Graduate Texts in Mathematics

Gert K. Pedersen

Analysis Now

现代分析 [英]

Springer-Verlag
World Publishing Corp

Gert K. Pedersen

Analysis Now

Springer-Verlag
World Publishing Corp

Gert K. Pedersen
Københavns Universitets Matematiske Institut
2100 København Ø
Denmark

Mathematics Subject Classification (1980): 46-01, 46-C99

Library of Congress Cataloging-in-Publication Data
Pedersen, Gert Kjærgård.
Analysis now / Gert K. Pedersen.
p. cm.—(Graduate texts in mathematics; 118)
Bibliography: p.
Includes index.
1. Functional analysis. I. Title. II. Series.
QA320.P39 1988
515.7—dc19 88-22437

Printed on acid-free paper.

Reprinted by World Publishing Corporation, Beijing, 1990
for distribution and sale in The People's Republic of China only

ISBN 7-5062-0740-0

ISBN 0-387-96788-5 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-96788-5 Springer-Verlag Berlin Heidelberg New York

Graduate Texts in Mathematics 118

For
Oluf and *Cecilie*,
innocents at home

Preface

Mathematical method, as it applies in the natural sciences in particular, consists of solving a given problem (represented by a number of observed or observable data) by neglecting so many of the details (these are afterward termed "irrelevant") that the remaining part fits into an axiomatically established model. Each model carries a theory, describing the implicit features of the model and its relations to other models. The role of the mathematician (in this oversimplified description of our culture) is to maintain and extend the knowledge about the models and to create new models on demand.

Mathematical analysis, developed in the 18th and 19th centuries to solve dynamical problems in physics, consists of a series of models centered around the real numbers and their functions. As examples, we mention continuous functions, differentiable functions (of various orders), analytic functions, and integrable functions; all classes of functions defined on various subsets of euclidean space $\mathbf{R}^n$, and several classes also defined with vector values. Functional analysis was developed in the first third of the 20th century by the pioneering work of Banach, Hilbert, von Neumann, and Riesz, among others, to establish a model for the models of analysis. Concentrating on "external" properties of the classes of functions, these fit into a model that draws its axioms from (linear) algebra and topology. The creation of such "super-models" is not a new phenomenon in mathematics, and, under the name of "generalization," it appears in every mathematical theory. But the users of the original models (astronomers, physicists, engineers, et cetera) naturally enough take a somewhat sceptical view of this development and complain that the mathematicians now are doing mathematics for its own sake. As a mathematician my reply must be that the abstraction process that goes into functional analysis is necessary to survey and to master the enormous material we have to handle. It is not obvious, for example, that a differential equation, a system of linear equations, and a problem in the calculus of variations have anything in common. A knowledge of operators on topological vector spaces gives, however, a basis of reference, within which the concepts of kernels, eigenvalues, and inverse transformations can be used on all three problems. Our critics, especially those well-meaning pedagogues, should come to realize that mathematics becomes simpler only through abstraction. The mathematics that represented the conceptual limit for the minds of Newton and Leibniz is taught regularly in our high schools, because we now have a clear (i.e. abstract) notion of a function and of the real numbers.

When this defense has been put forward for official use, we may admit in private that the wind is cold on the peaks of abstraction. The fact that the objects and examples in functional analysis are themselves mathematical theories makes communication with nonmathematicians almost hopeless and deprives us of the feedback that makes mathematics more than an aesthetical play with axioms. (Not that this aspect should be completely neglected.) The dichotomy between the many small and directly applicable models and the large, abstract supermodel cannot be explained away. Each must find his own way between Scylla and Charybdis.

The material contained in this book falls under Kelley's label: What Every Young Analyst Should Know. That the young person should know more (e.g. more about topological vector spaces, distributions, and differential equations) does not invalidate the first commandment. The book is suitable for a two-semester course at the first year graduate level. If time permits only a one-semester course, then Chapters 1, 2, and 3 is a possible choice for its content, although if the level of ambition is higher, 4.1–4.4 may be substituted for 3.3–3.4. Whatever choice is made, there should be time for the student to do some of the exercises attached to every section in the first four chapters. The exercises vary in the extreme from routine calculations to small guided research projects. The two last chapters may be regarded as huge appendices, but with entirely different purposes. Chapter 5 on (the spectral theory of) unbounded operators builds heavily upon the material contained in the previous chapters and is an end in itself. Chapter 6 on integration theory depends only on a few key results in the first three chapters (and may be studied simultaneously with Chapters 2 and 3), but many of its results are used implicitly (in Chapters 2–5) and explicitly (in Sections 4.5–4.7 and 5.3) throughout the text.

This book grew out of a course on the Fundamentals of Functional Analysis given at The University of Copenhagen in the fall of 1982 and again in 1983. The primary aim is to give a concentrated survey of the tools of modern analysis. Within each section there are only a few main results—labeled theorems—and the remaining part of the material consists of supporting lemmas, explanatory remarks, or propositions of secondary importance. The style of writing is of necessity compact, and the reader must be prepared to supply minor details in some arguments. In principle, though, the book is "self-contained." However, for convenience, a list of classic or established textbooks, covering (parts of) the same material, has been added. In the Bibliography the reader will also find a number of original papers, so that she can judge for herself "wie es eigentlich gewesen."

Several of my colleagues and students have read (parts of) the manuscript and offered valuable criticism. Special thanks are due to B. Fuglede, G. Grubb, E. Kehlet, K.B. Laursen, and F. Topsøe.

The title of the book may convey the feeling that the message is urgent and the medium indispensable. It may as well be construed as an abbreviation of the scholarly accurate heading: Analysis based on Norms, Operators, and Weak topologies.

Copenhagen Gert Kjærgård Pedersen

Contents

CHAPTER 1

General Topology

General or set-theoretical topology is the theory of continuity and convergence in analysis. Although the theory draws its notions and fundamental examples from geometry (so that the reader is advised always to think of a topological space as something resembling the euclidean plane), it applies most often to infinite-dimensional spaces of functions, for which geometrical intuition is very hard to obtain. Topology allows us to reason in these situations as if the spaces were the familier two- and three-dimensional objects, but the process takes a little time to get used to.

The material presented in this chapter centers around a few fundamental topics. For example, we only introduce Hausdorff and normal spaces when separation is discussed, although the literature operates with a hierarchy of more than five distinct classes. A mildly unusual feature in the presentation is the central role played by universal nets. Admittedly they are not easy to get aquainted with, but they facilitate a number of arguments later on (giving, for example, a five-line proof of Tychonoff's theorem). Since universal nets entail the blatant use of the axiom of choice, we have included (in the regie of naive set theory) a short proof of the equivalence among the axiom of choice, Zorn's lemma, and Cantor's well-ordering principle. All other topics from set theory, like ordinal and cardinal numbers, have been banned to the exercise sections. A fate they share with a large number of interesting topological concepts.

1.1. Ordered Sets

Synopsis. The axiom of choice, Zorn's lemma, and Cantor's well-ordering principle, and their equivalence. Exercises.

1.1.1. A binary relation in a set X is just a subset R of $X \times X$. It is customary, though, to use a relation sign, such as $\leq$, to indicate the relation. Thus $(x, y) \in R$ is written $x \leq y$.

An *order* in X is a binary relation, written $\leq$, which is *transitive* ($x \leq y$ and $y \leq z$ implies that $x \leq z$), *reflexive* ($x \leq x$ for every x), and *antisymmetric* ($x \leq y$ and $y \leq x$ implies $x = y$). We say that $(X, \leq)$ is an *ordered set*. Without the antisymmetry condition we have a *preorder*, and much of what follows will make sense also for preordered sets.

An element x is called a *majorant* for a subset Y of X, if $y \leq x$ for every y in Y. *Minorants* are defined analogously. We say that an order is *filtering upward*, if every pair in X (and, hence, every finite subset of X) has a majorant. Orders that are *filtering downward* are defined analogously. If a pair x, y in X has a smallest majorant, relative to the order $\leq$, this element is denoted $x \vee y$. Analogously, $x \wedge y$ denotes the largest minorant of the pair x, y, if it exists. We say that $(X, \leq)$ is a *lattice*, if $x \vee y$ and $x \wedge y$ exist for every pair x, y in X. Furthermore, $(X, \leq)$ is said to be *totally ordered* if either $x \leq y$ or $y \leq x$ for every pair x, y in X. Finally, we say that $(X, \leq)$ is *well-ordered* if every nonempty subset Y of X has a smallest element (a minorant for Y belonging to Y). This element we call the first element in Y.

Note that a well-ordered set is totally ordered (put $Y = \{x, y\}$), that a totally ordered set is a (trivial) lattice, and that a lattice order is both upward and downward filtering. Note also that to each order $\leq$ corresponds a reverse order $\geq$, defined by $x \geq y$ iff $y \leq x$.

1.1.2. Examples of orderings are found in the number systems, with their usual orders. Thus, the set $\mathbf{N}$ of natural numbers is an example of a well-ordered set. (Apart from simple repetitions, $\mathbf{N} \cup \mathbf{N} \cup \cdots$, this is also the only concrete example we can write down, despite 1.1.6.) The sets $\mathbf{Z}$ and $\mathbf{R}$ are totally ordered, but not well-ordered. The sets $\mathbf{Z} \times \mathbf{Z}$ and $\mathbf{R} \times \mathbf{R}$ are lattices, but not totally ordered, when we use the *product order*, i.e. $(x_1, x_2) \leq (y_1, y_2)$ whenever $x_1 \leq y_1$ and $x_2 \leq y_2$. [If, instead, we use the *lexicographic order*, i.e. $(x_1, x_2) \leq (y_1, y_2)$ if either $x_1 < y_1$, or $x_1 = y_1$ and $x_2 \leq y_2$, then the sets become totally ordered.]

An important order on the system $\mathscr{S}(X)$ of subsets of a given set X is given by inclusion; thus $A \leq B$ if $A \subset B$. The inclusion order turns $\mathscr{S}(X)$ into a lattice with $\emptyset$ as first and X as last elements. In applications it is usually the reverse inclusion order that is used, i.e. $A \leq B$ if $A \supset B$. For example, taking X to be a sequence (x_n) of real numbers converging to some x, and putting $T_n = \{x_k | k \geq n\}$, then clearly it is the reverse inclusion order on the tails T_n that describe the convergence of (x_n) to x.

1.1.3. The *axiom of choice*, formulated by Zermelo in 1904, states that for each nonempty set X there is a (choice) function

$$c: \mathscr{S}(X) \backslash \{\emptyset\} \to X,$$

satisfying $c(Y) \in Y$ for every Y in $\mathscr{S}(X) \backslash \{\emptyset\}$.

Using this axiom Zermelo was able to give a satisfactory proof of Cantor's *well-ordering principle*, which says that every set X has an order $\leq$, such that $(X, \leq)$ is well-ordered.

The well-ordering principle is a necessary tool in proofs "by induction," when the set over which we induce is not a segment of $\mathbb{N}$ (so-called *transfinite induction*). More recently, these proofs have been replaced by variations that pass through the following axiom, known in the literature as Zorn's lemma (Zorn 1935, but used by Kuratowski in 1922). Let us say that $(X, \leq)$ is *inductively ordered* if each totally ordered subset of X (in the order induced from X), has a majorant in X. *Zorn's lemma* then states that every inductively ordered set has a maximal element (i.e. an element with no proper majorants).

1.1.4. Let $(X, \leq)$ be an ordered set and assume that c is a choice function for X. For any subset Y of X, let $\text{maj}(Y)$ and $\min(Y)$, respectively, denote the sets of proper majorants and minorants for Y in X. Thus $x \in \text{maj}(Y)$ if $y < x$ for every y in Y, where the symbol $y < x$ of course means $y \leq x$ and $y \neq x$.

A subset C of X is called a *chain* if it is well-ordered (relative to $\leq$) and if for each x in C we have

$$c(\text{maj}(C \cap \min\{x\})) = x. \qquad (*)$$

Note that $c(X)$ is the first element in any chain and that $\{c(X)\}$ is a chain (though short).

1.1.5. Lemma. *If C_1 and C_2 are chains in X such that $C_1 \not\subset C_2$, there is an element x_1 in C_1 such that*

$$C_2 = C_1 \cap \min\{x_1\}.$$

PROOF. Since $C_1 \setminus C_2 \neq \emptyset$ and C_1 is well-ordered, there is a first element x_1 in $C_1 \setminus C_2$. By definition we therefore have

$$\text{(i)} \qquad C_1 \cap \min\{x_1\} \subset C_2.$$

If the inclusion in (i) is proper, the set $C_2 \setminus (C_1 \cap \min\{x_1\})$ has a first element x_2, since C_2 is well-ordered. By definition, therefore,

$$\text{(ii)} \qquad C_2 \cap \min\{x_2\} \subset C_1 \cap \min\{x_1\}.$$

If the inclusion in (ii) is proper, the set $(C_1 \cap \min\{x_1\}) \setminus \min\{x_2\}$ (contained in $C_1 \cap C_2$) has a first element y. By definition

$$\text{(iii)} \qquad C_1 \cap \min\{y\} \subset C_2 \cap \min\{x_2\}.$$

However, if $y \leq x$ for some x in $C_2 \cap \min\{x_2\}$, then $y \in C_2 \cap \min\{x_2\}$, contradicting the choice of y. Since both x and y belong to the well-ordered, hence totally ordered, set C_2, it follows that $x < y$ for every x in $C_2 \cap \min\{x_2\}$. Thus in (iii) we actually have equality. Since both C_1 and C_2 are chains (relative to the same ordering and the same choice function), it follows from the chain condition $(*)$ in 1.1.4 that $y = x_2$. But $y \in C_1 \cap \min\{x_1\}$ while

$x_2 \notin C_1 \cap \min\{x_1\}$. To avoid a contradiction we must have equality in (ii). Applying the chain condition to (ii) gives $x_1 = x_2$ in contradiction with $x_1 \notin C_2$ and $x_2 \in C_2$. Consequently, we have equality in (i), which is the desired result. □

1.1.6. Theorem. *The following three propositions are equivalent:*

(i) *The axiom of choice.*
(ii) *Zorn's lemma.*
(iii) *The well-ordering principle.*

Proof (i) ⇒ (ii). Suppose that $(X, \leq)$ is inductively ordered, and by assumption let c be a choice function for X. Consider the set $\{C_j | j \in J\}$ of all chains in X and put $C = \bigcup C_j$. We claim that for any x in C_j we have

$$C \cap \min\{x\} = C_j \cap \min\{x\}. \qquad (**)$$

For if y belongs to the first (obviously larger) set, then $y \in C_i$ for some i in J. Either $C_i \subset C_j$, in which case $y \in C_j$, or $C_i \not\subset C_j$. In that case there is by 1.1.5 an x_i in C_i such that $C_j = C_i \cap \min\{x_i\}$. As $y < x < x_i$, we again see that $y \in C_j$.

It now follows easily that C is well-ordered. For if $\emptyset \neq Y \subset C$, there is a j in J with $C_j \cap Y \neq \emptyset$. Taking y to be the first element in $C_j \cap Y$ it follows from (**) that y is the first element in all of Y. Condition (**) also immediately shows that C satisfies the chain condition (*) in 1.1.4. Thus C is a chain, and it is clearly the longest possible. Therefore, $\text{maj}(C) = \emptyset$. Otherwise we could take

$$x_0 = c(\text{maj}(C)) \in \text{maj}(C),$$

and then $C \cup \{x_0\}$ would be a chain [(*) in 1.1.4 has just been satisfied for x_0] effectively longer than C.

Since the order is inductive, the set C has a majorant x_ω in X. Since $\text{maj}(C) = \emptyset$, we must have $x_\omega \in C$, i.e. x_ω is the largest element in C. But then x_ω is a maximal element in X, because any proper majorant for x_ω would belong to $\text{maj}(C)$.

(ii) ⇒ (iii). Given a set X consider the system M of well-ordered, nonempty subsets $(C_j, \leq_j)$ of X. Note that $M \neq \emptyset$, the one-point sets are trivial members. We define an order $\leq$ on M by setting $(C_i, \leq_i) \leq (C_j, \leq_j)$ if either $C_i = C_j$ and $\leq_i = \leq_j$, or if there is an x_j in C_j such that

$$C_i = \{x \in C_j | x \leq_j x_j\} \quad \text{and} \quad \leq_i = \leq_j | C_i. \qquad (***)$$

The claim now is that $(M, \leq)$ is inductively ordered. To prove this, let N be a totally ordered subset of M and let C be the union of all C_j in N. Define $\leq$ on C by $x \leq y$ whenever $\{x, y\} \subset C_j \in N$ and $x \leq_j y$. Note that if $\{x, y\} \subset C_i \in N$, then $x \leq_i y$ iff $x \leq_j y$ because of the total ordering of N, so that $\leq$ is a well-defined order on C. Exactly as in the proof of (i) ⇒ (ii) one shows that if $x \in C_j$, then

$$C \cap \min\{x\} = C_j \cap \min\{x\} \qquad (**)$$

(the result of 1.1.5 has been built into the order on M). As before, this implies that $(C, \leq)$ is well-ordered. The conclusion that $(C, \leq)$ is a majorant for N is trivial if N has a largest element (which then must be C). Otherwise, each $(C_i, \leq_i)$ has a majorant $(C_j, \leq_j)$ in N and is thus of the form (***) relative to C_j; and, as (**) shows, also of the form (***) relative to C. We conclude that $(C, \leq)$ is a majorant for N, which proves that M is inductively ordered.

Condition (ii) now implies that M has a maximal element $(X_\omega, \leq_\omega)$. If $X_\omega \neq X$, we choose some x_ω in $X \setminus X_\omega$ and extend $\leq_\omega$ to $X_\omega \cup \{x_\omega\}$ by setting $x \leq_\omega x_\omega$ for every x in X_ω. This gives a well-ordered set $(X_\omega \cup \{x_\omega\}, \leq_\omega)$ that majorizes $(X_\omega, \leq_\omega)$ in the ordering in M, contradicting the maximality of $(X_\omega, \leq_\omega)$. Thus $X = X_\omega$ and is consequently well-ordered.

(iii) $\Rightarrow$ (i). Given a nonempty set X, choose a well-order $\leq$ on it. Now define $c(Y)$ to be the first element in Y for every nonempty subset Y of X. □

1.1.7. Remark. The subsequent presentation in this book builds on the acceptance of the axiom of choice and its equivalent forms given in 1.1.6. In the intuitive treatment of set theory used here, according to which a set is a properly determined collection of elements, it is not possible precisely to explain the role of the axiom of choice. For this we would need an axiomatic description of set theory, first given by Zermelo and Fraenkel. In 1938 Gödel showed that if the Zermelo–Fraenkel system of axioms is consistent (that in itself an unsolved question), then the axiom of choice may be added without violating consistency. In 1963 Cohen showed further that the axiom of choice is independent of the Zermelo–Fraenkel axioms. This means that our acceptance of the axiom of choice determines what sort of mathematics we want to create, and it may in the end affect our mathematical description of physical realities. The same is true (albeit on a smaller scale) with the parallel axiom in euclidean geometry. But as the advocates of the axiom of choice, among them Hilbert and von Neumann, point out, several key results in modern mathematical analysis [e.g. the Tychohoff theorem (1.6.10), the Hahn–Banach theorem (2.3.3), the Krein–Milman theorem (2.5.4), and Gelfand theory (4.2.3)] depend crucially on the axiom of choice. Rejecting it, one therefore loses a substantial part of mathematics, and, more important, there seems to be no compensation for the abstinence.

Exercises

E 1.1.1. A subset $\mathfrak{R}$ of a real vector space $\mathfrak{X}$ is called a cone if $\mathfrak{R} + \mathfrak{R} \subset \mathfrak{R}$ and $\mathbb{R}_+\mathfrak{R} = \mathfrak{R}$. If in addition $-\mathfrak{R} \cap \mathfrak{R} = \{0\}$ and $\mathfrak{R} - \mathfrak{R} = \mathfrak{X}$, we say that $\mathfrak{R}$ generates $\mathfrak{X}$. Show that the relation in $\mathfrak{X}$ defined by $x \leq y$ if $y - x \in \mathfrak{R}$ is an order on $\mathfrak{X}$ if $\mathfrak{R}$ is a generating cone. Find the set $\{x \in \mathfrak{X} \mid x \geq 0\}$, and discuss the relations between the order and the vector space structure. Find the condition on $\mathfrak{R}$ that makes the order total. Describe some cones in $\mathbb{R}^n$ for $n = 1, 2, 3$.