NORTH-HOLLAND
MATHEMATICAL LIBRARY

Theory of Linear Operations

S. BANACH †

North-Holland

Theory of Linear Operations

S. BANACH †

English translation by

F. JELLETT
London, United Kingdom

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD · TOKYO

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

This volume is a translation of:
Théorie des Operations Linéaires
© PWN- Państwowe Wydawnictwo Naukowe, Warsaw, Poland, 1979
and includes comments by:
A. Pelczyński and Cz. Bessaga *under the title* "Some Aspects of the Present Theory of Banach Spaces"

First edition 1987
Second impression 2000

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

ISBN: 0 444 70184 2

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

THEORY OF LINEAR OPERATIONS

North-Holland Mathematical Library

VOLUME 38

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD · TOKYO

Preface

The theory of operators, created by V. Volterra, has as its object the study of functions defined on infinite-dimensional spaces. This theory has penetrated several highly important areas of mathematics in an essential way: suffice it to recall that the theory of integral equations and the calculus of variations are included as special cases within the main areas of the general theory of operators. In this theory the methods of classical mathematics are seen to combine with modern methods in a remarkably effective and quite harmonious way. The theory often makes possible altogether unforeseen interpretations of the theorems of set theory or topology. Thus, for example, the topological theorem on fixed points may be translated, thanks to the theory of operators (as has been shown by Birkhoff and Kellogg) into the classical theorem on the existence of solutions of differential equations. There are important parts of mathematics which cannot be understood in depth without the help of the theory of operators. Contemporary examples are: the theory of functions of a real variable, integral equations, the calculus of variations, etc.

This theory, therefore, well deserves, for its aesthetic value as much as for the scope of its arguments (even ignoring its numerous applications) the interest that it is attracting from more and more mathematicians. The opinion of J. Hadamard, who considers the theory of operators one of the most powerful methods of contemporary research in mathematics, should come as no surprise.

The present book contains the basics of the algebra of operators. It is devoted to the study of so-called *linear* operators, which corresponds to that of the linear forms $a_1x_1 + a_2x_2 + \dots + a_nx_n$ of algebra.

The notion of linear operator can be defined as follows. Let E and E_1 be two abstract sets, each endowed with an associative addition operation as well as a zero element. Let $y = U(x)$ be a function (operator, transformation) under which an element y of E_1 corresponds to each element x of E (in the special case where E_1 is the space of real numbers, this function is also known as a *functional*). If, for any x_1 and x_2 of E, we have $U(x_1 + x_2) = U(x_1) + U(x_2)$, the operator U is said to be *additive*. If, in addition, E and E_1 are *metric* spaces, that is to say that in each space the *distance* between pairs of elements is defined, one can consider *continuous* operators U. Now operators which are both additive and continuous are called *linear*.

In this book, I have elected, above all, to gather together results concerning linear operators defined in general spaces of a certain kind, principally in the so-called *B-spaces* (i.e. *Banach spaces* [trans.]), examples of which are: the space of continuous functions, that of the p^{th}-power-summable functions, Hilbert space, etc.

I also give the interpretation of the general theorems in various mathematical areas, namely, group theory, differential equations, integral equations, equations with infinitely many unknowns, functions of a real variable, summation methods, orthogonal series, etc. It is interesting to see certain theorems giving results in such widely varying fields. Thus, for example, the theorem on the extension of an additive functional settles simultaneously the general problem of measure, the moment problem and the question of the existence of solutions of a system of linear equations in infinitely many unknowns.

Along with algebraic tools, the methods are principally those of general set theory, which in this book are to the fore in gaining, for this theory, several new applications. Also to be found in various chapters of this book are some new general theorems. In particular, in the last two chapters and the appendix: no part of the results included therein has been published before. They constitute an outline of the study of invariants with respect to linear transformations (of B-spaces). In particular, Chapter XII includes the definition and analysis of the properties of *linear dimension*, which in these spaces plays a rôle analogous to that of dimension in the usual sense in euclidean spaces.

Results and problems, which, for want of space, have not been considered, are discussed briefly in the Remarks at the end of the book. Some further references are also to be found there. In general, except in the Introduction or, rather, its accompanying Remarks at the end of the book, I do not indicate the origin of theorems which either I consider too elementary or else are proved here for the first time.

Some more recent work has appeared and continues to appear in the periodical *Studia Mathematica*, whose primary purpose is to present research in the area of functional analysis and its applications.

I intend to devote a second book, which will be the sequel to the present work, to the theory of other kinds of functional operators, using topological methods extensively.

In conclusion, I would like to express my sincere gratitude to all those who have assisted me in my work, in undertaking the translation of my Polish manuscript, or helping me in my labours with their valuable advice. Most particularly, I thank H. Auerbach for his collaboration in the writing of the Introduction and S. Mazur for his general assistance as well as for his part in the drafting of the final remarks.

Stefan Banach

Lwów, July 1932

Contents

Some aspects of the present theory of Banach spaces

Introduction

A. THE LEBESGUE - STIELTJES INTEGRAL

We assume the reader is familiar with measure theory and the Lebesgue integral.

§1. Some theorems in the theory of the Lebesgue integral.

If the measurable functions $x_n(t)$ form a (uniformly) bounded sequence and the sequence $(x_n(t))$ converges almost everywhere in a closed interval $[a,b]$ to the function $x(t)$, then

(1) $$\lim_{n\to\infty} \int_a^b x_n(t)\,dt = \int_a^b x(t)\,dt.$$

More generally, if there exists a summable function $\phi(t) \geq 0$ such that $|x_n(t)| \leq \phi(t)$ for $n=1,2,\ldots$, the limit function is also summable and (1) is still satisfied.

If the functions $x_n(t)$ are summable in $[a,b]$ and form a non-decreasing sequence which converges to the function $x(t)$, then (1) holds, when the function $x(t)$ is summable, and

$$\lim_{n\to\infty} \int_a^b x_n(t)\,dt = +\infty$$

otherwise.

If the sequence $(x_n(t))$ of p^{th}-power summable functions ($p \geq 1$) converges almost everywhere to the function $x(t)$ and if

$$\int_a^b |x_n(t)|^p\,dt < K \quad \text{for } n=1,2,\ldots,$$

the function $x(t)$ is also p^{th}-power summable.

§2. Some inequalities for p^{th}-power summable functions.

The class of functions which are p^{th}-power summable ($p > 1$) in $[a,b]$ will be denoted by L^p. To the number p, there corresponds the number q, connected with p by the equation $\frac{1}{p} + \frac{1}{q} = 1$, and known as the *conjugate exponent of* p. For $p = 2$, we have equally $q = 2$.

If $x(t) \in L^p$ and $y(t) \in L^q$, the function $x(t)y(t)$ is summable and its integral obeys the inequality

$$\left|\int_a^b xy\,dt\right| \leq \left(\int_a^b |x|^p\,dt\right)^{\frac{1}{p}} \left(\int_a^b |y|^q\,dt\right)^{\frac{1}{q}} .$$

In particular, we therefore have for $p = 2$:

$$\left|\int_a^b xy\,dt\right| \leq \left(\int_a^b x^2\,dt\right)^{\frac{1}{2}} . \left(\int_a^b y^2\,dt\right)^{\frac{1}{2}} .$$

If the functions $x(t)$ and $y(t)$ belong to L^p, so does the function $x(t)+y(t)$ and we have:

$$\left(\int_a^b |x+y|^p dt\right)^{\frac{1}{p}} \leq \left(\int_a^b |x|^p dt\right)^{\frac{1}{p}} + \left(\int_a^b |y|^p dt\right)^{\frac{1}{p}}$$

These inequalities are analogues of the following arithmetic inequalities:

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |b_i|^q\right)^{\frac{1}{q}},$$

$$\left(\sum_{i=1}^{n} |a_i+b_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |b_i|^p\right)^{\frac{1}{p}},$$

of which the first yields, for $p=2$, the well-known Schwarz inequality:

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leq \left(\sum_{i=1}^{n} a_i^2\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n} b_i^2\right)^{\frac{1}{2}}.$$

For every p^{th}-power summable function $(p \geq 1)$ and every $\varepsilon > 0$ there exists a continuous function $\phi(t)$ such that

$$\int_a^b |x-\phi|^p < \varepsilon.$$

§3. Asymptotic convergence.

The sequence $(x_n(t))$ of measurable functions defined on some set is said to be *asymptotically convergent* (or *convergent in measure*) to the function $x(t)$ defined on the same set, if for each $\varepsilon > 0$

$$\lim_{n\to\infty} m\left(\{t: |x_n(t) - x(t)| > \varepsilon\}\right) = 0,$$

where $m(A)$ stands for the (Lebesgue) measure of the set A.

A sequence $(x_n(t))$ which is asymptotically convergent to the function $x(t)$ always has a subsequence which converges pointwise to this function almost everywhere.

For a sequence $(x_n(t))$ to be asymptotically convergent, it is necessary and sufficient that, for each $\varepsilon > 0$,

$$\lim_{i,k\to\infty} m\left(\{t: |x_i(t) - x_k(t)| > \varepsilon\}\right) = 0.$$

§4. Mean convergence.

A sequence $(x_n(t))$ of p^{th}- power summable functions $(p \geq 1)$ in $[a,b]$ is said to be p^{th} - power *mean convergent* to the p^{th} - power summable function $x(t)$ if

$$\lim_{n\to\infty} \int_a^b |x_n(t) - x(t)|^p dt = 0.$$

A necessary and sufficient condition for such a function $x(t)$ to exist is that

$$\lim_{i,k\to\infty} \int_a^b |x_i(t) - x_k(t)|^p dt = 0.$$

The function $x(t)$ is then uniquely defined in $[a,b]$, up to a set of measure zero.

A sequence of functions which converges in mean to a function $x(t)$ is also asymptotically convergent to this function and therefore (c.f. §3) has a subsequence which converges pointwise to the same function almost everywhere.

§5. The Stieltjes Integral.

Let $x(t)$ be a continuous function and $\alpha(t)$ a function of bounded variation in $[a,b]$. By taking a partition of the interval $[a,b]$ into subintervals, using the numbers

$$a = t_0 < t_1 < t_2 < \dots < t_n = b$$

and choosing an arbitrary number θ_i in each of these subintervals, we can, by analogy with the definition of the Riemann integral, form the sum

$$S = \sum_{i=1}^{n} x(\theta_i)[\alpha(t_i) - \alpha(t_{i-1})] \qquad \text{where } t_i \geq \theta_i \geq t_{i-1}.$$

One shows that for every sequence of subdivisions, for which the length of the largest subinterval tends to 0, the sums S converge to a limit which is the same for all such sequences; this limit is denoted by

$$\int_a^b x(t)\,d\alpha(t)$$

and is called a *Stieltjes integral*.

This integral has the following properties:

$$\int_a^b x(t)\,d\alpha(t) = -\int_b^a x(t)\,d\alpha(t),$$

$$\int_a^b x(t)\,d\alpha(t) + \int_b^c x(t)\,d\alpha(t) = \int_a^c x(t)\,d\alpha(t),$$

$$\int_a^b [x_1(t) + x_2(t)]\,d\alpha(t) = \int_a^b x_1(t)\,d\alpha(t) + \int_a^b x_2(t)\,d\alpha(t).$$

The first mean value theorem here takes the form of the inequality

$$\left|\int_a^b x(t)\,d\alpha(t)\right| \leq MV,$$

where M denotes the supremum of the absolute value $|x(t)|$ and V the total variation of the function $\alpha(t)$ in $[a,b]$.

If the function $\alpha(t)$ is absolutely continuous, the Stieltjes integral can be expressed as a Lebesgue integral as follows:

$$\int_a^b x(t)\,d\alpha(t) = \int_a^b x(t)\alpha'(t)\,dt.$$

If $\alpha(t)$ is an increasing function (i.e. $\alpha(t') < \alpha(t'')$ whenever $a \leq t' < t'' \leq b$) and if, for each number $s \in [\alpha(a), \alpha(b)]$, one puts

$$\beta(s) = \sup(\{t : s \geq \alpha(t)\}),$$

one obtains:

$$(2) \qquad \int_a^b x(t)\,d\alpha(t) = \int_{\alpha(a)}^{\alpha(b)} x[\beta(s)]\,ds.$$

Proof. We have, by definition of $\beta(s)$:

$$(3) \qquad \beta[\alpha(t)] = t \quad \text{for } a \leq t \leq b.$$

Since $\beta(s)$ is increasing, by hypothesis, and takes all values in the interval $[a,b]$ where, by (3), $a = \beta[\alpha(a)]$ and $b = \beta[\alpha(b)]$, it is a continuous function. It follows that the function $x[\beta(s)]$ is continuous as well.

Consider a subdivision δ of $[a,b]$ given by the numbers $a = t_0 < t_1 < \dots < t_n = b$ and put $\alpha(t_i) = \theta_i$ for $i=1,2,\dots,n$. We have

$$I_i = \int_{\theta_{i-1}}^{\theta_i} x[\beta(s)]\,ds = (\theta_i - \theta_{i-1})\, x(\theta_i'),$$

where $\theta_i' = \beta(s_i')$ and $\theta_{i-1} \leq \theta_i' \leq \theta_i$. Clearly $\beta(\theta_{i-1}) \leq \beta(s_i') = \theta_i' \leq \beta(\theta_i)$. By (3) we have $\beta(\theta_{i-1}) = \beta[\alpha(t_{i-1})] = t_{i-1}$ and similarly $\beta(\theta_i) = t_i$. Consequently

$$t_{i-1} \leq \theta_i' \leq t_i,$$

so that

$$I_i = x(\theta_i')\,[\alpha(t_i) - \alpha(t_{i-1})],$$

whence

$$(4) \qquad \int_{\alpha(a)}^{\alpha(b)} x[\beta(s)]\,ds = \sum_{i=1}^{n} I_i = \sum_{i=1}^{n} x(\theta_i')\,[\alpha(t_i) - \alpha(t_{i-1})].$$

Now, since this last sum tends to $\int_a^b x(t)\,d\alpha(t)$ when the maximum length of the intervals of the subdivision δ tends to 0, the equality (4) yields (2), q.e.d.

This established, we now allow $\alpha(t)$ to be any function of bounded variation. Such a function $\alpha(t)$ can always be written as a difference $\alpha_1(t) - \alpha_2(t)$ of two increasing functions $\alpha_1(t)$ and $\alpha_2(t)$; denoting as before the corresponding functions by $\beta_1(s)$ and $\beta_2(s)$, we obtain

$$\int_a^b x(t)\,d\alpha(t) = \int_a^b x(t)\,d\alpha_1(t) - \int_a^b x(t)\,d\alpha_2(t) = \int_{\alpha_2(a)}^{\alpha_1(b)} x[\beta_1(s)]\,ds - \int_{\alpha_2(a)}^{\alpha_2(b)} x[\beta_2(s)]\,ds.$$

If the functions $x_n(t)$ are continuous and uniformly bounded and if the sequence $(x_n(t))$ converges everywhere (pointwise) to a continuous function $x(t)$, we have, for every function $\alpha(t)$ of bounded variation

$$\lim_{n\to\infty} \int_a^b x_n(t)\,d\alpha(t) = \int_a^b x(t)\,d\alpha(t),$$

because

$$\lim_{n\to\infty} \int_{\alpha_1(a)}^{\alpha_1(b)} x_n[\beta_1(s)]\,ds = \int_{\alpha_2(a)}^{\alpha_1(b)} x[\beta_1(s)]\,ds,$$

and

$$\lim_{n\to\infty} \int_{\alpha_2(a)}^{\alpha_2(b)} x_n[\beta_2(s)]\,ds = \int_{\alpha_2(a)}^{\alpha_2(b)} x[\beta_2(s)]\,ds.$$

§6. Lebesgue's theorem.

Let us note the following theorem, due to H. Lebesgue (*Annales de Toulouse* 1909).

For a sequence $(x_n(t))$ *of summable functions over* $[0,1]$ *to satisfy*

$$\lim_{n\to\infty} \int_0^1 \alpha(t) x_n(t)\,dt = 0$$

for every bounded measurable function $\alpha(t)$ *on* $[0,1]$, *it is necessary and sufficient that the following three conditions be simultaneously satisfied:*

1° *the sequence* $\left(\int_0^1 |x_n(t)|\,dt\right)$ *is bounded,*

2° *for every* $\varepsilon > 0$ *there exists an* $\eta > 0$ *such that for every subset* H *of* $[0,1]$ *of measure* $<\eta$, *the inequality* $\left|\int_H x_n(t)\,dt\right| \leqq \varepsilon$ *holds for* $n=1,2,\ldots,$

3° $\lim_{n\to\infty} \int_0^u x_n(t)\,dt = 0$ *for every* $0 \leqq u \leqq 1$.

We shall become acquainted with other theorems of this kind later in the book.

B. (B)-MEASURABLE SETS AND OPERATORS IN METRIC SPACES.

§7. Metric spaces

A non-empty set E is called a *metric space* or D-space when to each ordered pair (x,y) of its elements there corresponds a number $d(x,y)$ satisfying the conditions:

1) $d(x,x) = 0$, $d(x,y) > 0$ when $x \neq y$,

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

3) $d(x,z) \leqq d(x,y) + d(y,z)$.

The function d is called a *metric* and the number $d(x,y)$ is called the *distance* between the *points* (elements) x,y. A sequence of points (x_n) is said to be *convergent*, when

$$\lim_{p,q\to\infty} d(x_p,x_q) = 0; \tag{5}$$

the sequence (x_n) is said to be *convergent to the point* x_0, and we write $\lim_{n\to\infty} x_n = x_0$, when

$$\lim_{n\to\infty} d(x_n,x_0) = 0. \tag{6}$$

The point x_0 is then known as the *limit* of the sequence (x_n).

Remark. Sequences which are *convergent* in this sense are more usually known as *Cauchy* sequences. [Trans.]

It is easy to see that (6) implies (5), since we always have

$$d(x_p,x_q) \leqq d(x_p,x_0) + d(x_0,x_q).$$

Consequently, a sequence convergent to a point is convergent for this reason; of course, the converse is not always true.

A metric space with the property that every convergent sequence in it converges to some point is said to be *complete*.

A metric space with the property that every (infinite) sequence of its points has a subsequence convergent to some point is said to be *compact*.

The euclidean spaces constitute examples of complete metric spaces. We shall now describe some other important examples.

1. *The set S of measurable functions* in the interval [0,1]. For each ordered pair (x,y) of elements of this set, put

$$d(x,y) = \int_0^1 \frac{|x(t) - y(t)|}{1 + |x(t) - y(t)|}\, dt.$$

It is easily verified that conditions 1) - 3) above are satisfied. In fact, it is clear that conditions 1) and 2) are satisfied, (we do not distinguish between functions which only differ on a set of measure zero) and to see that condition 3) also holds, it is enough to remark that for every pair of real numbers a,b one has:

$$\frac{|a+b|}{1 + |a+b|} \leq \frac{|a|}{1 + |a|} + \frac{|b|}{1 + |b|}\,.$$

Thus "metrised", the set S therefore becomes a metric space; this space is complete, since convergence of a sequence (x_n) of its points (to a point x_0) means convergence in measure of the sequence of functions $\left(x_n(t)\right)$ (to the function $x_0(t)$) in [0,1].

2. *The set s of all sequences of numbers.* For each ordered pair (x,y) of its elements, put

$$d(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{|\xi_n - \eta_n|}{(1 + |\xi_n - \eta_n|)}\,,$$

where, as in all the examples of sequence spaces, $x = (\xi_n)$ and $y = (\eta_n)$.

The set s then becomes a complete metric space. In fact, convergence of a sequence of points (x_m) and its convergence to a point x_0 here mean (putting $x_m = \left(\xi_n^{(m)}\right)$ and $x_0 = (\xi_n)$) that for each natural number n, each of the sequences $\left(\xi_n^{(m)}\right)$ is convergent, and is convergent to ξ_n, respectively, as m tends to infinity.

3. *The set M of bounded measurable functions* in [0,1]. If one puts, for each pair x,y of its elements

$$d(x,y) = \operatorname*{ess\,sup}_{0\leq t\leq 1} |x(t) - y(t)|,$$

one obtains a complete metric space. Convergence of a sequence of points (x_n) (to a point x_0, respectively) here means uniform convergence almost everywhere in [0,1] of the sequence of functions $\left(x_n(t)\right)$ (to the function $x_0(t)$).

4. *The set m of bounded sequences of numbers.* Putting

$$d(x,y) = \sup_{1\leq n} |\xi_n - \eta_n|$$

one clearly obtains from m a complete metric space.

5. *The set C of continuous functions* in [0,1]. For each pair x,y of its elements put

$$d(x,y) = \max_{0\leq t\leq 1} |x(t) - y(t)|.$$