trim

TEXTS AND READINGS IN MATHEMATICS 52

Functional Analysis

S. Kesavan

HINDUSTAN BOOK AGENCY

TEXTS AND READINGS IN MATHEMATICS **52**

Functional Analysis

Functional Analysis

S. Kesavan
Institute of Mathematical Sciences
Chennai

HINDUSTAN
BOOK AGENCY

Published in India by

Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: info@hindbook.com
http://www.hindbook.com

ISBN 978-81-85931-87-6 ISBN 978-93-86279-42-2 (eBook)
DOI 10.1007/978-93-86279-42-2

Preface

Functional analysis studies linear spaces provided with suitable topological structures and (continuous) linear transformations between such spaces. It has far reaching applications in several disciplines. For instance, the modern theory of (partial) differential equations and the numerical approximation of their solutions rely a lot on functional analytic techniques.

Functional analysis is now an integral part of the curriculum in any post graduate course in Mathematics. Ideally it should be taught after having covered courses in linear algebra, real and complex analysis, and topology. An introduction to the theory of measure and integration is also helpful since the richest examples in functional analysis come from function spaces whose study demands such a knowledge.

The present book grew out of notes prepared by myself while lecturing to graduate students at the Tata Institute of Fundamental Research (Bangalore Centre) and the Institute of Mathematical Sciences, Chennai. The material presented in this book is standard and is ideally suited for a course which can be followed by masters students who have covered the necessary prerequisites mentioned earlier. While covering all the standard material, I have also tried to illustrate the use of various theorems via examples taken from differential equations and the calculus of variations, either through brief sections or through the exercises. In fact, this book is well suited for students who would like to pursue a research career in the applications of mathematics. In particular, familiarity with the material presented in this book will facilitate studying my earlier book, published nearly two decades ago, 'Topics in Functional Analysis and Applications' (Wiley Eastern, now called New Age International), which serves as a functional analytic introduction to the theory of partial differential equations.

The first chapter of the present book gives a rapid revision of linear algebra, topology and measure theory. Important definitions, examples and results are recalled and no proofs are given. At the end of each section, the reader is referred to a standard text on that topic. This chapter has been included only for reference purposes and it is not intended that it be covered in a course based on this book.

Chapter 2 introduces the notion of a normed linear space and that of continuous linear transformations between such spaces.

Chapter 3 studies the analytic and geometric versions of the Hahn-Banach theorem and presents some applications of these.

Chapter 4 is devoted to the famous 'trinity' in functional analysis - the Banach-Steinhaus, the open mapping and the closed graph theorems, which are all consequences of Baire's theorem for complete metric spaces. Several applications are discussed. The notion of an 'unbounded linear mapping' is introduced.

In my opinion, most texts do not emphasize the importance of weak topologies in a course on functional analysis. That a bounded sequence in a reflexive space admits a weakly convergent sequence is the corner stone of many an existence proof in the theory of (partial) differential equations. These topologies also provide nice counter-examples to show the inadequacy of sequences in a general topological space. For instance, we will see that two topologies on a set could be different while having the same convergent sequences. We will also see an example of a compact topological space in which a sequence does not have any convergent subsequence. Chapter 5 deals with weak and weak* topologies and their applications to the notions of reflexivity, separability and uniform convexity.

Chapter 6 introduces the Lebesgue spaces and also presents the theory of one of the simplest classes of Sobolev spaces.

Chapter 7 is devoted to the study of Hilbert spaces.

Chapter 8 studies compact operators and their spectra.

Much of the fun in learning mathematics comes from actually doing it! Every chapter from Chapters 2 through 8 has a fairly large collection of exercises at the end. These illustrate the results of the text, show the optimality of the hypotheses of the various theorems via various examples or counterexamples, or develop simple versions of theories not elaborated upon in the text. They are of varying degrees of difficulty.

Occasionally, some hints for the solution are provided. It is hoped that the students will benefit by solving them.

Since this is meant to be a first course on functional analysis, I have kept the bibliographic references to a minimum, merely citing important texts where the reader may find further details of the topics covered in this book.

No claim of originality is made with respect to the material presented in this book. My treatment of this subject has been influenced by the writings of authors like Simmons (whose book was my first introduction to Functional Analysis) and Rudin. These works figure in the bibliography of this book. I would also like to mention a charming book, hardly known outside the francophonic mathematical community, *viz.* 'Analyse Fonctionnelle' by H. Brézis.

The preparation of this manuscript would not have been possible but for the excellent facilities provided by the Institute of Mathematical Sciences, Chennai, and I wish to place on record my gratitude. I also thank the Hindustan Book Agency and the editor of the TRIM Series, Prof. R. Bhatia, for their kind cooperation in bringing out this volume. I must thank the anonymous referees who painstakingly went through the first draft of the manuscript. I have tried to incorporate many of their constructive suggestions in the current version.

Finally, I thank my family for its constant support and encouragement.

Chennai, **S. Kesavan**
May 2008.

Notations

Certain general conventions followed throughout the text regarding notations are described below. All other specific notations are explained as and when they appear in the text.

- The set of natural numbers is denoted by the symbol $\mathbb{N}$, the integers by $\mathbb{Z}$, the rationals by $\mathbb{Q}$, the reals by $\mathbb{R}$ and the complex numbers by $\mathbb{C}$.

- Sets (including vector spaces and their subspaces) and also linear transformations between vector spaces are denoted by upper case latin letters.

- Elements of sets (and, therefore, vectors as well) are denoted by lower case latin letters.

- Scalars are denoted by lower case greek letters.

- To distinguish between the scalar zero and the null (or zero) vector, the latter is denoted by the zero in boldface, *i.e.* $\mathbf{0}$. This is also used to denote the zero linear transformation and the the zero linear functional.

- Column vectors in Euclidean space are denoted by lower case latin letters in boldface and matrices are denoted by upper case latin letters in boldface.

- Elements of L^p spaces (cf. Chapter 6) are equivalence classes of functions under the equivalence relation of 'equality almost everywhere'. To emphasize this fact, all elements of L^p spaces (and hence of Sobolev spaces as well) are denoted by lower case latin letters in the sanserif font. A generic representative of that equivalence class is denoted by the same lower case latin letter (in italics). Thus, if $\mathsf{f} \in L^1(0,1)$, a generic representative of this class will be denoted by f and will feature in all computations involving this element. For instance, we have

$$\|\mathsf{f}\|_1 = \int_0^1 |f(t)|\, dt.$$

- The norm in a normed linear space V will be denoted by $\|.\|$, or by $\|.\|_V$, if we wish to distinguish it from other norms that may be entering the argument. Similarly the innerproduct in a Hilbert space H will, in general, be denoted by $(.,.)$ (or by $(.,.)_H$ in case we wish to stress the role played by H).

Contents

1 Preliminaries **1**

1.1 Linear Spaces . . . 1

1.2 Topological Spaces . . . 11

1.3 Measure and Integration . . . 18

2 Normed Linear Spaces **26**

2.1 The Norm Topology . . . 26

2.2 Examples . . . 28

2.3 Continuous Linear Transformations . . . 36

2.4 Applications to Differential Equations . . . 52

2.5 Exercises . . . 57

3 Hahn-Banach Theorems **69**

3.1 Analytic Versions . . . 69

3.2 Geometric Versions . . . 77

3.3 Vector Valued Integration . . . 81

3.4 An Application to Optimization Theory . . . 85

3.5 Exercises . . . 92

4 Baire's Theorem and Applications **97**

4.1 Baire's Theorem . . . 97

4.2 Principle of Uniform Boundedness . . . 98

4.3 Application to Fourier Series . . . 101

4.4 The Open Mapping and Closed Graph Theorems . . . 104

4.5 Annihilators . . . 108

4.6 Complemented Subspaces . . . 113

4.7 Unbounded Operators, Adjoints . . . 116

4.8 Exercises . . . 125

5 Weak and Weak* Topologies **132**
5.1 The Weak Topology 132
5.2 The Weak* Topology 139
5.3 Reflexive Spaces 145
5.4 Separable Spaces 147
5.5 Uniformly Convex Spaces 152
5.6 Application: Calculus of Variations 156
5.7 Exercises 158

6 L^p Spaces **162**
6.1 Basic properties 162
6.2 Duals of L^p Spaces 168
6.3 The Spaces $L^p(\Omega)$ 170
6.4 The Spaces $W^{1,p}(a,b)$ 178
6.5 Exercises 188

7 Hilbert Spaces **195**
7.1 Basic Properties 195
7.2 The Dual of a Hilbert Space 202
7.3 Application: Variational Inequalities 207
7.4 Orthonormal Sets 210
7.5 Exercises 223

8 Compact Operators **231**
8.1 Basic Properties 231
8.2 Riesz-Fredhölm Theory 238
8.3 Spectrum of a Compact Operator 244
8.4 Compact Self-Adjoint Operators 251
8.5 Exercises 258

Bibliography **265**

Index **266**

Chapter 1

Preliminaries

1.1 Linear Spaces

Functional Analysis is the study of vector spaces endowed with topological structures (that are compatible with the linear structure of the space) and of (linear) mappings between such spaces. Throughout this book we will be working with vector spaces whose underlying field is the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$.

For completeness, we will recall some basic definitions and some important results.

Definition 1.1.1 *A* **vector space** *or a* **linear space** *over a field* $\mathbb{F}$ *(whose elements are called* **scalars***) is a set* V*, whose elements are called* **vectors***, on which two operations -* **addition** *and* **scalar multiplication** *- are defined such that the following properties hold:*
Addition: $(x, y) \in V \times V \mapsto x + y \in V$ *such that*
(i) *(commutativity) for all* x *and* $y \in V$*, we have*

$$x + y = y + x;$$

(ii) *(associativity) for all* x, y *and* $z \in V$*, we have*

$$x + (y + z) = (x + y) + z;$$

(iii) *there exists a unique vector* $\mathbf{0} \in V$*, called the* **zero** *or the* **null vector***, such that, for every* $x \in V$*,*

$$x + \mathbf{0} = x;$$

(iv) *for every* $x \in V$*, there exists a unique vector* $-x \in V$ *such that*

$$x + (-x) = \mathbf{0}.$$

Scalar Multiplication: $(\alpha, x) \in \mathbb{F} \times V \mapsto \alpha x \in V$ *such that*
(v) *for every* $x \in V$, $1x = x$ *where* 1 *is the multiplicative identity in* $\mathbb{F}$;
(vi) *for all* α *and* $\beta \in \mathbb{F}$ *and for every* $x \in V$,

$$\alpha(\beta x) = (\alpha\beta)x;$$

(vii) *for all* α *and* $\beta \in \mathbb{F}$ *and for every* $x \in V$,

$$(\alpha + \beta)x = \alpha x + \beta x;$$

(viii) *for every* $\alpha \in \mathbb{F}$ *and for all* x *and* $y \in V$,

$$\alpha(x + y) = \alpha x + \alpha y. \blacksquare$$

Remark 1.1.1 The conditions (i) - (iv) above imply that V is an abelian group with respect to vector addition. The conditions (vii) and (viii) above are known as the *distributive laws.* ■

Example 1.1.1 Let $N \geq 1$ be a positive integer. Define

$$\mathbb{R}^N = \{x = (x_1, ..., x_N) \mid x_i \in \mathbb{R} \text{ for all } 1 \leq i \leq N\}.$$

We define addition and scalar multiplication componentwise, *i.e.* if $x = (x_1, ..., x_N)$ and $y = (y_1, ..., y_N)$ are elements of $\mathbb{R}^N$ and if $\alpha \in \mathbb{R}$, we define

$$x + y = (x_1 + y_1, ..., x_N + y_N)$$

and

$$\alpha x = (\alpha x_1, ..., \alpha x_N).$$

It is now easy to see that $\mathbb{R}^N$ is a vector space over $\mathbb{R}$ with the zero vector being that element in $\mathbb{R}^N$ with all its components zero. In the same way, we can define $\mathbb{C}^N$ as a vector space over $\mathbb{C}$.

Setting $N = 1$, we see that $\mathbb{R}$ (respectively $\mathbb{C}$) is a vector space over itself, the scalar multiplication being the usual multiplication operation. ■

Definition 1.1.2 *Let* V *be a vector space and let* $W \subset V$. *Then* W *is said to be a* **subspace** *of* V *if* W *is a vector space in its own right for the same operations of addition and scalar multiplication.* ■

Definition 1.1.3 *Let* V *be a vector space and let* $x_1, ..., x_n$ *be vectors in* V. *A* **linear combination** *of these vectors is any vector of the form* $\alpha_1 x_1 + ... + \alpha_n x_n$, *where the* α_i, $1 \leq i \leq n$ *are scalars. A* **linear relation** *between these vectors is an equation of the form*

$$\alpha_1 x_1 + ... + \alpha_n x_n = \mathbf{0}. \blacksquare$$

Definition 1.1.4 *A finite set of vectors in a vector space is said to be* **linearly independent** *if there does not exist any linear relation between them other than the trivial one, i.e, when all the scalar coefficients are zero. If there exists a non-trivial linear relation, then the set of vectors is said to be* **linearly dependent**. *An infinite set of vectors is said to be linearly independent if there does not exist any* finite *linear relation amongst vectors in that set.* ■

Given a vector space V and a set S of vectors, the collection of all finite linear combinations of vectors from S will form a subspace of V. In fact, it is clear that this subspace is the smallest subspace containing S and is called the *linear span* of the set S or the *subspace generated by* S. This subspace is denoted by $\text{span}\{S\}$.

Definition 1.1.5 *A maximal linearly independent subset of a vector space is called a* **basis**. ■

In other words, a basis is a linearly independent subset such that, if any other vector is adjoined to the set, the enlarged set becomes linearly dependent. This implies that every vector in the space can be expressed as a (finite) linear combination of the members of the basis. Thus, the vector space is generated by its basis.

Proposition 1.1.1 (i) *Every vector space has a basis.*
(ii) *Any two bases of a vector space have the same cardinality.* ■

The above proposition leads us to the following definition.

Definition 1.1.6 *A vector space V is said to be* **finite dimensional** *if it admits a basis with a finite number of elements. Otherwise, it is said to be* **infinite dimensional**. *The* **dimension** *of a vector space is the number of elements in a basis, if it is finite dimensional, and infinity if it is infinite dimensional and is denoted* $\dim(V)$. ■

Example 1.1.2 The space $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) has a basis which is defined as follows. Let $1 \leq i \leq N$. Let $\mathbf{e}_i$ be the vector whose i-th component is unity and all other components are zero. Then $\{\mathbf{e}_1, ..., \mathbf{e}_N\}$ is a basis for $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$) and is called the **standard basis**. ■

Example 1.1.3 Let $\mathcal{P}$ denote the collection of all polynomials in one variable with real coefficients. Let $\mathsf{p}(x) = \sum_{i=0}^{n} a_i x^i$ and $\mathsf{q}(x) = \sum_{i=0}^{m} b_i x^i$

where the a_i and b_i are real numbers and x is the variable. Let $\alpha \in \mathbb{R}$. Assume, without loss of generality, that $m \leq n$. Define

$$(\mathsf{p}+\mathsf{q})(x) \;=\; \sum_{i=0}^{n}(a_i+b_i)x^i$$

where we set $b_i = 0$ for $m < i \leq n$ if $m < n$. Define

$$(\alpha\mathsf{p})(x) \;=\; \sum_{i=0}^{n}\alpha a_i x^i.$$

With these operations, $\mathcal{P}$ becomes a vector space over $\mathbb{R}$. It is easy to check that the collection of monomials $\{\mathsf{p}_i\}_{i=0}^{\infty}$ where $\mathsf{p}_0(x) \equiv 1$ and $\mathsf{p}_i(x) = x^i$ for $i > 1$, forms a basis for $\mathcal{P}$. Thus, $\mathcal{P}$ is an infinite dimensional vector space. ■

We will come across numerous examples of infinite dimensional vector spaces in the sequel (cf. Section 2.2, for instance).

Let V be a vector space and let W_i, $1 \leq i \leq n$ be subspaces. The span of the W_i is the subspace of all vectors of the form

$$v = w_1 + \ldots + w_n$$

where $w_i \in W_i$ for each $1 \leq i \leq n$. The spaces are said to be independent if an element in the span is zero if, and only if each, $w_i = \mathbf{0}$. In particular, it follows that, if the W_i are independent, then for all $1 \leq i, j \leq n$ such that $i \neq j$, we have $W_i \cap W_j = \{\mathbf{0}\}$. Further, every element in the span will have a *unique* decomposition into vectors from the spaces W_i.

Definition 1.1.7 *Let V be a vector space and let W_i, $1 \leq i \leq n$ be subspaces. Then, V is said to be the* **direct sum** *of the W_i if the spaces W_i are independent and their span is the space V. In this case we write*

$$V \;=\; W_1 \oplus W_2 \oplus \ldots \oplus W_n \;=\; \oplus_{i=1}^{n} W_i. \;■$$

We now study mappings between vector spaces which preserve the linear structure.

Definition 1.1.8 *Let V and W be vector spaces (over the same base field). A* **linear transformation**, *or* **linear operator**, *is a mapping $T : V \to W$ such that for all x and $y \in V$ and for all scalars α and β, we have*

$$T(\alpha x + \beta y) \;=\; \alpha T(x) + \beta T(y).$$

If W is the base field (which is a vector space over itself), then a linear transformation from V into W is called a **linear functional** *on V.* ■

Definition 1.1.9 *Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. The image of T is a subspace of W and is called the* **range** *of T. The dimension of the range is called the* **rank** *of T. The set*

$$\{x \in V \,|T(x) = \mathbf{0}\}$$

is a subspace of V and is called the **null space** *or* **kernel** *of T.* ■

Definition 1.1.10 *Let V and W be vector spaces and T a linear transformation between them. The transformation is said to be* **invertible** *if T is a bijection.* ■

It is easy to see that a linear transformation which is an injection maps a linearly independent set onto a linearly independent set. In particular, if $T : V \to W$ is an injection, then, necesarily, $\dim(V) \leq \dim(W)$. On the other hand, if $T : V \to W$ is a surjection, clearly, $\dim(V) \geq \dim(W)$. Thus, if T is invertible, then the two spaces must have the same dimension.

We now focus our attention on finite dimensional spaces. Let V be a space of dimension n with basis $\{v_1, \ldots, v_n\}$ and W a space of dimension m with basis $\{w_1, \ldots, w_m\}$. A linear map $T : V \to W$ is completely defined, once it is defined on a basis. So let us write

$$T(v_j) = \sum_{i=1}^{m} t_{ij} w_i, \; 1 \leq j \leq n. \tag{1.1.1}$$

The coefficients (t_{ij}) in the above relation form a **matrix** with m rows and n columns. Such a matrix is referred to as an $m \times n$ matrix. The j-th column of the matrix represents the coefficients in the expansion of $T(v_j)$ in terms of the basis $\{w_i\}_{i=1}^m$ of W. Of course, if we change the bases for V and W, the same linear transformation will be given by another matrix. In particular, let $\dim(V) = n$ and let $T : V \to V$ be a linear operator. Let T be represented by the $n \times n$ matrix (also known as a square matrix of order n) $\mathbf{T} = (t_{ij})$ with respect to a given basis. If we change the basis, then T will be represented by another $n \times n$ matrix $\widetilde{\mathbf{T}} = (\widetilde{t_{ij}})$ and the two will be connected by a relation of the form:

$$\mathbf{T} = \mathbf{P}\widetilde{\mathbf{T}}\mathbf{P}^{-1}$$

where $\mathbf{P}$ is called the change of basis matrix and represents the linear transformation which maps one basis to another. The matrix $\mathbf{P}^{-1}$ represents the inverse of this change of basis mapping and is the inverse matrix of $\mathbf{P}$. In this case, the matrices $\mathbf{T}$ and $\widetilde{\mathbf{T}}$ are said to be **similar**. The identity matrix $\mathbf{I}$ represents the identity mapping $x \mapsto x$ for all $x \in V$ for any fixed basis of V. For a given basis, if $T : V \to V$ is invertible, then the matrix representing T^{-1} will be the inverse of the matrix representing T.

A square matrix is said to be **diagonal** if all its off-diagonal entries are zero. A $n \times n$ square matrix $\mathbf{A} = (a_{ij})$ is said to be **upper triangular** (repectively, **lower triangular**) if $a_{ij} = 0$ for all $1 \leq j < i \leq n$ (repectively, $a_{ij} = 0$ for all $1 \leq i < j \leq n$). It can be shown that every matrix is similar to an upper triangular matrix. A matrix is said to be **diagonalizable** if it is similar to a diagonal matrix.

Given an $m \times n$ matrix and two vector spaces of dimensions n and m respectively along with a basis for each of them, the matrix can be used, as in relation (1.1.1), to define a linear transformation beteween these two spaces. Thus, there is a one-to-one correspondence between matrices and linear transformations between vector spaces of appropriate dimension, once the bases are fixed.

Definition 1.1.11 *If $\mathbf{T} = (t_{ij})$ is an $m \times n$ matrix, then the $n \times m$ matrix $\mathbf{T}' = (t_{ji})$, formed by interchanging the rows and the columns of the matrix $\mathbf{T}$, is called the* **transpose** *of the matrix $\mathbf{T}$. If $\mathbf{T} = (t_{ij})$ is an $m \times n$ matrix with complex entries, then the $n \times m$ matrix $\mathbf{T}^* = (t^*_{ij})$ where $t^*_{ij} = \bar{t}_{ji}$ (the bar denoting complex conjugation), is called the* **adjoint** *of the matrix $\mathbf{T}$.* ■

If x and $y \in \mathbb{R}^n$ (respectively, $\mathbb{C}^n$), then $y'x$ (respectively, y^*x) represents the 'usual' scalar product of vectors in $\mathbb{R}^n$ (respectively, $\mathbb{C}^n$) given by $\sum_{i=1}^n x_i y_i$ (respectively, $\sum_{i=1}^n x_i \bar{y}_i$). If the scalar product is zero, we say that the vectors are **orthogonal** to each other and write $x \perp y$. If W is a subspace and x is a vector orthogonal to all vectors in W, we write $x \perp W$.

Definition 1.1.12 *Let $\mathbf{T}$ be an $m \times n$ matrix. Then its* **row rank** *is defined as the number of linearly independent row vectors of the matrix and the* **column rank** *is the number of independent column vectors of the matrix.* ■

The column rank is none other than the rank of the linear transformation defined by $\mathbf{T}$ and the row rank is the rank of the transformation defined by the transpose. We have the following important result.

Proposition 1.1.2 *For any matrix, the row and column ranks are equal and the common value is called the* **rank** *of the matrix.* ■

Definition 1.1.13 *The* **nullity** *of a matrix is the dimension of the null space of the linear transformation defined by the matrix.* ■

Proposition 1.1.3 *Let* $\mathbf{T}$ *be an* $m \times n$ *matrix. The the sum of the rank of* $\mathbf{T}$ *and the nullity of* $\mathbf{T}$ *is equal to* n.

Corollary 1.1.1 *An* $n \times n$ *matrix is invertible if, and only if, its nullity is zero or, equivalently, its rank is* n. *Equivalently, a linear operator on a finite dimensional space is one-to-one if, and only if, it is onto.* ■

Definition 1.1.14 *Let* $\mathbf{T}$ *be an* $n \times n$ *matrix with complex entries and* $\mathbf{T}^*$ *its adjoint. The matrix is said to be* **normal** *if*

$$\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T}.$$

The matrix is said to be **unitary** *if*

$$\mathbf{T}\mathbf{T}^* = \mathbf{T}^*\mathbf{T} = \mathbf{I}.$$

The matrix is said to be **self-adjoint** *or* **hermitian** *if* $\mathbf{T} = \mathbf{T}^*$. ■

Definition 1.1.15 *Let* $\mathbf{T}$ *be an* $n \times n$ *matrix with complex entries. It is said to be* **positive semi-definite** *if for every* $n \times 1$ *matrix with complex entries, i.e, a column vector,* $\mathbf{x}$, *we have*

$$\mathbf{x}^*\mathbf{T}\mathbf{x} \geq 0.$$

The matrix $\mathbf{T}$ *is said to be* **positive definite** *if, in addition, the above inequality is strict if* $\mathbf{x} \neq \mathbf{0}$. ■

Remark 1.1.2 A hermitian matrix is equal to its adjoint and the inverse of a unitary matrix is its adjoint. A matrix $\mathbf{T}$, with real entries, which is equal to its transpose is called **symmetric** and one whose inverse is its own transpose is called **orthogonal**. In case the matrix $\mathbf{T}$ has real entries, we can still define positive semi-definiteness (or positive definiteness) by considering real column vectors $\mathbf{x}$ in the above definition. ■

We now introduce an important notion, *viz.* that of the determinant. Before we do this, we need some notation. Let $\mathcal{S}_n$ denote the set of all permutations of n symbols. A *transposition* is a permutation wherein two symbols exchange places with each other and all other symbols are left invariant. It is known that every permutation is the product (in the sense of composition of mappings) of transpositions. A permutation is *even* if it is the product of an even number of transpositions and *odd* if it is the product of an odd number of transpositions. The *signature* of a permutation σ, denoted sgn σ, is $+1$ if the permutation is even and -1 if it is odd.

Definition 1.1.16 *Let* $\mathbf{T} = (t_{ij})$ *be an* $n \times n$ *matrix. The* **determinant** *of* $\mathbf{T}$*, denoted* $\det(\mathbf{T})$*, is given by the formula*

$$\det(\mathbf{T}) = \sum_{\sigma \in \mathcal{S}_n} (\text{sgn } \sigma) t_{1,\sigma(1)} \cdots t_{n,\sigma(n)}. \quad \blacksquare$$

We list below the important properties of the determinant.

Proposition 1.1.4 (i) *If* $\mathbf{I}$ *is the identity matrix of order* n*, then* $\det(\mathbf{I}) = 1$.
(ii) *If* $\mathbf{T}$ *and* $\mathbf{S}$ *are two* $n \times n$ *matrices, then*

$$\det(\mathbf{ST}) = \det(\mathbf{S}).\det(\mathbf{T}).$$

In particular, $\mathbf{T}$ *is invertible if, and only if,* $\det(\mathbf{T}) \neq 0$ *and*

$$\det(\mathbf{T}^{-1}) = (\det(\mathbf{T}))^{-1}.$$

(iii) *If* $\mathbf{T}$ *is an* $n \times n$ *matrix, then*

$$\det(\mathbf{T}') = \det(\mathbf{T}). \quad \blacksquare$$

Definition 1.1.17 *An invertible matrix is also said to be a* **non-singular** *matrix. Otherwise, it is said to be* **singular**. ■

If $\mathbf{T}$ is a non-singular matrix of order n, then, given any $n \times 1$ column vector $\mathbf{b}$, there exists a unique $n \times 1$ column vector $\mathbf{x}$ such that

$$\mathbf{Tx} = \mathbf{b}.$$

This is because the corresponding linear transformation is invertible and hence onto (which gives the existence of the solution) and one-to-one (which gives the uniqueness of the solution). If $\mathbf{T}$ is singular, this is no longer the case and we have the following result.

Proposition 1.1.5 *(***Fredhölm Alternative***) Let* $\mathbf{T}$ *be a singular matrix of order* n *and let* $\mathbf{b}$ *be an* $n \times 1$ *column vector. Then, either the system of* n *linear equations in* n *unknowns (written in matrix notation)*

$$\mathbf{Tx} = \mathbf{b}$$

has no solution or has an infinite number of solutions. The latter possibility occurs if, and only if

$$\mathbf{b}'\mathbf{u} = 0$$

for all (column) vectors $\mathbf{u}$ *such that* $\mathbf{T}'\mathbf{u} = \mathbf{0}$. ■

We now come to a very important notion in linear algebra and functional analysis.

Definition 1.1.18 *Let* $\mathbf{T}$ *be a square matrix of order* n *with complex entries. A complex number* λ *is said to be an* **eigenvalue** *of* $\mathbf{T}$ *if there exists an* $n \times 1$ *vector* $\mathbf{u} \neq \mathbf{0}$ *such that*

$$\mathbf{Tu} = \lambda\mathbf{u}.$$

Such a vector $\mathbf{u}$ *is called an* **eigenvector** *of* $\mathbf{T}$ *associated to the eigenvalue* λ*. The set of all eigenvectors associated to an eigenvalue* λ *is a subspace of* $\mathbb{C}^n$ *and is called the* **eigenspace** *associated to* λ. ■

From the above definition, we see that $\lambda \in \mathbb{C}$ is an eigenvalue of a square matrix $\mathbf{T}$ if, and only if the matrix $\mathbf{T} - \lambda\mathbf{I}$ is *not* invertible. Thus, λ will be an eigenvalue of $\mathbf{T}$ if, and only if,

$$\det(\mathbf{T} - \lambda\mathbf{I}) = 0. \tag{1.1.2}$$

The expression on the left-hand side of (1.1.2) is a polynomial of degree equal to the order of $\mathbf{T}$ and is called the **characteristic polynomial** of $\mathbf{T}$. Thus, every eigenvalue is a root of the characteristic polynomial and so every matrix of order n has a non-empty set of at most n distinct eigenvalues. Counting multiplicity, there are exactly n eigenvalues. The equation (1.1.2) is called the **characteristic equation** of the matrix $\mathbf{T}$.

Definition 1.1.19 *Let* λ *be an eigenvalue of a matrix* $\mathbf{T}$*. Its* **algebraic multiplicity** *is its multiplicity as a root of the characteristic polynomial. Its* **geometric multiplicity** *is the dimension of the eigenspace associated to it.* ■

Proposition 1.1.6 *The geometric multiplicity of an eigenvalue does not exceed its algebraic multiplicity.* ■

Definition 1.1.20 *The set of all eigenvalues of a matrix is called its* **spectrum**. *The maximum of the absolute values of the eigenvalues is called its* **spectral radius**. ■

The sum of the diagonal entries of a square matrix $\mathbf{T}$ is called its **trace** and is denoted by $\mathrm{tr}(\mathbf{T})$. It is easy to see from the characteristic equation that the trace is the sum of all the eigenvalues taking into account their multiplicities. Similarly the determinant of a matrix is the product of all its eigenvalues (again counting multiplicity).

Proposition 1.1.7 (i) *The eigenvalues of the adjoint of a matrix are the complex conjugates of the eigenvalues of the original matrix.*
(ii) *The eigenvalues of a hermitian matrix are all real.*
(iii) *The eigenvalues of a hermitian and positive definite matrix are positive.*
(iv) *The eigenvalues of a unitary matrix lie on the unit circle of the complex plane.* ■

The eigenvalues of a hermitian matrix admit a 'variational characterization'. Let $\mathbf{T}$ be a hermitian matrix of order n. Its eigenvalues are all real and let them be numbered in increasing order as follows:

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

Let v_i, $1 \leq i \leq n$ be a collection of eigenvectors, where v_i is associated to λ_i. Set $V_0 = \{\mathbf{0}\}$ and

$$V_i = \mathrm{span}\{v_1, \ldots, v_i\}$$

for $1 \leq i \leq n$.

We define the *Rayleigh quotient* associated to the matrix $\mathbf{T}$ as follows:

$$R_{\mathbf{T}}(x) = \frac{x^*\mathbf{T}x}{x^*x}, \; x \neq \mathbf{0},$$

which is real valued, since $\mathbf{T}$ is hermitian.

Proposition 1.1.8 (i) *The eigenvectors v_i can be chosen such that*

$$v_i^* v_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

(ii) *For each* $1 \leq i \leq n$, *we have*

$$\begin{aligned} \lambda_i &= R_{\mathbf{T}}(v_i) \\ &= \max_{x \in V_i} R_{\mathbf{T}}(x) \\ &= \min_{x \perp V_{i-1}} R_{\mathbf{T}}(x) \\ &= \min_{W \subset \mathbb{C}^n,\ \dim(W)=i} \max_{x \in W} R_{\mathbf{T}}(x). \end{aligned}$$

In particular,

$$\lambda_1 = \min_{x \in \mathbb{C}^n} R_{\mathbf{T}}(x), \quad \textit{and} \quad \lambda_n = \max_{x \in \mathbb{C}^n} R_{\mathbf{T}}(x). \ \blacksquare$$

By analogy, if $\mathbf{T}$ is a real symmetric matrix, a similar result holds with the adjoint being replaced by the transpose in the definition of the Rayleigh quotient and $\mathbb{C}^n$ being replaced by $\mathbb{R}^n$ in the formulae above.

Remark 1.1.3 A matrix is diagonalizable if, and only if, it admits a basis of eigenvectors. In particular, if all the eigenvalues of a matrix are distinct, it is diagonalizable. All normal matrices are diagonalizable. In particular, hermitian matrices are diagonalizable. ■

For more details on linear spaces, the reader is referred to, for instance, Artin [1].

1.2 Topological Spaces

In this section, we recall the important definitions and results of topology which will be used in the sequel.

Definition 1.2.1 *A* **topology** *on a set* X *is a collection* $\mathcal{T}$ *of subsets of* X *such that*
(i) X *and* $\emptyset$ *are in* $\mathcal{T}$;
(ii) *the union of any collection of members in* $\mathcal{T}$ *is a member of* $\mathcal{T}$;
(iii) *the intersection of any* finite *collection of members of* $\mathcal{T}$ *is a member of* $\mathcal{T}$.
The pair $\{X, \mathcal{T}\}$ *is said to be a* **topological space** *and the members of the topology* $\mathcal{T}$ *are called* **open sets**. *The complements of open sets are called* **closed sets**. *If* $x \in X$, *then a* **neighbourhood** *of* x *is any open set containing* x. ■

It is clear from the above definition that X and $\emptyset$ are both open and closed. Further, a finite union and an arbitrary intersection of closed sets is closed. It then follows that given any set $A \subset X$, there is a smallest closed set containing it. This is called the **closure** of the set A and is usually denoted by $\overline{A}$. If $\overline{A} = X$, we say that A is **dense** in X. Similarly, given any set $A \subset X$, there is a largest open set contained in A. This set is called the **interior** of A and is denoted by A°. A set $A \subset X$ is said to be **nowhere dense** if $(\overline{A})^\circ = \emptyset$.

If $\{X, \mathcal{T}\}$ is a topological space and if $Y \subset X$, then Y inherits a natural topology from that of X. The open sets are those of the form $U \cap Y$, where U is open in X.

Definition 1.2.2 *A topological space* $\{X, \mathcal{T}\}$ *is said to be* **Hausdorff** *if for every pair of distinct elements* x *and* y *in* X, *there exist* disjoint *open sets* U *and* V *such that* $x \in U$ *and* $y \in V$. ■

Definition 1.2.3 *A* **metric** *on a set* X *is a function* $d : X \times X \to [0, \infty)$ *such that*
(i) $d(x, y) = 0$ *if, and only if,* $x = y$*;*
(ii) *for all* x *and* $y \in X$, *we have*

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

(iii) *for all* x, y *and* $z \in X$, *we have*

$$d(x, z) \leq d(x, y) + d(y, z). \tag{1.2.1}$$

The pair $\{X, d\}$ *is called a* **metric space.** ■

Remark 1.2.1 The notion of a metric generalizes that of a distance, as we know it, in the Euclidean spaces $\mathbb{R}^2$ and $\mathbb{R}^3$. The inequality (1.2.1) is just an abstract version of a familiar theorem from Euclidean plane geometry which states that the sum of the lengths of two sides of a triangle is greater than the length of the third side. For this reason, it is called the *triangle inequality.* Of all the conditions on a metric, this one will need the most non-trivial verification. ■

If $\{X, d\}$ is a metric space, then it is easy to see that we have a topology (called the metric topology) induced on X by the metric d which is defined as follows. A non-empty set $U \subset X$ is open if, and only if, for every $x \in U$, there exists $r > 0$ such that

$$B(x; r) \stackrel{\text{def}}{=} \{y \in X \mid d(x, y) < r\} \subset U.$$

The set $B(x;r)$ described above is called the (open) ball centred at x and of radius r. It is a simple exercise to check that open balls themselves are open sets. It is also immediate to see that this topology is Hausdorff.

On $\mathbb{R}$ or $\mathbb{C}$, we have the 'usual' metric defined by

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

The topology induced by this metric will be called the 'usual' topology on $\mathbb{R}$ or $\mathbb{C}$, as the case may be. Similarly, on $\mathbb{R}^N$ (or $\mathbb{C}^N$), we have the 'usual' Euclidean distance which defines a metric on that space: if $x = (x, ..., x_N)$ and $y = (y_1, ..., y_N)$ are vectors in $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$), then

$$d(x,y) = \left(\sum_{i=1}^{N} |x_i - y_i|^2\right)^{\frac{1}{2}}$$

The topology induced by this metric will be referred to as the 'usual' topology on $\mathbb{R}^N$ (respectively, $\mathbb{C}^N$).

In the metric topology defined above, we see that every open set is the union of open balls. In a general topological space a collection $\mathcal{B}$ of open sets is called a **base** for the topology if every open set can be expressed as the union of members of $\mathcal{B}$.

Definition 1.2.4 *Let $\{X, \mathcal{T}\}$ be a topological space and let $\mathcal{S}$ be a collection of open sets in X. We say that $\mathcal{S}$ is a* **subbase** *for the topology $\mathcal{T}$ if every open set can be expressed as unions of* finite *intersections of members of $\mathcal{S}$.* ■

Clearly any topology containing $\mathcal{S}$ will have to contain $\mathcal{T}$. Thus $\mathcal{T}$ is the smallest topology containing $\mathcal{S}$. The set of finite intersections of members of $\mathcal{S}$ form a base for the topology.

Definition 1.2.5 *Let $\{X, \mathcal{T}\}$ be a topological space and let A be an arbitrary subset of X. A point $x \in X$ is said to be a* **limit point** *of A if every neighbourhood of x contains a point of A (different from x, in case $x \in A$).* ■

Definition 1.2.6 *Let $\{X, \mathcal{T}\}$ be a topological space and let $\{x_n\}$ be a sequence of elements in X. We say that the sequence* **converges** *to a point $x \in X$ if for every neighbourhood U of x, we can find a positive integer N (depending on U) such that $x_k \in U$ for all $k \geq N$. In this case, we write $x_n \to x$ in X.* ■

Definition 1.2.7 *Let $\{X_i, \mathcal{T}_i\}$, $i = 1, 2$, be two topological spaces and let $f : X_1 \to X_2$ be a given function. We say that f is* **continuous** *if $f^{-1}(U)$ is an open set in X_1 for every open set U in X_2. If f is a bijection such that both f and f^{-1} are continuous, then f is said to be a* **homeomorphism** *and the two topological spaces are said to be homeomorphic to each other.* ■

The following propositions are easy to prove.

Proposition 1.2.1 *Let $\{X, d\}$ be a metric space and let $\{x_n\}$ be a sequence in X. Then $x_n \to x$ in X if, and only if, for every $\varepsilon > 0$, there exists a positive integer N such that*

$$d(x_k, x) \ < \ \varepsilon \ \ \textit{for every} \ \ k \geq N. \ ■$$

In particular, every convergent sequence is **bounded**, *i.e.*, it can be contained in a (sufficiently large) ball.

Proposition 1.2.2 *Let $\{X_i, d_i\}, i = 1, 2$, be metric spaces and let $f : X_1 \to X_2$ be a given function. The following are equivalent:*
(i) *f is continuous.*
(ii) *For every $x \in X$ and for every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $d_1(x, y) < \delta$, we have $d_2(f(x), f(y)) < \varepsilon$.*
(iii) *If $x_n \to x$ in X_1, then $f(x_n) \to f(x)$ in X_2.* ■

Remark 1.2.2 When property (ii) (or, equivalently, (iii)) holds for a particular point $x \in X$, we say that f is *continuous at* x. Thus, f is continuous if, and only if, it is continuous at each point of X. If f is continuous and, for a given $\varepsilon > 0$, the $\delta > 0$ described in statement (ii) above does not depend on the point x, then the function is said to be *uniformly continuous* on X. ■

Let $\{X, d\}$ be a metric space and let E be an arbitrary subset of X. Let $x \in X$. Define

$$d(x, E) \ = \ \inf_{y \in E} d(x, y).$$

This is called the distance of the point x from the set E. The following proposition is easy to prove.

Proposition 1.2.3 *Let $\{X, d\}$ be a metric space and let $E \subset X$. Then*
(i) *for all x and $y \in X$, we have*

$$|d(x, E) - d(y, E)| \ \leq \ d(x, y).$$

Thus, the function $x \mapsto d(x,E)$ is a uniformly continuous function on X;
(ii) *if E is a closed set, then, $d(x,E) = 0$ if, and only if, $x \in E$; more generally, if E is any subset of X, we have*

$$\overline{E} \;=\; \{x \in X \mid d(x,E) = 0\}.$$

Definition 1.2.8 *Let $\{X,d\}$ be a metric space. A sequence $\{x_n\}$ in X is said to be* **Cauchy** *if, for every $\varepsilon > 0$, there exists a positive integer N such that*

$$d(x_k, x_l) \;<\; \varepsilon$$

for every $k \geq N, l \geq N$. ■

It is simple to verify that every Cauchy sequence is bounded. It is also easy to see that every convergent sequence in a metric space is Cauchy. The converse is not true and this leads to the following important definition.

Definition 1.2.9 *A metric space is said to be* **complete** *if every Cauchy sequence is convergent.* ■

With their usual metric, the spaces $\mathbb{R}$ and $\mathbb{C}$ are complete.

We now introduce an important notion which we will study in detail in later chapters.

Definition 1.2.10 *Let $\mathcal{I}$ be an arbitrary indexing set. Let X be a set and $\{X_i, \mathcal{T}_i\}_{i\in\mathcal{I}}$ be topological spaces. Let $f_i : X \to X_i$ be given functions. The* **weak topology** *generated by the functions f_i, $i \in \mathcal{I}$, is the smallest topology on X such that all the f_i are continuous.* ■

From the above definition it follows that a subbase for the weak topology generated by the f_i is the collection of all sets of the form $f_i^{-1}(U)$ where U is an arbitrary open set in X_i and the index i ranges over the indexing set $\mathcal{I}$. A typical neighbourhood of a point $x \in X$ will therefore be a *finite* intersection of sets of the form $f_i^{-1}(U_i)$ where U_i is a neighbourhood of $f_i(x)$ in X_i.

Definition 1.2.11 *Let $\mathcal{I}$ be an indexing set and let $\{X_i, \mathcal{T}_i\}$, $i \in \mathcal{I}$ be topological spaces. Set $X = \Pi_{i\in\mathcal{I}} X_i$. Let $x = (x_i)_{i\in\mathcal{I}}$. Let $p_i : x \in X \mapsto x_i \in X_i$ be the i-th coordinate projection. The* **product topology** *on X is the weak topology generated by the coordinate projections, i.e. it is the smallest topology such that the projections are all continuous.* ■

Thus, sets of the form $\Pi_{i \in \mathcal{I}} U_i$, where $U_i = X_i$ for all $i \neq i_0$ (an arbitrary element of $\mathcal{I}$) and U_{i_0} is open in X_{i_0}, form a subbase for the product topology. A base for the topology is the collection of all sets of the form $\Pi_{i \in \mathcal{I}} U_i$ where $U_i = X_i$ for all but a finite number of indices and, for those indices, U_i is an open set in X_i.

Definition 1.2.12 *Let $\{X, \mathcal{T}\}$ be a topological space and let $\emptyset \neq K \subset X$. A collection of open sets $\mathcal{F}$ is said to be an* **open cover** *of K if the union of the members of $\mathcal{F}$ contains K. A* **subcover** *of $\mathcal{F}$ is a subcollection of members of $\mathcal{F}$ which is also an open cover of K.* ∎

Definition 1.2.13 *Let $\{X, \mathcal{T}\}$ be a topological space and let $\emptyset \neq K \subset X$. The set K is said to be a* **compact set** *if every open cover of K admits a* finite *subcover.* ∎

If X is itself a compact set, we say that it is a compact space. We can also describe compactness via closed sets.

Definition 1.2.14 *A collection $\mathcal{A}$ of subsets of a set X is said to have* **finite intersection property** *if every finite subcollection has non-empty intersection.* ∎

The following proposition is easily proved.

Proposition 1.2.4 *A non-empty subset K of a topological space is compact if, and only if, every collection of closed sets in K having finite intersection property has non-empty intersection.* ∎

Definition 1.2.15 *Let $\{X, \mathcal{T}\}$ be a topological space and let $\emptyset \neq K \subset X$. The set K is said to be* **sequentially compact** *if every sequence in K has a convergent subsequence.* ∎

We list important facts about compact sets in the following proposition.

Proposition 1.2.5 (i) *Every continuous image of a compact set is compact.*
(ii) *The product of compact sets is compact.*
(iii) *Compact subsets of Hausdorff spaces are closed.*
(iv) *A compact subset of a metric space is closed and bounded.*
(v) *For the usual topology on $\mathbb{R}^N$, a subset is compact if, and only if, it is closed and bounded.*

(vi) *Every continuous real valued function on a compact space is bounded and attains its maximum and minimum values in that set.*
(vii) *Every continuous real valued function on a compact metric space is uniformly continuous.* ■

Compact metric spaces are very special. In order to characterize them, we need the following notion.

Definition 1.2.16 *A metric space* $\{X,d\}$ *is said to be* **totally bounded** *if, for every* $\varepsilon > 0$*, there exists* finite *set of points* $\{x_i\}_{i=1}^{k(\varepsilon)}$ *such that*

$$X \subset \cup_{i=1}^{k(\varepsilon)} B(x_i;\varepsilon).\ ■$$

Proposition 1.2.6 *Let* $\{X,d\}$ *be a metric space. The following statements are equivalent:*
(i) X *is compact.*
(ii) X *is sequentially compact.*
(iii) *Every infinite subset of* X *has a limit point.*
(iv) X *is complete and totally bounded.* ■

We conclude this section with one final important topological notion.

Definition 1.2.17 *Let* $\{X,\mathcal{T}\}$ *be a topological space. We say that* X *is* **connected** *if there* do not *exist* non-empty *open sets* U *and* V *such that* $X = U \cup V$ *and* $U \cap V = \emptyset$*. A subset* $A \subset X$ *is said to be connected if there do not exist disjoint open sets* U *and* V *such that* $A = A \cap (U \cup V)$, $A \cap U \neq \emptyset$, $A \cap V \neq \emptyset$.■

Definition 1.2.18 *A non-empty subset* A *of a topological space is said to be* **path connected** *if given any pair of points* x *and* y *in* A*, there exists a continuous function* $\gamma : [0,1] \to A$ *such that* $\gamma(0) = x$ *and* $\gamma(1) = y$. ■

Proposition 1.2.7 (i) *Every continuous image of a connected set is connected.*
(ii) *The product of connected sets is connected.*
(iii) *Every path connected set is connected.* ■

In particular, every ball in a metric space is path connected, and hence, connected.

The only connected sets in $\mathbb{R}$ are intervals.

For a detailed study of topological spaces, the reader is referred to, for instance, Simmons [9].