OXFORD APPLIED MATHEMATICS
AND COMPUTING SCIENCE SERIES

A First Look at Numerical Functional Analysis

W. W. SAWYER

W. W. SAWYER
Professor Emeritus, University of Toronto

A first look at Numerical functional analysis

1478
1978
CLARENDON PRESS · OXFORD

Oxford University Press, Walton Street, Oxford OX2 6DP

OXFORD LONDON GLASGOW NEW YORK
TORONTO MELBOURNE WELLINGTON CAPE TOWN
IBADAN NAIROBI DAR ES SALAAM LUSAKA ADDIS ABABA
KUALA LUMPUR SINGAPORE JAKARTA HONG KONG TOKYO
DELHI BOMBAY CALCUTTA MADRAS KARACHI

British Library Cataloguing in Publication Data

Sawyer, Walter Warwick
A first look at numerical functional analysis.
—(Oxford applied mathematics and computing science series).
1. Numerical analysis 2. Functional analysis
I. Title II. Series
511'.7'015157 QA297 77-30689

ISBN 0-19-859628-6
ISBN 0-19-859629-4 Pbk

Typeset in Northern Ireland
at The Universities Press, Belfast
Printed in Great Britain
by Richard Clay & Co. Ltd., Bungay

Preface

THE AIM of this book is to provide an intelligible introduction to functional analysis by giving samples of its applications to numerical analysis. I have been very much helped in my attempt to select topics of practical value by discussions and correspondence with mathematicians in the University of Lancaster, Brunel University, the University of Bristol, the University of Glasgow, the University of Strathclyde, and the University of Exeter. For this help I should like to express my sincere thanks.

34 *Pretoria Road,*
Cambridge.
July 1977

Notes on symbols and terminology

FOR CONCEPTS not mentioned here, see the general index.

$\rightarrow$, (i) 'tends to the limit', (ii) 'maps to', e.g. 'input $\rightarrow$ output'.

$\equiv$, 'identically equal'. An *identity*, $f(x) \equiv g(x)$, indicates that the functions have the same values for all relevant x.

$\simeq$, 'approximately equal'.

$\sim$, see section 8.4.

!, factorial; $n! = n(n-1)(n-2)\ldots 3 \times 2 \times 1$.

Δ, finite difference operator, 'the change in'.

D, differentiation operator, d/dx.

ln, natural logarithm, $\log_e$.

$P_n(x)$, Legendre polynomial, section 8.4.

$T_n(x)$, Chebyshev polynomial, section 8.4.

$[a, b]$, the closed interval, corresponding to $a \leqslant x \leqslant b$, in contrast to the open interval, (a, b), corresponding to $a < x < b$.

$\mathbb{R}$, the real numbers

$\mathbb{C}$, the complex numbers.

$\in$, 'belongs to'. Thus, $x \in \mathbb{R}$ means 'x is a real number'. Note that $\in$ is not the same symbol as the Greek epsilon, ε, traditional symbol for a small, positive number.

$\{:\}$, a collection of objects satisfying a condition. Thus $\{x : x > 0\}$ means 'the positive numbers' and $\{(x, y) : x^2 + y^2 = 1\}$ means 'the collection of points for which $x^2 + y^2 = 1$', that is, the unit circle.

$\max\{:\}$, the maximum, the largest number in the collection indicated.

$\min\{:\}$, the minimum, the smallest number in the collection.

$\sup\{:\}$, the supremum, the 'ceiling' of the collection, that is, the smallest number not exceeded by any other in the collection. Thus $\sup\{1/2, 2/3, 3/4, \ldots, n/(n+1), \ldots\}$ is 1. We cannot use 'max' here, since there is no largest number in this collection.

$\inf\{:\}$, infimum, the 'floor' of the collection.

majorant, something known to be larger than the object under investigation.

$\Rightarrow$, 'implies', e.g. '$x=-5 \Rightarrow x^2=25$' means 'if $x=-5$, then $x^2=25$'.

$\Leftrightarrow$, connects statements, each of which follows from the other, e.g. $2x=10 \Leftrightarrow x=5$.

$\ell_1, \ell_2, \ell_\infty$, see section 2.3.

ℓ_p, section 4.1.

$\mathscr{L}_2$, section 8.3.

$\mathscr{C}[a, b]$, section 3.3.

$\mathscr{X}, \mathscr{Y}, \mathscr{Z}$, symbols for spaces.

$\mathscr{B}(\mathscr{X}, \mathscr{Y})$, the space of bounded, linear operators, $\mathscr{X} \to \mathscr{Y}$, section 5.3.

$\mathscr{X}^*$, the space dual to $\mathscr{X}$, section 10.1.

$\mathbb{R}^n$, vector space of n dimensions with real co-ordinates. If $u=(u_1, \ldots, u_n)$ and $v=(v_1, \ldots, v_n)$ are in $\mathbb{R}^n$, then $u+v=(u_1+v_1, \ldots, u_n+v_n)$ and, for any number, k, $ku=(ku_1, \ldots, ku_n)$.

$\mathscr{E}^n$, Euclidean space of n dimensions.

$u \cdot v$, scalar product in $\mathscr{E}^3$, sections 8.1 and 9.1.

(u, v), generalization of $u \cdot v$, chapter 8.

Perpendicular projection. If $PM \perp OMQ$, the vector OM is called the perpendicular projection of the vector OP onto the line OQ.

$|\;|$, absolute value; the magnitude of a number, irrespective of its sign, e.g. $|-7|=|7|=7$.

$\|PQ\|$, the length of the line segment, PQ. Its generalizations; vector norm, $\|v\|$, section 3.2; $\|v\|_p$, norm of a vector in ℓ_p; operator norm, $\|A\|$, section 5.2; function norm, $\|f\|$, section 3.3; norm in $\mathscr{L}_2$, (Hilbert space), section 8.3; $\|A\|_p$, norm of an operator A, $\ell_p \to \ell_p$ or $\ell_p \to \mathbb{R}$.

$d(P, Q)$, the distance of P from Q.

$S(P, r)$, the sphere, $B(P, r)$ the open ball.

$\bar{B}(P, r)$, the closed ball, in each case with centre P and radius r. See section 2.3.

I, the identity operator or matrix, 'leave everything as it is'.

A^{-1}, the operator or matrix inverse to A. When A^{-1} exists, $A^{-1}A = AA^{-1} = I$. See section 5.6.

Invariant line. A line such that every point on it maps, under a specified linear transformation, M, to a point of the same line.

Eigenvector. A non-zero vector, v, lying in an invariant line; it satisfies the equation, $Mv = \lambda v$, for some number λ.

Eigenvalue. The number, λ, mentioned in the definition of eigenvector.

Contents

1 A first course in functional analysis

In the introduction to his book, *Functional analysis and numerical analysis*, L. Collatz said that numerical analysis had been revolutionized by two things—the electronic computer and the use of functional analysis.

This statement is striking not only for its emphatic tone but also for the contrast of mathematical epochs it involves. Numerical analysis is an earthy subject concerned with questions involving numbers. Even if it uses very sophisticated modern equipment, essentially it is concerned with arithmetic, the oldest branch of mathematics, whose beginnings lie in prehistory. Functional analysis, on the other hand, while it has roots in nineteenth century mathematics, is a product of the present century.

In trying to cope with such a modern branch of mathematics, a student of numerical analysis encounters two difficulties. Any part of modern mathematics is the end-product of a long history. It has drawn on many branches of earlier mathematics, it has extracted various essences from them and has been reformulated again and again in increasingly general and abstract forms. Thus a student may not be able to see what it is all about, in much the same way that a caveman confronted with a vitamin pill would not easily recognize it as food.

The second difficulty is that functional analysis was not created with numerical applications in view. It arose from a great variety of sources—from the calculus of variations, from integral equations, from Fourier series, from mechanics, from the theory of real and complex variables, from number theory, and from other topics. It therefore has a great range of possible applications and a student cannot assume, because a theorem in functional analysis is generally regarded by mathematicians as of great importance, that it will necessarily help us in problems of numerical analysis.

The aim of this book is to make some contribution towards overcoming these two difficulties—by explaining the ideas of the subject and, as far as possible, by emphasising those ideas that have proved useful to numerical analysts. In the main the concepts of functional analysis will be introduced by discussing problems in numerical analysis from which these concepts could arise. References will be given so that readers can go back to original sources to clear up any obscurity or to judge for themselves the relevance of the topic to their own interests. A reference such as (Smith, (3), p. 42) would indicate page 42 of the third paper by Smith in the list of references at the back of the book.

Also at the front of the book will be found a list of symbols and words used, with their meanings.

This list should be consulted if any unfamiliar word or symbol is encountered. It seemed better to arrange things this way, rather than to risk wearying readers by explaining in the body of the book matters which were already familiar to them. Indeed at present it is extremely difficult to know what is familiar. This is partly owing to the separation of school programs into modern and traditional, and, at a deeper level, to the fact that mathematics is becoming so varied that it is no longer possible to assume that some agreed core is known to all students of mathematics.

In many schools, a modern course, such as S.M.P., is supplemented by more traditional material, designed to give facility in algebraic manipulation. This is probably the ideal background for our subject. This book does assume some acquaintance with vectors and matrices. Students with a traditional background (which they will find an asset in this work) may find it helpful to consult the S.M.P. 'transitional' books, published at a time when schools were just beginning to move from the older to the newer syllabus. The introduction to matrices in S.M.P. Book T.4 is perhaps better than anything in the later S.M.P. publications.

It is not too difficult for someone who has a good understanding and command of traditional algebra and calculus to acquire the more modern concepts. This is perhaps because the later developments grew on the foundation of the earlier ones. Students who are in the opposite position, of having the more recent concepts, but not feeling comfortable with the older topics and skills, will probably find it harder to adjust and establish a correct balance of new and old. Students, who find difficulty with the manipulative aspects of the work, should consult, and above all work exercises from, books of an earlier period when such skills were emphasized—often to excess.

1.1. 'Soft and hard' analysis

There is a custom among mathematicians of referring to classical analysis as 'hard' and modern analysis as 'soft'. Students may not be surprised to hear nineteenth-century analysis described as hard, but one would expect the twentieth-century analysis that followed it to be more advanced and even harder. The explanation of this terminology is along the following lines. Classical analysis frequently involved long chains of reasoning and calculation. Mathematicians tried to simplify it by disentangling the various strands involved in it, in order to find simplicity underlying the complexity. The real numbers have many properties;—algebraic properties, properties concerned with limits, order properties such as that 3 comes before 10 when we count. An effort was made to separate these and to see what could be said about a mathematical structure if you knew only its algebraic properties, or its order properties, or its behaviour in regard to limits. Such studies, concerned only with one aspect of the real numbers, were naturally simpler than the older work that dealt with all aspects simultaneously. Sometimes they did not involve calculation at all. As P. J. Davis said, functional analysis 'loves soft analysis and avoids hard analysis like the plague; its ideal proof is wholly verbal' (see Hayes, p. 160).

The advantage of soft analysis, besides its relative simplicity, is its generality. One of the main features of mathematics in the last century has been the gradual realization that the theorems and procedures of algebra and calculus apply not only to the real and complex numbers, but to a wide variety of other objects, including several that are of great interest to numerical analysts. All the work of this book will illustrate this theme.

Ideas of great generality are extremely valuable, but they are hardly ever sufficient, by themselves, for dealing with a particular, concrete situation. Soft analysis therefore is a supplement to, not a substitute for, hard analysis. Soft analysis, as we have just seen, grew from classical analysis and revealed wider applications of the older ideas. It therefore binds together our knowledge of different topics and reduces the strain on the memory. Instead of learning disconnected facts about unrelated objects, we can take a classical theorem or procedure and see it operating again and again in different environments. The organization of this book corresponds to that idea. We shall take in turn concepts or theorems of classical mathematics, briefly review their role in their original setting (real or complex numbers) and then see to what other situations they can be applied.

Functional analysis helps us also by providing a way of visualizing what we are doing. The whole language of the subject is in terms of 'spaces'. This means that we are able to use geometrical and pictorial imagery in situations that, at first sight, appear to have nothing to do with geometry.

What has been discussed so far may be called the inherent benefits of functional analysis. Functional analysis has also acquired a social value—it has become necessary for communication. Books on numerical analysis, even when using classical methods, often express these in modern terminology. For instance Cheney, on page 85 of his well known book, *Introduction to approximation theory*, discusses the problem of finding a polynomial that closely approximates a given continuous function. The argument is purely classical, involving inequalities for real numbers. The terminology however involves two items that are not classical; functions are defined on 'a compact metric space' and the algebra repeatedly uses $\|P\|$, the symbol for a norm. To follow the argument on this page, a student needs only to know the meaning of these concepts and their most elementary properties.

2 Old ideas in new contexts

2.1. Examples of iteration

As HAS already been indicated, we plan to extend several concepts of classical analysis so that they become available for a wider range of applications. A frequent object of such applications is the process of iteration, one of the most valuable and widely used methods of calculation. It is therefore helpful to look at a few examples of iteration, in order to see something of the variety that this process covers.

Example 1. The idea of iteration is extremely old and is implicit in Zeno's paradox of Achilles and the tortoise (B.C. 500). The tortoise is given unit distance start but Achilles moves ten times as fast. When Achilles has covered a distance x, the tortoise is at the point t, where $t = 1 + (0.1)x$. Initially $x = 0$, $t = 1$. When Achilles has reached the tortoise's initial position, $x = 1$, but by then $t = 1.1$. When x is raised to 1.1, t has become 1.11 and so the argument continues. Its effect is to produce a sequence of x-values, say $x_1, x_2, x_3, \ldots$ with $x_{n+1} = 1 + (0.1)x_n$. As $n \to \infty$, x_n approaches the solution of $x = 1 + (0.1)x$.

Example 2. In any application of iteration convergence has to be considered. We would of course get disastrous results if we tried to solve $x = 1 + 3x$ by the iteration $x_{n+1} = 1 + 3x_n$, which would lead to $+\infty$ rather than -0.5.

Example 3. It is not necessary that the equation should be linear. If we take $x_{n+1} = 3/(x_n + 10)$ with $x_0 = 0$, six iterations are sufficient to give 0.291 502 622 as a solution of $x^2 + 10x - 3 = 0$.

Example 4. If we perform the iteration

$$x_{n+1} = x_n y_n + x_n + 0.07$$
$$y_{n+1} = x_n^2 + y_n^2 + y_n - 0.41$$

with initial values $x_0 = 0$, $y_0 = 0$, about twenty iterations are sufficient to give $x = 0.111\,002\,285$, $y = -0.630\,617\,546$ as an intersection of the curves $xy = -0.07$, $x^2 + y^2 = 0.41$.

Needless to say, this is not offered as an example of practical value. The examples in this book have been kept within the capacity of a pocket-size programmable computer. In a serious, real-life application there might well be a hundred equations in a hundred unknowns.

Example 5. The function $f(x)=e^{-x}$ satisfies the integral equation

$$f(x)=1-\int_0^x f(t)\,dt.$$

It can be obtained by the iteration

$$f_{n+1}(x)=1-\int_0^x f_n(t)\,dt.$$

If $f_0(x)=0$, $f_1(x)=1$, $f_2(x)=1-x$, $f_3(x)=1-x+(x^2/2)$ and so on. The terms of the series for e^{-x} gradually appear.

Here again we have taken a very simple example from a very wide class. A classical method for dealing with integral equations of the form

$$f(x)=g(x)+\int_a^b K(x,y)f(y)\,dy,$$

where $g(x)$ and $K(x, y)$ are given, is to apply iteration to determine $f(x)$. The solution then appears as a series known as the Neumann series.

Example 6. The integral equation in example 5 is a linear integral equation. It is not necessary to restrict ourselves to this type. If we have the equation

$$f(x)=x+\int_0^x [f(t)]^2\,dt$$

and use the iteration

$$f_{n+1}(x)=x+\int_0^x [f(t)]^2\,dt$$

with $f_0(x)=0$ initially, we find $f_1(x)=x$,

$$f_2(x)=x+(x^3/3),\quad f_3(x)=x+(x^3/3)+(2/15)x^5+(1/63)x^7.$$

In fact $f(x) = \tan x$ is the solution of our equation, and in the nth iterate, $f_n(x)$, the first n terms coincide with the Taylor series for $\tan x$.

These six examples differ in the material they involve. In the first three examples we are concerned with a single number, x; example 4 deals with a pair of numbers (x, y) which may be considered as representing a point in a plane; examples 5 and 6 are concerned with functions of a single variable. But it is clear these examples have something in common, and this is particularly evident if we imagine the iteration being carried out by a computer. In each case there is a sub-routine which is applied again and again, the output of each stage becoming the input of the next. All our examples can be expressed by a single symbolic form. If T denotes the sub-routine and ϕ_0 the initial input, we have successive outputs ϕ_1, ϕ_2, $\phi_3 \ldots$ with $\phi_1 = T\phi_0$, $\phi_2 = T\phi_1$, and so on. We may write $\phi_n = T^n\phi_0$ to indicate that the nth output is obtained by applying n times the process T to the initial input ϕ_0.

When we study iteration from the viewpoint of functional analysis, we are not concerned about the nature of the objects ϕ_n. They may be numbers, vectors, matrices, functions, or maybe something else altogether. We are trying to discover properties that relate to the process of iteration and that can be used, in a wide variety of circumstances, to distinguish cases, like example 2 above, in which iteration leads to a disaster, from those in which it proves reliable. You may well doubt whether such a theory is possible and indeed if we took it in its most general form—a study of all repeatable operations—there would indeed be little to say. What has happened historically is that particular examples of successful iteration were found before any general theory was envisaged. From this experience of successful iterations mathematicians managed to winnow certain general theorems and principles to guide them in their future work. It is with such results that we shall be concerned.

2.2. Functions, ancient and modern

The present use of the word *function* is in most respects much wider than it was in former times, though in one respect it is narrower. Three centuries ago an equation $y = f(x)$ would have implied that x and y were numbers, and that the calculation giving y in terms of x belonged to a class that was conventionally accepted.

In nineteenth-century work on complex variables, a large, and very beautiful, part was played by *many-valued functions*, for example $y = f(x)$ being defined by $y^2 = x$. At the present time the first type of restriction has disappeared entirely. It is no longer required that x and y in $y = f(x)$ be numbers, nor is there any restriction on the kind of calculation or rule that leads from x to y. But the term *many-valued function* has been discarded; the one stipulation that is made is that to an acceptable input, x, there is one and only one output, y.

In all the six examples of iteration, the symbol T thus represents a function, for in each of these examples the input determines without any uncertainty the output.

At some stage of life this must involve a student in some mental readjustment. A student in the sixth form, who has been asked to give an example of a function, might mention $f(x) = \sin x$ or $f(x) = x^2$. It would probably cause some surprise if the student mentioned the operations of integration and differentiation. Yet, with modern usage, such a suggestion might well be justified. Suppose, in order to avoid analytical complications, we agree that the acceptable inputs are to be polynomials. To any input, $P(x)$, the operation of differentiation makes correspond one clearly defined output, $P'(x)$. Equally, if by integration we understand the calculation of some definite integral, this operation too defines a function; to any polynomial input, $P(x)$, there corresponds without any uncertainty or ambiguity the output $\int_0^1 P(x)\,dx$.

Functions are often indicated by the symbolism of an arrow. Thus instead of $f(x) = x^2$ we may write $f: x \to x^2$. This indicates that when the input is any number x, the output is the square of that number, x^2. This type of symbolism is also used for another purpose, to indicate the kind of things that occur as inputs and outputs. Thus, if our example of squaring is concerned with real numbers, we may write $f: \mathbb{R} \to \mathbb{R}$, where $\mathbb{R}$ is a rather pretty symbol used to indicate the real number system. This indicates that the input and output are real numbers.

In the same way, for the function D that represents differentiation, we may define D by writing $D: P(x) \to P'(x)$. We may also indicate that the input and output are assumed to be polynomials by writing D: polynomial $\to$ polynomial.

Some authors like to use different kinds of arrows for these two purposes. This will not be done in this book. Notation becomes an obstacle rather than an aid to learning when it becomes fussy and