I. J. MADDOX

Elements of Functional Analysis

CAMBRIDGE UNIVERSITY PRESS

ELEMENTS OF FUNCTIONAL ANALYSIS

I. J. MADDOX
B.A., Ph.D., D.Sc.
Professor of Pure Mathematics
in the Queen's University of Belfast

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON · NEW YORK · MELBOURNE

Published by the Syndics of the Cambridge University Press
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
Bentley House, 200 Euston Road, London NW1 2DB
32 East 57th Street, New York, NY 10022, USA
296 Beaconsfield Parade, Middle Park, Melbourne 3206, Australia

Library of Congress catalogue card number: 77-6575

ISBN 0 521 07617 X hard covers
ISBN 0 521 29266 2 paperback

First published 1970
First paperback edition 1977

Printed in Great Britain
at the University Press, Cambridge

This book is dedicated to the memory of

EDITH MADDOX (1877–1958)

May it serve as an epitaph

ERRATA

p. 33, question 11 — *for* $|a_{ij}|$ *read* $|a_{ij}-b_{ij}|$
p. 48, line 17 — *for* is it *read* it is
p. 72, line 21 — *for* x_p *read* x_k
p. 72, line 25 — *for* λ_1 *read* λ
p. 72, line 31 — *delete* . *after* simple
p. 74, line 20 — *for* subset M *read* subspace M
p. 88, line 16 — *for* x_k *read* b_k
p. 101, question 6 — *for* $\frac{1^n}{2}$ *read* $\left(\frac{1}{2}\right)^n$
p. 107 — *Replace the last six lines by*
Now $\|A_n(x)-A_m(x)\| \leqslant \epsilon\|x\|$ for each $x \in X$ and all $n, m \geqslant N(\epsilon)$. Letting $m \to \infty$ we get
$$\|A_n(x)-A(x)\| \leqslant \epsilon\|x\|,$$
so $\|A_n - A\| \leqslant \epsilon$ for all $n \geqslant N(\epsilon)$, whence $A_n \to A$ in the norm of $B(X, Y)$.
p. 109, line 10 — *for* $\{e, e_1, e_2, \ldots\}$ *read* $(e, e_1, e_2, \ldots)$
p. 120, line 18 — *for* injection *read* bijection
p. 124, line 2 up — *for* $[0, 1]$ *read* $(0, 1]$
p. 125, line 3 — *for* Now define *read* Now define $z_n(0) = x(0)$, and for $0 < u \leqslant 1$ define
p. 125, line 5 — *For* $[0, 1]$ *read* $(0, 1]$. *Also, for* $((r-1)/n, r/n)$ *read* $((r-1)/n, r/n]$
p. 126, lines 8, 9, 10 — *for* 0 *in the lower limit of integration read* c
p. 129, line 26 — *for* 5 *read* 3
p. 134, line 2 — *for* a_{kj} *read* b_{kj}
p. 135, line 26 — *for* $\bar{x}\bar{y}$ *read* $\overline{xy}$.
p. 136, question 8 — *for* $\overline{xy} = \bar{y}\bar{x}$ *read* $\overline{xy} = \bar{y}.\bar{x}$
p. 137, line 23 — *for* $z \neq \theta$ *read* $f(z) \neq 0$
p. 139, line 4 — *for* coefficients *read* coefficients such that $|\lambda_1| + |\lambda_2| > 0$
p. 139, line 5 — *Delete* question 2.
p. 139, line 17 — *for* $xi \in I$. Such a set I is called a *left ideal* *read* $xi \in I$ and $ix \in I$. Such a set I is called an *ideal*
p. 139, line 26 — *for* is a left ideal *read* is an ideal
p. 139, line 33 — *for* an ideal, i.e. $x \in X$ *read* an ideal, i.e. rad (X) is a linear subspace, and $x \in X$
p. 141, line 1 — *for* x *read* A
p. 142, line 12 — *for* $\sum_0^\infty (e-x)^k$ *read* $\sum_1^\infty (e-x)^k$

ERRATA

p. 152, line 4 up	*for* $y = px$ *read* $x = py$
p. 155, line 7 up	*for* Example 3 *read* Example 4
p. 157, line 9	*for* orthogonal *read* orthonormal
p. 168, line 4 up	the word *implies* should be in roman type
p. 175, line 7	*for* e^{-k} *read* e^{-t}
p. 198, question 1	*P. Enflo, in Acta Math.* (1973), *proved that there is a separable Banach which has no basis.*
p. 198, question 5	*for* isomorphic *read* continuously isomorphic
p. 198, question 7	*for* X *read* s. *Also, for* be a sequence space *read* be the space of all sequences

PREFACE

There are several excellent books which deal with the subject of functional analysis. Few can be regarded as really elementary or introductory. As a beautiful theory in its own right and for its richness in applications, functional analysis in some shape or form is now taught to second and third year mathematics undergraduates at several British universities. My experience in teaching such students has indicated that they need quite a gentle introduction—largely due to two things: that their analytical abilities are not sufficiently developed and that they are unused to 'abstract' reasoning. In my view, the field of elementary functional analysis is the ideal place in which to learn some abstract structural mathematics and to develop analytical technique.

It is my hope that this book may provide a really introductory, though non-trivial, course on functional analysis for undergraduates who have completed basic courses on real and complex variable theory. Although primarily addressed to students of mathematics it is expected that the approach is basic enough to enable students of physics and engineering to get something of the flavour of the subject.

Of the several excellent books mentioned above, the master work of Banach: *Théorie des opérations Linéaires* (1932) must stand first. Every serious student of analysis should regard his education incomplete until he has read something of this remarkable germinal book.

There is one feature of the present work which we should perhaps mention. Much of the theory is illustrated by examples involving sequence spaces rather than integration spaces. This is partly because most results for sequence spaces will fairly readily generalize to integration spaces, but mainly because the student to whom this book is addressed is unlikely to be sufficiently familiar with integrals of the depth of Lebesgue to enable him to really appreciate examples involving them. However, it has been thought advisable to prove the completeness of the important L_p spaces, referring to works on integration for the relevant theorems on interchange of limit and integral.

Chapter 1 of the book is absolutely fundamental, though extremely elementary. Some may wish to omit it and proceed to the next chapter on metric and topological spaces. In my view it would be best to make certain of the material in chapter 1 before attempting the rest of the text. There are over 300 exercises in the book, many of which are quite routine, though just a few, which appear at the end of the last chapter, are quite difficult. It is recommended that most of the exercises should be attempted—to learn mathematics one must do it.

The final chapter of the book concerns an area of Mathematics which is of special interest to me. Those students who wish to begin graduate work in this field may find it a useful introduction. Readers who are not so inclined may, nevertheless, see functional analysis at work in a fairly concrete situation.

Debts of gratitude are several. At the undergraduate level my interest in analysis was stimulated by Professor D. C. Russell. As a research student I was greatly influenced by my supervisor, Dr B. Kuttner. A number of my colleagues at the University of Lancaster have made helpful comments on the book, during many conversations. I am especially indebted to P. L. Walker for his careful scrutiny of the typescript and for numerous valuable suggestions. Useful assistance was also rendered by J. W. Roles and C. G. Lascarides.

The manuscript was expertly typed by Mrs Sylvia Brennan and Miss June Unsworth, and I gratefully acknowledge their help.

I. J. MADDOX

University of Lancaster, 1969

CONTENTS

1

BASIC SET THEORY AND ANALYSIS

1. Sets and functions

The great German mathematician G. Cantor (1845–1918) is usually regarded as the creator of the theory of sets. As our starting point for this book we shall take Cantor's definition of a set: 'A set is any collection of definite, distinguishable objects of our thought, to be conceived as a whole.' The objects mentioned in the definition are called the elements or members of the set. Usually we denote sets by capital letters and elements by lower case letters. If X is a set then we write $x \in X$ to mean that x is an element of X. When an object x is not an element of a set X, we write $x \notin X$.

In what follows we shall take for granted the following sets, which occur throughout mathematics:

$N = \{1, 2, 3, \ldots\}$, the set of all positive integers,

$Z = \{0, 1, -1, \ldots\}$, the set of all integers,

Q, the set of all rational numbers,

R, the set of all real numbers,

C, the set of all complex numbers.

The notation arises as follows: N for natural numbers, Z for Zahlen (German for integers), Q for quotient. The notation R and C seems to require no explanation.

Usually, in a given discussion, we take a fixed set and everything is carried out with reference to it alone. In such a case the fixed set is called the universe of discourse. For example, in number theory the universe of discourse is Z. Within a universe of discourse X a common way of generating a set is to take an object in X of a certain type and then to consider the set of all such objects. For example, having defined an object in Z called a prime number we may then consider the set of all prime numbers.

In a work of the present nature we are primarily concerned with the manipulation of sets, rather than with their deeper properties. To this end we now introduce notation and definitions, and observe some simple results.

First, there is no way, in general, of explicitly writing down all the elements of a set. For example, it is in the nature of the positive integers N that they cannot all be explicitly exhibited. We have to be content to write $N = \{1, 2, 3, \ldots\}$; the three dots leaving much to the imagination. Generally, we use the curly bracket notation for sets either writing down the first few elements and then some dots which we agree is to tell us that the law of formation of the elements is well-known or obvious, or we put in the law of formation. For example, $\{x | x \in N \text{ and } x > 8\}$ is read as 'the set of all x such that x is a positive integer and x is greater than 8'. The vertical bar following x is read as 'such that'. Thus we could write this last set as $\{9, 10, 11, \ldots\}$. Again, $\{x | x \in R \text{ and } x > 0\}$ denotes the set of all strictly positive real numbers. In this case it is not possible to write down the elements explicitly, or even in such a way as to indicate the law of formation, such is the nature of the real numbers. In fact it will be seen later that the set $\{x | x \in R \text{ and } x > 0\}$ is uncountable, so that the elements cannot even be exhibited as an infinite sequence $x_1, x_2, x_3, \ldots$. We remark that the order of the elements in a set is generally irrelevant. For example, N is the same set as $\{2, 1, 4, 3, 6, 5, \ldots\}$.

If A, B are sets then the notation $A \subset B$ means that every element of A is also an element of B. If $A \subset B$ then we say that A is a subset of B, B is a superset of A, A is included in B and also B includes A. The notation $B \supset A$ is regarded as equivalent to $A \subset B$. We define $A = B$ if and only if $A \subset B$ and $B \subset A$. Also, we say A is a proper subset of B if and only if $A \subset B$ but $A \neq B$. For example, the set of odd integers is a proper subset of Z. We remark that some writers use the notation $A \subseteq B$, which allows equality, and reserve $A \subset B$ for proper subsets. On occasion we shall also say that '$A \subset B$, strictly', meaning that A is a proper subset of B.

Two simple properties of the set inclusion $\subset$ are:

(i) $A \subset A$,

(ii) $A \subset B$ and $B \subset C$ imply $A \subset C$.

If A is a given set let us consider that subset of A defined as $\{x \in A | x \neq x\}$. This set has no elements and is known as the *empty set*. It is denoted by $\varnothing$ and has the property that $\varnothing \subset A$ for every set A. Each set $A \neq \varnothing$ has at least two distinct subsets, A and $\varnothing$. If A has only these two subsets then A must be a one element set, $A = \{a\}$, say, where a is the sole element of A. Note that $\varnothing$ has no elements but that the one element set $\{\varnothing\}$ is not empty.

Unions and intersections of sets

Given sets A, B we may form two new sets from them:

$$A \cup B = \{x | x \text{ belongs to at least one of } A \text{ and } B\},$$
$$A \cap B = \{x | x \in A \text{ and } x \in B\}.$$

We call $A \cup B$ the *union* and $A \cap B$ the *intersection*, of A and B. For example, $\{1, 2, 3\} \cup \{1, 4, 3\} = \{1, 2, 3, 4\}$; $\{2, 3\} \cap \{1, 3, 2\} = \{2, 3\}$. It is trivial that $A \cap B \subset A \subset A \cup B$ for any sets A and B. If $A \cap B = \varnothing$, then we say that A and B are disjoint.

We shall often want to form the union or intersection of a whole class (or collection) of sets. Let $\mathscr{S}$ be a class of sets A. Then we define

$$\cup \{A | A \in \mathscr{S}\} = \{x | x \in A \text{ for at least one } A \in \mathscr{S}\},$$
$$\cap \{A | A \in \mathscr{S}\} = \{x | x \in A \text{ for all } A \in \mathscr{S}\}.$$

Sometimes we write $\cup A_\alpha$, $\cap A_\alpha$, where we think of α as running through some indexing set. If α runs through N we usually write

$$\cup \{A_n | n \in N\} = \bigcup_{n=1}^{\infty} A_n,$$

and similarly for $\bigcap_{n=1}^{\infty} A_n$. The '$\infty$' in this notation is conventional, but superfluous, not to say confusing. It is emphasized that A_∞ is not in the collection $\{A_n | n \in N\}$. Observe also that no limiting process is involved in the above. Thus, for example, to say that $x \in \bigcup_{n=1}^{\infty} A_n$, is to say that there is a positive integer p such that $x \in A_p$.

Example 1. Let A_n be the interval $[0, 1 + 1/n)$ on the real line, i.e. $A_n = \{x \in R | 0 \leqslant x < 1 + 1/n\}$, $n = 1, 2, \ldots$. Then

$$\bigcap_{n=1}^{\infty} A_n = [0, 1] = \{x \in R | 0 \leqslant x \leqslant 1\}.$$

To show this we first prove $[0, 1] \subset \cap A_n$ and then prove $\cap A_n \subset [0, 1]$. Now $x \in [0, 1]$ implies $0 \leqslant x \leqslant 1 < 1 + 1/n$, for all $n \in N$, i.e. $x \in A_n$ for all $n \in N$, i.e. $x \in \cap A_n$. Conversely, $x \in \cap A_n$ implies $0 \leqslant x < 1 + 1/n$, for all $n \in N$, whence $0 \leqslant x \leqslant 1$ (either letting $n \to \infty$, or supposing $x > 1$ and obtaining a contradiction to $x < 1 + 1/n$ for all $n \in N$).

Cover for a set

Let $\mathscr{S}$ be a class of sets A. Then the class $\mathscr{S}$ is called a cover for a set X if and only if

$$X \subset \cup \{A \mid A \in \mathscr{S}\}.$$

Any subclass of $\mathscr{S}$ which also covers X is called a subcover of $\mathscr{S}$.

The notion of 'open' cover will be employed in chapter 2, in connection with compact sets. The 'open' here refers to the fact that the sets of the cover are open sets, in the sense of topology. For the moment we shall be content with a very simple example on covers.

Example 2. (i) Let I_n be the open interval

$$(n, n+1) = \{x \in R \mid n < x < n+1\}$$

on the real line. Then the class $\{I_n \mid n \in Z\}$ is not a cover for R, for no integer belongs to $\cup \{I_n \mid n \in Z\}$.

(ii) If $J_n = \{x \in R \mid n \leqslant x < n+1\} = [n, n+1)$, then the class $\{J_n \mid n \in Z\}$ is a cover for R.

(iii) Let $S[a, r] = \{z \in C \mid |z-a| \leqslant r\}$, where $a \in C$ and $r > 0$. Thus $S[a, r]$ is the closed disc of centre a and radius r in the complex plane. It is clear that the class $\{S[m+in, 1] \mid m, n \in Z\}$ is a cover for C.

Complementation

If X is our universe of discourse and $A, B \subset X$ then we define

$$A \sim B = \{x \in X \mid x \in A, x \notin B\}.$$

We call $A \sim B$ the *complement of B with respect to A*. By $\sim A$ we mean $X \sim A$, and we call $\sim A$ the complement of A. It is clear that $A \sim B = A \cap (\sim B)$, $\sim(\sim A) = A$, and that $A \subset B$ is equivalent to $\sim B \subset \sim A$.

The two following results concerning complementation are known as De Morgan's laws:

$$\sim \cup A_\alpha = \cap (\sim A_\alpha); \quad \sim \cap A_\alpha = \cup (\sim A_\alpha).$$

To prove the first of these, for example, we merely note that $x \in \sim A_\alpha$ for all α is equivalent to $x \notin A_\alpha$ for any α.

Some other properties of union and intersection which are easy to show are

(i) $\cap A_\alpha \subset A_\alpha \subset \cup A_\alpha$, for any α,

(ii) $A \cup (\cap A_\alpha) = \cap (A \cup A_\alpha)$,

(iii) $A \cap (\cup A_\alpha) = \cup (A \cap A_\alpha)$.

Ordered pair

Let x, y be any objects. Then the ordered pair (x, y) is defined as the set $\{\{x\}, \{x, y\}\}$. It is easy to check the fundamental property of ordered pairs: $(x, y) = (u, v)$ if and only if $x = u$ and $y = v$. More generally we may define in a similar way an ordered n-tuple $(x_1, \ldots, x_n)$ with the property $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ if and only if $x_1 = y_1, \ldots, x_n = y_n$.

Relation

A relation ρ is defined to be a set of ordered pairs. For example, $\rho = \{(1, 2), (a, b)\}$ is a relation.

Equivalent notation for $(x, y) \in \rho$ is $x\rho y$. Thus in our example we might write $1\rho 2$ instead of $(1, 2) \in \rho$.

An important type of relation is the

Cartesian product

Let X, Y be given sets. Then

$$X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}$$

is called the Cartesian product of X and Y.

Another important relation is the

Equivalence relation

Let X be a given set. A relation ρ is called an equivalence relation on X if and only if it is (i) *Reflexive, i.e. $x\rho x$ for all $x \in X$,* (ii) *Symmetric, i.e. $x\rho y$ implies $y\rho x$,* (iii) *Transitive, i.e. $x\rho y$ and $y\rho z$ imply $x\rho z$. It is usual to denote an equivalence relation by $\sim$ rather than ρ. There is little danger of confusion with complementation.*

Example 3. (i) Equality is obviously an equivalence relation on any set.

(ii) Let $X = \{(x, y) \mid x, y \in N\}$. Define $(x, y) \sim (u, v)$ to mean $xv = yu$. Then $\sim$ is an equivalence relation on X. For example, let us check transitivity: $(x,y) \sim (u,v)$ and $(u,v) \sim (z,w)$ imply $xv = yu$, $uw = vz$, whence $xvuw = yuvz$, and so $xw = yz$, i.e. $(x,y) \sim (z,w)$.

(iii) Define $x \sim y$ to mean $x-y$ is divisible by 2. It is easy to check that $\sim$ is an equivalence relation on Z.

Let us return to general relations. The *domain* of a relation is the set of all first co-ordinates of its members. The *range* is the set of all second co-ordinates. If $\sim$ is an equivalence relation on X, then we define $E_x = \{y \in X \mid y \sim x\}$ and call E_x the equivalence class containing the element x. For example, in example 3 (iii) we have

$$E_0 = \{0, \pm 2, \pm 4, \ldots\}.$$

In general, it is easy to check that $E_x = E_y$ if and only if $x \sim y$, and that $E_x \cap E_y = \varnothing$ if $E_x \neq E_y$. It is thus evident that $\{E_x \mid x \in X\}$ forms a partition of X, i.e. X is the union of the disjoint classes E_x. For example, in example 3 (iii) we have

$$Z = \{0, \pm 2, \pm 4, \ldots\} \cup \{\pm 1, \pm 3, \pm 5, \ldots\}.$$

Probably the most significant type of relation that occurs in mathematics is that which is called a function. The following definition of a function may seem rather strange to those who are used to books of analysis which extensively employ functions but never actually define them.

Function

A function f is defined to be a relation, such that if $(x,y) \in f$ and $(x,z) \in f$ then $y = z$. Four other terms for function are map, mapping, operator, and transformation.

Our concept of a function as a certain set of ordered pairs is what some would call the graph of a function, since they define a function as a 'rule' or some such. On occasion we shall use the term 'graph of a function', when this seems more expressive. However, to us, a function and its graph are exactly the same thing.

Example 4. (i) $\{(1,2),(2,2)\}$ and $\{(z, z+1) \mid z \in C\}$ are functions.

(ii) $\{(1,2),(1,4)\}$ and $\{(x^2, x) \mid x \in R\}$ are not functions. For example $(1,1)$ and $(1,-1)$ are in the second set.

(iii) $\{(x^2, x) \mid x \in R^+\}$ is a function. Here $R^+ = \{x \in R \mid x > 0\}$. In this case, if the first co-ordinates are equal, $x^2 = y^2$, then $(x-y)(x+y) = 0$, so $x = y$, i.e. the second co-ordinates are equal.

If f is a function and $(x, y) \in f$ then we write $y = f(x)$, which is the conventional notation for y as a function of x. We say that y is the value of f at x, or that y is the image of x under f.

The notation

$$f: X \to Y$$

is now widely used in mathematics. It is interpreted as 'f is a function from the set X into the set Y'. The meaning of $f: X \to Y$ is that X is the domain of f and that the range of f is a subset of Y, not necessarily the whole of Y.

If $f: X \to Y$ and $A \subset X$, then the function $g: A \to Y$, defined by $g(a) = f(a)$, for $a \in A$, is called the *restriction* of f to A.

Example 5. (i) Define f by $f(x) = e^x$, for $x \in R$, i.e. the domain of f is R and $f = \{(x, e^x) \mid x \in R\}$. The range of f is in fact R^+, as is well-known. We may write, with increasing accuracy, $f: R \to R$, and $f: R \to R^+$.

(ii) Define f by $f(z) = |z|$, for $z \in C$. Then, with increasing precision, we have $f: C \to C$, $f: C \to R$, and $f: C \to \{x \in R \mid x \geqslant 0\}$.

Bijective maps

Let $f: X \to Y$. Then f is called injective if $f(x_1) = f(x_2)$ implies $x_1 = x_2$, for every $x_1, x_2 \in X$. If the range of f is the whole of Y, then f is called surjective. A mapping which is both injective and surjective is called bijective.

The terms 'one-one', 'onto' and 'one-one correspondence' are sometimes used instead of 'injective', 'surjective' and 'bijective mapping' respectively.

Example 6. $f: R \to \{x \in R \mid x \geqslant 0\}$, defined by $f(x) = x^2$ is surjective but not injective. The same prescription for f, but with $f: R^+ \to R^+$, is bijective.

Inverse function

Let $f: X \to Y$ be bijective. Since f is surjective, if $y \in Y$ then there exists $x \in X$ such that $y = f(x)$. This x is unique, since f is injective. Hence there is an inverse function $g: Y \to X$ such that $g(f(x)) = x$, for all $x \in X$, and $f(g(y)) = y$ for all $y \in Y$. It is usual to write $g = f^{-1}$.