A can be written as

$$A = B + iC \quad \text{with} \quad B = \frac{A + A^*}{2}, \quad C = \frac{A - A^*}{2i},$$

where B and C are compact self-adjoint operators whenever A is a compact operator.

One may ask whether every compact operator A on a (not necessarily complex) Hilbert space X is a (norm) limit of a sequence of a finite rank operators. The answer is still in the affirmative, as can be seen from the following section.

13.2 Singular Value Representation of Compact Operators

Now we take up the problem of representing a compact operator between Hilbert spaces.

Let X and Y be Hilbert spaces and $T : X \to Y$ be a compact operator. Then we know that T^*T is a compact, self-adjoint and positive operator. Therefore, by Theorem 13.2, T^*T can be represented in terms of its nonzero eigenvalues which are positive reals. More precisely, we have

$$T^*Tx = \sum_{n\in\Delta} \sigma_n^2 \langle x, u_n\rangle u_n, \qquad x \in X,$$

where $\{\sigma_n : n \in \Delta\}$ is a set of positive real numbers and $\{u_n : n \in \Delta\}$ is an orthonormal basis of $N(T^*T)^{\perp}$. Here, $\Delta = \mathbb{N}$ if $\sigma(T^*T)$ is an infinite set, and $\Delta = \{1, \ldots, \ell\}$ for some $\ell \in \mathbb{N}$ if $\sigma(T^*T)$ is a finite set. In fact, if $\{\lambda_i : i \in \Delta_0\}$ is the set of all nonzero eigenvalues of T^*T, and if $\{u_1^{(i)}, \ldots, u_{m_i}^{(i)}\}$ is an orthonormal basis of $N(A - \lambda_i I)$ for each $i \in \Delta_0$, then

$$\sigma_{m_{i-1}+j}^2 = \lambda_i, \qquad u_{m_{i-1}+j} = u_j^{(i)}, \qquad j = 1, \ldots, m_i; \, i \in \Delta_0,$$

with $m_0 = 0$. Thus,

$$\Delta = \cup_{i\in\Delta_0} \{m_{i-1} + 1, m_{i-1} + 2, \ldots, m_{i-1} + m_i\}$$

Since $T^*Tu_n = \sigma_n^2 u_n$, it follows, by taking $v_n = \sigma_n^{-1} Tu_n$, that

$$Tu_n = \sigma_n v_n, \qquad T^*v_n = \sigma_n u_n \quad \forall n \in \Delta.$$

The scalars σ_n's are called the **singular values** of T, and the set

$$\{(\sigma_n, u_n, v_n) : n \in \Delta\}$$

is called a **singular system** for the compact operator T.

In order to represent T in terms of its singular system, we make use of the following result.

Proposition 13.3 *Let X, Y be Hilbert spaces and $T : X \to Y$ be a compact operator. Let $\{(\sigma_n, u_n, v_n) : n \in \Delta\}$ be a singular system of T. Then*

(i) $\{u_n : n \in \Delta\}$ *is an orthonormal basis for* $N(T)^{\perp}$,

(ii) $\{v_n : n \in \Delta\}$ *is an orthonormal basis for* $\overline{R(T)}$.

Proof. We know that $\{u_n : n \in \Delta\}$ is an orthonormal basis of $N(T^*T)^\perp$. Therefore, the result in (i) holds, as $N(T^*T)^\perp = N(T)^\perp$ (see Theorem 11.4(v)). To obtain the result in (ii), we note that

$$\langle v_i, v_j\rangle = \left\langle \frac{Tu_i}{\sigma_i}, \frac{Tu_j}{\sigma_j}\right\rangle = \frac{1}{\sigma_i\sigma_j}\langle T^*Tu_i, u_j\rangle = \frac{\sigma_i}{\sigma_j}\langle u_i, u_j\rangle = \delta_{ij}.$$

Thus, $\{v_n : n \in \Delta\}$ is an orthonormal set. To show that $\{v_n : n \in \Delta\}$ is an orthonormal basis of $\overline{R(T)}$, it is enough to prove that

$$y \in R(T), \qquad \langle y, v_n\rangle = 0 \quad \forall n \implies y = 0.$$

So, let $y \in R(T)$ be such that $\langle y, v_n\rangle = 0$ for all $n \in \Delta$, and let $x \in X$ be such that $y = Tx$. Then we have

$$\langle y, v_n\rangle = \langle Tx, v_n\rangle = \langle x, T^*v_n\rangle = \sigma_n\langle x, u_n\rangle.$$

Thus, $\langle x, u_n\rangle = 0$ for all $n \in \Delta$. Since $\{u_n : n \in \Delta\}$ is an orthonormal basis of $N(T)^\perp$, $x \in N(T)$. Consequently, $y = Tx = 0$. ∎

Theorem 13.4 *Let X and Y be Hilbert spaces, $T : X \to Y$ be a compact operator and $\{(\sigma_n, u_n, v_n) : n \in \Delta\}$ be a singular system for T. Then*

$$Tx = \sum_{n\in\Delta} \sigma_n\langle x, u_n\rangle v_n, \qquad T^*y = \sum_{n\in\Delta} \sigma_n\langle y, v_n\rangle u_n$$

for all $x \in X$, $y \in Y$. Besides, if

$$T_n x = \sum_{j=1}^{n} \sigma_j\langle x, u_j\rangle v_j, \qquad x \in X,$$

for $n < \operatorname{rank} T$, then

$$\|T - T_n\| \leq \max_{j>n} \sigma_j.$$

In particular, if T is of infinite rank, then

$$\|T - T_n\| \leq \max_{j>n} \sigma_j \to 0 \quad as \quad n \to \infty.$$

Proof. By Proposition 13.3, we know that $\{v_n : n \in \Delta\}$ is an orthonormal basis for $\overline{R(T)}$. Therefore, by Fourier expansion (Theorem 4.9(ii)),

$$Tx = \sum_j \langle Tx, v_j\rangle v_j = \sum_j \langle x, T^* v_j\rangle v_j = \sum_j \sigma_j \langle x, u_j\rangle v_j \quad \forall x \in X.$$

Also, since $\{u_n : n \in \Delta\}$ is an orthonormal basis for $N(T)^\perp$ and since $R(T^*) \subseteq N(T)^\perp$, we have

$$T^* y = \sum_j \langle T^* y, u_j\rangle u_j = \sum_j \langle y, Tu_j\rangle u_j = \sum_j \sigma_j \langle y, v_j\rangle u_j \quad \forall y \in Y.$$

Now, let

$$T_n x = \sum_{j=1}^{n} \sigma_j \langle x, u_j\rangle v_j, \quad x \in X,$$

for $n < \operatorname{rank} T$. Then, for every $x \in X$, we have

$$\|(T - T_n)x\|^2 = \sum_{j>n} |\sigma_j|^2 |\langle x, u_j\rangle|^2 \le \max_{j>n} \sigma_j^2 \|x\|^2$$

Thus,

$$\|T - T_n\| \le \max_{j>n} \sigma_j.$$

In case T is of infinite rank, then $\{\lambda \in \mathbb{K} : \lambda = \sigma_n^2 \text{ for some } n \in \Delta\}$ is a denumerable set of eigenvalues of T^*T so that $\sigma_n \to 0$ as $n \to \infty$. Hence, in this case, we have

$$\|T - T_n\| \le \max_{j>n} \sigma_j \to 0 \quad \text{as} \quad n \to \infty.$$

This completes the proof. ∎

Exercise 13.1 Prove the above proposition and theorem using the spectral representation of the compact self-adjoint operator TT^* instead of T^*T. □

13.3 Spectral Representation of Self-Adjoint Operators

To motivate the spectral representation of a general bounded self-adjoint operator, we recall again the finite dimensional situation.

Suppose X is a finite dimensional inner product space and A is a self-adjoint operator on X. Then we have seen that A has the representation

$$A = \sum_{i=1}^{k} \lambda_i P_i,$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A, and for each $i \in \{1, \ldots, k\}$, P_i is the orthogonal projection onto $N(A - \lambda_i I)$. Without loss of generality, we may assume that $\lambda_1 < \lambda_2 < \ldots < \lambda_k$. Writing

$$E_0 = 0, \quad E_1 = P_1 \quad \text{and} \quad E_i = P_1 + \cdots + P_i, \quad i = 2, \ldots, k,$$

we have

$$A = \sum_{i=1}^{k} \lambda_i (E_i - E_{i-1}).$$

It can be seen (*Verify*) that each E_i is an orthogonal projection onto the span of $\cup_{j=1}^{i} R(P_j)$, and

$$0 = E_0 \leq E_1 \leq \ldots \leq E_k = I.$$

The representation of A given above can be viewed as a Riemann-Stieltjes integral with respect to a projection-valued function as follows.

For $\lambda \in \mathbb{R}$, define

$$E(\lambda) = \begin{cases} 0 & \text{if } \lambda < \lambda_1, \\ P_1 + \cdots + P_i & \text{if } \lambda_i \leq \lambda < \lambda_{i+1},\ i = 1, \ldots, k-1, \\ I & \text{if } \lambda \geq \lambda_k. \end{cases}$$

Then, taking $t_1^{(n)}, \ldots, t_n^{(n)}$ such that

$$a < \lambda_1 < t_1^{(n)} < t_2^{(n)} < \ldots < t_n^{(n)} = \lambda_k = b,$$

$$\max\{t_j^{(n)} - t_{j-1}^{(n)} : j = 1, \ldots, n\} \to 0 \quad \text{as} \quad n \to \infty,$$

we see (*Verify*) that

$$A = \sum_{i=1}^{k} \lambda_i [E(\lambda_i) - E(\lambda_{i-1})] = \lim_{n\to\infty} \sum_{j=1}^{n} \tau_j^{(n)} [E(t_j^{(n)}) - E(t_{j-1}^{(n)})],$$

where $\tau_1^{(n)}, \ldots, \tau_n^{(n)}$ satisfy $t_{j-1}^{(n)} \leq \tau_j^{(n)} \leq t_j^{(n)}$ $j = 1, \ldots, n$.

Thus, analogous to the definition of the Riemann-Stieltjes integral, we may write

$$A = \int_a^b \lambda \, dE(\lambda).$$

The above observation is the motivation to look for an integral representation for a self-adjoint operator on an infinite dimensional Hilbert space as well. First we give a definition.

A family $\{E(\lambda) : \lambda \in \mathbb{R}\}$ of orthogonal projections on a Hilbert space X is said to be a **normalized resolution of identity** on an interval $[a, b]$ if

(a) $E(\lambda) = O$ for every $\lambda < a$,

(b) $E(\lambda) \leq E(\mu)$ whenever $\lambda \leq \mu$,

(c) $E(\lambda) = I$ for every $\lambda \geq b$, and

(d) $E(\lambda + 1/n)x \to E(\lambda)x$ as $n \to \infty$ for every $x \in X$.

We observe that the family $\{E(\lambda) : \lambda \in \mathbb{R}\}$ defined earlier in the case of finite dimensional X is a normalized resolution of identity on a closed interval containing the eigenvalues of the operator.

Now suppose that $\{E(\lambda) : \lambda \in \mathbb{R}\}$ is a resolution of identity on $[a, b]$. Since $E(\lambda) \leq E(\mu)$ whenever $\lambda \leq \mu$, we see that for each $x \in X$, the function

$$\lambda \mapsto \langle E(\lambda)x, x\rangle, \qquad \lambda \in [a, b],$$

is monotonically increasing. Hence, using the polarization identity (Theorem 2.20), we see that for each $x, y \in X$, the function

$$\nu_{x,y} : \lambda \mapsto \langle E(\lambda)x, y\rangle, \qquad \lambda \in [a, b]$$

is of bounded variation. Therefore, for every continuous function $f : [a, b] \to \mathbb{K}$, the Riemann-Stieltjes integral

$$\int_a^b f(\lambda) \, d\nu_{x,y}(\lambda)$$

is well defined for every $x, y \in X$. We may also define the integral

$$\int_a^b f(\lambda) \, dE(\lambda),$$

in the sense of Riemann-Stieltjes as follows: For a partition $\Delta : a = \lambda_0 < \lambda_1 < \ldots < \lambda_n = b$, and for $\xi_j \in [\lambda_{j-1}, \lambda_j]$, consider the operator

$$S_n = \sum_{j=1}^{n} f(\xi_j) \left[E(\lambda_j) - E(\lambda_{j-1}) \right].$$

If the sequence (S_n) converges as $\max_{1 \leq j \leq n} (\lambda_j - \lambda_{j-1}) \to 0$, then the resulting limit is denoted by

$$\int_a^b f(\lambda)\, dE(\lambda).$$

Now we state the spectral theorem for a self-adjoint operator. The proof is too much involved to be included in a first course on Functional Analysis. However, we shall indicate the main steps involved in the proof. For a detailed proof, one may refer to Bachmann and Narici [6].

For a self-adjoint operator $A \in \mathcal{B}(X)$, recall the definitions

$$\alpha_A = \inf w(A), \qquad \beta_A = \sup w(A),$$

where $w(A) := \{\langle Ax, x\rangle : \|x\| = 1\}$ is the numerical range of A.

Theorem 13.5 *Let X be a Hilbert space and $A \in \mathcal{B}(X)$ be a self-adjoint operator. Then there exists a normalized resolution of identity $\{E(\lambda) : \lambda \in \mathbb{R}\}$ on $[\alpha_A, \beta_A]$ such that*

$$\langle Ax, y\rangle = \int_{\alpha_A}^{\beta_A} \lambda\, d\langle E(\lambda)x, y\rangle \quad \forall x, y \in X.$$

Moreover,

$$A = \int_{\alpha_A}^{\beta_A} \lambda\, dE(\lambda).$$

In fact, for every real-valued $f \in C[\alpha_A, \beta_A]$,

$$f(A) = \int_{\alpha_A}^{\beta_A} f(\lambda)\, dE(\lambda).$$

Main steps involved in the proof. For $x, y \in X$, define the function $F_{x,y} : C[\alpha_A, \beta_A] \to \mathbb{K}$ by

$$F_{x,y}(f) = \langle f(A)x, y\rangle, \qquad f \in C[\alpha_A, \beta_A].$$

The proofs of the statements (1) – (3) below are omitted.

(1) *$F_{x,y} \in (C[\alpha_A, \beta_A])'$ for all $(x, y) \in X \times X$.*

Hence, by Theorem 8.8, there exists a unique $v_{x,y} \in NBV[\alpha_A, \beta_A]$ such that

$$F_{x,y}(f) = \int_{\alpha_A}^{\beta_A} f(t)\, dv_{x,y}(t), \qquad f \in C[\alpha_A, \beta_A].$$

For each $t \in [\alpha_A, \beta_A]$, define $\Psi_t : X \times X \to \mathbb{K}$ by

$$\Psi_t(x, y) = v_{x,y}(t), \qquad (x, y) \in X \times X.$$

(2) *Ψ_t is a continuous sesquilinear functional and*

$$\Psi_t(y, x) = \overline{\Psi_t(x, y)} \quad \forall\, t \in [\alpha_A, \beta_A] \quad \forall\, (x, y) \in X \times X.$$

Hence, by Theorem 11.6, for each $t \in [\alpha_A, \beta_A]$, there exists a unique $E(t) \in \mathcal{B}(X)$ such that

$$\Psi_t(x, y) = \langle E(t)x, y\rangle, \qquad (x, y) \in X \times X.$$

Thus, for every $t \in [\alpha_A, \beta_A]$,

$$v_{x,y}(t) = \langle E(t)x, y\rangle, \qquad x, y \in X.$$

Besides, the property $\Psi_t(y, x) = \overline{\Psi_t(x, y)}$ of Ψ shows that $E(t)$ is self-adjoint for each $t \in [\alpha_A, \beta_A]$.

(3) *The family $\{E(\lambda) : \alpha_A \leq \lambda \leq \beta_A\}$ is a normalized resolution of identity.*

Then, for every $f \in C[\alpha_A, \beta_A]$, we have

$$\langle f(A)x, y\rangle = F_{x,y}(f) = \int_{\alpha_A}^{\beta_A} f(t)\, dv_{x,y}(t) = \int_{\alpha_A}^{\beta_A} f(t)\, d\langle E(t)x, y\rangle.$$

In particular, with $f(t) = t$, $t \in [\alpha_A, \beta_A]$, we have

$$\langle Ax, y\rangle = \int_{\alpha_A}^{\beta_A} t\, d\langle E(t)x, y\rangle, \qquad x, y \in X.$$

Let $f \in C[\alpha_A, \beta_A]$ be real-valued. Then, we know by Theorem 12.14(vii) that $f(A)$ is a self-adjoint operator. Consider a partition

$\alpha_A = \lambda_0 < \lambda_1 < \ldots < \lambda_n = \beta_A$ of $[\alpha_A, \beta_A]$, and $\xi_j \in [\lambda_{j-1}, \lambda_j]$ for $j = 1, \ldots, n$. Let

$$S_n = \sum_{j=1}^{n} f(\xi_j)\,[E(\lambda_j) - E(\lambda_{j-1})].$$

Then, for every $x \in X$, we have

$$\langle (f(A) - S_n)x, x\rangle = \sum_{j=1}^{n} \int_{\lambda_{j-1}}^{\lambda_j} [f(\lambda) - f(\xi_j)]\, d\nu_{x,x}(\lambda).$$

Now, let $\varepsilon > 0$ and let $n \in \mathbb{N}$ be such that

$$\max_{1 \le j \le n} (\lambda_j - \lambda_{j-1}) < \varepsilon \quad \forall n \ge N.$$

Then we have

$$\langle (f(A) - S_n)x, x\rangle = \sum_{j=1}^{n} \int_{\lambda_{j-1}}^{\lambda_j} [f(\lambda) - f(\xi_j)]\, d\nu_{x,x}(\lambda).$$

Now, using the fact that

$$\nu_{x,x}(\alpha_A) = \langle E(\alpha_A)x, x\rangle = 0, \qquad \nu_{x,x}(\beta_A) = \langle E(\beta_A)x, x\rangle = \langle x, x\rangle$$

and uniform continuity of f, we get

$$\langle (f(A) - S_n)x, x\rangle \le \varepsilon \sum_{j=1}^{n} \int_{\lambda_{j-1}}^{\lambda_j} d\nu_{x,x}(\lambda) = \varepsilon \int_{\alpha_A}^{\beta_A} d\nu_{x,x}(\lambda) = \varepsilon \langle x, x\rangle$$

for all $x \in X$. Similarly, we have

$$\langle (f(A) - S_n)x, x\rangle \ge -\varepsilon \langle x, x\rangle \quad \forall x \in X.$$

Thus,

$$|\langle (f(A) - S_n)x, x\rangle| \le \varepsilon \|x\|^2 \quad \forall x \in X.$$

Hence,

$$\|f(A) - S_n\| = \sup\{|\langle (f(A) - S_n)x, x\rangle| : \|x\| = 1\} \to 0$$

as $n \to \infty$ showing that

$$f(A) = \int_{\alpha_A}^{\beta_A} f(\lambda)\, dE_\lambda.$$

In particular, by taking $f(t) = t$ for every $t \in [\alpha_A, \beta_A]$, we get

$$A = \int_{\alpha_A}^{\beta_A} \lambda\, dE_\lambda. \quad \blacksquare$$

PROBLEMS

In the following, X and Y denote Hilbert spaces.

1. Let $A : X \to X$ be a nonzero compact self-adjoint operator and $\lambda_1, \lambda_2, \ldots$ be its nonzero eigenvalues. For each $j = 1, 2, \ldots$, let $\{u_1^{(j)}, \ldots, u_{m_j}^{(j)}\}$ be an orthonormal basis of $N(A - \lambda_j I)$. Show that $\cup_j \{\{u_1^{(j)}, \ldots, u_{m_j}^{(j)}\}$ is an orthonormal basis of $N(A)^\perp$.

2. Show that a compact self-adjoint operator $A : X \to X$ is of finite rank if and only if it has only a finite number of nonzero eigenvalues.

3. If $A = \sum_j \lambda_j P_j$ is the spectral representation of a compact self-adjoint operator $A : X \to X$ on a Hilbert space X, then show that

$$p(A) = \sum_j p(\lambda_j) P_j$$

for every polynomial $p(t)$.

4. If $A = \sum_j \lambda_j P_j$ is the spectral representation of a positive compact self-adjoint operator $A : X \to X$ on a Hilbert space X, and if $B = \sum_j \sqrt{\lambda_j} P_j$, then show that $B^2 = A$.

5. Suppose (u_n) and (v_n) are orthonormal sequences in X and Y, respectively, and (σ_n) is a sequence of positive real numbers that converge to zero. Then show that $T : X \to Y$ defined by

$$Tx = \sum_n \sigma_n \langle x, u_n \rangle v_n, \qquad x \in X,$$

is a compact operator from X to Y, and $\{(\sigma_n, u_n, v_n) : n = 1, 2, \ldots\}$ is the singular system of T.

6. Let $T : X \to Y$ be a nonzero compact operator and $\zeta(T)$ be the set of singular values of T. Show that

(a) $\zeta(T) = \{\sigma > 0 : \sigma^2 \in \sigma_{\text{eig}}(T^*T)\}$,

(b) $\zeta(T) \neq \varnothing$,

(c) $\|T\| = \sup\{\sigma : \sigma \in \zeta(T)\}$, and

(d) $\zeta(T)$ is a finite set if and only if T is of finite rank.

7. Let $A = \sum_{j=1}^{k} \lambda_j P_j$ be the spectral representation of a self-adjoint operator $A : X \to X$ on a finite dimensional inner product space X, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A, and for each j, P_j is the orthogonal projection onto the eigenspace $N(A - \lambda_j I)$. Show that, if $E_0 = 0$, and $E_j = P_1 + \cdots + P_j$ for $j = 1, \ldots, k$, then each E_j is an orthogonal projection onto span $[\cup_{i=1}^{j} R(P_i)]$, and $E_i E_j = E_i$ for $i \leq j$.

8. If $\{E(\lambda) : \lambda \in \mathbb{R}\}$ is a resolution of identity on $[a, b]$, then show the following:

(a) $E(\lambda)E(\mu) = E(\lambda)$ for $\lambda \leq \mu$.

(b) $R(E(\mu))$ is the closure span $[\cup_{\lambda \leq \mu} R(E(\lambda))]$ for each $\mu \in \mathbb{R}$.

(c) $\lambda \mapsto \langle E(\lambda)x, y\rangle$ is a function of bounded variation on $[a, b]$ for every $x, y \in X$.

[*Hint* : For proving (c), one may use the polarization identity (Theorem 2.20).]

14

Solution of Operator Equations

Many of the problems which occur in science and engineering can be modelled as an operator equation

$$Tx = y, \tag{14.1}$$

where $T : X \to Y$ is a linear operator between normed linear spaces X and Y. Equation (14.1) is said to be a 'well-posed equation' if for every $y \in Y$, there exists a unique $x \in X$ satisfying (14.1), and the solution x depends continuously on the *data* y, i.e., if (y_n) is a sequence in Y which converges to y, and if $x_n \in X$ is such that $Tx_n = y_n$ for all large enough $n \in \mathbb{N}$, then (x_n) converges to x. Equation (14.1) is said to be an 'ill-posed equation' if it is not a well-posed equation.

We consider equations of the form

$$Ax - \lambda x = y, \qquad Tx = y,$$

where $A : X \to X$ and $T : X \to Y$ are bounded operators between Banach spaces, and λ is a nonzero scalar which does not belong to the spectrum of A. Recall that if A is a compact operator, then the requirement that $\lambda \neq 0$ is not in the spectrum of A is equivalent to the fact that $\lambda \neq 0$ is not an eigenvalue of A (Theorem 10.23). Also, if T is a compact operator of infinite rank, then its range is not closed (Theorem 9.11), and even if the equation $Tx = y$ has a unique solution for a given $y \in R(T)$, the solution does not depend continuously on the data y (Theorem 9.8). Thus, when A and T are compact operators, the above two equations are prototypes of well-posed equations and ill-posed equations, respectively.

In this chapter we only indicate how some of the results that we have already proved in the earlier chapters can be applied to discuss the problem of solving operator equations of the above forms.

A detailed discussion of such equations is beyond the scope of this course. For a detailed study of well-posed and ill-posed equations, one may refer to the monographs by Atkinson [4] and Engl, Hanke and Neubauer [13], respectively.

In Section 14.1, we consider the well-posed operator equation $Ax - \lambda x = y$, where $A : X \to X$ is a compact self-adjoint operator on a Hilbert space. We make use of the spectral theorem for a compact self-adjoint operator (Theorem 13.2) for representing the solution in terms of 'known' entities. In Section 14.2, we discuss the ill-posed equation $Tx = y$. A modified form of solution, namely, the *generalized solution* of such equations, is discussed for any bounded operator $T : X \to Y$ between Hilbert spaces. In order to illustrate the ill-posedness of the equations, we restrict $T : X \to Y$ to be a compact operator. In this case, we rely on its singular value representation (Theorem 13.4) for representing and approximating a solution. In Section 14.3, we shall indicate a general procedure for obtaining approximate solutions for the well-posed equation when the operator involved is approximated by a sequence of operators. This is done in a more general setting in which X and Y are Banach spaces.

14.1 Well-Posed Operator Equations

Consider the operator equation

$$Ax - \lambda x = y, \tag{14.2}$$

where $A : X \to X$ is a compact self-adjoint operator on a Hilbert space X and λ is a nonzero scalar which is not an eigenvalue of A. We have already pointed out that the assumption on λ implies that equation (14.2) has a unique solution x for every $y \in X$ (cf. Theorem 10.23).

Let

$$Ax = \sum_j \mu_j \langle x, u_j \rangle u_j \quad \forall x \in X$$

be the spectral representation of A (Theorem 13.2), where $u_1, u_2, \ldots$ are eigenvectors of A which form an orthonormal basis for $N(A)^\perp$, and the scalars $\mu_1, \mu_2, \ldots$, possibly repeated, are the positive eigenvalues of A. Then, for $y \in X$, and for $x \in X$ satisfying 14.2), we

have
$$x = \frac{1}{\lambda}(-y + Ax) = \frac{1}{\lambda}\big[-y + \sum_j \mu_j \langle x, u_j\rangle u_j\big].$$

Since
$$\langle Ax, u_n\rangle - \lambda\langle x, u_n\rangle = \langle y, u_n\rangle,$$
$$\langle Ax, u_n\rangle = \langle x, Au_n\rangle = \mu_n\langle x, u_n\rangle$$

for all n, it follows that
$$(\mu_n - \lambda)\langle x, u_n\rangle = \langle y, u_n\rangle$$

so that we have
$$x = -\frac{1}{\lambda}y + \frac{1}{\lambda}\sum_j \frac{\mu_j}{\mu_j - \lambda}\langle y, u_j\rangle u_j.$$

Thus, we have expressed the solution of $Ax - \lambda x = y$ in terms of y, and the eigenvalues and eigenvectors of A.

In practice, it may not be possible to know all the eigenvalues and the corresponding eigenvectors. So, naturally, one has to satisfy by an approximation of the solution obtained using a few of the eigenvalues, say $\mu_1, \ldots, \mu_n$, and the corresponding eigenvectors $u_1, \ldots, u_n$. Thus, we consider
$$x_n = -\frac{1}{\lambda}y + \frac{1}{\lambda}\sum_{j=1}^{n} \frac{\mu_j}{\mu_j - \lambda}\langle y, u_j\rangle u_j.$$

The question now is how well this x_n approximates the solution x. To see this, we observe that
$$\|x - x_n\|^2 = \frac{1}{|\lambda|^2}\sum_{j=n+1}^{\infty}\frac{|\mu_j|^2}{|\mu_j - \lambda|^2}|\langle y, u_j\rangle|^2 \le \Big(\frac{\varepsilon_n}{|\lambda| d_n}\Big)^2 \sum_{j=n+1}^{\infty}|\langle y, u_j\rangle|^2,$$

where
$$\varepsilon_n = \max_{j>n}|\mu_j|, \qquad d_n = \inf\{|\lambda - \mu_j| : j > n\}.$$

Note that $\varepsilon_n \to 0$ as $n \to \infty$, and
$$d_n \ge d := \operatorname{dist}(\lambda, \{\mu_1, \mu_2, \ldots\}) \;\; \forall n.$$

Thus, we get
$$\|x - x_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

In fact, we have

$$\|x - x_n\| \leq c_n(y)\varepsilon_n \leq \frac{1}{d|\lambda|}\|y\|\varepsilon_n,$$

where

$$c_n(y) = \frac{1}{|\lambda|d_n}\sqrt{\sum_{j=n+1}^{\infty} |\langle y, u_j\rangle|^2} \to 0 \quad \text{as} \quad n \to \infty.$$

14.2 Ill-Posed Operator Equations

Let X and Y be Hilbert spaces and $T : X \to Y$ be a bounded linear operator. For $y \in Y$, consider the operator equation

$$Tx = y. \tag{14.3}$$

Clearly, the above equation has a unique solution for every $y \in Y$ if and only if T is bijective. We have already mentioned that if T is a compact operator, then, unlike (14.2), equation (14.3) need not have a solution for every $y \in Y$ (cf. Theorem 9.11). Even if a unique solution exists for some $y \in Y$, it need not depend continuously on the 'data' y (cf. Proposition 9.8). Thus, in such cases, (14.3) is an ill-posed equation.

14.2.1 Least Residual Norm Solution

For ill-posed equations, what one looks for is not a solution, but a *least-residual norm solution* or an *LRN solution*. By an **LRN-solution**, we mean an element $x \in X$ such that

$$\|Tx - y\| = \inf \{\|Tu - y\| : u \in X\}.$$

Clearly, if (14.3) has a solution, then it is an LRN solution. Note that, if an LRN solution exists, then its uniqueness is guaranteed only if $N(T) = \{0\}$. Let us give a necessary and sufficient condition for the existence of an LRN solution.

Theorem 14.1 *Let X and Y be Hilbert spaces, $T \in \mathcal{B}(X, Y)$ and $P : Y \to Y$ be the orthogonal projection onto $\overline{R(T)}$. Then the following are equivalent:*

(i) *$Tx = y$ has an LRN solution.*

(ii) *$Tx = Py$ has a solution.*

(iii) $T^*Tx = T^*y$ *has a solution.*

(iv) $y \in R(T) + R(T)^\perp$.

Proof. As a consequence of the projection theorem (see the remarks following Corollary 2.48),

$$\inf\{\|v - y\| : v \in \overline{R(T)}\} = \|Py - y\|$$

so that

$$\inf\{\|Tu - y\| : u \in X\} = \|Py - y\|.$$

Thus, an element $x \in X$ is an LRN solution if and only if

$$\|Tx - y\| = \|Py - y\|.$$

Now, using the fact that $R(P) \perp N(P)$, we have

$$\|Tx - y\|^2 = \|Tx - Py\|^2 + \|Py - y\|^2.$$

Thus, (14.3) has an LRN solution if and only if $Tx = Py$ has a solution, and the equation $Tx = Py$ has a solution if and only if $y \in R(T) + R(T)^\perp$. Also, we observe that

$$\begin{aligned} Tx = Py &\iff P(Tx - y) = 0 \\ &\iff Tx - y \in R(T)^\perp = N(T^*) \\ &\iff T^*Tx = T^*y. \end{aligned}$$

This complete the proof. ∎

14.2.2 Generalized Solution and Generalized Inverse

For $y \in Y$, let us denote the set of all LRN solutions of (14.3) by

$$S_y = \{x \in X : \|Tx - y\| \leq \|Tu - y\| \ \forall u \in X\}.$$

Note that

$$S_y \neq \varnothing \iff y \in R(T) + R(T)^\perp.$$

Suppose $y \in R(T) + R(T)^\perp$ and $x \in S_y$. Let $\hat{x} \in N(T)^\perp$ and $u \in N(T)$ be such that

$$x = \hat{x} + u.$$

Then it is obvious that

$$\hat{x} \in S_y, \qquad \|\hat{x}\| \leq \|x\|.$$

In fact, it can be seen (Problem 2) that if $y \in R(T) + R(T)^{\perp}$, then there exists a unique $\hat{x} \in S_y$ such that

$$\|\hat{x}\| \le \|x\| \quad \forall x \in S_y,$$

and this $\hat{x}$ belongs to $N(T)^{\perp}$.

For $y \in R(T) + R(T)^{\perp}$, the unique $\hat{x}$ obtained above is called the **LRN solution of minimal norm** of (14.3) or the **generalized solution** of (14.3). The map

$$T^{\dagger} : R(T) + R(T)^{\perp} \to X$$

which associates each $y \in R(T) + R(T)^{\perp}$ the generalized solution of (14.3), is called the **generalized (Moore-Penrose) inverse** of T.

The proof of the following theorem is left as an exercise for the reader (see Problem 3).

Theorem 14.2 *For $T \in \mathcal{B}(X,Y)$, the operator $T^{\dagger}$ is a closed linear operator, and it is continuous if and only if $R(T)$ is closed in Y.*

Picard's criterion

Now suppose that $y \in Y$ is such that $S_y \neq \varnothing$ and $T : X \to Y$ is a compact operator. Let us find a representation for the generalized solution of (14.3).

Let

$$Tx = \sum_n \sigma_n \langle x, u_n \rangle v_n, \quad x \in X$$

be the singular value representation (Theorem 13.4) of T corresponding to the singular system $\{(\sigma_n, u_n, v_n) : n = 1, 2, \ldots\}$ of T. Recall that $\{u_1, u_2, \ldots\}$ is an orthonormal basis of $N(T)^{\perp}$, and $\{v_1, v_2, \ldots\}$ is an orthonormal basis of $\overline{R(T)}$ (cf. Proposition 13.3).

From the above discussion we know that, if $x \in S_y$, then

$$x = \hat{x} + u \quad \text{with} \quad \hat{x} \in N(T)^{\perp}, \; u \in N(T),$$

where $\hat{x}$ is the generalized solution of (14.3). Since $\{u_n : n = 1, 2, \ldots\}$ is an orthonormal basis of $N(T)^{\perp}$, by Fourier expansion (Theorem 4.9(ii)), we have

$$\hat{x} = \sum_n \langle x, u_n \rangle u_n.$$

In view of Theorem 14.1, since $\hat{x} \in S_y$, we have $T\hat{x} = Py$ so that

$$\langle T\hat{x}, v_n \rangle = \langle Py, v_n \rangle = \langle y, v_n \rangle \quad \forall n,$$

where $P : Y \to Y$ is the orthogonal projection onto $\overline{R(T)}$. Also, note that

$$\langle T\hat{x}, v_n \rangle = \langle \hat{x}, T^* v_n \rangle = \sigma_n \langle \hat{x}, u_n \rangle.$$

Thus we have

$$\sigma_n \langle \hat{x}, v_n \rangle = \langle y, v_n \rangle, \qquad \langle \hat{x}, u_n \rangle = \frac{1}{\sigma_n} \langle y, v_n \rangle$$

so that

$$\hat{x} = \sum_n \frac{1}{\sigma_n} \langle y, v_n \rangle u_n.$$

This expression for $\hat{x}$ shows that the condition

$$\sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty,$$

called *Picard's criterion*, is necessary on y for the existence of an LRN solution for (14.3). Thus, by Theorem 14.1,

$$R(T) + R(T)^\perp \subseteq \left\{ y \in Y : \sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty \right\}.$$

Conversely, suppose that $y \in Y$ is such that the Picard's criterion is satisfied. Then it is seen (Problem 6(i)) that

$$\hat{x} := \sum_n \frac{1}{\sigma_n} \langle y, v_n \rangle u_n$$

is the generalized solution of (14.3).

Thus, we have proved the following theorem.

Theorem 14.3 *Let $T : X \to Y$ be a compact operator between Hilbert spaces, and $\{(\sigma_n, u_n, v_n) : n = 1, 2, \ldots\}$ be a singular system of T. Then the equation $Tx = y$ has an LRN solution if and only if*

$$\sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty,$$

and in that case,

$$\begin{aligned} R(T) + R(T)^\perp &= \left\{ y \in Y : \sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty \right\}, \\ T^\dagger y &= \sum_n \frac{1}{\sigma_n} \langle y, v_n \rangle u_n. \end{aligned}$$

Note that the representation of the generalized solution $\hat{x} := T^\dagger y$ given in the above theorem involves all the nonzero singular values of T. Suppose we have at our disposal, only a finite subcollection of $\{(\sigma_j, u_j, v_j) : j = 1, 2, \ldots\}$, say, $\{(\sigma_j, u_j, v_j) : j = 1, \ldots, n\}$. Then the question is whether the element

$$x_n = \sum_{j=1}^{n} \frac{1}{\sigma_j} \langle y, v_j \rangle u_j$$

obtained by truncating the series expansion of $\hat{x}$ is an approximation of $\hat{x}$. The answer is obvious by the fact that (u_n) is an orthonormal sequence. Indeed,

$$\|\hat{x} - x_n\|^2 = \sum_{j=n}^{\infty} \frac{|\langle y, v_j \rangle|^2}{|\sigma_j|^2} \to 0 \quad \text{as} \quad n \to \infty.$$

One important point to observe is that if we make a small error in the 'data' y having a nonzero component in the direction of v_n for large n, then the error in the solution can be very large. For example, suppose that T is of infinite rank, and suppose we take

$$y_n = y + \sqrt{\sigma_n} v_n$$

instead of y. Then we have $\|y - y_n\| = \sqrt{\sigma_n} \to 0$ as $n \to \infty$. But, note that

$$\hat{x}_n := T^\dagger y_n = \sum_j \frac{1}{\sigma_j} \langle y + \sqrt{\sigma_n} v_n, v_j \rangle u_j = \hat{x} + \frac{1}{\sqrt{\sigma_n}} u_n,$$

$$\|\hat{x} - \hat{x}_n\| = \frac{1}{\sqrt{\sigma_n}} \to \infty \quad \text{as} \quad n \to \infty.$$

The main idea used here is the fact that, if T is a compact operator of infinite rank, then 0 is an accumulation point of $\sigma(T^*T)$. This property is not special to compact operators alone; it holds good for every bounded operator on a Banach space with its range not closed (cf. Kulkarni and Nair [19]).

14.3 Approximate Solutions

Suppose X and Y are Banach spaces, and $A : X \to Y$ is a bijective bounded linear operator. Obviously, this assumption implies that for every $y \in Y$, there exists a unique $x \in X$ such that

$$Ax = y. \tag{14.4}$$

However, in practical problems, either the operator is not available exactly or while solving the equation numerically, one may approximate the operator A by another 'simpler' operator $\tilde{A}$. This can be true in the case of the 'data' y as well. Thus, what one solves is an equation of the form

$$\tilde{A}\tilde{x} = \tilde{y}, \tag{14.5}$$

where $(\tilde{A}, \tilde{y})$ is close to (A, y) in some sense. Let us assume, for the time being, that $\tilde{A} : X \to Y$ is invertible, and let x and $\tilde{x}$ be the solutions of (14.4) and (14.5), respectively. Then from the relations

$$\tilde{A}(x - \tilde{x}) = (\tilde{A} - A)x + (y - \tilde{y}), \tag{14.6}$$

$$A(\tilde{x} - x) = (A - \tilde{A})\tilde{x} + (\tilde{y} - y),$$

we have the error bounds

$$\begin{aligned} \|x - \tilde{x}\| &\leq 2\,\|\tilde{A}^{-1}\| \max\left\{\|(\tilde{A} - A)x\|, \|y - \tilde{y}\|\right\}, \\ \|x - \tilde{x}\| &\leq 2\,\|A^{-1}\| \max\left\{\|(A - \tilde{A})\tilde{x}\|, \|y - \tilde{y}\|\right\}. \end{aligned}$$

A common situation in numerical approximation is that we have a sequence (A_n) of operators in $\mathcal{B}(X, Y)$, which approximates A in some sense, A_n is bijective for all sufficiently large n, say for $n \geq N$, and $\{\|A_n\| : n \geq N\}$ is bounded. In such a case, one may take $\tilde{A}$ to be A_n for some $n \geq N$. Then the above discussion leads to the following theorem.

Theorem 14.4 *Suppose (A_n) in $\mathcal{B}(X, Y)$ is such that for some positive integer N, A_n is invertible for all $n \geq N$ and $\{\|A_n^{-1}\| : n \geq N\}$ is bounded. Let (y_n) be a sequence in Y, and let $x_n \in X$ be such that $A_n x_n = y_n$ for all $n \geq N$. Then*

$$\|x - x_n\| \leq c_1 \max\{\|(A - A_n)x\|, \|y_n - y\|\},$$

$$\|x - x_n\| \leq c_2 \max\{\|(A - A_n)x_n\|, \|y_n - y\|\}$$

for every $n \geq N$, where

$$c_1 \geq 2\,\|A_n^{-1}\| \quad \forall n \geq N, \qquad c_2 \geq 2\,\|A^{-1}\|.$$

If, in addition, $\|(A - A_n)x\| \to 0$ and $\|y - y_n\| \to 0$ as $n \to \infty$, then

$$\|x - x_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

Now let us consider a situation where the assumptions of the above theorem are satisfied: Suppose (A_n) is such that $\|A-A_n\| \to 0$ as $n \to \infty$. From the identity

$$A_n = [I - (A - A_n)A^{-1}]A$$

it follows, by Theorem 10.8, that if n is such that

$$\|A - A_n\| < \frac{1}{\|A^{-1}\|},$$

then A_n is invertible and

$$\|A_n^{-1}\| \le \frac{\|A^{-1}\|}{1 - \|A - A_n\| \, \|A^{-1}\|}.$$

Thus, if $N \in \mathbb{N}$ is such that $\|A - A_n\| \, \|A^{-1}\| \le 1/2$, then A_n is invertible for all $n \ge N$ and $\|A_n^{-1}\| \le 2\|A^{-1}\|$ for all $n \ge N$.

14.3.1 Projection Methods

Now let us describe a situation which occurs often in numerical analysis. Suppose $Y = X$ and $A = K - \lambda I$, where $K : X \to X$ is a compact operator and $\lambda \neq 0$ is not an eigenvalue of K. Thus, in this case, (14.4) is

$$Kx - \lambda x = y. \tag{14.7}$$

Suppose (P_n) is a sequence of projection operators in $\mathcal{B}(X)$ such that

$$\|P_n v - v\| \to 0 \quad \text{as} \quad n \to \infty$$

for every $v \in X$. In practical situations, each P_n may be a finite rank projection such that $\cup_{n=1}^{\infty} R(P_n)$ is dense in X. Now, by Corollary 6.6, it follows that

$$\|K - P_n K\| \to 0 \quad \text{as} \quad n \to \infty.$$

Galerkin approximation

By taking

$$A_n = P_n K - \lambda I, \qquad y_n = P_n y,$$

we have

$$\|A - A_n\| = \|K - P_n K\| \to 0, \qquad \|y - y_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

The approximate equation to be solved in this case is

$$K_n x_n - \lambda x_n = y_n, \tag{14.8}$$

where $K_n = P_n K$. Note that the solution x_n of the above equation belongs to $R(P_n)$. Suppose $N \in \mathbb{N}$ is such that A_n is invertible for all $n \geq N$. In this case, the error bounds in Theorem 14.4 take the form

$$\|x - x_n\| \leq c_1 \max\{\|(I - P_n)Kx\|, \|(I - P_n)y\|\},$$

$$\|x - x_n\| \leq c_2 \max\{\|(I - P_n)Kx_n\|, \|(I - P_n)y\|\}.$$

In fact, in this special case, we can derive a slightly better error estimate, namely,

$$\|x - x_n\| \leq c\,\|P_n x - x\| \quad \forall n \geq N,$$

where $c \geq |\lambda|\|(P_n K - \lambda I)^{-1}\|$ for all $n \geq N$. This follows from the relation

$$\begin{aligned}
(P_n K - \lambda I)(x - x_n) &= (P_n K - \lambda I)x - (P_n K - \lambda I)x_n \\
&= (P_n K - \lambda I)x - P_n y \\
&= P_n(Kx - y) - \lambda x \\
&= \lambda(P_n x - x).
\end{aligned}$$

The above estimate also shows that the solution x_n of the equation (14.8) has the best possible order of convergence in the sense that no approximation from $R(P_n)$ can give higher order. To see this, note that

$$\|x - P_n x\| = \|(I - P_n)(x - u)\| \leq \|I - P_n\|\,\|x - u\| \quad \forall u \in R(P_n)$$

so that taking infimum over all $u \in R(P_n)$,

$$\|x - P_n x\| \leq \|I - P_n\| \operatorname{dist}(x, R(P_n)).$$

Now, since (P_n) converges pointwise to the identity operator, by Banach-Steinhaus theorem (Theorem 6.3), $(\|P_n\|)$ is bounded, so that

$$\|x - P_n x\| \leq (1 + c_0) \operatorname{dist}(x, R(P_n)),$$

where $c_0 \geq \|P_n\|$ for all $n \in \mathbb{N}$.

The above procedure for obtaining an approximation (x_n) using the sequence (P_n) of projections is called the **Galerkin method**.

Now, let us look at the computational implications of the Galerkin method, and an iterated form of the Galerkin method.

If $R(P_n)$ is finite dimensional, say with basis $\{u_1, \ldots, u_n\}$ and if $f_1, \ldots, f_n$ are continuous linear functionals on X such that $f_i(u_j) = \delta_{ij}$, then

$$P_n v = \sum_{j=1}^{n} f_j(v) u_j \quad \forall v \in X$$

so that what we are looking for is an n-tuple $(\alpha_1, \ldots, \alpha_n)$ of scalars such that $x_n := \sum_{j=1}^{n} \alpha_j u_j$ satisfies the equation

$$K_n x_n - \lambda x_n = y_n,$$

where $K_n = P_n K$ and $y_n = P_n y$. Equivalently, we are looking for scalars $\alpha_1, \ldots, \alpha_n$ which satisfy the simultaneous equations

$$\sum_{j=1}^{n} a_{ij}\alpha_j - \lambda \alpha_i = \beta_i, \qquad i = 1, \ldots, n,$$

where $a_{ij} = \langle Au_j, u_i \rangle$ and $\beta_i = \langle y, u_i \rangle$ for $i, j = 1, \ldots, n$.

Iterated Galerkin approximation

We may observe from (14.7) that

$$x = \frac{1}{\lambda}(Kx - y).$$

We see (*Verify*) that the sequence $(\tilde{x}_n)$ defined by

$$\tilde{x}_n = \frac{1}{\lambda}(Kx_n - y)$$

satisfies

$$\|x - \tilde{x}_n\| \leq \frac{1}{|\lambda|}\|K(x - x_n)\|, \tag{14.9}$$

so that $(\tilde{x}_n)$ is an approximation of x.

The above approximation $(\tilde{x}_n)$ is called the **iterated Galerkin approximation** of the solution of the equation (14.7).

We may observe that $\tilde{x}_n$ satisfies the equation

$$\tilde{K}_n \tilde{x}_n - \lambda \tilde{x}_n = y,$$

where $\tilde{K}_n = KP_n$, and $\tilde{K}_n - \lambda I$ is invertible for all $n \geq N$ (*Verify*). Now, taking $\tilde{A} = \tilde{K}_n - \lambda I$ in (14.6), it follows that

$$x - \tilde{x}_n = (\tilde{K}_n - \lambda I)^{-1} K(P_n - I)x.$$

This, together with the fact that $(I - P_n)x_n = 0$, yields an error estimate

$$\|x - \tilde{x}_n\| \leq \|(\tilde{K}_n - \lambda I)^{-1}\| \, \|K(P_n - I)(x - x_n)\|. \tag{14.10}$$

Since $(\tilde{K}_n - \lambda I)^{-1} y \to (K - \lambda I)^{-1} y$ for every $y \in X$, it follows by Banach-Steinhaus theorem (Theorem 6.3) that $(\|(\tilde{K}_n - \lambda I)^{-1}\|)$ is bounded. Hence, from (14.9) and (14.10), we have

$$\|x - \tilde{x}_n\| \leq \kappa \min \{\|K(x - x_n)\|, \|K(P_n - I)\| \, \|x - x_n\|\},$$

where

$$\kappa \geq \max\{\|K\|/|\lambda|, \|(\tilde{K}_n - \lambda I)^{-1}\|\} \quad \forall n \geq N.$$

The above error bound shows that if $\|K(I - P_n)\| \to 0$, then the iterated Galerkin method provides a better approximation for the solution of (14.7) than by the Galerkin method.

14.3.2 Nyström Method

In this section we describe an approximation procedure for solving the integral equation

$$\int_a^b k(s,t)x(t)\,dt - \lambda x(s) = y(s), \qquad a \leq s \leq b,$$

419

where $y(\cdot)$ and $k(\cdot,\cdot)$ are continuous functions. The above equation can be written as

$$Kx - \lambda x = y,$$

where

$$(Kx)(s) = \int_a^b k(s,t)x(t)\,dt, \qquad x \in C[a,b],$$

defines a compact operator on $(C[a,b], \|\cdot\|_\infty)$. For approximating K, we consider the **Nyström approximation** (K_n) of K, defined by

$$(K_n x)(s) = \sum_{j=1}^{n} k(s, t_{i,n}) x(t_{i,n}) w_{i,n}, \qquad x \in C[a,b],$$

where $t_{1,n}, t_{2,n}, \ldots, t_{n,n}$ are the 'nodes' and $w_{1,n}, w_{2,n}, \ldots, w_{n,n}$ are the 'weights' of a convergent quadrature formula:

$$Q_n(u) = \sum_{i=1}^{n} u(t_{i,n})w_{i,n}, \qquad u \in C[a,b], \quad n \in \mathbb{N}.$$

It was an exercise earlier (Chapter 9, Problem 5) that $\|Kx - K_n x\| \to 0$ for every $x \in C[a,b]$ and

$$\|(K - K_n)K\| \to 0, \qquad \|(K - K_n)K_n\| \to 0$$

as $n \to \infty$. Using these results and the identity

$$[\lambda I - (K - K_n)](\lambda I - K_n) = \lambda(\lambda I - K) + (K - K_n)K_n,$$

it follows (*How?*) that $K_n - \lambda I$ is invertible for all large enough n, say for $n \geq N$, and $\{\|(K_n - \lambda I)^{-1}\| : n \geq N\}$ is bounded. Thus, Theorem 14.4 gives an error estimate for the approximation x_n which is the unique solution of

$$K_n x_n - \lambda x_n = y,$$

namely,

$$\|x - x_n\| \leq c \min\{\|(K - K_n)x\|, \|(K - K_n)x_n\|\}.$$

by the definition of K_n, we have

$$\begin{aligned} \lambda x_n(s) &= K_n x_n(s) - y(s) \\ &= \sum_{j=1}^{n} k(s, t_{i,n}) x_n(t_{i,n}) w_{i,n} - y(s). \end{aligned}$$

Thus, the approximation x_n is completely determined by the scalars $\alpha_i := x_n(t_{i,n})$, $i = 1, \ldots, n$ which are obtained by solving the system of equations

$$\sum_{j=1}^{n} k(t_{j,n}, t_{i,n}) w_{i,n} \alpha_j - \lambda \alpha_i = y(t_{i,n}), \qquad i = 1, \ldots, n.$$

For some other variants of the Galerkin method, one may refer to Nair and Anderssen [24], and for an extensive study of Nyström approximation, one may refer to Atkinson [4].

Finally, we remark that the approach adopted in this section is not only useful for obtaining approximations for well-posed equations, but also can be effectively used for approximating *regularized solutions* of ill-posed equations (see, for example, Engl, Hanke and Neubauer [13], Nair [23], and George and Nair [16]).

PROBLEMS

1. Let X be a Hilbert space and A be a compact self-adjoint operator with spectral representation

$$Ax = \sum_j \mu_j \langle x, u_j \rangle u_j, \qquad x \in X.$$

Let $\lambda \in \mathbb{K}$ such that $\lambda \neq \mu_j$ and $\alpha_j = 1/(\mu_j - \lambda)$ for every j.

(a) Suppose $y \in \overline{R(A)}$. Show that the equation $Ax - \lambda x = y$ has a solution if and only if

$$\sum_j |\alpha_j \langle y, u_j \rangle|^2 < \infty,$$

and in that case,

$$x := \sum_j \alpha_j \langle y, u_j \rangle u_j$$

is a solution of $Ax - \lambda x = y$.

(b) Show that the series $\sum_j \mu_j \alpha_j u_j$ converges for every $y \in X$ and

$$u := \sum_j \mu_j \alpha_j u_j$$

is a solution of $Au - \lambda u = Ay$.

(c) Suppose $\lambda \neq 0$ and u is as in (b) above. Show that $x := (u - y)/\lambda$ is the unique solution of $Ax - \lambda x = y$.

(d) Suppose $\lambda \neq 0$ and $A_n = P_n A|_{R(P_n)}$, where $P_n : X \to X$ is the orthogonal projection onto the span of $\{u_1, \ldots, u_n\}$. Show that the equation $A_n \tilde{x}_n - \lambda \tilde{x}_n = A_n y$ has a unique solution for every n, and $x_n := (\tilde{x}_n - y)/\lambda$ satisfies

$$\|x - x_n\| \leq \frac{1}{d|\lambda|} \|y\| \varepsilon_n,$$

where

$$\varepsilon_n := \max_{j>n} |\mu_j|, \qquad d := \operatorname{dist}(\lambda, \{\mu_1, \mu_2, \ldots\}).$$

2. Let X and Y be Hilbert spaces, and $T \in \mathcal{B}(X,Y)$. Show that if $y \in R(T) + R(T)^\perp$, then there exists a unique $\hat{x} \in S_y$ such that $\|\hat{x}\| \le \|x\|$ for all $x \in S_y$ and this $\hat{x}$ belongs to $N(T)^\perp$.

3. Prove Theorem 14.2.

4. Let X and Y be Hilbert spaces and $T \in \mathcal{B}(X,Y)$. Show that $TT^\dagger T = T$, $T^\dagger TT^\dagger = T^\dagger$. Also, show that $T^\dagger T$ is the projection onto $N(T)^\perp$.

5. Let X and Y be Hilbert spaces, $T \in \mathcal{B}(X,Y)$ and $\alpha > 0$. Show that the equation $(T^*T + \alpha I)x = T^*y$ is uniquely solvable for every $y \in Y$, and the solution x_α is the unique element in X such that

$$\|Tx_\alpha - y\|^2 + \alpha\|x_\alpha\|^2 = \inf_{u\in X} \|Tu - y\|^2 + \alpha\|u\|^2.$$

Also, show that if $y \in R(T)+R(T)^\perp$, then $\|T^\dagger y - x_\alpha\| \to 0$ as $\alpha \to 0$.

6. Let X and Y be Hilbert spaces and $T : X \to Y$ be a compact operator with the singular system $\{(\sigma_n, u_n, v_n) : n \in \mathbb{N}\}$. Prove the following:

(a) If $y \in Y$ such that

$$\sum_n \frac{|\langle y, v_n\rangle|^2}{\sigma_n^2} < \infty,$$

then

$$\hat{x} := \sum_n \frac{1}{\sigma_n} \langle y, v_n\rangle u_n$$

is convergent and it is the generalized solution of $Tx = y$.

(b) For $y \in R(T) + R(T)^\perp$, let $\hat{x} = T^\dagger y$, and let $\nu > 0$. Then show that

$$\hat{x} \in R((T^*T)^\nu) \iff \sum_n \frac{|\langle y, v_n\rangle|^2}{\sigma_n^{4\nu+2}} < \infty.$$

(c) Let $y \in Y$ and $\alpha > 0$. Show that the solution x_α of the equation $(T^*T + \alpha I)x = T^*y$ is given by

$$\sum_n \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle y, v_n\rangle u_n.$$

Also, show that if $y \in R(T) + R(T)^{\perp}$, then $x_\alpha \to T^\dagger y$ as $\alpha \to 0$.

7. Suppose X is a Banach space, $A \in \mathcal{B}(X)$ and (A_n) in $\mathcal{B}(X)$ is such that $\|A - A_n\| \to 0$ as $n \to \infty$. If A is bijective, then show that there exists $N \in \mathbb{N}$ such that A_n is bijective for every $n \geq N$, and $\{\|A_n^{-1}\| : n \geq N\}$ is bounded.

8. Show that the conclusions of Problem 7 above hold if the condition $\|A - A_n\| \to 0$ is replaced by the following conditions:

(a) $(\|A_n\|)$ bounded.

(b) $\|(A - A_n)A\| \to 0$.

(c) $\|(A - A_n)A_n\| \to 0$.

9. Show that if A is a compact operator, then the conditions (a), (b), and (c) in the above problem can be replaced by

(d) $\|Ax - A_n x\| \to 0$ for every $x \in X$ and

(e) $\|(A - A_n)^2\| \to 0$.

10. Suppose that X is a Banach space, $K : X \to X$ is a compact operator, and $\lambda \neq 0$ is not an eigenvalue of K. Let (P_n) be a sequence of projections on X such that $\|P_n x - x\| \to 0$ as $n \to \infty$. For $y \in X$, suppose $x \in X$ and $x_n \in R(P_n)$ are the unique solutions of

$$Kx - \lambda x = y, \qquad P_n K x_n - \lambda x_n = P_n y,$$

respectively. Show that

$$\|x - x_n\| \leq c\|(I - P_n)x\|$$

for some $c > 0$.

References

[1] Ahues, M., Largillier, A., and Limaye, B. V., *Spectral Computations for Bounded Operators*, CRC Press, London, 2001.

[2] Ahlfors, L. V., *Complex Analysis*, McGraw-Hill, 1978.

[3] Anselone, P. M., *Collectively Compact Operator Approximation Theory and Applications to Integral Equations*, Prentice-Hall, Englewood-Cliffs, New Jersey, 1971.

[4] Atkinson, K. E., *The Numerical Solution of Fredholm Integral Equations of the Second Kind*, Cambridge University Press, London, 1997.

[5] Axler, S., Down with determinants!, *Amer. Math. Month.*, February (1995), 139–154.

[6] Bachmann, G. and Narici, L., *Functional Analysis*, Academic Press, New York, 1966.

[7] Baker, G. B. and Mills, T. M., Lagrange Interpolation – Divergence, *Mathematical Chronicle*, **17** (1988), 1–18.

[8] de Barra, G., *Measure Theory and Integration*, Wiley Eastern, New Delhi, 1981.

[9] Bollabas, B., *Linear Analysis: An introductory course* , Cambridge Mathematical Textbooks, Cambridge University Press, New York, 1994.

[10] Chatelin, F., *Spectral Approximation of Linear Operators*, Academic Press, New York, 1983.

[11] Eberlein, W. F., Weak compactness in Banach spaces, *Proc. Nat. Acad. Sci. (USA)*, **33** (1947), 51–53.

[12] Enflo, P., A counterexample to the approximation problem in Banach spaces, *Acta Math.*, **130** (1973), 309–317.

[13] Engl, H. W., Hanke, M., and Neubauer, A., *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht (The Netherlands), 1996.

[14] Fabér, G., Über die interpolatorische Darstellung steitiger Functionen, *Jahresber der Deutschen Math.*, **23** (1914), 190–210.

[15] Goffman, C. and Pedrick, G., *First Course in Functional Analysis*, Prentice-Hall of India, New Delhi, 1995.

[16] Georege, S. and Nair, M. T., On a generalized Arcangeli's method for Tikhonov regularization with inexact data, *Numer. Funct. Anal. and Optimiz.*, **19** (1998), 773–787.

[17] Kato, T., *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1976.

[18] Kesavan, S., *Topics in Functional Analysis and Applications*, Wiley Eastern, New Delhi, 1989.

[19] Kulkarni, S. H. and Nair, M. T., A characterization of closed range operators, *Indian J. Pure and Appl. Math.*, **31** (2000), 353–361.

[20] Limaye, B. V., *Functional Analysis*, Wiley Eastern, New Delhi, 1981.

[21] Limaye, B. V., *Spectral Perturbation and Approximation with Numerical Experiments*, Proceedings of the Centre for Mathematical Analysis, Australian National University, Canberra, **13**, 1986.

[22] Lindenstrauss, J. and Tzafriri, L., On the complemented subspaces problem, *Israel J. Math.* **9** (1971), 263-269.

[23] Nair, M. T., A unified approach for regularized approximation methods for Fredholm integral equations of the first kind, *Numer. Funct. Anal. Optim.*, **15** (1989), 381–389.

[24] Nair, M. T. and Anderssen, R. S., Superconvergence of modified projection methods for integral equations of the second kind, *Integr. Equat. Appl.*, **3** (1991), 255–269.

[25] Phillips, R. S., On linear transformations, *Trans. Amer. Math. Soc.*, **48** (1940), 516–541.

[26] Royden, H. L., *Real Analysis*, 3rd ed., Prentice-Hall of India, New Delhi, 1995.

[27] Rudin, W., *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York, 1976.

[28] Rudin, W., *Real and Complex Analysis*, 3rd ed., McGraw-Hill International Editions, New York, 1987.

[29] Shmulyan, V. L., Über lineare topologische Räume, *Math. Sobornik, N.S.*, **17** (1940), 425–448.

[30] Simmons, G. F., *Introduction to Topology and Modern Analysis*, McGraw-Hill Kogakusha, Tokyo, 1963.

[31] Sunder, V. S., *Functional Analysis: Spectral theory*, Hindustan Book Agency, New Delhi, 1997.

[32] Taylor, A. E. and Lay, D. C., *Introduction to Functional Analysis*, 2nd ed., Wiley, New York, 1980.

[33] Weinberger, H. F., *Partial Differential Equations*, Blaisdell Publishing Co., New York, 1965.

[34] Wendroff, B., *Theoretical Numerical Analysis*, Academic Press, New York, 1961.

[35] Whitely, R. J., Projecting m onto c_0, *Amer. Math. Monthly*, **73** (1966), 285–286.

[36] Yosida, K., *Functional Analysis*, Springer-Verlag, Berlin, 1974.

Index

Functional Analysis

A FIRST COURSE

M. Thamban Nair

Intended as an introductory-level text on Functional Analysis for the postgraduate students in Mathematics, this compact and well-organized text covers all the topics considered essential to the subject. In so doing, it provides a very good understanding of the subject to the reader.

The book begins with a review of linear algebra, and then it goes on to give the basic notion of a norm on linear space (proving thereby most of the basic results), progresses gradually, dealing with operators, and proves some of the basic theorems of Functional Analysis. Besides, the book analyzes more advanced topics like dual space considerations, compact operators, and spectral theory of Banach and Hilbert space operators.

The text is so organized that it strives, particularly in the last chapter, to apply and relate the basic theorems to problems which arise while solving operator equations.

KEY FEATURES

- Plenty of examples have been worked out in detail, which not only illustrate a particular result, but also point towards its limitations so that subsequent stronger results follow.
- Exercises, which are meant to aid understanding and to promote mastery of the subject, are interspersed throughout the text.

This student-friendly text, with its clear exposition of concepts, should prove to be a boon to the beginner aspiring to have an insight into Functional Analysis.

M. THAMBAN NAIR, Ph.D., Post Doctoral Fellow, University of Grenoble (France), is Professor of Mathematics at Indian Institute of Technology Madras, Chennai. He has held visiting positions at Australian National University, Canberra, and University of Kaiserslautern (Germany) on several occasions. Besides, he has given many invited talks at various institutions in India and abroad.

A vigorous researcher in the area of Numerical Functional Analysis, Dr. Nair has published a number of research papers in reputed national and international journals, including Journal of Indian Mathematical Society, Numerical Functional Analysis and Optimization, Proceedings of American Mathematical Society, Integral Equations and Operator Theory and Studia Mathematica.

ISBN:978-81-203-1735-2

9 788120 317352

www.phindia.com