Dolciani Mathematical Expositions #49
MAA Guides #9

A Guide to Functional Analysis

Steven G. Krantz

MAA
Mathematical Association of America

A Guide to Functional Analysis

Library of Congress Catalog Card Number 2013935093

Print Edition ISBN 978-0-88385-357-3

Electronic Edition ISBN 978-1-61444-213-4

Printed in the United States of America

Current Printing (last digit):
10 9 8 7 6 5 4 3 2 1

The Dolciani Mathematical Expositions
NUMBER FORTY-NINE

MAA Guides # 49

A Guide to Functional Analysis

Steven G. Krantz
Washington University in St. Louis

Published and Distributed by
The Mathematical Association of America

The DOLCIANI MATHEMATICAL EXPOSITIONS series of the Mathematical Association of America was established through a generous gift to the Association from Mary P. Dolciani, Professor of Mathematics at Hunter College of the City University of New York. In making the gift, Professor Dolciani, herself an exceptionally talented and successful expositor of mathematics, had the purpose of furthering the ideal of excellence in mathematical exposition.

The Association, for its part, was delighted to accept the gracious gesture initiating the revolving fund for this series from one who has served the Association with distinction, both as a member of the Committee on Publications and as a member of the Board of Governors. It was with genuine pleasure that the Board chose to name the series in her honor.

The books in the series are selected for their lucid expository style and stimulating mathematical content. Typically, they contain an ample supply of exercises, many with accompanying solutions. They are intended to be sufficiently elementary for the undergraduate and even the mathematically inclined high-school student to understand and enjoy, but also to be interesting and sometimes challenging to the more advanced mathematician.

1. *Mathematical Gems,* Ross Honsberger
2. *Mathematical Gems II,* Ross Honsberger
3. *Mathematical Morsels,* Ross Honsberger
4. *Mathematical Plums,* Ross Honsberger (ed.)
5. *Great Moments in Mathematics* (*Before 1650*), Howard Eves
6. *Maxima and Minima without Calculus,* Ivan Niven
7. *Great Moments in Mathematics* (*After 1650*), Howard Eves
8. *Map Coloring, Polyhedra, and the Four-Color Problem,* David Barnette
9. *Mathematical Gems III,* Ross Honsberger
10. *More Mathematical Morsels,* Ross Honsberger
11. *Old and New Unsolved Problems in Plane Geometry and Number Theory,* Victor Klee and Stan Wagon
12. *Problems for Mathematicians, Young and Old,* Paul R. Halmos
13. *Excursions in Calculus: An Interplay of the Continuous and the Discrete,* Robert M. Young
14. *The Wohascum County Problem Book,* George T. Gilbert, Mark Krusemeyer, and Loren C. Larson
15. *Lion Hunting and Other Mathematical Pursuits: A Collection of Mathematics, Verse, and Stories by Ralph P. Boas, Jr.,* edited by Gerald L. Alexanderson and Dale H. Mugler
16. *Linear Algebra Problem Book,* Paul R. Halmos
17. *From Erdős to Kiev: Problems of Olympiad Caliber,* Ross Honsberger
18. *Which Way Did the Bicycle Go? ... and Other Intriguing Mathematical Mysteries,* Joseph D. E. Konhauser, Dan Velleman, and Stan Wagon
19. *In Pólya's Footsteps: Miscellaneous Problems and Essays,* Ross Honsberger
20. *Diophantus and Diophantine Equations,* I. G. Bashmakova (Updated by Joseph Silverman and translated by Abe Shenitzer)
21. *Logic as Algebra,* Paul Halmos and Steven Givant
22. *Euler: The Master of Us All,* William Dunham
23. *The Beginnings and Evolution of Algebra,* I. G. Bashmakova and G. S. Smirnova (Translated by Abe Shenitzer)
24. *Mathematical Chestnuts from Around the World,* Ross Honsberger
25. *Counting on Frameworks: Mathematics to Aid the Design of Rigid Structures,* Jack E. Graver
26. *Mathematical Diamonds,* Ross Honsberger
27. *Proofs that Really Count: The Art of Combinatorial Proof,* Arthur T. Benjamin and Jennifer J. Quinn
28. *Mathematical Delights,* Ross Honsberger
29. *Conics,* Keith Kendig
30. *Hesiod's Anvil: falling and spinning through heaven and earth,* Andrew J. Simoson
31. *A Garden of Integrals,* Frank E. Burk

32. *A Guide to Complex Variables* (MAA Guides #1), Steven G. Krantz
33. *Sink or Float? Thought Problems in Math and Physics*, Keith Kendig
34. *Biscuits of Number Theory*, Arthur T. Benjamin and Ezra Brown
35. *Uncommon Mathematical Excursions: Polynomia and Related Realms*, Dan Kalman
36. *When Less is More: Visualizing Basic Inequalities*, Claudi Alsina and Roger B. Nelsen
37. *A Guide to Advanced Real Analysis* (MAA Guides #2), Gerald B. Folland
38. *A Guide to Real Variables* (MAA Guides #3), Steven G. Krantz
39. *Voltaire's Riddle: Micromégas and the measure of all things*, Andrew J. Simoson
40. *A Guide to Topology*, (MAA Guides #4), Steven G. Krantz
41. *A Guide to Elementary Number Theory*, (MAA Guides #5), Underwood Dudley
42. *Charming Proofs: A Journey into Elegant Mathematics*, Claudi Alsina and Roger B. Nelsen
43. *Mathematics and Sports*, edited by Joseph A. Gallian
44. *A Guide to Advanced Linear Algebra*, (MAA Guides #6), Steven H. Weintraub
45. *Icons of Mathematics: An Exploration of Twenty Key Images*, Claudi Alsina and Roger B. Nelsen
46. *A Guide to Plane Algebraic Curves*, (MAA Guides #7), Keith Kendig
47. *New Horizons in Geometry,* Tom M. Apostol and Mamikon A. Mnatsakanian
48. *A Guide to Groups, Rings, and Fields,* (MAA Guides #8), Fernando Q. Gouvêa
49. *A Guide to Functional Analysis,* (MAA Guides #9), Steven G. Krantz

MAA Service Center
P.O. Box 91112
Washington, DC 20090-1112
1-800-331-1MAA FAX: 1-301-206-9789

To the memory of Stefan Banach.

Contents

Preface

Functional analysis is a child of the twentieth century. Rapid developments in the theory of differential equations and especially in harmonic analysis (the theory of Fourier series) made it desirable to study entire spaces of functions. These were usually infinite dimensional spaces, which revealed new worlds of harmony and truth. Functional analysis gave analysis a new set of techniques and an entirely new way of looking at things. It created the idea of "soft" analysis (as opposed to "hard" analysis). It often was able to prove in a few lines results that were hard work to verify by classical means.

Functional analysis is abstract mathematics at its best. It requires a good deal of the reader, and particularly of the end user. It is a demanding discipline, but one which yields many fruits.

Most graduate students are required to learn some functional analysis as part of the qualifying exam system. Working analysts have to have some functional analysis under their belts. It is part of our toolkit, just as Galois theory is for the algebraist.

The creators of functional analysis are also legend. Stefan Banach and Stanislaw Ulam, to name just two, were part of the Scottish Cafe team in pre-war Poland, and they helped to set a standard for how mathematics is practiced today. A bit later, John von Neumann played a critical role in establishing the importance of Hilbert space theory, both in mathematics and in physics. Some of the most important and powerful mathematicians today are functional analysts.

The purpose of this book is to introduce the reader with minimal background to the basic scripture of functional analysis. Readers should know some real analysis and some linear algebra. Measure theory rears its ugly head in some of the examples and also in the treatment of spectral theory. The latter is unavoidable and the former allows us to present a rich variety of examples. The nervous reader may safely skip any of the measure theory and still derive a lot from the rest of the book. Apart from this caveat, the book is almost completely self-contained; in a few instances we mention easily accessible references.

A feature that sets this book apart from most other functional analysis texts is that it has a *lot* of examples and a *lot* of applications. This helps to make the material more concrete, and relates it to ideas that the reader has already seen. It also makes the book more accessible to a broader audience.

I thank Don Albers for being a worldly and gentle editor. I thank Jerry Folland for helpful comments at various junctures. And I thank the staff at the MAA for another delightful publishing experience.

St. Louis, Missouri Steven G. Krantz

CHAPTER 1

FUNDAMENTALS

1.1 WHAT IS FUNCTIONAL ANALYSIS?

The mathematical analysts of the nineteenth century (Cauchy, Riemann, Weierstrass, and others) contented themselves with studying one function at a time. As a sterling instance, the Weierstrass nowhere differentiable function is a world-changing example of the real function theory of "one function at a time." Some of Riemann's examples in Fourier analysis give other instances. This was the world view 150 years ago. To be sure, Cauchy and others considered sequences and series of functions, but the end goal was to consider the single limit function.

A major paradigm shift took place, however, in the early twentieth century. For then people began to consider *spaces* of functions. By this we mean a linear space, equipped with a norm. The process began slowly. At first people considered very specific spaces, such as the square-integrable real functions on the unit circle. Much later, people branched out to more general classes of spaces. An important feature of the spaces under study was that they must be complete. For we want to pass to limits, and completeness guarantees that this process is reliable.

Thus was born the concepts of Hilbert space and Banach space. People like to joke that, in the early 1940s, Hilbert went to one of his colleagues in Göttingen and asked, "What is a Hilbert space?" Perhaps he did. For it was a new idea at the time, and not well established. Banach spaces took even longer to catch on. But indeed they did. Later came topological vector spaces. These all proved to be powerful and flexible tools that provide new insights and new power to the study of classical analysis. They also afford a completely different point of view in the subjects of real and complex analysis. Functional analysis is a lovely instance of how mathematical abstraction enables one to see new things, and see them very clearly.

The purpose of this book is to introduce the reader to the wealth of ideas that is functional analysis. This will not be a thorough grounding, but rather a taste of what the subject is like. We shall make a special effort to provide examples and concrete applications of the abstract ideas, just so that the neophyte can get a concrete grip on the techniques. Certainly we provide references to more advanced and more comprehensive texts.

As readers work through the book, they may find it useful to refer to some of the great classic texts, such as [DUS], [RES], [RUD2], and [YOS].

1.2 Normed Linear Spaces

Let X be a collection of objects equipped with a binary operation $+$ of addition and also with a notion of scalar multiplication. Thus, if $x, y \in X$, then $x + y \in X$. Also, if $x \in X$ and $c \in \mathbb{C}$ then $cx \in X$. (The scalar field can be the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. For us it will usually be $\mathbb{C}$, but there will be exceptions. When we want to refer to the scalar field generically, we use the letter k.) We equip X with a norm; thus, if $x \in X$, then $\|x\| \in \mathbb{R}^+ \equiv \{t \in \mathbb{R} : t \geq 0\}$. We demand that

1. $\|x\| \geq 0$,
2. $\|x\| = 0$ if and only if $x = 0$,
3. If $x \in X$ and $c \in \mathbb{C}$ then $\|cx\| = |c| \cdot \|x\|$,
4. If $x, y \in X$, then $\|x + y\| \leq \|x\| + \|y\|$.

We call X a *normed linear space* (or NLS).

Notice that X as described above is naturally equipped with balls. If $x \in X$ and $r > 0$ then

$$B(x, r) = \{t \in X : \|x - t\| < r\}$$

is the (open) *ball with center x and radius r*. We may think of the collection of balls as the subbasis for a topology on X. Concomitantly, we say that a sequence $\{x_j\} \subseteq X$ *converges* to $x \in X$ if $\|x_j - x\| \to 0$ as $j \to \infty$. The sequence $\{x_j\}$ is *Cauchy* if, for any $\epsilon > 0$, there is a J so large that $j, k > J$ implies $\|x_j - x_k\| < \epsilon$.

We use the notation $\overline{B}(x, r) \equiv \{t \in X : \|x - t\| \leq r\}$ to denote the *closed ball* with center x and radius r. It is worth commenting that this closed ball is not necessarily the closure of the open ball (exercise).

DEFINITION. Let X be a normed linear space. We say that X is a *Banach space* if X is complete in the topology induced by the norm. That is to say, if $\{x_j\}$ is a Cauchy sequence in X, then there is a limit element $x \in X$ such that $x_j \to x$ as $j \to \infty$.

EXAMPLE. Let $X = \mathbb{R}^N$ equipped with the usual norm: If $x = (x_1, x_2, \dots, x_N)$ is a point of $\mathbb{R}^N$, then

$$\|x\| = \left(\sum_{j=1}^{N} x_j^2\right)^{1/2} .$$

We certainly know that this norm satisfies the axioms for a norm. It is a standard fact that $\mathbb{R}^N$, equipped with the topology coming from this norm, is complete. So $\mathbb{R}^N$ is a Banach space.

EXAMPLE. Let

$$X = \{f : f \text{ is a continuous function on the unit interval } [0, 1] \text{ with values in } \mathbb{R}\} .$$

We equip X with the norm, for $f \in X$, given by

$$\|f\| = \max_{t \in [0,1]} |f(t)| .$$

It is straightforward to verify that this norm satisfies the four axioms.

Furthermore, if $\{f_j\}$ is a Cauchy sequence in this norm, then in fact $\{f_j\}$ is *uniformly Cauchy*. It is a standard result from real analysis (see [KRA1]) that such a sequence has a continuous limit function f. Hence our space is complete. And X is therefore a Banach space. We usually denote this space by $C([0, 1])$.

EXAMPLE. Let us consider the space $X = \ell^1$ of sequences $\alpha = \{a_j\}$ of real numbers with the property that $\sum_j |a_j| < \infty$. The norm on this space is

$$\|\alpha\| \equiv \sum_j |a_j| .$$

It is easy to check the four axioms of a norm. Addition is defined componentwise, as is scalar multiplication.

If $\alpha^j = \{a_\ell^j\}_{\ell=1}^\infty$ is a Cauchy sequence of elements of X then let $\epsilon > 0$. Choose $K > 0$ such that, if $j, k > K$ then $\|\alpha^j - \alpha^k\| < \epsilon/5$. It follows for such j, k and any index ℓ that

$$|a_\ell^j - a_\ell^k| \le \|\alpha^j - \alpha^k\| < \frac{\epsilon}{5} .$$

By the completeness of the real numbers, we find for each ℓ that the sequence $\{a_\ell^j\}_{j=1}^\infty$ converges to a real limit a'_ℓ. We claim that the sequence $A \equiv \{a'_\ell\}$ lies in ℓ^1 and is the limit in norm of the original Cauchy sequence $\{\alpha^j\}$.

Choose K as above. Select L so large that $\sum_{m=L}^\infty |\alpha_m^K| < \epsilon/5$. If $n > K$ then

$$\sum_{m=L}^\infty |a_m^n| \le \sum_{m=L}^\infty |a_m^n - a_m^K| + \sum_{m=L}^\infty |a_m^K| < \frac{2\epsilon}{5}. \tag{1}$$

As a result,

$$\sum_{m=1}^\infty |a_m^n - a'_m| \le \sum_{m=1}^{L-1} |a_m^n - a'_m| + \sum_{m=L}^\infty |a_m^n| + \sum_{m=L}^\infty |a'_m| < \frac{\epsilon}{5} + \frac{2\epsilon}{5} + \frac{2\epsilon}{5} = \epsilon\,.$$

Here we use the fact that $a_m^n \to a'_m$, each m, so the first sum is less than $\epsilon/5$ if n is large enough. That the last sum does not exceed $2\epsilon/5$ follows from (1) by letting $n \to \infty$. Therefore the α^j converge to A as desired.

We see that X is complete, so it is a Banach space. We usually denote this space by ℓ^1.

EXAMPLE. It is a fact, and we shall not provide all the details here, that if $1 \le p < \infty$, then the collection of sequences $\alpha = \{a_j\}$ such that $\sum_j |a_j|^p < \infty$ forms a Banach space. The appropriate norm is

$$\|\alpha\| \equiv \left(\sum_j |a_j|^p\right)^{1/p}.$$

We usually denote this space by ℓ^p.

For $p = \infty$, the appropriate space is that of all *bounded* sequences $\alpha = \{a_j\}$ of real numbers. The right norm is

$$\|\alpha\| = \sup_j |a_j|\,.$$

We denote this space by ℓ^∞. See [RUD2] for a thorough treatment of these spaces.

As previously indicated, the balls $B(x, r)$ in a normed linear space X may be taken to be a subbasis for the topology on X.

PROPOSITION 1.1. *The topology on a normed linear space is Hausdorff.*

Proof. Let $x, y \in X$ be distinct elements. Let $\|x - y\| = \delta > 0$. Then, by the triangle inequality, the balls $B(x, \delta/3)$ and $B(y, \delta/3)$ are disjoint. Hence the space is Hausdorff. $\square$

1.3 FINITE-DIMENSIONAL SPACES

The examples of the previous section indicate that there are many interesting norms on infinite-dimensional spaces. Such is not the case in finite dimensions.

Recall that a linear space is *finite dimensional* if it has a basis with finitely many elements. It is an easy, basic result that, in this circumstance, any two bases for the space have the same number of elements N. We call N the *dimension* of the space.

PROPOSITION 1.2. *Let X be a finite-dimensional space. Then any two norms $\| \ \|_1$ and $\| \ \|_2$ on X are equivalent in the sense that there is a constant $C > 1$ so that, for any $x \in X$,*

$$\frac{1}{C} \cdot \|x\|_1 \leq \|x\|_2 \leq C \cdot \|x\|_1 .$$

Proof. The unit sphere in the norm $\| \ \|_1$ is closed and bounded (in the $\| \ \|_1$ topology), so it is a compact set. The function

$$x \longmapsto \|x\|_2$$

is a continuous, nonvanishing function on that unit sphere. So it has a positive minimum m and a positive maximum M. Thus

$$m \leq \|x\|_2 \leq M$$

for all x in the $\| \ \|_1$ unit sphere. In other words,

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

for all x in the $\| \ \|_1$ unit sphere. If x is *any element* of X then we apply the last set of inequalities to $x/\|x\|_1$. The result is

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

for all elements $x \in X$. We simply take C to be the maximum of M and $1/m$. □

REMARK. Notice that the compactness of the unit sphere in the $\| \ \|_1$ norm played a key role in the last proof. Pick one of the examples in the last section and show that the unit sphere in that infinite-dimensional space is *not* compact.

The upshot of this last proposition is that, for a given dimension N, the only normed linear space with that dimension is $\mathbb{R}^N$. This is a frequently useful observation.

An important piece of information for us is the following:

PROPOSITION 1.3. *The norm on a normed linear space X is continuous. That is to say, the function*

$$X \ni x \longmapsto \|x\|$$

is a continuous function from X to $\mathbb{R}$.

Proof. From the triangle inequality,

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\| \,.$$

That shows that the function is in fact Lipschitz. □

1.4 Linear Operators

All of modern mathematics is formulated in the language of sets and functions. In the subject of functional analysis, the most important functions are the linear operators (linear transformations).

If X and Y are normed linear spaces and

$$\Lambda : X \longrightarrow Y$$

is a function such that

$$\Lambda(c_1 x_1 + c_2 x_2) = c_1 \Lambda(x_1) + c_2 \Lambda(x_2)$$

for all $x_1, x_2 \in X$ and scalars c_1, c_2, then we call Λ a *linear operator.* In the special case that Y is the scalar field k then we call Λ a *linear functional* on X. The collection of continuous (or bounded), linear functionals on X, denoted by X^*, is a very important space in itself that carries a great deal of powerful information about X.

EXAMPLE. Let X be the space $C([0,1])$ and define the linear operator Λ by

$$C([0,1]) \ni f \longmapsto \int_0^1 f(x)\,dx \in \mathbb{R}\,.$$

This is a linear functional on X.

EXAMPLE. Let X be the space $C([0,1])$ and define the linear operator Λ by

$$C([0,1]) \ni f \longmapsto x^2 \cdot f(x) \in C([0,1])\,.$$

This is a linear operator from X to itself.

EXAMPLE. Let X be the space $C([0,1])$ and define, for $f \in X$ and $j \in \mathbb{Z}$,

$$\widehat{f}(j) = \int_0^1 f(t)e^{-2\pi ijt}\,dt\,.$$

It is easy to see that $|\widehat{f}(j)| \le \|f\|_{C([0,1])}$.

Consider the linear operator from $C([0,1])$ to ℓ^∞ given by

$$C([0,1]) \ni f \longmapsto \{\widehat{f}(j)\}_{j=-\infty}^{\infty}\,.$$

This is an important operator in Fourier analysis.

DEFINITION. Let $T : X \to Y$ be a linear operator. The *norm* of the operator T is defined to be

$$\|T\| = \|T\|_{\mathrm{op}} = \sup_{\|x\|_X \le 1} \|Tx\|_Y\,.$$

We sometimes denote this norm by $\|T\|_{X,Y}$ or $\|T\|_{\mathrm{op}}$.

In case $S : X \to \mathbb{C}$ is a linear functional, then the norm is

$$\|S\| = \sup_{\|x\|_X \le 1} |Sx|\,.$$

The most important linear operators, and particularly linear functionals, are the continuous ones. It turns out that these are particularly easy to recognize and to deal with.

THEOREM 1.4. *Let X and Y be normed linear spaces and let $T : X \to Y$ be a linear map. Then the following statements are equivalent:*

(a) *T is continuous,*

(b) *T is continuous at $\mathbf{0}$,*

(c) *T is bounded (i.e., there is a $C > 0$ so that $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$).*

Proof. That (a) implies (b) is trivial. Assume (b). Then there is a neighborhood U of $\mathbf{0}$ such that $T(U) \subseteq \{y \in Y : \|y\| \le 1\}$. Also U must contain a ball $\overline{B}(\mathbf{0},\delta)$ about $\mathbf{0}$. Hence $\|Tx\|_Y \le 1$ when $\|x\|_X \le \delta$. Since T commutes with scalar multiplication, we see that $\|Tx\|_Y \le a\delta^{-1}$ whenever $\|x\|_X \le a$. That is to say, $\|Tx\|_Y \le \delta^{-1}\|x\|_X$. So we see that (b) implies (c).

Finally, if $\|Tx\|_Y \le C\|x\|_X$ for all x, then $\|Tx_1 - Tx_2\|_Y = \|T(x_1 - x_2)\|_Y \le \epsilon$ whenever $\|x_1 - x_2\|_X \le C^{-1}\epsilon$. Hence T is continuous and (a) holds. □

1.5 The Baire Category Theorem

Recall that a *metric space* is a set E together with a distance function $d : E \times E \to \mathbb{R}^+$ satisfying these axioms:

(a) $d(x, y) = d(y, x)$ for all $x, y \in E$,

(b) $d(x, y) = 0$ if and only if $x = y$,

(c) $d(x, y) \le d(x, z) + d(z, y)$ for all $x, y, z \in E$.

We say that the metric space (E, d) is *complete* if, whenever $\{x_j\}$ is a Cauchy sequence (defined in the obvious fashion) in E, then that sequence has a limit in E. The Baire category theorem is one of the most important applications of the idea of completeness.

THEOREM 1.5. *If E is a complete metric space, then the intersection of any countable collection of dense, open subsets of E is also dense in E (it is not necessarily open).*

REMARK. The theorem is plainly false without the hypothesis of completeness. For example, let E be the rational numbers $\mathbb{Q}$ equipped with the usual Euclidean metric. Let $\{q_j\}$ be an enumeration of the rationals. And define, for $j = 1, 2, \ldots,$

$$S_j = E \setminus \{q_j\}.$$

Then each S_j is open and dense in E. But the intersection of the sets S_j is the empty set.

REMARK. The contrapositive statement of the Baire theorem is also of interest: If F_j are closed, nowhere dense sets in E, then the union of the F_j is the complement of a dense set. In particular, it cannot be the whole space.

Proof of the Baire Category Theorem. This is a quite standard argument that can be found in many texts, such as [RUD1] or [RUD3] or [KRA1].

Suppose that $V_1, V_2, \ldots$ are dense, open sets in E. Let U be *any* nonempty open set in E. The property of a set S being dense means that S has nontrivial intersection with any nonempty open set. Thus our job is to show that $(\cap_j V_j) \cap U \ne \emptyset$.

Let $B(x,r)$ be the metric balls in E. Let $\overline{B}(x,r)$ be the closed ball with center x and radius r. Since V_1 is dense in E, $U \cap V_1$ is a nonempty open set. So it contains a ball: We can find $x_1 \in E$ and $r_1 > 0$ such that

$$\overline{B}(x_1, r_1) \subseteq U \cap V_1 . \tag{2}$$

Inductively, if $n \geq 2$ and x_{n-1}, r_{n-1} have been chosen, the denseness of V_n tells us that $V_n \cap B(x_{n-1}, r_{n-1}) \neq \emptyset$. So we can find a ball

$$\overline{B}(x_n, r_n) \subseteq V_n \cap B(x_{n-1}, r_{n-1}) , \tag{3}$$

with $0 < r_n < 1/n$.

This inductive process produces a sequence $\{x_n\}$ in E. If $j, k > n$, then we see that x_j and x_k both lie in $B(x_n, r_n)$. Hence $d(x_j, x_k) < 2r_n < 2/n$. We see then that $\{x_j\}$ is a Cauchy sequence. Since E is complete, we conclude that there is a point $x \in E$ such that $x_j \to x$.

Since x_j lies in $\overline{B}(x_n, r_n)$ if $j > n$, we may conclude now that x lies in each $\overline{B}(x_n, r_n)$, and (3) shows that x lies in each V_n. Line (2) shows that x lies in U. We have shown, as advertised, that $(\cap_j V_j) \cap U \neq \emptyset$. □

A lot of terminology has grown up around the Baire theorem, and we shall not belabor it here. Suffice it to say that a set in E is called a G_δ *set* if it is the intersection of countably many open sets. So Baire's theorem says that the intersection of dense G_δs is also a dense G_δ.

Further, a set G in E is said to be of *first category* if it is the countable union of nowhere dense sets. A subset that is not of the first category is said to be of the *second category*. Baire's theorem says that no complete metric space is of the first category.

1.6 The Three Big Results

Elementary Banach space theory boils down to three big theorems. We shall enunciate them and prove them, and then give a number of remarkable examples to illustrate their importance.

In what follows we shall use the notion of semicontinuity. Recall that a function $f : X \to \mathbb{R}$ is *upper semicontinuous* if, for each real β, $\{x \in X : f(x) < \beta\}$ is open. The function is *lower semicontinuous* if, for each real α, $\{x \in X : f(x) > \alpha\}$ is open. Clearly, if f is both lower and upper semicontinuous, then f is continuous.

If $T_\alpha : X \to Y$ is a collection of operators then we say that $\{T_\alpha\}$ is *uniformly bounded* if there is a constant $M > 0$ such that $\|T_\alpha\|_{\mathrm{op}} \leq M$.

We say that the T_α *blow up* on a dense G_δ set E if $\sup_{\alpha \in A} \|T_\alpha x\| = \infty$ for every x in E.

THEOREM 1.6 (The Banach-Steinhaus Theorem). *Suppose that X is a Banach space and that Y is a normed linear space (not necessarily complete). Assume that $\{T_\alpha\} : X \to Y$ is a collection of bounded linear operators, for α in some index set A. Then either the $\{T_\alpha\}$ are uniformly bounded or else they blow up on a dense G_δ.*

REMARK. The Banach-Steinhaus theorem is sometimes also called the *uniform boundedness principle*. The uniform boundedness aspect is the first option and the denial of uniform boundedness is the second option. This result is a lovely application of the Baire category theorem.

Proof of the Banach-Steinhaus Theorem. This is a straightforward application of the Baire category theorem and elementary analysis. The proof may be found in many textbooks, including [RUD1].

Set, for $x \in X$,

$$\phi(x) = \sup_{\alpha \in A} \|T_\alpha x\| .$$

Put

$$V_n = \{x \in X : \phi(x) > n\}$$

for $n = 1, 2, 3, \ldots$. Since each T_α is continuous and since the norm function on Y is continuous, each function $x \to \|T_\alpha x\|$ is continuous on X. Thus ϕ, the supremum of such functions, is lower semicontinuous. We conclude then that each V_n is open. There are now two cases:

(a) Suppose that every V_n is dense in X. In this case, $\cap_n V_n$ is a dense G_δ in X by the Baire category theorem. Since $\phi(x) = \infty$ for every $x \in \cap_n V_n$, we see that the blow up holds.

(b) If instead some V_m fails to be dense in X, then there is a point $x_0 \in X$ and an $r > 0$ such that $\|x\| \leq r$ implies that $x_0 + x \notin V_m$ (in other words, V_m misses a ball). We see then that $\phi(x_0 + x) \leq m$, or

$$\|T_\alpha(x_0 + x)\| \leq m$$

for all $\alpha \in A$ and all x with $\|x\| \leq r$. We have

$$\|T_\alpha x\| = \|T_\alpha((x_0 + x) + (-x_0))\| \leq \|T_\alpha(x_0 + x)\| + \|T_\alpha(x_0)\| \leq 2m .$$

We conclude that uniform boundedness holds with $M = 2m/r$. $\square$

THEOREM 1.7 (The Open Mapping Principle). *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded, surjective, linear transformation. Then the image of any open set is open. That is to say, if $U \subseteq X$ is open, then $\Lambda(U) \equiv \{\Lambda u : u \in U\} \subseteq Y$ is open.*

REMARK. An immediate consequence of the theorem is that there is a $\delta > 0$ so that

$$\Lambda(B_X) \supset \delta B_Y ,$$

where B_X, B_Y are the open unit balls in X and Y respectively. In case Λ is both injective and surjective, we thus see that Λ^{-1} is continuous (see more on this point below). This is important information.

Proof of the Open Mapping Principle. See [RUD1] for the ideas behind this proof.

Continue to use the notation B_X, B_Y for the open unit balls in X and Y, respectively. Let $y \in Y$. Then there is an $x \in X$ such that $\Lambda x = y$. For any $k > 0$, if $\|x\| < k$ then $y \in \Lambda(kB_X)$. Thus Y is the union of the sets $\Lambda(kU)$, $k = 1, 2, 3, \ldots$. Since Y is complete, the Baire category theorem then tells us that there is a nonempty open set W in the closure of some $\Lambda(kB_X)$. Therefore every point of W is the limit of a sequence Λx_j, where $x_j \in kB_X$ for some k. Fix this k and this W.

Select $y_0 \in W$, and choose $\eta > 0$ such that $y_0 + y \in W$ for all $\|y\| < \eta$. For any such point y there are sequences $\{a_j\}$ and $\{b_j\}$ in kU such that

$$\Lambda a_j \to y_0 \quad \text{and} \quad \Lambda b_j \to y_0 + y$$

as $j \to \infty$. Set $x_j = b_j - a_j$. Then we have $\|x_j\| < 2k$ and $\Lambda x_j \to y$. This is true for every y with $\|y\| < \eta$. So the linearity of Λ tells us that, if $\delta \equiv \eta/2k$ then, for every $y \in Y$ and every $\epsilon > 0$, there is an $x \in X$ so that

$$\|x\| \leq \frac{1}{\delta} \cdot \|y\| \quad \text{and} \quad \|y - \Lambda x\| < \epsilon . \tag{4}$$

We want to shrink ϵ to 0.

Fix $y \in \delta V$ and $\epsilon > 0$. By (4), there is an $x_1 \in X$ with $\|x_1\| < 1$ and

$$\|y - \Lambda x_1\| < \frac{1}{2}\delta\epsilon .$$

Suppose now that $x_1, x_2, \ldots, x_k$ are chosen so that

$$\|y - \Lambda x_1 - \Lambda x_2 - \cdots - \Lambda x_k\| < 2^{-k}\delta\epsilon . \tag{5}$$

Use (4), with y replaced by $y - \Lambda x_1 - \Lambda x_2 - \cdots - \Lambda x_k$, to obtain an x_{k+1} such that (5) holds with $k+1$ in place of k and such that

$$\|x_{k+1}\| < 2^{-k}\epsilon \tag{6}$$

for $k = 1, 2, \ldots$.

If we set $s_k = x_1 + x_2 + \cdots + x_k$, then (6) shows that $\{s_k\}$ is a Cauchy sequence in X. Since X is complete, there is therefore an $x \in X$ such that $s_k \to x$. The inequality $\|x_1\| < 1$, along with (6), shows that $\|x\| < 1 + \epsilon$. Since Λ is continuous, $\Lambda s_k \to \Lambda x$. By (5), $\Lambda s_k \to y$. We conclude that $\Lambda x = y$.

So we have proved that

$$\Lambda((1+\epsilon)B_X) \supset \delta B_Y$$

or, dividing by $(1+\epsilon)$,

$$\Lambda(B_X) \supset (1+\epsilon)^{-1}\delta B_Y \tag{7}$$

for every $\epsilon > 0$. The union of the sets on the right-hand side of (7), taken over all $\epsilon > 0$, is δB_Y. This proves that $\Lambda(B_X) \supset \delta B_Y$. It follows naturally that the image of any open set is open, because any open set is the union of balls. $\square$

COROLLARY 1.8. *Let X and Y be Banach spaces. Let $\Lambda : X \to Y$ be a univalent, surjective, bounded linear operator. Then there is a $\delta > 0$ such that*

$$\|\Lambda x\|_Y \geq \delta \|x\|_X$$

for all $x \in X$—that is, $\|\Lambda x\|$ is bounded below. In other words, Λ^{-1} is a bounded linear operator from Y to X.

REMARK. Notice that the inequality in the conclusion of the corollary is just the *opposite* of the inequality that gives boundedness of the operator Λ.

Proof of the Corollary. Choose δ as in the statement of the open mapping principle. Then the conclusion of that theorem, together with the fact that Λ is one-to-one, shows that $\|\Lambda x\| < \delta$ implies that $\|x\| < 1$. Taking contrapositives, we see that $\|x\| \geq 1$ implies that $\|x\| \geq \delta$. $\square$

The next result is the third and last of our "big three." It is unusual in that it does not require the space in question to be complete. In some sense it is more of a logic result than a functional analysis result.

Historically, the Hahn-Banach theorem was first proved for real normed linear spaces. It was Bohnenblust [BOH] who determined how to extend the result to complex normed linear spaces.

If X is a normed linear space and E a subspace, then let $\lambda : E \to \mathbb{R}$ be a bounded linear functional. We call $\widehat{\lambda} : X \to \mathbb{R}$ an *extension* of λ to X if $\widehat{\lambda}\,\big|_E = \lambda$ and $\|\widehat{\lambda}\| = \|\lambda\|$.

THEOREM 1.9 (The Hahn-Banach Theorem). *Let X be a normed linear space and E a (not necessarily closed) subspace. Let λ be a bounded linear functional on E. Then there exists an extension $\widehat{\lambda}$ of λ to X.*

REMARK. It is worth noting here that E need not be closed. And X need not be complete. These are signals that the tools of analysis will not play a significant role in the proof.

It should also be noted that the extension is not usually unique. In many cases there will be uncountably many distinct extensions.

REMARK. In fact the Hahn-Banach theorem is false for linear operators (rather than linear functionals). We leave the details of this assertion for the interested reader. Or the reader might refer to [KAK] or [SOB] or [KEL].

Proof of the Hahn-Banach theorem. This quite standard proof may be found in [RUD1].

First assume that X is a *real* normed linear space, and that λ is a real-linear, real-valued, bounded linear functional on E. If $\|\lambda\| = 0$ then the desired extension $\widehat{\lambda}$ is is the zero-functional. Forgetting this trivial case, we may after a renormalization assume that $\|\lambda\| = 1$.

Select $x_0 \in X$ with $x_0 \notin E$. Let E_1 be the linear space spanned by E and x_0. Thus E_1 consists of all vectors of the form $x + \mu x_0$, where $x \in E$ and μ is a real scalar. Let us define $T_1(x + \mu x_0) = \lambda(x) + \mu\alpha$, where α is some fixed real number (to be specified later). We see that T_1 is a real linear functional on E_1. The catch is that we need to choose α so that the extended functional still has norm 1. This will hold provided that

$$|\lambda(x) + \mu\alpha| \le \|x + \mu x_0\| \tag{8}$$

for $x \in E$ and μ real. We replace x by $-\mu x$ and divide both sides of (8) by $|\mu|$. The requirement becomes

$$|\lambda(x) - \alpha| \le \|x - x_0\|$$

for $x \in E$.

It is convenient now to rewrite our condition as

$$\alpha_x \le \alpha \le \beta_x ,$$

where

$$\alpha_x = \lambda(x) - \|x - x_0\| \quad \text{and} \quad \beta_x = \lambda(x) + \|x - x_0\| . \tag{9}$$

Such an α exists if and only if all the intervals $[\alpha_x, \beta_x]$ for $x \in E$ have a common point. That is to say, if and only if

$$\alpha_x \le \beta_y \tag{10}$$

for all $x \in E$, $y \in E$. But

$$\lambda(x) - \lambda(y) = \lambda(x - y) \le \|x - y\| \le \|x - x_0\| + \|y - x_0\|$$

or

$$\lambda(x) - \|x - x_0\| \le \lambda(y) + \|y - x_0\| .$$

Hence

$$\alpha_x \le \beta_y$$

by (9).

Thus we have shown that there is a norm-preserving extension T_1 of λ from E to E_1.

We come to the logic part of the proof. Let $\mathcal{R}$ be the collection of all ordered pairs (E', λ'), where E' is a subspace of X that contains E and where λ' is a real-linear extension of λ to E' with $\|\lambda'\| = 1$. We may partially order $\mathcal{R}$ by $(E', \lambda') \le (E'', \lambda'')$ if and only if $E' \subset E''$ and $\lambda''(x) = \lambda'(x)$ for all $x \in E'$. We note that the axioms of a partial ordering are definitely satisfied and also that $\mathcal{R}$ is not empty since (E, λ) lies in $\mathcal{R}$. We may thus apply the Hausdorff maximality theorem (or Zorn's lemma) to conclude that there is a maximal totally ordered subcollection Π of $\mathcal{R}$.

Let Ψ be the collection of all E' such that $(E', \lambda') \in \Pi$ for some linear functional λ'. Then Ψ is a totally ordered by set inclusion and hence the union $\widetilde{E}$ of all members of Ψ is a subspace[1] of X. If $x \in \widetilde{E}$ then $x \in E'$ for some $E' \in \Psi$. Define $\widetilde{\lambda}(x) = \lambda'(x)$, where λ' is the functional that occurs in the pair $(E', \lambda') \in \Psi$. Our definition of the partial order in Π shows that it does not matter which $E' \in \Pi$ we choose to define $\widetilde{\lambda}$ as long as E' contains x.

[1] We need to be careful here, because the union of two subspaces need not be a subspace. For example, the union of the x-axis and the y-axis in the Euclidean plane is not a subspace.

We easily check that $\widetilde{\lambda}$ is a linear functional on $\widetilde{E}$, with $\|\widetilde{\lambda}\| = 1$. If $\widetilde{E}$ is a proper subspace of X, then the first part of the proof would give us a further extension of $\widetilde{\lambda}$, and that would contradict the maximality of Π. Thus $\widetilde{E} = X$ and we have completed the proof in the case of real scalars and real linear functionals.

Next we treat Bohnenblust's contribution. If now λ is a complex-linear functional on the subspace E of the complex normed linear space X, let η be the real part of λ. Use the real Hahn-Banach theorem, which we have just proved, to extend η to a real-linear functional $\widetilde{\mu}$ on X with $\|\widetilde{\eta}\| = \|\eta\|$. Define

$$\widetilde{\lambda}(x) = \widetilde{\eta}(x) - i\widetilde{\eta}(ix)$$

for $x \in X$. We may check directly that $\widetilde{\lambda}$ is a complex-linear extension of λ and that

$$\|\widetilde{\lambda}\| = \|\widetilde{\eta}\| = \|\eta\| = \|\lambda\| . \qquad \square$$

1.7 APPLICATIONS OF THE BIG THREE

We shall spend some considerable time examining applications of the big three theorems. This will help us to put these important results into perspective, and also help us to understand what they say and what they mean.

The applications of these three important theorems are broad and diverse. They come from partial differential equations, Fourier analysis, and many other parts of mathematics. Some of them require a bit of background, which we cannot provide in a book of this brevity. But we shall provide easily accessible references.

It is useful in what follows to define the *circle group* $\mathbb{T}$ to be the interval $[0, 2\pi]$ with the endpoints identified. This identification is nicely effected by the map

$$[0, 2\pi] \ni \theta \mapsto e^{i\theta} .$$

We continue to do arithmetic (and analysis) on $[0, 2\pi]$ as usual, just remembering that 0 and 2π are the same point. In particular, a function f is continuous on $\mathbb{T}$ if it is continuous on $[0, 2\pi]$ in the usual sense and also $f(0) = f(2\pi)$.

EXAMPLE. This example concerns convergence and divergence of Fourier series. See [KAT] or [KRA3] for the chapter and verse in this matter.

In what follows, we use the notation L^p to denote functions that are p^{th} power integrable (see [RUD1]). If f is an L^1 function on $\mathbb{T}$, we define the

n^{th} Fourier coefficient of f to be

$$\widehat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int}\, dt\,.$$

The N^{th} *partial sum* of the Fourier series of f is defined to be

$$S_N f(e^{i\theta}) = \sum_{n=-N}^{N} \widehat{f}(n) e^{in\theta}\,.$$

In point of fact, one can easily calculate (again see [KAT] or [KRA3]) that

$$S_N f(e^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} f(t) D_N(\theta - t)\, dt\,,$$

where

$$D_N(t) \equiv \frac{\sin\left(N + \frac{1}{2}\right) t}{\sin \frac{1}{2} t}\,.$$

The function D_N is known as the *Dirichlet kernel*.

Our goal in this example is to show that there exists a broad class of functions with divergent Fourier series (in a sense to be specified).

It will be crucial for us to know the L^1 norm of D_N. In fact

$$\begin{aligned}
\int_0^{2\pi} |D_N(t)|\, dt &= \int_{-\pi}^{\pi} |D_N(t)|\, dt \\
&\geq 2 \int_0^{\pi} \frac{\left|\sin\left(N + \frac{1}{2}\right) t\right|}{t}\, dt \\
&= 2 \int_0^{(N+1/2)\pi} \frac{|\sin t|}{t}\, dt \\
&> 2 \sum_{n=1}^{(N+1/2)\pi} \int_{(n-1)\pi}^{n\pi} \frac{|\sin t|}{n\pi}\, dt \\
&\geq \frac{4}{\pi} \sum_{n=1}^{N} \frac{1}{n} \approx \frac{4}{\pi} \ln N\,.
\end{aligned}$$

We see that $\|D_N\|_{L^1} \to \infty$ as $N \to \infty$.

We consider the operators S_N operating on the Banach space

$$X = \left\{ f \text{ continuous on } [0, 2\pi] \text{ such that } f(0) = f(2\pi) \right\}.$$

Of course the norm on X is the supremum norm.

We claim that

$$\|S_N\| = \|D_N\|_{L^1} . \tag{11}$$

If we can prove (11), then $\|S_N\| \to \infty$. The uniform boundedness principle then tells us that there is a dense subset $\mathcal{D}$ of X so that, for every $f \in \mathcal{D}$, the Fourier series of f at 0 diverges (since $S_N f(e^{i0}) = (1/[2\pi] \int_0^{2\pi} f(t) D_N(t)\, dt)$. So all that remains is to prove the claim.

On the one hand, if $f \in X$, then

$$|S_N(f)| \le \frac{1}{2\pi} \int_0^{2\pi} |f(s)||D_N(s)|\, ds \le \|f\|_{\sup} \cdot \|D_N\|_{L^1} .$$

That is one half of our task.

On the other hand, consider the closed set

$$E = \{t \in [0, 2\pi] : D_N(t) \ge 0\} .$$

For $n = 1, 2, \dots$, let

$$f_n(t) = \frac{1 - nd(t, E)}{1 + nd(t, E)} .$$

Here $d(t, E)$ denotes the distance of the point t to the set E. Then

(a) $\|f_n\|_{\sup} \le 1$,

(b) $f_n(t) = 1$ for $t \in E$,

(c) $f_n(t) \to -1$ for $t \notin E$.

The Lebesgue dominated convergence theorem tells us that

$$S_N(f_n) \longrightarrow \int_0^{2\pi} |D_N(s)|\, ds$$

as $n \to \infty$. We conclude that $\|S_N\| = \|D_N\|_{L^1}$ and our proof is complete.

It is well to recall here the Stone-Weierstrass theorem:

THEOREM 1.10. *Let A be an algebra*[2] *of continuous functions on a compact, Hausdorff space X. Suppose that*

(a) *A separates points (i.e., if $x \neq y$ are elements of X then there is an $f \in A$ such that $f(x) \neq f(y)$),*

[2] An algebra is a collection of objects equipped with binary operations of addition and multiplication, and also with a notion of scalar multiplication.

(b) *A vanishes nowhere (i.e., if $x \in X$ then there is an $f \in A$ such that $f(x) \neq 0$),*

(c) *In case A and $C([0,1])$ consist of complex-valued functions, then A is closed under complex conjugation.*

Then A is dense in $C([0,1])$.

The Stone-Weierstrass theorem is a generalization of the classical Weierstrass approximation theorem.

EXAMPLE. Let us take another look at the partial summation operators for Fourier series. It is a matter of great interest to know whether, if $f \in L^p(\mathbb{T})$, then does it follow that $S_N f \to f$ in L^p norm?

If q is a trigonometric polynomial, then $S_N q = q$ as soon as N exceeds the degree of q. So the convergence of $S_N q$ to q is trivial.

Fix $1 \leq p < \infty$ and suppose that we know that $\|S_N\|_{L^p, L^p} \leq C$, with the estimate independent of N. If $f \in L^p$ is arbitrary, let $\epsilon > 0$. By the Stone-Weierstrass theorem, select a trigonometric polynomial q such that $\|f - q\|_{L^p} < \epsilon$. Select N so large that N exceeds the degree of q. Then we have

$$\begin{aligned}
\|S_N f - f\|_{L^p} &\leq \|S_N(f-q)\|_{L^p} + \|S_N q - q\|_{L^p} + \|q - f\|_{L^p} \\
&\leq C\|f - q\|_{L^p} + 0 + \epsilon \\
&\leq C\epsilon + \epsilon \\
&= (C+1)\epsilon\,.
\end{aligned}$$

So we see that a uniform bound on the operator norms of the summation operators S_N gives us the desired L^p convergence of Fourier series.

For the converse, suppose it is known that $S_N f \to f$ in L^p norm for every $f \in L^p(\mathbb{T})$. Then the uniform boundedness principle tells us that there is a constant C so that $\|S_N\|_{L^p, L^p} \leq C$. So we have a necessary and sufficient condition for L^p convergence of Fourier series.

A classical calculation, amounting mainly to algebraic trickery (see [KRA3]), in fact reduces the question of the uniform bound on the operator norms of the S_N to a single bound on a single operator called the *Hilbert transform*. The Hilbert transform is given by the singular integral

$$f \longmapsto \int \frac{f(t)}{x-t}\, dt\,.$$

The Hilbert transform is arguably the most important linear operator in all of analysis, and this discussion gives an indication of one of the reasons why.

EXAMPLE. This is another application to Fourier series.

First we recall the Riemann-Lebesgue lemma. It says this:

Lemma: *Let $f \in L^1(\mathbb{T})$. Define the Fourier coefficients as usual by*

$$\widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt\,.$$

Then

$$\lim_{n\to\pm\infty} |\widehat{f}(n)| = 0\,.$$

The reason for this result is simplicity itself. First, in the special case that f is a trigonometric polynomial

$$f(t) = \sum_{j=-M}^{M} a_j e^{ijt}\,,$$

then $\widehat{f}(n) = 0$ as soon as $|n| > M$. If f is an arbitrary L^1 function then let $\epsilon > 0$ and use the Stone-Weierstrass theorem to select a trigonometric polynomial p so that $\|f - p\|_{L^1} < \epsilon$. Let M be the degree of p and let $|n| > M$. Then

$$|\widehat{f}(n)| \le |\widehat{(f-p)}(n)| + |\widehat{p}(n)| \le \|f-p\|_{L^1} + 0 < \epsilon\,.$$

That establishes the result.

The question that we want to consider now is whether the converse of the Riemann-Lebesgue lemma is true. That is to say, if $\{a_j\}$ is a doubly infinite sequence of complex numbers that vanishes at $\pm\infty$ then is there an $f \in L^1$ such that $\widehat{f}(n) = a_n$ for all $n \in \mathbb{Z}$? The answer is "no," and we shall see this using a little functional analysis.

Let us define c_0 to be the space of doubly infinite sequences of complex numbers which vanish at $\pm\infty$. The norm on this space is the supremum norm. We have the operator

$$T : L^1([0, 2\pi]) \longrightarrow c_0$$

given by

$$T(f) = \{\widehat{f}(n)\}_{n=-\infty}^{\infty} \equiv \widehat{f}\,.$$

We first show that T is one-to-one. Suppose that $f \in L^1$ and $Tf = 0$. Then $\widehat{f}(n) = 0$ for every n. It follows that, for any trigonometric polynomial p,

$$\int_0^{2\pi} f(t)p(t)\,dt = 0\,.$$

By the Stone-Weierstrass theorem, we may conclude than that

$$\int_0^{2\pi} f(t)g(t)\,dt = 0$$

for any $g \in C([0, 2\pi])$ with $g(0) = g(2\pi)$. A simple approximation argument allows us to conclude that

$$\int_0^{2\pi} f(t)\chi(t)\,dt = 0$$

for any χ the characteristic function of a disjoint union of intervals in $[0, 2\pi]$. It follows that $f \equiv 0$. Thus T is univalent.

Seeking a contradiction, we suppose that the range of T is all of c_0. Then the corollary to the open mapping principle would say that there is a $\delta > 0$ such that

$$\|\widehat{f}\|_{\ell^\infty} \geq \delta \|f\|_{L^1} \tag{12}$$

for all $f \in L^1([0, 2\pi])$. Let D_n be the Dirichlet kernel as in the last example but one. Then $D_n \in L^1$, $\|\widehat{D_n}\|_{\ell^\infty} = 1$, and $\|D_n\|_{L^1} \to \infty$ as $n \to \infty$. So there cannot be a $\delta > 0$ so that

$$\|\widehat{D_n}\|_{\ell^\infty} \geq \delta \|D_n\|_{L^1}$$

for every n. That is a contradiction.

REMARK. It is actually quite difficult to give a "constructive" proof of the last result. And certainly functional analysis gives it to us rather easily.

EXAMPLE. We shall take a moment to discuss the so-called closed graph theorem. This is an extremely useful criterion for telling when a linear operator is continuous. The statement is as follows:

> **Theorem:** Suppose that X and Y are Banach spaces. Let $\Lambda : X \to Y$ be a linear mapping. Assume that the graph $G = \{(x, \Lambda x) : x \in X\}$ is a closed set in $X \times Y$. Then Λ is continuous.

For the proof, we begin by noticing that $X \times Y$ is a vector space if we define addition and scalar multiplication componentwise. We define a norm on $X \times Y$ by

$$\|(x, y)\|_{X \times Y} \equiv \|x\| + \|y\|\,.$$

Then it is straightforward to check that, so equipped, $X \times Y$ is a Banach space.

The graph G of Λ is the set of ordered pairs $(x, \Lambda x) \in X \times Y$. It is, by hypothesis, closed. So G is itself a Banach space. Also the mapping

$$\pi_1 : (x, \Lambda x) \longmapsto x$$

is a continuous, one-to-one, linear mapping of G onto X. We also define

$$\pi_2 : (x, y) \longmapsto y$$

from $X \times Y$ to Y.

The open mapping theorem guarantees that $\pi_1^{-1} : X \to G$ is continuous. Trivially π_2 is continuous. Therefore

$$\Lambda = \pi_2 \circ \pi_1^{-1}$$

is continuous, as was to be proved.

We close this discussion by noting that a common, and commonly used, formulation of the closed graph theorem is this:

> **Theorem:** Let Λ be a linear mapping of Banach space X to Banach space Y. Assume that, for each sequence $\{x_j\}$ in X for which $x = \lim_{j\to\infty} x_j$ and $y = \lim_{j\to\infty} \Lambda x_j$ exist, it is true that $y = \Lambda x$. Then Λ is continuous.

We leave it to the reader to check that our two formulations of the closed graph theorem are equivalent.

EXAMPLE. It is a basic fact from harmonic analysis that any smoothly bounded domain in $\mathbb{R}^N$ has a Green's function. See [KRA2] for the details. In fact the argument in that source depends on Stokes's theorem in an essential way. We provide here, for planar domains, an alternative construction due to Peter Lax [LAX]. The main tool is the Hahn-Banach theorem.

Let $\Omega \subseteq \mathbb{R}^2$ be a smoothly bounded domain. This means that $\partial\Omega$ consists of finitely many smooth, disjoint, closed curves. Let $w \in \Omega$. A *Green's function* for Ω with singularity at w is a function $G(\cdot, w)$ on $\overline{\Omega}$ such that

(a) $G(\cdot, w)$ is continuous on $\overline{\Omega} \setminus \{w\}$,

(b) $G(\cdot, w)$ vanishes on $\partial\Omega$,

(c) $G(z, w) + \log|z - w|/(2\pi)$ is harmonic (in the z variable) on Ω.

It is clear that, if a Green's function with singularity at w exists, then it is unique (apply the maximum principle).

We work in this example with real Banach spaces over the real field $\mathbb{R}$. Let X be the space $C(\partial\Omega)$ of real-valued functions continuous on $\partial\Omega$. Let Y be the subspace consisting of those continuous functions that have a harmonic extension to the interior of the domain Ω. Clearly Y is a linear subspace[3] of X. We remind the reader that a C^2 (twice continuously differentiable) function u is *harmonic* on a domain Ω if

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) u(x, y) \equiv 0$$

on Ω.

Obviously, if $u \in Y$, then the harmonic extension $\hat{u}$ to Ω is unique. Fix $w \in \Omega$. Consider the functional

$$\varphi_w : Y \to \mathbb{R}$$

given by $\varphi_w(u) = \hat{u}(w)$ for $u \in Y$. Then φ_w is a linear functional on Y and $\|\varphi_w\| = 1$ by the maximum principle. By the Hahn-Banach theorem, there is an extension $\widehat{\varphi_w}$ of φ_w to X. And of course $\widehat{\varphi_w}$ will also have norm 1.

Let $z \in \mathbb{R}^2 \setminus \partial\Omega$. For $t \in \partial\Omega$, consider the function

$$\psi_z(t) = \frac{1}{2\pi} \cdot \log|z - t| .$$

Then certainly $\psi_z \in X$. If $z \notin \overline{\Omega}$, then

$$\varphi_w(\psi_z) = \widehat{\psi_z}(w) = \frac{1}{2\pi} \cdot \log|z - w| .$$

If instead $z \in \Omega$, then we may define

$$H(z, w) = \widehat{\varphi_w}(\psi_z) .$$

Then we take

$$G(z, w) = H(z, w) - \frac{1}{2\pi} \cdot \log|z - w| .$$

We claim that $G(\,\cdot\,, w)$ is the Green's function for Ω with singularity at w.

For this purpose it suffices to show that $H(z, w)$ is harmonic in $z \in \Omega$ and that $H(z, w)$ tends to $\log|t_0 - w|/(2\pi)$ as z tends to any $t_0 \in \partial\Omega$.

[3]This is a fascinating instance of formal reasoning. In the end, we know that X and Y are the same space. That is to say, the Dirichlet problem can be solved for any continuous boundary datum φ. But the proof of that fact involves the Poisson kernel, and that is constructed using the Green's function. What we are doing here is constructing the Green's function. So we treat X and Y as formally distinct.

Of course $\widehat{\varphi_w}$ is continuous and linear. Also

$$\lim_{h \to 0} \frac{\log|z+h-t| - \log|z-t|}{h}$$

exists uniformly for $t \in \partial\Omega$ (as long as z, $z+h$ are inside Ω). Thus $\widehat{\varphi_w}$ commutes with differentiation with respect to x or y. As a result, for $z \in \Omega$,

$$\begin{aligned}
\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) H(z,w) &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \widehat{\varphi_w}(\psi_z) \\
&= \widehat{\varphi_w}\left(\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\psi_z\right) \\
&= \widehat{\varphi_w}(0) \\
&= 0\,,
\end{aligned}$$

since ψ_z is harmonic in Ω.

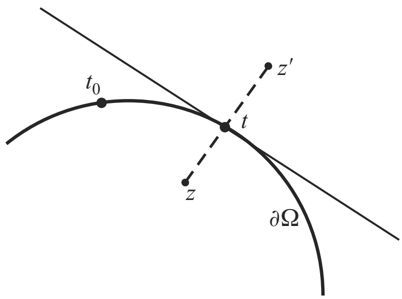

FIGURE 1.1. The reflected point z'.

Finally, if z is near to $t_0 \in \partial\Omega$, then let t be the point of $\partial\Omega$ that is nearest to z. Let z' be the mirror image of z in the tangent line to $\partial\Omega$ at t (see Figure 1.1). Since $\partial\Omega$ is smooth, if z is sufficiently near to t_0, then $z' \notin \Omega$ and z' is also near to t_0. Hence

$$\widehat{\varphi_w}(\psi_{z'}) = \varphi_w(\psi_{z'}) = \frac{1}{2\pi} \cdot \log|z' - w|\,,$$

and this tends to $\log|t_0 - w|/(2\pi)$ as $z \to t_0$. On the other hand,

$$|\widehat{\varphi_w}(\psi_z) - \widehat{\varphi_w}(\psi_{z'})| \le \|\psi_z - \psi_{z'}\|_\infty\,,$$

and this expression tends to 0 as $t \to t_0$. Hence $\widehat{\varphi}_w(\psi_z) \to \log|t_0 - w|/(2\pi)$ as $z \to t_0$. This proves that our $G(\cdot\,, w)$ is indeed the Green's function for Ω with singularity at $w \in \Omega$.

EXAMPLE. It is a classical construction (see [KRA2]) to obtain the Poisson kernel for a domain by calculating the unit outward normal derivative of the Green's function at the boundary. In what follows we take a more abstract approach—using the Hahn-Banach theorem—to derive the Poisson integral formula.

Let Ω be a smoothly bounded domain in the plane and $\overline{\Omega}$ its closure. Of course $\partial\Omega$ denotes the boundary of the domain.

Let Z be the space of those functions continuous on $\overline{\Omega}$ and harmonic in the interior. Using the supremum norm, we see that this is a Banach space. We may also consider the space Y of restrictions of elements of Z to $\partial\Omega$, and equip Y with the supremum norm also. By the maximum principle, we see that, for $f \in Z$, we have

$$\|f\|_Z = \|f\|_Y .$$

We also know that, if $z \in \overline{\Omega}$, then

$$|f(z)| \le \|f\|_Y .$$

In particular, if $f \in Z$ and $f(z) = 0$ for all $z \in \partial\Omega$ then $f \equiv 0$. In other words, if $f_1, f_2 \in Z$ and $f_1 = f_2$ on $\partial\Omega$ then $f_1 = f_2$.

Summarizing what we have learned, if $f \in Y$ then there is a unique extended function (still denoted by f) on $\overline{\Omega}$ so that $f \in Z$ and the restriction of the extension to $\partial\Omega$ equals the original function f.

Fix a point $z \in \overline{\Omega}$. We know that

$$\lambda : Y \ni f \longmapsto f(z)$$

is a bounded linear functional of norm 1. The Hahn-Banach theorem tells us that there is a an extension $\widehat{\lambda}$ of this functional to $C(\partial\Omega)$. Certainly

$$\widehat{\lambda}(1) = 1 \quad \text{and} \quad \|\widehat{\lambda}\| = 1 . \tag{13}$$

We claim that these facts imply that $\widehat{\lambda}$ is a *positive linear functional* on $C(\partial\Omega)$.

To see this, let $f \in C(\partial\Omega)$ with $0 \le f \le 1$ and put $g = 2f - 1$. We write

$$\widehat{\lambda} g = \alpha + i\beta ,$$

where α and β are real. Notice that $-1 \le g \le 1$. Thus

$$|g + ir|^2 \le 1 + r^2$$

for any real constant r. Thus (13) implies that

$$(\beta + r)^2 \le |\alpha + i(\beta + r)|^2 = |\widehat{\lambda}(g + ir)|^2 \le 1 + r^2 .$$

We conclude that

$$\beta^2 + 2r\beta \le 1$$

for every real r. This forces $\beta = 0$. Since $\|g\|_Y \le 1$, we conclude next that $|\alpha| \le 1$. As a result,

$$\widehat{\lambda} f = \frac{1}{2}\lambda(1 + g) = \frac{1}{2}(1 + \alpha) \ge 0 .$$

We may apply the Riesz representation theorem (see [RUD1, Theorem 2.14]). It tells us that there is a regular, positive Borel measure μ_z on $\partial\Omega$ such that

$$\widehat{\lambda} f = \int_{\partial\Omega} f \, d\mu_z$$

for $f \in C(\partial\Omega)$. In particular, we see that

$$f(z) = \int_{\partial\Omega} f \, d\mu_z \tag{14}$$

for $f \in z$.

Let us summarize what we have proved:

> **Theorem:** To each $z \in \overline{\Omega}$ there corresponds a positive measure μ_z on the boundary $\partial\Omega$ that "represents" z in the sense that (14) holds for every $f \in Z$.

Notice that $\widehat{\lambda}$ determines μ_z uniquely. But, in general, the Hahn-Banach extension is certainly not unique. So, thus far, we cannot say anything about the uniqueness of the representing measure.

Let us specialize down. Let Ω be the unit disc D in the complex plane and let $z \in D$. Write $z = re^{i\theta}$. Define

$$u_n(w) = w^n$$

for $n = 0, 1, 2, \ldots$. Then $u_n \in Z$. Thus

$$r^n e^{in\theta} = \int_{\partial\Omega} u_n d\mu_z .$$

Since $u_{-n} = \overline{u_n}$ on $\partial\Omega$, we find that

$$\int_{\partial\Omega} u_n d\mu_z = r^{|n|} e^{in\theta} \tag{15}$$

for $n = 0, \pm 1, \pm 2, \ldots$. Thus it makes sense to examine the real function

$$P_r(\theta - t) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(\theta - t)} \tag{16}$$

for t real. Notice that

$$\frac{1}{2\pi} \int_0^{2\pi} P_r(\theta - t) e^{int}\, dt = r^{|n|} e^{int} \tag{17}$$

for $n = 0, \pm 1, \pm 2, \ldots$.

We see that the series (16) is dominated by the convergent geometric series $\sum r^{|n|}$, so it is legitimate to insert the series into the integral (17) and to switch the sum and the integral (so that we integrate term by term). That is what gives (17). Comparison of (17) and (15) gives

$$\int_{\partial\Omega} f\, d\mu_z = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) P_r(\theta - t)\, dt$$

for $f = u_n$, hence for any trigonometric polynomial p. Again, by the Stone-Weierstrass theorem, we see that this last line holds for any continuous f on ∂D. So now we see that μ_z was uniquely determined.

In particular, the last line holds for $f \in Z$. So we now have the representation

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) P_r(\theta - t)\, dt$$

for $f \in Z$. The series (16) can be summed explicitly, since it is the real part of

$$1 + 2\sum_1^{\infty} (ze^{-it})^n = \frac{e^{it} + z}{e^{it} - z} = \frac{1 - r^2 + 2ir\sin(\theta - t)}{|1 - ze^{-it}|^2}\,.$$

Thus we see that

$$P_r(\theta - t) = \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2}\,.$$

This is the *Poisson kernel* for the unit disc. We have $P_r(\theta - t) \geq 0$ if $0 \leq r < 1$.

CHAPTER 2

ODE TO THE DUAL SPACE

2.1 INTRODUCTION

If X is a normed linear space, then the collection of bounded linear functionals on X is called its *dual space* and is denoted X^*. Given a Banach space X, we frequently want to calculate X^*. Knowledge of X^* can give us a good deal of information about X itself.

We note that, when X is a Banach space, then X^* will also be a Banach space. If $\alpha \in X^*$ then

$$\|\alpha\|_{X^*} \equiv \sup_{\|x\| \leq 1} |\alpha(x)| .$$

When $\alpha \in X^*$ and $x \in X$ then we sometimes write $\alpha(x)$ for the action of α on x and sometimes write $\langle \alpha, x \rangle$ or $\langle x, \alpha \rangle$ (in order to emphasize the dual nature of the action).

EXAMPLE. Let the Banach space X be $\mathbb{R}^N$ equipped with the usual norm. Of course any linear functional on X is automatically bounded, and it is given by inner product with an N-vector. So the space X^* of bounded linear functionals on $\mathbb{R}^N$ is just $\mathbb{R}^N$ itself.

EXAMPLE. Let X be the Banach space ℓ^1, which is the collection of all summable sequences of complex numbers. We calculate X^*. Let $\alpha \in X^*$.

Set $\{u_j\}$ to be the sequence which has a 1 in the j^{th} slot and 0s in all the other slots. Certainly $u_j \in \ell^1$. Set $a_j = \alpha(u_j)$. Consider the sequence $\{a_j\}$. We claim that $\{a_j\}$ is in fact a bounded sequence. If not then, given $M > 0$, there is an a_j with $|a_j| > M$. But then

$$\alpha(u_j) = a_j$$

and we see that

$$|\alpha(u_j)| > M\|u_j\| .$$

Since this is true for any $M > 0$, we find that α is unbounded. Contradiction. Thus we have shown that any bounded linear functional α on ℓ^1 is associated to a bounded sequence $\{a_j\}$ (i.e., an element of ℓ^∞), and the functional is given by

$$\alpha(\{x_j\}) = \sum_j a_j x_j .$$

Conversely, if $\{b_j\}$ is any bounded sequence, then the mapping

$$\{x_j\} \to \sum_j b_j x_j$$

defines a bounded linear functional on ℓ^1.

We have proved that the dual space $(\ell^1)^*$ is ℓ^∞.

EXAMPLE. We cannot provide all the details here (see [RUD1] for the chapter and verse), but we note that if $X = C([0,1])$, then X^* is the collection of regular Borel measures on $[0,1]$. This is the celebrated Riesz representation theorem.

EXAMPLE. Let X be the space $L^1(\mathbb{T})$ of Lebesgue integrable functions on the circle group $\mathbb{T}$. Then it can be shown (see [RUD2]) that X^* is the space $L^\infty(\mathbb{T})$ of essentially bounded functions on $\mathbb{T}$. However we must note that the dual of $L^\infty(\mathbb{T})$ is *not* $L^1(\mathbb{T})$. In fact it is impossible to give a compact characterization of the dual of $L^\infty(\mathbb{T})$.

EXAMPLE. If $1 \le p < \infty$ then let the conjugate exponent q be defined by

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \text{or} \quad q = \frac{p}{p-1} .$$

Define the Banach space ℓ^p by

$$\ell^p = \left\{ \{a_j\}_{j=1}^\infty : \left(\sum_{j=1}^\infty |a_j|^p \right)^{1/p} \equiv \|\{a_j\}\|_{\ell^p} < \infty \right\} .$$

It is a fact, which we shall not prove here, that $(\ell^p)^* = \ell^q$. Indeed, we have treated the case $p = 1$ in the second example. Given this information, one might wonder whether $(\ell^\infty)^* = \ell^1$. In fact this is not the case, and it is essentially impossible to describe the dual of ℓ^∞ in any compact fashion.

Let us treat this matter now. Let Y be the Banach subspace of ℓ^∞ consisting of those sequences $\{a_j\}$ such that $\lim_{j\to+\infty} a_j$ exists. Define a bounded linear functional λ on Y by

$$\lambda(\{a_j\}) = \lim_{j\to+\infty} a_j \,.$$

We may apply the Hahn-Banach theorem to elicit a bounded linear functional $\widehat{\lambda}$, still having norm 1, on ℓ^∞ that extends λ. If $(\ell^\infty)^* = \ell^1$, then there would be an ℓ^1 sequence $\beta = \{b_j\}$ that represents $\widehat{\lambda}$. But then we can apply β to the sequence that equals 1 in the first N entries and 0 elsewhere to see that all the partial sums of β are 0. It follows that β is identically 0, and that is a contradiction. Therefore the dual of ℓ^∞ is *not* ℓ^1.

2.2 CONSEQUENCES OF THE HAHN-BANACH THEOREM

The Hahn-Banach theorem gives us important information about the dual space. We record some of it here.

PROPOSITION 2.1. *Let X be a normed linear space. Pick $x_0 \in X$, $x_0 \neq 0$. Then there is a bounded linear functional ρ on X with $\|\rho\| = 1$ so that $\rho(x_0) = \|x_0\|$.*

Proof. Let $E = \{tx_0 : t \in \mathbb{R}\}$. Define a linear functional r on E by $r(tx_0) = t\|x_0\|$. Then r is a bounded linear functional with norm 1. Apply the Hahn-Banach theorem to obtain the functional ρ on X with the required properties. $\square$

REMARK. Proposition 2.1 shows that X^* has a great many nontrivial elements.

COROLLARY 2.2. *Let X be a normed linear space and let $x \in X$. Then*

$$\|x\| = \sup\{|\lambda(x)| : \lambda \in X^*, \|\lambda\| = 1\}\,.$$

Proof. Clearly

$$|\lambda(x)| \leq \|\lambda\| \cdot \|x\| \leq \|x\|\,.$$

That proves half of the statement.

For the converse, we may assume that $x \neq 0$. Use the proposition to find a linear functional ρ of norm 1 so that $\rho(x) = \|x\|$. It follows that

$$\|x\| = |\rho(x)| \leq \sup\{|\lambda(x)| : \lambda \in X^*, \|\lambda\| = 1\}\,. \qquad \square$$

PROPOSITION 2.3. *Let X be a normed linear space and E a linear subspace. Fix a point $x_0 \in X$. Then x_0 is in the closure $\overline{E}$ of E if and only if there is no bounded linear functional λ on X so that $\lambda(x) = 0$ for all $x \in E$ but $\lambda(x_0) \neq 0$.*

REMARK. For clarity, it is worth formulating the contrapositive statement of this new proposition:

> **Proposition.** Let X be a normed linear space and E a linear subspace. Fix a point $x_0 \in X$. There is a bounded linear functional λ on X with the property that $\lambda(x) = 0$ for all $x \in E$ while $\lambda(x_0) \neq 0$ if and only if $x_0 \notin \overline{E}$.

Proof of the Proposition. Certainly if $x_0 \in \overline{E}$, and if λ is a bounded linear functional on X, and if we further know that $\lambda(x) = 0$ for all $x \in E$, then the continuity of λ forces $\lambda(x_0) = 0$.

Conversely, suppose that $x_0 \notin \overline{E}$. Then there is a $\delta > 0$ such that $\|x - x_0\| > \delta$ for all $x \in E$. Let E' be the subspace of X generated by E and x_0. So a typical element of E' has the form $x + tx_0$ for $x \in E$ and $t \in \mathbb{R}$. Let us define

$$\lambda(x + tx_0) = t .$$

Since

$$\delta|t| \leq |t|\|x_0 + t^{-1}x\| = \|tx_0 + x\| ,$$

we see that λ is a linear functional on E' whose norm is at most δ^{-1}. Moreover, $\lambda(x) = 0$ for $x \in E$ and $\lambda(x_0) = 1$.

Finally, we apply the Hahn-Banach theorem to obtain a bounded linear functional, also of norm at most δ^{-1}, which extends λ to all of X. $\square$

EXAMPLE. Let us consider the Banach space $X = L^2([0, 1])$ consisting of all square-integrable, measurable functions on $[0, 1]$. The norm on L^2 is

$$\|f\|_{L^2} \equiv \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2} .$$

Let $\{t_1, t_2, \ldots\}$ be any countable, dense subset of $[0, 1]$. For $j = 1, 2, \ldots$ we set

$$\varphi_j(t) = \begin{cases} 1 & \text{if} \quad 0 \leq t \leq t_j \\ 0 & \text{if} \quad t_j < t \leq 1 . \end{cases}$$

We shall show that every $f \in L^2([0, 1])$ can be approximated in norm by linear combinations of the φ_j.

To see this, let $\lambda \in X^*$. It is known (see [RUD1]) that L^2 is its own dual. So there is an element $y \in X$ such that, for all $x \in X$,

$$\lambda(x) = \int_0^1 x \cdot y \, ds \, .$$

Assume that $\lambda(\varphi_j) = 0$ for $j = 1, 2, \ldots$. Thus

$$\int_0^{t_j} y(s) \, ds = 0$$

for all j. For any $t \in [0, 1]$, let

$$z(t) = \int_0^t y(s) ds \, .$$

Then $z \in C([0, 1])$ and $z(t_j) = 0$ for all $j = 1, 2, \ldots$. Since $\{t_1, t_2, \ldots\}$ is dense in $[0, 1]$, we see therefore that $z \equiv 0$ on $[0, 1]$. Therefore $y = 0$ almost everywhere. In conclusion, the functional $\lambda = 0$.

Proposition 2.3 tells us that, if we set Y equal to the span of the φ_j, then every element of $L^2([0, 1])$ is in the closure of Y.

CHAPTER 3

Hilbert Space

3.1 Introduction

If we place some additional structure on a Banach space, then a richer theory results. That is what we now do.

The key idea is to equip the Banach space with an inner product. That gives us notions of orthogonality and projection, and the result is a beautiful and coherent theory.

Definition. A complex vector space H is called an *inner product space* if, to each ordered pair of vectors $x, y \in H$, there is associated a complex number $\langle x, y\rangle$. This number is called the *inner product* or *scalar product* of x and y. In the language of algebra, we may think of this inner product as a bilinear operator from $H \times H$ to the scalar field $\mathbb{C}$. In order to be useful, the inner product must satisfy the following properties:

(a) $\langle x, y\rangle = \overline{\langle y, x\rangle}$,

(b) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,

(c) $\langle ax, y\rangle = a\langle x, y\rangle$ for $a \in \mathbb{C}$,

(d) $\langle x, x\rangle \geq 0$,

(e) $\langle x, x\rangle = 0$ if and only if $x = 0$.

Example. If H is $\mathbb{C}^n$ then a convenient inner product is the Hermitian inner product given by

$$\langle (z_1, z_2, \ldots, z_n), (w_1, w_2, \ldots, w_n)\rangle = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n\ .$$

EXAMPLE. If H is the Lebesgue space of square integrable, measurable functions on the interval $[0, 1]$, then a useful inner product is

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\, dx\,.$$

There are a number of easy consequences of the inner product axioms that are worth noting explicitly.

(i) (c) implies that $\langle \mathbf{0}, y\rangle = 0$,

(ii) (b) and (c) imply that, for a fixed $y \in H$, the mapping $x \mapsto \langle x, y\rangle$ is a linear functional on H,

(iii) (a) and (c) imply that $\langle x, ay\rangle = \overline{a}\langle x, y\rangle$,

(iv) (a) and (b) give the distributive law

$$\langle z, x + y\rangle = \langle z, x\rangle + \langle z, y\rangle\,;$$

(v) (d) allows us to define $\|x\| = \sqrt{\langle x, x\rangle}$. Thus

$$\|x\|^2 = \langle x, x\rangle\,.$$

We have the fundamental Schwarz inequality:

PROPOSITION 3.1 (Schwarz inequality). *If $x, y \in H$ then*

$$|\langle x, y\rangle| \le \|x\| \cdot \|y\|\,.$$

REMARK. The Schwarz inequality is a fairly simple idea, but it has profound consequences. For instance, it tells us that the projection of an element x of H will have length not exceeding $\|x\|$. It also gives important information about the trigonometry of space.

Proof of the Proposition. This argument is inspired by the proof in [RUD1]. It is quite standard.

Set $A = \|x\|^2$, $B = |\langle x, y\rangle|$, and $C = \|y\|^2$. Choose a complex number α such that $|\alpha| = 1$ and $\alpha\langle y, x\rangle = B$. If r is any real number then we have

$$\langle x - r\alpha y, x - r\alpha y\rangle = \langle x, x\rangle - r\alpha\langle y, x\rangle - r\overline{\alpha}\langle x, y\rangle + r^2\langle y, y\rangle\,. \quad (1)$$

The expression on the left-hand side of (1) is real and nonnegative. Thus

$$A - 2Br + Cr^2 \ge 0$$

for any real r.

If $C = 0$ then the last inequality forces $B = 0$ so our inequality is trivially true. If $C > 0$ then take $r = B/C$ in the last inequality to derive $B^2 \leq A \cdot C$. $\square$

PROPOSITION 3.2 (The Triangle Inequality). *If $x, y \in H$, then*

$$\|x + y\| \leq \|x\| + \|y\| .$$

REMARK. It is noteworthy that the triangle inequality is a consequence of Schwarz's lemma. This is the standard means of deriving the triangle inequality in the abstract setting of a normed linear space.

Proof of the Proposition. We apply the Schwarz inequality as follows:

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 .\end{aligned}$$ $\square$

We call an inner product space a *Hilbert space* if it is complete.

EXAMPLE. Let X be $\mathbb{C}^n$ and let the inner product be

$$\langle (z_1, z_2, \ldots, z_n), (w_1, w_2, \ldots, w_n)\rangle \equiv \sum_{j=1}^{n} z_j \overline{w_j} .$$

Then it is straightforward to check that this is an inner product space. Since it is finite dimensional, it is also easy to check that it is complete. So it is a Hilbert space.

EXAMPLE. Let $L^2([0, 1])$ be the measurable, square integrable, complex-valued functions on the interval $[0, 1]$. Equip this space with the inner product

$$\langle f, g\rangle \equiv \int_0^1 f(s)\overline{g(s)}\, ds .$$

Then this is an inner product space. It is a nontrivial theorem (see [RUD1]) that the space is complete. So it is a Hilbert space.

EXAMPLE. As usual, let ℓ^2 be the set of square-summable complex sequences. Define an inner product by

$$\big\langle \{a_j\}, \{b_j\}\big\rangle \equiv \sum_j a_j \overline{b_j} .$$

Then it is straightforward to check that this is an inner product. And one can actually show that the space is complete. So this is a Hilbert space.

EXAMPLE. Consider the space of all piecewise linear functions on the interval [0, 1]. This is certainly a linear space, and we can equip it with the inner product

$$\langle f, g\rangle = \int_0^1 f(s)\overline{g(s)}\,ds\,.$$

The space is not complete, however, as any measurable, square-integrable function can be approximated by piecewise linear functions.

DEFINITION. Let H be an inner product space. We say that a subset $E \subseteq H$ is a *subspace* if it is itself a vector space with the inherited operations of addition and scalar multiplication. One checks that E is a subspace by noting that E is closed under addition and scalar multiplication, and that E contains the zero element.

EXAMPLE. Let H be the inner product space consisting of all measurable, square-integrable functions on [0, 1] equipped with the usual inner product

$$\langle f, g\rangle = \int_0^1 f(s)\overline{g(s)}\,ds\,.$$

Then the set E of continuous functions on [0, 1] forms a subspace.

In this book we shall be particularly interested in *closed* subspaces.

3.2 The Geometry of Hilbert Space

DEFINITION. If x is an element of an inner product space H, then $x^\perp$ denotes the set of all y such that $\langle x, y\rangle = 0$. We call $x^\perp$ the *orthogonal space* to x or the *orthogonal complement of* x.

If $E \subseteq H$ is either a subspace or a subset, then we let $E^\perp$ denote the set of all y such that $\langle e, y\rangle = 0$ for all $e \in E$. We call $E^\perp$ the *orthogonal complement of* E.

It is worth noting explicitly that $x^\perp$, as just defined, is itself an inner product space—in particular, it is closed under addition and scalar multiplication. Likewise, $E^\perp$ is an inner product space.

DEFINITION. We say that a set F in a linear space V is *convex* if, whenever $x, y \in F$ and $0 \le t \le 1$, then

$$(1-t)x + ty \in F\,.$$

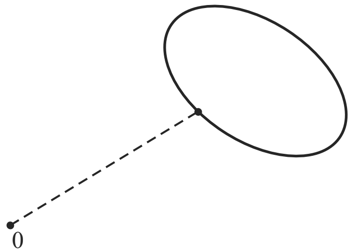

FIGURE 3.1. Unique element of least norm.

As t ranges from 0 to 1, the displayed expression describes a segment from x to y.

Any subspace of V is convex. Any translate of a convex set is convex. If A is a convex set then the *Minkowski functional* μ_A of A is defined by

$$\mu_A(x) = \inf\{t > 0 : t^{-1}x \in A\} .$$

A set $B \subseteq X$ is said to be *balanced* if $\alpha B \subseteq B$ for every α in the scalar field with $|\alpha| \leq 1$.

LEMMA 3.3 (The Parallelogram Law). *We have the identity*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 .$$

Proof. We calculate that

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \left(\|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2\right) \\ &\qquad + \left(\|x\|^2 - \langle x, y\rangle - \langle y, x\rangle + \|y\|^2\right) \\ &= 2\|x\|^2 + 2\|y\|^2 . \end{aligned}$$

□

PROPOSITION 3.4. *Each nonempty, closed, convex set E in a Hilbert space H contains a unique element of least norm. (See Figure 3.1.) That is to say, there is an element $y_0 \in E$ such that*

$$\|y_0\| \leq \|e\|$$

for all $e \in E$. Furthermore, there is only one element $y_0 \in E$ with this property.

REMARK. It is helpful to think of a closed disc in the plane that does not contain the origin. A segment from the origin to the center of the disc will intersect the boundary in the point of least norm.

Proof of the Proposition. This is a standard result with a standard proof—see, for instance, [RUD1], [RUD2].

Let $\delta = \inf\{\|x\| : x \in E\}$. So δ is the "least norm" that we are discussing. For any $x, y \in E$, we apply Proposition 3.3 to $(1/2)x$ and $(1/2)y$ to obtain

$$\frac{1}{4}\|x-y\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 - \left\|\frac{x+y}{2}\right\|^2 .$$

Since E is convex, $(x+y)/2 \in E$. Therefore

$$\|x-y\|^2 \leq 2\|x\|^2 + 2\|y\|^2 - 4\delta^2 . \tag{2}$$

If it happens to be the case that both x and y are elements of smallest norm, then $\|x\| = \delta$ and $\|y\| = \delta$. Hence the last line tells us that $x = y$. That takes care of the uniqueness part of the proposition.

Certainly the definition of δ tells us that there is a sequence $\{y_j\}$ in E so that $\|y_j\| \to \delta$ as $j \to \infty$. Apply (2) to y_j and y_k. As $j, k \to \infty$, it then follows that the right-hand side of (2) will tend to 0. Thus $\{y_j\}$ is a Cauchy sequence. Since H is a Hilbert space, it is complete. So there is a $y_0 \in H$ such that $y_j \to y_0$. In other words, $\|y_j - y_0\| \to 0$ as $j \to \infty$. Since $y_j \in E$ and E is closed, we see that $y_0 \in E$. Since the norm is a continuous function on H, we conclude that

$$\|y_0\| = \lim_{j\to\infty} \|y_j\| = \delta .$$

Thus y_0 is the element of least norm that we seek. □

We develop now the properties of Hilbert space projections.

THEOREM 3.5. *Let H be a Hilbert space. Let E be a closed subspace. Then there is a unique pair of linear mappings P and Q such that $P : H \to E$, $Q : H \to E^{\perp}$, and*

$$x = Px + Qx \tag{3}$$

for all $x \in H$. These mappings have the following additional properties:

(4) *If $x \in E$, then $Px = x$ and $Qx = 0$. If $x \in E^{\perp}$, then $Px = 0$ and $Qx = x$,*

(5) $\|x - Px\| = \inf\{\|x-y\| : y \in E\}$,

(6) $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$.

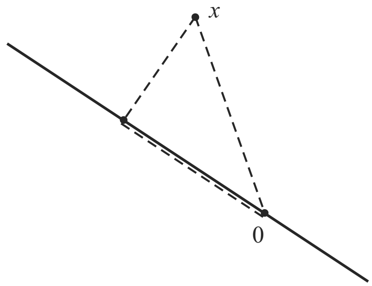

FIGURE 3.2. Orthogonal projection.

We call P and Q the orthogonal projections of H onto E and $E^{\perp}$ respectively. See Figure 3.2.

REMARK. This result once again illustrates the dictum that Hilbert space is an infinite-dimensional generalization of finite-dimensional Euclidean space. For, certainly in $\mathbb{R}^N$, if E is a subspace then there is a projection onto E and there is a projection onto the subspace F perpendicular to E. In fact let $\{e_1, e_2, \ldots, e_k\}$ be an orthonormal basis for E and let $\{f_1, f_2, \ldots, f_{N-k}\}$ be an orthonormal basis for F. Then the projection P onto E is given by

$$Px = \langle x, e_1\rangle e_1 + \langle x, e_2\rangle e_2 + \cdots \langle x, e_k\rangle e_k$$

and the projection Q onto F is given by

$$Qx = \langle x, f_1\rangle f_1 + \langle x, f_2\rangle f_2 + \cdots \langle x, f_k\rangle f_k\,.$$

Think about this simple example as you read the proof of the theorem.

Proof of the Theorem. See [RUD1], [RUD2] for the background.

If $x \in H$, then the set $x + E = \{x + e : e \in E\}$ is closed and convex. It is, after all, a linear space. Define Qx to be the unique element of smallest norm (using Proposition 3.4) in $x + E$. Define $Px = x - Qx$. Then (3) is automatically true. Since $Qx \in x + E$, it follows that $Px \in E$. So P maps H into E.

Next we show that $\langle Qx, e\rangle = 0$ for all $e \in E$. We may assume that $\|e\| = 1$, and we put $z = Qx$. The minimal property of Qx now shows that

$$\langle z, z\rangle = \|z\|^2 \le \|z - \alpha e\|^2 = \langle z - \alpha e, z - \alpha e\rangle$$

for any scalar α. This simplifies to

$$0 \le -\alpha\langle e, z\rangle - \overline{\alpha}\langle z, e\rangle + |\alpha|^2\,.$$

Take $\alpha = \langle z, e \rangle$. This gives

$$0 \leq -|\langle z, e \rangle|^2 ,$$

so that $\langle z, e \rangle = 0$. Hence Q maps H into $E^\perp$.

Write $x = x_E + x_{E^\perp}$, with $x_E \in E$ and $x_{E^\perp} \in E^\perp$. So we have

$$x_E - Px = Qx - x_{E^\perp} .$$

Since $x_E - Px \in E$ and $Qx - x_{E^\perp} \in E^\perp$, and since $E \cap E^\perp = \{\mathbf{0}\}$, we see that $x_E = Px$ and $x_{E^\perp} = Qx$. This proves the uniqueness assertions.

To see that P and Q are linear, we apply (3) to x, y, and $\alpha x + \beta y$ to obtain

$$P(\alpha x + \beta y) - \alpha Px - \beta Py = \alpha Qx + \beta Qy - Q(\alpha x + \beta y) .$$

The left-hand side is in E, and the right-hand side is in $E^\perp$. Thus both are $\mathbf{0}$, so P and Q are linear.

Property (4) follows from (3). Property (5) was used to define P. Also (6) follows from (3) since $\langle Px, Qx \rangle = 0$. □

COROLLARY 3.6. *If $E \subseteq H$ is a closed subspace and $E \neq H$ then there is a $y \in H$, $y \neq \mathbf{0}$, such that $y \perp E$.*

Proof. Take $x \in H$, $x \notin E$. Set $y = Qx$. Since $x \neq Px$, it follows that $y \neq \mathbf{0}$. □

We next treat one of the most central and significant results of Hilbert space theory. Often called the *Riesz representation theorem*, it says that every bounded linear functional on a Hilbert space H is given by inner product with an element of H.

THEOREM 3.7. *If λ is a bounded linear functional on the Hilbert space H, then there is a unique element $y \in H$ such that*

$$\lambda x = \langle x, y \rangle$$

for all $x \in H$.

Proof. If $\lambda x = 0$ for all x, then we may simply take $y = \mathbf{0}$. Otherwise we define

$$E = \{x \in H : \lambda x = 0\} .$$

The linearity of λ shows that E is a subspace, and the fact that λ is continuous shows that E is closed. Since $\lambda h \neq 0$ for some $h \in H$, we know that $E^{\perp}$ does not consist of $\mathbf{0}$ alone.

Clearly the y that we seek lives in $E^{\perp}$. Also $\lambda y = \langle y, y \rangle$.

Choose a nonzero $z \in E^{\perp}$. Then $z \notin E$, so $\lambda z \neq 0$. Set $y = \alpha z$, where $\overline{\alpha} = (\lambda z)/\langle z, z \rangle$. Then $y \in E^{\perp}$, $\lambda y = \langle y, y \rangle$, and $y \neq \mathbf{0}$.

For any $x \in H$, we put

$$x' = x - \frac{\lambda x}{\langle y, y \rangle}\, y \quad \text{and} \quad x'' = \frac{\lambda x}{\langle y, y \rangle}\, y\,.$$

Then $\lambda x' = 0$, hence $x' \in E$. Thus $\langle x', y \rangle = 0$. Therefore

$$\langle x, y \rangle = \langle x'', y \rangle = \lambda x\,.$$

This gives the representation of λ that we seek.

For the uniqueness of y, assume that $\langle x, y \rangle = \langle x, y' \rangle$ for all $x \in H$. Set $z = y - y'$. Then $\langle x, z \rangle = 0$ for all $x \in H$. In particular, $\langle z, z \rangle = 0$ so that $z = 0$. $\square$

We turn to a consideration of orthogonal systems in Hilbert space. It is this set of ideas that really shows us the structure of the Hilbert space, and how Hilbert space is very naturally a generalization of finite-dimensional Euclidean space.

DEFINITION. We say that a collection of elements $\{u_\alpha\}_{\alpha \in A}$ in a Hilbert space H is *orthonormal* if each vector has norm 1 and if $\langle u_\alpha, u_\beta \rangle = 0$ whenever $\alpha \neq \beta$. We may also write these conditions as

$$\langle u_\alpha, u_\beta \rangle = \begin{cases} 1 & \text{if} \quad \alpha = \beta\,, \\ 0 & \text{if} \quad \alpha \neq \beta\,. \end{cases}$$

If $\{u_\alpha\}_{\alpha \in A}$ is an orthonormal set, or orthonormal system, then we associate with each $x \in H$ a function $\widehat{x}$ on the index set A given by

$$\widehat{x}(\alpha) = \langle x, u_\alpha \rangle \quad \text{for } \alpha \in A\,.$$

We sometimes call the numbers $\widehat{x}(\alpha)$ the *Fourier coefficients* of x relative to the orthonornal system $\{u_\alpha\}$.

PROPOSITION 3.8. *Let $u_1, u_2, \ldots, u_k$ be an orthonormal set in the Hilbert space H. Set*

$$x = \sum_{j=1}^{k} c_j u_j\,.$$

Then

$$c_j = \langle x, u_j \rangle \quad \text{for } 1 \le j \le k\,;$$
$$\|x\|^2 = \sum_{j=1}^{k} |c_j|^2\,.$$

Proof. Just calculate. □

A fundamental problem in modern analysis is that of best approximation. As a specific instance, given an orthonormal set $u_1, u_2, \ldots, u_k$ in a Hilbert space H and an element $x \in H$, we wish to choose coefficients $c_1, c_2, \ldots, c_k$ so as to minimize the expression

$$\left\| x - \sum_{j=1}^{k} c_j u_j \right\|\,. \tag{7}$$

THEOREM 3.10. *The choice of c_j that minimizes the expression*

$$\left\| x - \sum_{j=1}^{k} c_j u_j \right\|$$

is

$$c_j = \langle x, u_j \rangle\,.$$

The vector

$$\sum_{j=1}^{k} \langle x, u_j \rangle u_j$$

is in fact the orthogonal projection of x into the subspace generated by $\{u_1, u_2, \ldots, u_k\}$. If δ is the distance of x to this subspace, then

$$\sum_{j=1}^{k} |\langle x, u_j \rangle|^2 = \|x\|^2 - \delta^2\,. \tag{8}$$

Proof. We seek the element of the space spanned by $\{u_1, u_2, \ldots, u_k\}$ that is nearest to x. That element will be the projection of x into the space, and that is given by

$$\sum_{j=1}^{k} \langle x, u_j \rangle u_j\,.$$

The equation (8) is just the Pythagorean theorem. □

We note in passing that, if $\{u_\alpha\}$ is any orthonormal set in a Hilbert space H, then it is contained in a maximal orthonormal set. This follows by partially ordering all orthonormal sets in H by containment and then applying Zorn's lemma in a standard way. A maximal orthonormal set $\{v_\beta\}_{\beta \in B}$ has the property that if $x \in H$ and $\langle x, v_\beta \rangle = 0$ for all β then $x = \mathbf{0}$. A maximal orthonormal set is commonly called a *complete orthonormal system.*

COROLLARY 3.11 (Bessel's Inequality). *If $\{u_\alpha : \alpha \in A\}$ is any orthonormal set in the Hilbert space H, and if $\widehat{x}(\alpha) \equiv \langle x, u_\alpha \rangle$ for each α, then*

$$\sum_{\alpha \in A} |\widehat{x}(\alpha)|^2 \leq \|x\|^2 .$$

Proof. This is immediate from the last identity in the statement of the theorem. □

We shall say a few words about the important Riesz-Fischer theory.

THEOREM 3.12 (The Riesz-Fischer Theorem). *Let $\{u_\alpha\}_{\alpha \in A}$ be a complete orthonormal system in H. Let $\varphi \in \ell^2(A)$. Then $\varphi = \widehat{x}$ for some $x \in H$.*

Proof. For $n = 1, 2, \ldots$, let $A_n = \{\alpha \in A : |\varphi(\alpha)| > 1/n\}$. Since $\varphi \in \ell^2$, we see that each A_n is a finite set. Set

$$x_n = \sum_{\alpha \in A_n} \varphi(\alpha) u_\alpha .$$

Then $\widehat{x_n} = \varphi \cdot \chi_{A_n}$ (where χ_S denotes the characteristic function of the set S). Hence $\widehat{x_n}(\alpha) \to \varphi(\alpha)$ for each $\alpha \in A$. Also $\|\varphi - \widehat{x_n}\|^2 \leq \|\varphi\|^2$. Thus, by the dominated convergence theorem (see [FOL]), $\|\varphi - \widehat{x_n}\|_{\ell^2} \to 0$.

We conclude that $\{\widehat{x_n}\}$ is a Cauchy sequence in $\ell^2(A)$. Since the sets A_n are finite, Proposition 3.8 now shows that $\|x_n - x_m\| = \|\widehat{x_n} - \widehat{x_m}\|_{\ell^2}$. Thus $\{x_n\}$ is a Cauchy sequence in H. Since H is complete, there is an x that is the limit of the x_n in H.

For any $\alpha \in A$ we then have

$$\widehat{x}(\alpha) = \langle x, u_\alpha \rangle = \lim_{n \to \infty} \langle x_n, u_\alpha \rangle = \lim_{n \to \infty} \widehat{x_n}(\alpha) = \varphi(\alpha) . \qquad \square$$

THEOREM 3.13. *Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set in the Hilbert space H. The following statements are equivalent:*

(a) *$\{u_\alpha\}$ is a complete orthonormal system in H,*

(b) *The set S of all finite linear combinations of members of $\{u_\alpha\}$ is dense in H,*

(c) *For every* $x \in H$, *we have* $\|x\|^2 = \sum_{\alpha \in A} |\widehat{x}(\alpha)|^2$,

(d) *If* $x, y \in H$, *then* $\langle x, y\rangle = \sum_{\alpha \in A} \widehat{x}(\alpha)\overline{\widehat{y}(\alpha)}$.

Proof. The scheme of our proof is

$$(a) \Longrightarrow (b) \Longrightarrow (c) \Longrightarrow (d) \Longrightarrow (a).$$

Let E be the closure of S (see part (b)). Since S is a subspace, so is E. If S is not dense in H, then $E \neq H$, so that $E^{\perp}$ contains some nonzero element. Therefore $\{u_\alpha\}$ is not maximal if S is not dense, and we see that (a) implies (b).

Suppose that (b) holds. Fix $x \in H$ and $\epsilon > 0$. Since S is dense, there is a finite set $u_{\alpha_1}, u_{\alpha_2}, \ldots, u_{\alpha_k}$ such that some linear combination of these vectors has distance less than ϵ from x. This approximation can only be improved if we take $\widehat{x}(\alpha_j)$ for the coefficients of u_{α_j}. Thus if

$$z = \widehat{x}(\alpha_1)u_{\alpha_1} + \cdots + \widehat{x}(\alpha_k)u_{\alpha_k},$$

then we have $\|x - z\| < \epsilon$. Therefore $\|x\| < \|z\| + \epsilon$. So Corollary 3.11 tells us that

$$(\|x\| - \epsilon)^2 < \|z\|^2 = |\widehat{x}(\alpha_1)|^2 + \cdots + |\widehat{x}(\alpha_k)|^2 \leq \sum_{\alpha \in A} |\widehat{x}(\alpha)|^2. \tag{9}$$

Since $\epsilon > 0$ was arbitrary, we see that (c) follows from (9) and the Bessel inequality.

We can write the equation in (c) as

$$\langle x, x\rangle = \langle \widehat{x}, \widehat{x}\rangle,$$

where the inner product on the right is that in ℓ^2. Fix $x, y \in H$. If (c) holds, then

$$\langle x + \mu y, x + \mu y\rangle = \langle \widehat{x} + \mu\widehat{y}, \widehat{x} + \mu\widehat{y}\rangle$$

for every scalar μ. Thus

$$\overline{\mu}\langle x, y\rangle + \mu\langle y, x\rangle = \overline{\mu}\langle \widehat{x}, \widehat{y}\rangle + \mu\langle \widehat{y}, \widehat{x}\rangle.$$

We take $\mu = 1$ and also $\mu = i$ in the last displayed line. This information tells us that $\langle x, y\rangle$ and $\langle \widehat{x}, \widehat{y}\rangle$ have the same real and imaginary parts. Hence they are equal. Thus we have that (c) implies (d).

Finally, if (a) is false, then there exists a $0 \neq w \in H$ such that $\langle w, u_\alpha\rangle = 0$ for all $\alpha \in A$. If $x = y = w$, then $\langle x, y\rangle = \|w\|^2 \neq 0$. But $\widehat{x}(\alpha) = 0$ for all $\alpha \in A$. So (d) fails. We see that (d) implies (a). $\square$

CHAPTER 4

The Algebra of Operators

4.1 Preliminaries

Before we begin the subject proper of this chapter, we need briefly to discuss some topological issues. In particular, we must treat some topologies on X and X^*.

DEFINITION. Let X be a Banach space and X^* its dual. We say that a sequence $\{x_j\}$ in X converges to $x \in X$ in the *weak topology* if $\varphi(x_j) \to \varphi(x)$ for every $\varphi \in X^*$.

DEFINITION. Let X be a Banach space and X^* its dual. We say that a sequence $\{\varphi_j\}$ in X^* converges to $\varphi \in X^*$ in the *weak-$*$ topology* if $\varphi_j(x) \to \varphi(x)$ for every $x \in X$.

We can see that, in a certain sense, weak convergence and weak-$*$ convergence are dual notions. Weak-$*$ convergence is pointwise convergence for functionals. Weak convergence is the dual of this idea.

EXAMPLE. Let $X = \ell^1$ so that its dual $X^* = \ell^\infty$. Let $\varphi_j \in X^*$ be the sequence that has a 1 in the j^{th} position and 0s elsewhere. If $f \in X$ then $\langle \varphi_j, f \rangle \to 0$. Hence the sequence $\{\varphi_j\}$ converges to 0 in the weak-$*$ topology. But notice that $\{\varphi_j\}$ does *not* converge in norm.

EXAMPLE. Let X be $L^1(\mathbb{T})$ so that its dual X^* is $L^\infty(\mathbb{T})$. Let $x_j = e^{ijt} \in X$. Then the x_j *do not* converge in norm. Indeed they are not Cauchy.

However, if $y \in X^*$ is any element then

$$\langle y, x_j \rangle \to 0$$

as $j \to \pm\infty$. This is just the Riemann-Lebesgue lemma. Hence the sequence $\{x_j\}$ converges to 0 in the weak topology.

THEOREM 4.1 (Banach-Alaoglu). *Let X be a Banach space and let B be the closed unit ball in X^*. Then B is weak-$*$ compact.*

REMARK. The proof of this fundamental result relies on Tychonoff's theorem from topology, which in turn depends on the axiom of choice. Basically we think of an element of B in terms of its graph, and we think of that graph as living in the product of closed unit discs. We need to know that that product is compact.

Proof of the Theorem. Let V be any neighborhood of $\mathbf{0}$ in X. Let $x \in X$. Then there is a positive, finite number $\gamma(x)$ such that $x \in \gamma(x)V$. Therefore

$$|Lx| \le \gamma(x)$$

for every $L \in B$ and $x \in X$.

Let D_x be the set of all scalars α such that $|\alpha| \le \gamma(x)$. Let $\mathcal{P}$ be the product of all the D_x, and let τ be the product topology on $\mathcal{P}$. Since each D_x is compact, then so is $\mathcal{P}$.

We may think of an element of $\mathcal{P}$ as a function f on X (linear or not) satisfying $|f(x)| \le \gamma(x)$. As a result, $B \subseteq X^* \cap \mathcal{P}$. So B inherits two topologies:

(a) The weak-$*$ topology,

(b) The topology τ from $\mathcal{P}$.

We shall show that these two topologies coincide on B, and that B is then a closed subset of $\mathcal{P}$. Since $\mathcal{P}$ is compact (by Tychonoff's theorem), it will then follow that B is τ-compact and therefore that B is weak-$*$ compact.

Fix an $L_0 \in B$. Choose $x_j \in X$ for $j = 1, 2, \ldots, n$. Also select $\delta > 0$. Set

$$W_1 = \{L \in X^* : |Lx_j - L_0 x_j| < \delta \text{ for } 1 \le j \le n\}$$

and

$$W_2 = \{f \in \mathcal{P} : |f(x_j) - L_0 x_j| < \delta \text{ for } 1 \le j \le n\}.$$

We let n, x_j, and δ range over all possible values. This generates a family of sets W_1 and a family of sets W_2. The sets W_1 form a local basis for the weak-$*$ topology of X^* at L_0. The sets W_2 form a local basis for the product topology of $\mathcal{P}$ at L_0. Since $B \subseteq \mathcal{P} \cap X^*$, we see that

$$W_1 \cap B = W_2 \cap B .$$

That proves that the two topologies coincide on B.

Assume that f_0 is in the τ-closure of B. Choose $x, y \in X$, scalars α, β, and $\epsilon > 0$. The set of all $f \in \mathcal{P}$ satisfying $|f - f_0| < \epsilon$ at x is a τ-neighborhood of f_0. Likewise, the set of all $f \in \mathcal{P}$ satisfying $|f - f_0| < \epsilon$ at y is a τ-neighborhood of f_0. And the same for the set of all such f satisfying $|f - f_0| < \epsilon$ at $\alpha x + \beta y$. Thus B contains such an f. Since f is linear, we have

$$\begin{aligned} &f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y) \\ &\qquad = (f_0 - f)(\alpha x + \beta y) + \alpha(f - f_0)(x) + \beta(f - f_0)(y) . \end{aligned}$$

Hence

$$|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)| < (1 + |\alpha| + |\beta|)\epsilon .$$

Since $\epsilon > 0$ was arbitrary, we find that f_0 is linear.

Finally, if $x \in V$ and $\epsilon > 0$, then the same argument shows that there is an $f \in B$ such that $|f(x) - f_0(x)| < \epsilon$. Since $|f(x)| \leq 1$, the definition of B tells us that $|f_0(x)| \leq 1$. We conclude that $f_0 \in B$. This proves that B is a closed subset of $\mathcal{P}$. □

4.2 THE ALGEBRA OF BOUNDED LINEAR OPERATORS

There is considerable interest in studying the algebra of bounded linear operators from a Hilbert space H to itself. This is in part because John von Neumann taught us that this is the right device for studying quantum mechanics (a *state* is such a bounded operator). Today these so-called *von Neumann algebras* are studied in their own right, and are the source of considerable fascinating mathematics.

We shall study these matters in a slightly more general context. Given Banach spaces X and Y, we shall consider $\mathcal{B}(X, Y)$—the bounded linear operators from X to Y.

THEOREM 4.2. *Associate to each $L \in \mathcal{B}(X, Y)$ the number*

$$\|L\| = \sup\{\|Lx\|_Y : x \in X, \|x\| \leq 1\} .$$

This is a norm on $\mathcal{B}(X, Y)$. Since Y is a Banach space, then so is $\mathcal{B}(X, Y)$.

Proof. A subset of a normed linear space is bounded if and only if it lies in some multiple of the unit ball. It follows then that $\|L\| < \infty$ for every $L \in \mathcal{B}(X, Y)$. If α is a scalar, then $(\alpha L)(x) = \alpha \cdot Lx$, so that $\|\alpha L\| = |\alpha| \cdot \|L\|$.

The triangle inequality in Y shows that

$$\begin{aligned}\|(L_1 + L_2)x\| &= \|L_1x + L_2x\| \\ &\leq \|L_1x\| + \|L_2x\| \\ &\leq (\|L_1\| + \|L_2\|)\|x\| \\ &\leq \|L_1\| + \|L_2\|\end{aligned}$$

for every $x \in X$ with $\|x\| \leq 1$. Thus

$$\|L_1 + L_2\| \leq \|L_1\| + \|L_2\| .$$

If $L \neq 0$, then $Lx \neq 0$ for some $x \in X$. Therefore $\|L\| > 0$. So $\mathcal{B}(X, Y)$ is a normed space.

Assume that Y is complete and that $\{L_n\}$ is a Cauchy sequence in $\mathcal{B}(X, Y)$. Since

$$\|L_nx - L_mx\| \leq \|L_n - L_m\| \cdot \|x\| , \tag{1}$$

and, since we assume that $\|L_n - L_m\| \to 0$ as $n, m \to \infty$, we see that $\{L_nx\}$ is a Cauchy sequence in Y for each $x \in X$. Therefore

$$Lx \equiv \lim_{n\to\infty} L_nx$$

exists. Clearly $L : X \to Y$ is linear.

If $\epsilon > 0$, then the right-hand side of (1) does not exceed $\epsilon\|x\|$ provided that n, m are sufficiently large. Thus

$$\|Lx - L_mx\| \leq \epsilon\|x\|$$

for all large m. As a result, $\|Lx\| \leq (\|L_m\| + \epsilon)\|x\|$. Thus $L \in \mathcal{B}(X, Y)$ and $\|L - L_m\| \leq \epsilon$. We see then that $L_m \to L$ in the norm of $\mathcal{B}(X, Y)$. So we have established the completeness of $\mathcal{B}(X, Y)$. □

Let X be a Banach space and X^* its dual. We let x be any element of X and x^* any element of X^*. Instead of writing $x^*(x)$, we frequently write

$$\langle x, x^* \rangle .$$

This is because we can think of X^* acting on X or we can think of X acting on X^*. The following result clarifies this situation.

THEOREM 4.3. *Let X be a Banach space and B the closed unit ball of X. If $x^* \in X^*$, then define*

$$\|x^*\| = \sup\{|\langle x, x^* \rangle| : x \in B\} .$$

Then we have the following properties:

(a) *The given norm makes X^* into a Banach space.*

(b) *If B^* is the closed unit ball of X^* then, for every $x \in X$,*

$$\|x\| = \sup\{|\langle x, x^*\rangle| : x^* \in B^*\}.$$

As a result, the mapping $x^* \mapsto \langle x, x^*\rangle$ is a bounded linear functional on X^* having norm $\|x\|$.

(c) *The closed unit ball B^* is weak-$*$ compact.*

Proof. Since $\mathcal{B}(X, Y) = X^*$ when Y is the scalar field, assertion (a) is obvious.

Fix $x \in X$. The Hahn-Banach theorem tells us that there is a $y^* \in B^*$ such that

$$\langle x, y^*\rangle = \|x\|.$$

But we also know that

$$|\langle x, x^*\rangle| \leq \|x\| \cdot \|x^*\| \leq \|x\|$$

for every $x^* \in B^*$. Part (b) follows from these two facts.

Finally, the open ball U of X is dense in B by definition. So the definition of $\|x^*\|$ shows that $x^* \in B^*$ if and only if $|\langle x, x^*\rangle| \leq 1$ for every $x \in U$. Part (c) now follows from the Banach-Alaoglu theorem. □

If $T \in \mathcal{B}(X, Y)$, then we associate to it its *adjoint* operator $T^* \in \mathcal{B}(Y^*, X^*)$. In fact we define T^* by way of the equation

$$\langle T^* y^*, x\rangle = \langle y^*, Tx\rangle.$$

This definition bears a moment's thought. Given a $y^* \in Y^*$, we want $T^* y^*$ to lie in X^*. And in fact it is uniquely determined by this last equation. For if $\langle \alpha, x\rangle = \langle \beta, x\rangle$ for all $x \in X$ then $\langle \alpha - \beta, x\rangle = 0$ for all $x \in X$. Hence $\alpha - \beta \equiv 0$.

Proposition 4.4. *We have that T^* is a linear mapping and*

$$\|T^*\| = \|T\|.$$

Proof. If $y^* \in Y^*$ and $T \in \mathcal{B}(X, Y)$ then define

$$T^* y^* = y^* \circ T.$$

Certainly then $T^* y^*$ so defined lies in X^*. Also

$$\langle x, T^* y^*\rangle = [T^* y^*](x) = y^*(Tx) = \langle Tx, y^*\rangle.$$

This shows that the definition of T^* that we have given here is consistent with the one that we presented before the proposition.

If $y_1^*, y_2^* \in Y^*$, then

$$\begin{aligned}\langle x, T^*(y_1^* + y_2^*)\rangle &= \langle Tx, y_1^* + y_2^*\rangle \\ &= \langle Tx, y_1^*\rangle + \langle Tx, y_2^*\rangle \\ &= \langle x, T^*y_1^*\rangle + \langle x, T^*y_2^*\rangle \\ &= \langle x, T^*y_1^* + T^*y_2^*\rangle\end{aligned}$$

for every $x \in X$. Therefore

$$T^*(y_1^* + y_2^*) = T^*y_1^* + T^*y_2^* .$$

Similarly, $T^*(\alpha y^*) = \alpha T^* y^*$. Hence $T^* : Y^* \to X^*$ is linear.

Finally, part (b) of Theorem 4.3 tells us that

$$\begin{aligned}\|T\| &= \sup\left\{ |\langle Tx, y^*\rangle| : \|x\| \le 1, \|y^*\| \le 1\right\} \\ &= \sup\left\{ |\langle x, T^*y^*\rangle| : \|x\| \le 1, \|y^*\| \le 1\right\} \\ &= \sup\left\{ \|T^*y^*\| : \|y^*\| \le 1\right\} \\ &= \|T^*\| .\end{aligned}$$

□

If T is an operator then we let $\mathcal{N}(T)$ denote its null space and $\mathcal{R}(T)$ its range. Of course I is the identity operator. □

PROPOSITION 4.5. *Let X and Y be Banach spaces and $T \in \mathcal{B}(X, Y)$. Then*

$$\mathcal{N}(T^*) = \mathcal{R}(T) \quad \text{and} \quad \mathcal{N}(T) = \mathcal{R}(T^*)^{\perp} .$$

Proof. Exercise. Or see [RUD2, p. 94]. □

4.3 COMPACT OPERATORS

In classical analysis, a compact set is a (potentially) infinite set that behaves in many ways like a finite set.

In functional analysis, a compact operator is an operator with (potentially) infinite-dimensional range that behaves in many ways like an operator with finite-dimensional range.

We shall give three equivalent definitions of compact operator. Along the way, we shall use the concept of "totally bounded." A set E in a metric space X is said to be *totally bounded* if, for each $\epsilon > 0$, there is a finite collection of balls of radius ϵ so that E lies in the union of those balls. Equivalently, E in a Banach space X is totally bounded if, for any neighborhood V of $\mathbf{0}$ in X, there is a finite set $F \subseteq X$ so that $E \subseteq F + V$.

We give three equivalent definitions of compact operator:

DEFINITION. A linear operator $T \in \mathcal{B}(X, Y)$ is said to be *compact* if one of the following equivalent conditions holds:

(a) If U is the unit ball in X, then the closure of $T(U)$ is compact in Y.

(b) $T(U)$ is totally bounded.

(c) Every sequence $\{x_j\} \subseteq X$ contains a subsequence $\{x_{j_k}\}$ such that $\{Tx_{j_k}\}$ converges to a point $y \in Y$.

We leave it as an exercise for the reader to verify that these three conditions are equivalent. We develop some properties of compact operators.

If X, Y are Banach spaces, we have the notation $\mathcal{B}(X, Y)$ to denote the bounded linear operators from X to Y. We also write $\mathcal{B}(X)$ for the bounded linear operators from X to X. This latter space is an algebra[1] with the binary operations of composition and addition. Note that, if $S, T \in \mathcal{B}(X)$, then

$$\|S \circ T\| \leq \|S\|\|T\| .$$

An operator $T \in \mathcal{B}(X)$ is *invertible* if there is a $S \in \mathcal{B}(X)$ such that $ST = TS = I$. We write $S = T^{-1}$. Notice that, by the open mapping theorem, T is invertible if and only if the kernel of T is $\{\mathbf{0}\}$ and the range of T is X.

The *spectrum* $\sigma(T)$ of an operator $T \in \mathcal{B}(X)$ is the set of all scalars λ such that $T - \lambda I$ is *not* invertible. Thus $\lambda \in \sigma(T)$ if and only if at least one of these statements is true:

(i) The range of $T - \lambda I$ is not all of X.

(ii) $T - \lambda I$ is not one-to-one.

When (ii) holds, λ is said to be an *eigenvalue* of the operator T. The corresponding eigenspace is the null space of $T - \lambda I$. Each nonzero element of that null space is an *eigenvector* of T. It will satisfy the equation

$$Tx = \lambda x .$$

[1]An algebra is a collection of objects equipped with binary operations of addition and multiplication, and also with a notion of scalar multiplication.

We have

PROPOSITION 4.6. *Let X and Y be Banach spaces. Then*

(a) *If $T \in \mathcal{B}(X)$ and $\dim \mathcal{R}(T) < \infty$, then T is compact.*

(b) *If $T \in \mathcal{B}(X)$, T is compact, and $\mathcal{R}(T)$ is closed, then $\dim \mathcal{R}(T) < \infty$.*

(c) *The compact operators form a closed subspace of $\mathcal{B}(X, Y)$ in the norm topology.*

Proof. These are all exercises in the definitions. The result may be found in many standard texts, including [RUD2].

Statement (a) is trivial.

If $\mathcal{R}(T)$ is closed, then $\mathcal{R}(T)$ is complete. Hence T is an open mapping of X onto $\mathcal{R}(T)$. If T is compact, then we see that $\mathcal{R}(T)$ is locally compact. Hence (b) follows from the fact that every locally compact normed linear space is finite dimensional (see [RUD2]).

It remains to treat part (c). If S and T are two compact operators from X into Y, then so is $S+T$, just because the sum of any two compact subsets of Y is compact. Hence the compact operators form a subspace Ξ of $\mathcal{B}(X, Y)$. We now show that Ξ is closed. That will complete the verification of (c).

Let $T \in \mathcal{B}(X, Y)$ be in the closure of Ξ. Select $r > 0$ and let U be the open unit ball in X. There exists an $S \in \Xi$ with $\|S - T\| < r$. Since $S(U)$ is totally bounded, there are points $x_1, x_2, \ldots, x_k$ in U such that $S(U)$ is covered by the balls of radius r having centers at the points Sx_j. Since $\|Sx - Tx\| < r$ for every $x \in U$, we conclude that $T(U)$ is covered by the balls of radius $3r$ with centers at the points Tx_j. Thus $T(U)$ is totally bounded, proving that $T \in \Xi$. □

PROPOSITION 4.7. *Let X and Y be Banach spaces. Then*

(a) *If $T \in \mathcal{B}(X)$, T is compact, and $\lambda \neq 0$, then $\dim \mathcal{N}(T - \lambda I) < \infty$.*

(b) *If $\dim X = \infty$, $T \in \mathcal{B}(X)$, and T is compact, then $0 \in \sigma(T)$.*

(c) *If $S \in \mathcal{B}(X)$, $T \in \mathcal{B}(X)$, and T is compact, then so are ST and TS.*

Proof. Put $Y = \mathcal{N}(T - \lambda I)$ in part (a). The restriction of T to Y is a compact operator whose range is Y. So (a) and (b) both follow from (b) of the last proposition because, if $0 \notin \sigma(T)$, then $\mathcal{R}(T) = X$. The proof of (c) is trivial. □

PROPOSITION 4.8. *Let X, Y be Banach spaces. Assume that $T \in \mathcal{B}(X, Y)$. Then T is compact if and only if T^* is compact.*

Proof. This is just a matter of understanding the definition of compactness. See [RUD2].

Assume that T is compact. Let $\{y_j^*\}$ be a sequence in the unit ball of Y^*. Define

$$f_j(y) = \langle y, y_j^* \rangle$$

for $y \in Y$. Since $|f_j(y) - f_j(y')| \le \|y - y'\|$, we see that $\{f_j\}$ is equicontinuous. Since $T(U)$ has compact closure in Y (since T is compact), the Ascoli–Arzela theorem guarantees that $\{f_j\}$ has a subsequence $\{f_{n_j}\}$ that converges uniformly on $T(U)$. Since

$$\begin{aligned} \|T^* y_{n_j}^* - T^* y_{n_k}^*\| &= \sup_{x \in U} |\langle Tx, y_{n_j}^* - y_{n_k}^* \rangle| \\ &= \sup_{x \in U} |f_{n_j}(Tx) - f_{n_k}(Tx)|, \end{aligned}$$

the completeness of X^* implies that $\{T^* y_{n_j}^*\}$ converges. Therefore T^* is compact.

For the converse direction, imitate the proof just given. $\square$

DEFINITION. Let M be a closed subspace of a Banach space X. If there is another closed subspace N of X such that

$$X = M + N \quad \text{with} \quad M \cap N = \{\mathbf{0}\},$$

then M is said to be *complemented* in X. We further say that X is the *direct sum* of M and N and we write

$$X = M \oplus N .$$

The dimension of X/M is called the *codimension* of M in X.

REMARK. It is a remarkable result of Lindenstrauss and Tzafriri [LIT] that the only Banach space in which every subspace is complemented is Hilbert space.

The idea of complemented subspace is closely related to the question of producing a Hahn-Banach type theorem for linear operators (as opposed to linear functionals). See [KAK] and [SOB] for these ideas.

LEMMA 4.9. *Let M be a closed subspace of a Banach space X.*

(a) *If $\dim M < \infty$, then M is complemented in X.*

(b) *If $\dim (X/M) < \infty$, then M is complemented in X.*

REMARK. The statement of this result is algebraic in nature, so it is no surprise that the proof is a combination of algebra and logic.

Proof of the Lemma. For (a), let $\{e_1, e_2, \ldots, e_k\}$ be a basis for M. Each $x \in M$ then has a unique representation

$$x = \alpha_1(x)e_1 + \alpha_2(x)e_2 + \cdots + \alpha_k(x)e_k \, .$$

Each α_j is a bounded linear functional on M. So, by the Hahn-Banach theorem, each has a continuous extension $\widehat{\alpha_j}$ to a member of X^*. Let N be the intersection of all the null spaces of all the $\widehat{\alpha_j}$. Then $X = M \oplus N$. That proves (a).

For (b), let $\pi : X \to X/M$ be the quotient map. Let $\{e_1, e_2, \ldots, e_m\}$ be a basis for X/M. Choose $x_j \in X$ so that $\pi x_j = e_j$, each j. Let N be the vector space spanned by $\{x_1, x_2, \ldots, x_m\}$. Then $X = M \oplus N$. □

LEMMA 4.10. *Let M be a subspace of the Banach space X. If M is not dense in X, and if $r > 1$, then there is an $x \in X$ such that*

$$\|x\| < r \quad \textit{but} \quad \|x - y\| \geq 1 \quad \textit{for all } y \in M \, .$$

Proof. There is an $x_1 \in X$ whose distance from M is 1. That is to say,

$$\inf\{\|x_1 - y\| : y \in M\} = 1 \, .$$

Select $y_1 \in M$ such that $\|x_1 - y_1\| < r$. Set $x = x_1 - y_1$. □

PROPOSITION 4.11. *Suppose that X is a Banach space. Let $T \in \mathcal{B}(X)$ be compact. If $\lambda \neq 0$, then $T - \lambda I$ has closed range.*

REMARK. This result is essential to the study of eigenvalues and eigenvectors.

Proof of the Proposition. This is a straightforward application of the ideas—see [RUD2].

By part (a) of Proposition 4.7, $\dim \mathcal{N}(T - \lambda I) < \infty$. By part (a) of Lemma 4.9, X is the direct sum of $\mathcal{N}(T - \lambda I)$ and a closed subspace M.

Define an operator $S \in \mathcal{B}(M, X)$ by

$$Sx = Tx - \lambda x \, .$$

Then S is one-to-one on M. Also $\mathcal{R}(S) = \mathcal{R}(T - \lambda I)$. To show that $\mathcal{R}(S)$ is closed, it is enough to show the existence of an $r > 0$ such that

$$r\|x\| \leq \|Sx\| \tag{2}$$

for all $x \in M$. For if this last inequality holds, and if $\{Sx_j\}$ is a Cauchy sequence, then so is $\{x_j\}$. Thus the completeness of $\mathcal{R}(S)$ will follow.

Seeking a contradiction, we suppose that (2) fails for every $r > 0$. So there exists a sequence $\{x_j\}$ in M such that $\|x_j\| = 1$, $Sx_j \to 0$, and (after passing to a subsequence) $Tx_j \to x_0$ for some $x_0 \in X$. We conclude that $\lambda x_j \to x_0$. So $x_0 \in M$ and

$$Sx_0 = \lim_{j\to\infty} (\lambda S x_j) = 0\,.$$

Since S is one-to-one, we find that $x_0 = 0$. But $\|x_j\| = 1$ for every j, and $x_0 = \lim \lambda x_j$, hence $\|x_0\| = |\lambda| > 0$. This is a contradiction. Hence (2) follows for some $r > 0$. □

THEOREM 4.12. *Suppose that X is a Banach space. Let $T \in \mathcal{B}(X)$ be compact. Let $r > 0$ and let E be a set of eigenvalues λ of T such that $|\lambda| > r$ for each $\lambda \in E$. Then*

(a) *For each $\lambda \in E$, $\mathcal{R}(T - \lambda I) \neq X$,*

(b) *E is a finite set.*

Proof. Our strategy is to show that if either (a) or (b) is false then there are closed subspaces M_j of X and scalars $\lambda_j \in E$ so that

(3) $M_1 \subset M_2 \subset \cdots$ with $M_j \neq M_{j+1}$ for all j;

(4) $T(M_j) \subset M_j$ for $j \geq 1$;

(5) $(T - \lambda_j I)(M_j) \subset M_j$ for $j \geq 2$.

The proof is brought to closure by showing that this information contradicts the compactness of T.

Suppose that (a) is false. Then $\mathcal{R}(T - \lambda_0 I) = X$ for some $\lambda_0 \in E$. Set $S = T - \lambda_0 I$ and define M_j to be the null space of S^j. Since λ_0 is an eigenvalue of T, there is an $x_1 \in M_1$, $x_1 \neq 0$. Since $\mathcal{R}(S) = X$, there is a sequence $\{x_j\} \subseteq X$ so that $Sx_{j+1} = x_j$ for $j = 1, 2, \ldots$. Then we have

$$S^j x_{j+1} = x_1 \neq 0 \quad \text{but} \quad S^{j+1} x_{j+1} = Sx_1 = 0\,.$$

Hence M_j is a proper, closed subspace of M_{j+1}. We conclude that (3)–(5) hold with $\lambda_j = \lambda_0$.

Assume that (b) is false. Then E contains a sequence $\{\lambda_j\}$ of distinct eigenvalues of T. Choose corresponding eigenvectors e_j; let M_k be the (finite-dimensional, hence closed) subspace of X spanned by $e_1, e_2, \ldots, e_k$.

Since the λ_j are distinct, $\{e_1, e_2, \ldots, e_k\}$ is a linearly independent set, so that M_{j-1} is a proper subspace of M_j. This proves (3).

If $x \in M_k$, then

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_k e_k ,$$

which shows that $Tx \in M_k$ and

$$(T - \lambda_k I)x = \alpha_1(\lambda_1 - \lambda_k)e_1 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)e_{k-1} \in M_{k-1} .$$

Hence (4) and (5) hold.

Since we have closed subspaces M_j satisfying (3) to (5), Lemma 4.9 gives us vectors $y_j \in M_j$ for $j = 2, 3, \ldots$ such that

$$\|y_j\| \leq 2 \quad \text{and} \quad \|y_j - x\| \geq 1 \quad \text{if} \quad x \in M_{j-1} . \tag{6}$$

If $2 \leq j < k$, then define

$$z = Ty_j - (T - \lambda_j I)y_j .$$

By (4) and (5), $z \in M_{k-1}$. So (6) tells us that

$$\|Ty_k - Ty_j\| = \|\lambda_k y_k - z\| = |\lambda_k| \|y_k - \lambda_k^{-1} z\| \geq |\lambda_k| > r .$$

The sequence $\{Ty_k\}$ has thus no convergent subsequences, even though $\{y_k\}$ is bounded. This contradicts the compactness of T. □

We come to the big structure theorem for compact operators.

Theorem 4.13. *Let X be a Banach space and $T \in \mathcal{B}(X)$. Assume that T is compact. Then*

(a) *If $\lambda \neq 0$, then the four numbers*

$$\begin{aligned} \alpha &= \dim \mathcal{N}(T - \lambda I) \\ \beta &= \dim X/\mathcal{R}(T - \lambda I) \\ \alpha^* &= \dim \mathcal{N}(T^* - \lambda I) \\ \beta^* &= \dim X^*/\mathcal{R}(T^* - \lambda I) \end{aligned}$$

are equal and finite.

(b) *If $\lambda \neq 0$ and $\lambda \in \sigma(T)$, then λ is an eigenvalue of T and also of T^*.*

(c) *$\sigma(T)$ is compact, at most countable, and has at most one limit point at 0.*

Proof. For simplicity, write $S = T - \lambda I$.

Let M_0 be a closed subspace of a Banach space Y. Let k be a positive integer such that $k \leq \dim Y/M_0$. Then there are vectors $y_1, y_2, \ldots, y_k$ in Y such that the vector space M_j generated by M_0 and $y_1, y_2, \ldots, y_j$ contains M_{j-1} as a proper subspace. By a standard result in the theory of normed linear spaces, each M_j is closed. By the Hahn-Banach theorem, there are continuous linear functionals $L_1, L_2, \ldots, L_k$ on Y such that $L_j y_j = 1$ but $L_j y = 0$ for all $y \in M_{j-1}$. These functionals are plainly linearly independent. We may thus conclude that

> If Ξ denotes the space of all continuous linear functionals on Y that annihilate M_0, then
> $$\dim Y/M_0 \leq \dim \Xi\,. \tag{7}$$

We apply this result with $Y = X$, $M_0 = \mathcal{R}(S)$. By Proposition 4.11, $\mathcal{R}(S)$ is closed. Also $\Xi = \mathcal{R}(S)^\perp = \mathcal{N}(S^*)$, so that (7) becomes

$$\beta \leq \alpha^*\,. \tag{8}$$

Next we take $Y = X^*$ with the weak-$*$ topology. Let $M_0 = \mathcal{R}(S^*)$. It follows that $\mathcal{R}(S^*)$ is weak-$*$ closed. Since Ξ consists of all weak-$*$ continuous linear functionals on X^* that annihilate $\mathcal{R}(S^*)$, we see that Ξ is isomorphic to $\mathcal{R}(S^*)^\perp = \mathcal{N}(S)$ (see Proposition 4.5) and (7) becomes

$$\beta^* \leq \alpha\,. \tag{9}$$

Our next task is to prove that

$$\alpha \leq \beta\,. \tag{10}$$

Once we have established that, then

$$\alpha^* \leq \beta^* \tag{11}$$

follows, since T^* is a compact operator (Proposition 4.7). Since $\alpha < \infty$ by (a) of Proposition 4.7, (a) of the present theorem is an obvious consequence of the inequalities (8), (9), (10), and (11).

So assume that (10) is false. Then $\alpha > \beta$. Since $\alpha < \infty$, Lemma 4.9 shows that X contains closed subspaces E and F such that $\dim F = \beta$ and

$$X = \mathcal{N}(S) \oplus E = \mathcal{R}(S) \oplus F\,.$$

Each $x \in X$ has a unique representation as $x = x_1 + x_2$ with $x_1 \in \mathcal{N}(S)$ and $x_2 \in E$. Define $\pi : X \to \mathcal{N}(S)$ by setting $\pi x = x_1$. The closed graph theorem tells us that π is continuous.

Since $\dim \mathcal{N}(S) > \dim F$, there is a linear mapping ϕ of $\mathcal{N}(S)$ onto F such that $\phi x_0 = 0$ for some $x_0 \neq 0$. Define

$$\Phi x = Tx + \phi\pi x$$

for $x \in X$. Then $\Phi \in \mathcal{B}(X)$. Since $\dim \mathcal{R}(\phi) < \infty$, $\phi\pi$ is a compact operator. Thus so is Φ (see Proposition 4.6(a)).

Notice that

$$\Phi - \lambda I = S + \phi\pi \,. \tag{12}$$

Since $x_0 \in \mathcal{N}(S)$, $\pi x_0 = x_0$, hence $\phi\pi x_0 = 0$. We may conclude then that λ is an eigenvalue of Φ (with eigenvector x_0). So

$$\mathcal{R}(\Phi - \lambda I) \neq X \tag{13}$$

by Theorem 4.12.

Since $\pi x = 0$ for every $x \in E$, (12) shows that

$$(\Phi - \lambda I)(E) = S(E) = S(X) = \mathcal{R}(S)\,.$$

If $x \in \mathcal{N}(S)$, then $\pi x = x$, and (12) gives

$$(\Phi - \lambda I)(\mathcal{N}(S)) = \phi(\mathcal{N}(S)) = F\,.$$

The last two displayed lines now tell us that

$$\mathcal{R}(\Phi - \lambda I) \supset \mathcal{R}(S) + F = X\,. \tag{14}$$

The evident contradiction between (13) and (14) shows that (10) is true. That completes our proof of (a).

Part (b) follows from (a) because, if λ is not an eigenvalue of T, then $\alpha(T) = 0$ and (a) implies that $\beta(T) = 0$. That is to say, $\mathcal{R}(T - \lambda I) = X$. Thus $T - \lambda I$ is invertible, so that $\lambda \notin \sigma(T)$.

Lastly, part (c) of Theorem 4.13 tells us that 0 is the only possible limit point of $\sigma(T)$, that $\sigma(T)$ is at most countable, and that $\sigma(T) \cup \{0\}$ is compact. If $\dim X < \infty$, then $\sigma(T)$ is finite. If $\dim X = \infty$, then $0 \in \sigma(T)$ by part (b) of Proposition 4.6. Hence $\sigma(T)$ is compact. This gives part (c) of the current theorem and completes the proof. $\square$

CHAPTER 5

Banach Algebra Basics

The idea of a Banach algebra was conceived by I. M. Gelfand in his Ph.D. thesis of 1936. It is a beautiful blend of functional analysis and classical hard analysis. Particularly striking is how quickly and easily it leads to profound and elegant results. We shall present some of these in the present chapter.

5.1 Introduction to Banach Algebras

An *algebra* is a collection of objects equipped with binary operations of addition and multiplication, and also with a notion of scalar multiplication. For example, the collection of polynomials $p(x)$ of one variable forms an algebra. This is clearly an algebraic idea. Gelfand's key insight was to combine this notion with some analysis.

DEFINITION. Let A be a vector space over the complex numbers $\mathbb{C}$ that is equipped with operations of multiplication and addition and also of scalar multiplication. We assume that, for $x, y, z \in A$,

(a) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$,

(b) $z \cdot (x + y) = z \cdot x + z \cdot y$,

(c) $\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$ for all scalars α.

Further suppose that A is equipped with a norm that satisfies the inequality

$$\|x \cdot y\| \leq \|x\| \, \|y\| \, .$$

We then say that A is a *normed complex algebra*. If, in addition, A is complete in the topology induced by this norm, then we call A a *Banach algebra*.

EXAMPLE. Let A be the collection of all functions continuous on the closure of the unit disc D in the complex plane and holomorphic on the interior of that disc. Define the norm

$$\|f\| = \max_{z \in \overline{D}(0,1)} |f(z)| \, .$$

Then it is easy to check that A is a Banach algebra.

EXAMPLE. Let A be the collection of all bounded holomorphic functions on the open unit disc D in the complex plane. Define the norm

$$\|f\| = \sup_{z \in D} |f(z)| \, .$$

Then it is straightforward to check that A is a Banach algebra.

EXAMPLE. Let A be the collection of all polynomials $p(z)$ of a single complex variable. Define the norm

$$\|p\| = \sup_{z \in \overline{D}} |p(z)| \, .$$

Then A is a normed complex algebra, but it is not complete. (For example, the function $f(z) = \sum_j 2^{-j} z^{2^j}$ is in the closure of this algebra, but it is not itself a polynomial.) So A is *not* a Banach algebra.

EXAMPLE. Let X be a compact, Hausdorff space. Let $C(X)$ denote the continuous, complex-valued functions on X. Equip $C(X)$ with the norm

$$\|f\| = \max_{x \in X} |f(x)| \, .$$

Then $C(X)$ is a Banach algebra. As we shall see below, it is in some sense the most important and the most typical Banach algebra.

PROPOSITION 5.1. *Multiplication in a Banach algebra is continuous.*

Proof. If $x_j \to x$ and $y_j \to y$ then

$$\|x_j y_j - xy\| \le \|x_j(y_j - y)\| + \|y(x_j - x)\| \, .$$

It is easy to see that $\{x_j\}$ and $\{y_j\}$ are bounded. Therefore the right-hand side may be majorized by

$$C\|y_j - y\| + C\|x_j - x\| \, ,$$

and this clearly tends to 0. $\square$

Although there has been considerable effort (see [RIC]) to study Banach algebras that have no unit element, we shall be able to streamline our studies considerably by assuming that there *is* a unit element. This is an element $e \in A$ such that $ex = xe = x$ for all $x \in A$. We shall further assume that $\|e\| = 1$.

We call an element $x \in A$ *invertible* if there is an element $x^{-1} \in A$ such that $x \cdot x^{-1} = x^{-1} \cdot x = e$. The unit in the Banach algebra A and the inverse of each invertible element are unique.

The *spectrum* of an element $x \in A$ is the set of all complex numbers λ such that $x - \lambda e$ is not invertible. We denote the spectrum of x by $\sigma(x)$.

Of great interest in the study of any particular Banach algebra A is the collection of its so-called *multiplicative linear functionals*.[1] These are bounded linear functionals φ that also respect multiplication:

$$\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y) .$$

PROPOSITION 5.2. *Let φ be a nontrivial multiplicative linear functional on the Banach algebra A with unit e. Then $\varphi(e) = 1$ and $\varphi(x) \neq 0$ for each invertible $x \in A$.*

Proof. Certainly

$$\varphi(e) = \varphi(e \cdot e) = \varphi(e) \cdot \varphi(e) .$$

It follows immediately that $\varphi(e) = 1$ (for if $\varphi(e) = 0$ then the functional is identically 0).

If x is invertible then

$$1 = \varphi(e) = \varphi(x \cdot x^{-1}) = \varphi(x) \cdot \varphi(x^{-1}) .$$

It follows that $\varphi(x) \neq 0$. □

PROPOSITION 5.3. *Let A be a Banach algebra, $x \in A$, and $\|x\| < 1$. Then*

(a) *$e - x$ is invertible,*

(b) $\|(e - x)^{-1} - e - x\| \leq \|x\|^2 \cdot (1 - \|x\|)^{-1}$,

(c) $|\varphi(x)| < 1$ *for every multiplicative linear functional φ on A.*

Proof. Since $\|x^n\| \leq \|x\|^n$ and $\|x\| < 1$, the elements

$$u_n = e + x + x^2 + \cdots + x^n \tag{1}$$

[1] A multiplicative linear functional is actually a ring or algebra homomorphism. Nevertheless, "multiplicative linear functional" is the most commonly used terminology.

form a Cauchy sequence in A. Since A is complete, there is an element $u \in A$ such that $u_n \to u$. Since $x^n \to \mathbf{0}$, and since

$$u_n \cdot (e - x) = e - x^{n+1} = (e - x) \cdot u_n ,$$

the continuity of multiplication implies that u is the inverse of $e - x$.

Line (1) implies that

$$\|u - e - x\| = \|x^2 + x^3 + \cdots\| \leq \sum_{n=2}^{\infty} \|x\|^n = \frac{\|x\|^2}{1 - \|x\|} .$$

That is assertion (b).

Finally, if $\lambda \in \mathbb{C}$ and $|\lambda| \geq 1$, then (a) implies that $e - \lambda^{-1}x$ is invertible. The preceding proposition then tells us that

$$1 - \lambda^{-1}\varphi(x) = \varphi(e - \lambda^{-1}x) \neq 0 .$$

Thus $\varphi(x) \neq \lambda$. □

An obvious consequence of Proposition 5.2 and part (c) of the last proposition is that any multiplicative linear functional has norm precisely 1.

DEFINITION. Let A be a Banach algebra and $x \in A$. The *spectral radius* of x is the number

$$\rho(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\} .$$

We now give a slightly different argument to again show that the norm of a multiplicative linear functional cannot exceed 1.

PROPOSITION 5.4. *Let φ be a multiplicative linear functional on the Banach algebra A. Then, for all $x \in A$,*

$$|\varphi(x)| \leq \|x\| .$$

Proof. Fix $x \in X$. If λ is a complex scalar with $|\lambda| > \|x\|$, then

$$\lambda e - x = \lambda(e - x/\lambda) .$$

The term $e - x/\lambda$ is invertible because its inverse is the Neumann series

$$e + x/\lambda + x^2/\lambda^2 + \cdots .$$

Thus $\lambda \notin \sigma(x)$.

In other words,

$$\sigma(x) \subseteq \{\lambda : |\lambda| \le \|x\|\}.$$

Let φ be a multiplicative linear functional on A. Let $s = \varphi(x)$. We claim that $s \in \sigma(x)$. This is so because $\varphi(se - x) = s - s = 0$, so $se - x$ cannot be invertible.

We conclude that

$$\{s : s \text{ is in the image of } x \text{ under some multiplicative} \\ \text{linear functional } \varphi\} \subseteq \{\lambda : |\lambda| \le \|x\|\}.$$

But this says that

$$|\varphi(x)| \le \|x\|. \qquad \square$$

5.2 THE STRUCTURE OF A BANACH ALGEBRA

We present some classical results that begin to lay out how a Banach algebra is put together.

PROPOSITION 5.5. *Let A be a Banach algebra. For each $x \in A$, $\sigma(x)$ is compact and nonempty.*

Proof. Notice that $\lambda \in \sigma(x)$ if and only if $x - \lambda e$ is not invertible. The complement of the invertible elements is closed (because the set of invertible elements is open—see below). Finally, the mapping

$$\lambda \longmapsto x - \lambda e$$

is a continuous mapping of the complex plane into A. So the inverse image of a closed set is closed. We already know that the spectrum of x is bounded by $\|x\|$. Hence the spectrum is compact.

Fix $x_0 \in A$ and choose $\lambda_0 \notin \sigma(x_0)$. Then $(x_0 - \lambda_0 e)^{-1} \neq 0$. Then the Hahn-Banach theorem guarantees the existence of a bounded linear functional Φ on A such that

$$f(\lambda_0) \equiv \Phi[(x_0 - \lambda_0 e)^{-1}]$$

is not zero.

One may calculate, at any $\lambda \notin \sigma(x_0)$, that

$$\lim_{\mu \to \lambda} \frac{f(\mu) - f(\lambda)}{\mu - \lambda} = \Phi[(x_0 - \lambda e)^{-2}].$$

Thus f is complex differentiable and hence holomorphic at any λ not in $\sigma(x_0)$.

If $\sigma(x_0)$ were empty, then $f(\lambda)$ would be an entire function that tends to 0 at ∞. It then follows that $f(\lambda) \equiv 0$ by Liouville's theorem. That contradicts the fact that $f(\lambda_0) \neq 0$. Hence $\sigma(x)$ is not empty. □

PROPOSITION 5.6. *Let A be a Banach algebra. Then the set of invertible elements in A is open.*

Proof. Let x be invertible in A and let $h \in A$ be small. Then

$$x + h = x \cdot (e + x^{-1}h)\,.$$

If h is sufficiently small in norm then $\|x^{-1}h\|$ is smaller than 1. Then the usual Neumann series argument shows that $e + x^{-1}h$ is invertible. And of course x is invertible. Hence $x + h$ is invertible. □

THEOREM 5.7 (Gelfand-Mazur). *If A is a complex Banach algebra with unit in which every nonzero element has an inverse, then A is isometrically isomorphic to the complex field $\mathbb{C}$.*

Proof. Let $x \in A$ and let λ_1, λ_2 be unequal complex numbers. Then at least one of $x - \lambda_1 e$ and $x - \lambda_2 e$ is nonzero. Hence one of them, by hypothesis, must be invertible. The last proposition now tells us that $\sigma(x)$ therefore consists of just one point. Call it $\lambda(x)$.

Since $x - \lambda(x)e$ is not invertible, it must be 0. Hence $x = \lambda(x)e$. Thus the mapping $x \to \lambda(x)$ is an isomorphism of A onto the complex field. It is also an isometry since $|\lambda(x)| = \|\lambda(x)e\| = \|x\|$ for every $x \in A$. □

THEOREM 5.8 (The Spectral Radius Theorem). *Let A be a Banach algebra. Let $x \in A$. Then*

$$\rho(x) = \lim_{n\to\infty} \|x^n\|^{1/n}\,.$$

REMARK. It is not even *a priori* clear that the limit exists. The existence of the limit is part of the assertion of the theorem, and we shall prove it.

This result is aesthetically very pleasing, but it is also a useful analytic tool.

Proof of the Theorem. Fix $x \in A$ and let n be a positive integer. Choose a complex number λ and assume that $\lambda^n \notin \sigma(x^n)$. Then

$$(x^n - \lambda^n e) = (x - \lambda e)(x^{n-1} + \lambda x^{n-2} + \cdots + \lambda^{n-1} e)\,.$$

Multiplying both sides by $(x^n - \lambda^n e)^{-1}$, shows that $(x - \lambda e)$ is invertible so that $\lambda \notin \sigma(x)$.

Thus, if $\lambda \in \sigma(x)$, then $\lambda^n \in \sigma(x^n)$ for $n = 1, 2, 3, \ldots$. By our standard bounds on the spectrum, we know that $|\lambda^n| \leq \|x^n\|$. Hence $|\lambda| \leq \|x^n\|^{1/n}$. As a result,

$$\rho(x) \leq \liminf_{n\to\infty} \|x^n\|^{1/n} .$$

If $|\lambda| > \|x\|$, it is then easy to check that

$$(\lambda e - x) \cdot \sum_{n=0}^{\infty} \lambda^{-n-1} x^n = e . \tag{2}$$

This last series is thus equal to $-(x - \lambda e)^{-1}$.

Let Φ be a bounded linear functional on A and define

$$f(\lambda) = \Phi[(x - \lambda e)^{-1}] .$$

By line (2), the expansion

$$f(\lambda) = -\sum_{n=0}^{\infty} \Phi(x^n)\lambda^{-n-1} \tag{3}$$

is valid for all λ with $|\lambda| > \|x\|$.

Arguing as in the last proof, we see that f is holomorphic outside $\sigma(x)$, hence certainly in the set $\{\lambda : |\lambda| > \rho(x)\}$. Thus the power series (3) converges for $|\lambda| > \rho(x)$. In particular

$$\sup_n |\Phi(\lambda^{-n} x^n)| < \infty \tag{4}$$

for all $|\lambda| > \rho(x)$ and for every bounded linear functional Φ on A.

The Hahn-Banach theorem tells us that the norm of any element of A is just the same as its norm as a linear functional on the dual space of A. Since (4) holds for every Φ, we can apply the uniform boundedness principle and conclude that, for each λ with $|\lambda| > \rho(x)$, there is a real number $C(\lambda)$ such that

$$\|\lambda^{-n} x^n\| \leq C(\lambda)$$

for every positive integer n. Multiply this last line by $|\lambda|^n$ and take nth roots. As a result,

$$\|x^n\|^{1/n} \leq |\lambda||C(\lambda)|^{1/n}$$

if $|\lambda| > \rho(x)$. Thus

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leq \rho(x) . \qquad \square$$

5.3 IDEALS

Let A be a Banach algebra. A set $\mathcal{I} \subseteq A$ is called an *ideal* if

(a) $\mathcal{I}$ is a linear subspace of A,

(b) If $x \in A$ and $y \in \mathcal{I}$, then $xy \in \mathcal{I}$.

EXAMPLE. Consider the Banach algebra $A = C([0,1])$. Then the set of functions f in A such that $f(1/2) = 0$ is an ideal.

Also the set of $f \in A$ such that $f(1/3) = f(1/2) = f(2/3) = 0$ is an ideal.

It is a theorem that any maximal ideal (see below) in $C([0,1])$ is a set of the form

$$\mathcal{I} = \{f \in A : f(x_0) = 0\}$$

for some $x_0 \in [0,1]$.

EXAMPLE. Consider the Banach algebra $A = H^\infty(D)$, the bounded analytic functions on the unit disc $D \subseteq \mathbb{C}$. Consider

$$\mathcal{I} = \{f \in A : f(0) = f'(0) = 0\}.$$

It is easy to check that $\mathcal{I}$ is an ideal in A.

DEFINITION. An ideal $\mathcal{I}$ in a Banach algebra A is called a *maximal ideal* if there is no ideal $\mathcal{J}$ that properly contains $\mathcal{I}$ and is properly contained in A.

EXAMPLE. Consider the Banach algebra $A = C([0,1])$. Let

$$\mathcal{I} = \{f \in A : f(1/2) = 0\}.$$

We claim that $\mathcal{I}$ is a maximal ideal.

If not, then there is an ideal $\mathcal{J}$ that lies between $\mathcal{I}$ and A. So there is an element $f \in \mathcal{J}$ such that $f(1/2) \neq 0$. Consider an element $g \in \mathcal{I}$ that vanishes at $1/2$ but nowhere else. Then

$$h = f^2 + g^2$$

lies in $\mathcal{J}$ and vanishes at no point of $[0,1]$. Then $m = 1/h$ lies in A so that $(1/h) \cdot h \equiv 1 \in \mathcal{J}$.

But if $1 \in \mathcal{J}$ then $A \subseteq \mathcal{J}$. So $\mathcal{J}$ is not a proper ideal.

PROPOSITION 5.9. *Let A be a Banach algebra. Let $\mathcal{I}$ be a proper ideal in A. Then $\mathcal{I}$ lies in a maximal ideal in A.*

Proof. This is just a straightforward application of Zorn's lemma. Simply partially order the collection of all ideals by containment. □

PROPOSITION 5.10. *Let A be a Banach algebra. Let $\mathcal{I}$ be a maximal ideal in A. Then $\mathcal{I}$ is closed in A.*

Proof. It is clear that the closure of an ideal is still an ideal.

Certainly $\mathcal{I}$ contains no invertible elements. And the set of all invertible elements is open (Proposition 5.6). So the closure of $\mathcal{I}$ will contain no invertible elements. Thus the closure of $\mathcal{I}$ will still be a proper ideal. Since $\mathcal{I}$ is a maximal ideal, it must then be that the closure of $\mathcal{I}$ equals $\mathcal{I}$. Hence $\mathcal{I}$ is closed. □

The most convenient way to identify, and to handle, maximal ideals is by way of a theorem that connects them up with multiplicative linear functionals.

THEOREM 5.11. *Let A be a Banach algebra and let φ be a multiplicative linear functional on A. Then the kernel of φ is a maximal ideal in A.*

Conversely, if $\mathcal{I}$ is a maximal ideal in A, then $\mathcal{I}$ is the kernel of some multiplicative linear functional φ on A.

We defer the proof of this result for a few moments while we develop some ancillary machinery that will be needed. We note that, because of this result, the collection of all multiplicative linear functionals on a Banach algebra A is often called the *maximal ideal space* of A. We sometimes denote this space by $\triangle$.

DEFINITION. If A is a Banach algebra and $\mathcal{I} \subseteq A$ an ideal then we define the *quotient* of A by $\mathcal{I}$ to be the collection of all cosets

$$x + \mathcal{I} \equiv \{x + i : i \in \mathcal{I}\} .$$

Denote such a coset by $\widehat{x}$. The collection of all cosets is denoted $A/\mathcal{I}$.

We see that

$$\widehat{x} + \widehat{y} = \{x + y + i : i \in \mathcal{I}\}$$

so that $\widehat{x} + \widehat{y} = \widehat{x + y}$. Likewise, $\widehat{x} \cdot \widehat{y} = \widehat{x \cdot y}$. Thus $A/\mathcal{I}$ is an algebra.

We define a norm on $A/\mathcal{I}$ by

$$\|\widehat{x}\|_{A/\mathcal{I}} \equiv \inf_{i \in \mathcal{I}} \|x + i\|_A .$$

THEOREM 5.12. *Let A be a Banach algebra and $\mathcal{J}$ a closed ideal in A. Then the quotient $A/\mathcal{J}$ has the following properties:*

(a) *$A/\mathcal{J}$ is a normed linear space,*

(b) *If A is a Banach space, then so is $A/\mathcal{J}$,*

(c) *If A is a commutative Banach algebra and $\mathcal{J}$ is a proper closed ideal in A, then $A/\mathcal{J}$ is a commutative Banach algebra.*

Proof. Let $\varphi : A \to A/\mathcal{J}$ be the usual quotient map. If $x \in \mathcal{J}$ then $\varphi(x) = \mathbf{0}$ so $\|\varphi(x)\| = 0$. If $x \notin \mathcal{J}$ then the fact that $\mathcal{J}$ is closed implies that $\|\varphi(x)\| > 0$. Clearly $\|\lambda\varphi(x)\| = |\lambda|\|\varphi(x)\|$. Thus if $x_1, x_2 \in A$ and $\epsilon > 0$ then there exists $y_1, y_2 \in \mathcal{J}$ such that

$$\|x_1 + y_1\| < \|\varphi(x_1)\| + \epsilon$$

and

$$\|x_2 + y_2\| < \|\varphi(x_2)\| + \epsilon\,.$$

Hence

$$\|\varphi(x_1 + x_2)\| \leq \|x_1 + x_2 + y_1 + y_2\| < \|\varphi(x_1)\| + \|\varphi(x_2)\| + 2\epsilon\,.$$

This gives the triangle inequality on the quotient space and proves (a).

Suppose that A is complete and let $\{\varphi(x_n)\}$ be a Cauchy sequence in $A/\mathcal{J}$. Then there is a subsequence $\{x_{n_j}\}$ so that

$$\|\varphi(x_{n_j}) - \varphi(x_{n_{j+1}})\| < 2^{-j}$$

for each positive j. Also there exist elements $z_j \in A$ so that $z_j - x_{n_j} \in \mathcal{J}$ and $\|z_j - z_{j+1}\| < 2^{-j}$. Thus $\{z_j\}$ is a Cauchy sequence in A. Since A is complete, there exists a $z \in A$ so that $\|z_j - z\| \to 0$. Therefore $\varphi(x_{n_j})$ converges to $\varphi(z)$ in $A/\mathcal{J}$. But, if a Cauchy sequence has a convergent subsequence, then the full sequence converges. So $A/\mathcal{J}$ is complete. We have proved (b).

To verify (c), select $x_1, x_2 \in A$ and $\epsilon > 0$ and choose $y_1, y_2 \in \mathcal{J}$ so that

$$\|x_j + y_j\| < \|\varphi(x_j)\| + \epsilon \tag{5}$$

holds for $j = 1, 2$. Notice that $(x_1 + y_1) \cdot (x_2 + y_2) \in x_1x_2 + \mathcal{J}$ so that

$$\|\varphi(x_1x_2)\| \leq \|(x_1 + y_1)(x_2 + y_2)\| \leq \|x_1 + y_1\| \cdot \|x_2 + y_2\|\,.$$

Line (5) tells us that

$$\|\varphi(x_1x_2)\| \leq \|\varphi(x_1)\| \cdot \|\varphi(x_2)\|\,.$$

If e is the unit in A, then we take $x_1 \notin \mathcal{J}$ and $x_2 = e$ in this last displayed line. This yields $\|\varphi(e)\| \geq 1$. But $e \in \varphi(e)$, and the definition of the quotient norm now shows us that $\|\varphi(e)\| \leq \|e\| = 1$. So $\|\varphi(e)\| = 1$. □

The next result sums up many of our key ideas. We have provided some separate arguments elsewhere, but repeat some of the key notions here.

THEOREM 5.13. *Let A be a Banach algebra. Then*

(a) *Every maximal ideal $\mathcal{M}$ of A is the kernel of some multiplicative linear functional on A,*

(b) *$\lambda \in \sigma(x)$ if and only if $f(x) = \lambda$ for some multiplicative linear functional f on A,*

(c) *An element x is invertible in A if and only if $f(x) \neq 0$ for every multiplicative linear functional f,*

(d) *$f(x) \in \sigma(x)$ for every $x \in A$ and every multiplicative linear functional f,*

(e) *$|f(x)| \leq \rho(x) \leq \|x\|$ for every $x \in A$ and every multiplicative linear functional f.*

Proof. Many arguments in Banach algebra theory use the ideas in this proof. It is well worth mastering. The ultimate resource in Banach algebra theory is [RIC]. A more modern text is [KAN].

If $\mathcal{M}$ is a maximal idea of A, then $A/\mathcal{M}$ is a field. Since $\mathcal{M}$ is closed, $A/\mathcal{M}$ is a Banach algebra. Let φ be the quotient map. The Gelfand-Mazur theorem then tells us that there is an isomorphism j of $A/\mathcal{M}$ onto the complex field $\mathbb{C}$. If we set $h = j \circ \varphi$, then h is a multiplicative linear functional on A and the kernel of h is $\mathcal{M}$ itself. This proves (a).

If $\lambda \in \sigma(x)$, then $x - \lambda e$ is not invertible. Therefore the set of all elements $(x - \lambda e)y$, where $y \in A$, is a proper ideal in A. This ideal lies in some maximal ideal $\mathcal{N}$ of A. By part **(a)**, there is a multiplicative linear functional h so that $h(x - \lambda e) = 0$. Since $h(e) = 1$, we see that $h(x) = \lambda$. That proves half of (b).

If instead $\lambda \notin \sigma(x)$, then there is a $y \in A$ so that $(x - \lambda e)y = e$. It follows that $h(x - \lambda e)h(y) = 1$ for every multiplicative linear functional h on A. Thus $h(x - \lambda e) \neq 0$ or $h(x) \neq \lambda$. That proves the rest of (b).

Since x is invertible if and only if $0 \notin \sigma(x)$, we see that (c) follows from (b).

Finally, (d) and (e) are immediate consequences of (b). □

REMARK. That the kernel of a multiplicative linear functional is a maximal ideal is now obvious. Thus the terminology "maximal ideal space" is fully justified.

The last general result for Banach algebras that we present is a powerful structure theorem. We begin with a little terminology.

A Banach algebra A equipped with an involution $x \mapsto x^*$ that satisfies

$$\|xx^*\| = \|x\|^2$$

for every $x \in A$ is called a B^*-algebra.

It holds that

$$\|x\|^2 = \|xx^*\| \leq \|x\|\|x^*\|$$

so that $\|x\| \leq \|x^*\|$. Also

$$\|x^*\| \leq \|x^{**}\| = \|x\| .$$

Thus

$$\|x\| = \|x^*\| .$$

It also follows that

$$\|xx^*\| = \|x\|\|x^*\| .$$

Let A be a Banach algebra and $\triangle$ the maximal ideal space. The mapping that assigns to each $x \in A$ a function $\widehat{x} : \triangle \to \mathbb{C}$ by way of the formula

$$\widehat{x}(h) = h(x)$$

is called the *Gelfand transform*. The Gelfand transform enables us to realize any Banach algebra A as an algebra of continuous functions on a compact Hausdorff space (the compact Hausdorff space being $\triangle$ itself equipped with the weak-$*$ topology). The Banach-Alaoglu theorem guarantees that $\triangle$ is compact.

The following theorem is generally considered to be one of the most central and profound results of this elegant theory.

THEOREM 5.14 (Gelfand-Naimark). *Let A be a commutative B^* algebra with maximal ideal space $\triangle$. The Gelfand transform is then an isometric isomorphism of A onto $C(\triangle)$ that has the additional property that*

$$h(x^*) = \overline{h(x)}$$

for $x \in A$ and $h \in \triangle$. [Here the overline denotes complex conjugation.] Equivalently, we assert that

$$\widehat{(x^*)} = \overline{\widehat{x}}$$

for $x \in A$. (Here $\widehat{\ }$ is the Gelfand transform: $\hat{x}(h) = h(x)$ for $x \in A$ and $h \in \triangle$.) In particular, x is hermitian (that is, $x^* = x$) if and only if $\hat{x}$ is a real-valued function.

REMARK. The theorem tells us that that virtually any Banach algebra is actually the continuous functions on some compact Hausdorff space. This gives a helpful way to think of a Banach algebra.

Many important theorems in analysis, such as the Stone-Weierstrass theorem, are formulated on compact Hausdorff spaces.

Proof of the Theorem. Assume first that $u \in A$ and $u = u^*$ (we will have an algebraic trick below for handling the more general case). Let $h \in \triangle$. We need to prove that $h(u)$ is real. For t real, set $z = u + ite$. Write $h(u) = \alpha + i\beta$, with α and β real. Then

$$h(z) = \alpha + i(\beta + t) \quad \text{and} \quad zz^* = u^2 + t^2 .$$

Thus

$$\alpha^2 + (\beta + t)^2 = |h(z)|^2 \leq \|z\|^2 = \|zz^*\| \leq \|u\|^2 + t^2 .$$

In other words,

$$\alpha^2 + \beta^2 + 2\beta t \leq \|u\|^2$$

for $t \in \mathbb{R}$. This last line tells us (since it is true for all $t \in \mathbb{R}$) that $\beta = 0$. So $h(u)$ is real.

If $x \in A$, then $x = u + iv$ with $u = u^*$ and $v = v^*$. Thus $x^* = u - iv$. Since $\hat{u}$ and $\hat{v}$ are real, the second statement of the theorem is thus proved.

Thus we see that $\widehat{A}$ is closed under complex conjugation. By the Stone-Weierstrass theorem, $\widehat{A}$ is therefore dense in $C(\triangle)$.

If $x \in A$ and $y = xx^*$, then $y = y^*$. Hence $\|y^2\| = \|y\|^2$. It follows, by induction on n, that $\|y^m\| = \|y\|^m$ for $m = 2^n$. Therefore $\|\hat{y}\|_\infty = \|y\|$ by the spectral radius theorem and part (e) of Theorem 5.13. Since $y = xx^*$, the second displayed line of the theorem tells us that $\hat{y} = |\hat{x}|^2$. Hence

$$\|\hat{x}\|_\infty^2 = \|\hat{y}\|_\infty = \|y\| = \|xx^*\| = \|x\|^2 ,$$

so that $\|\hat{x}\|_\infty = \|x\|$. Thus $x \leftrightarrow \hat{x}$ is an isometry. So $\widehat{A}$ is closed in $C(\triangle)$. Since $\widehat{A}$ is also dense in $C(\triangle)$, we conclude that $\widehat{A} = C(\triangle)$. $\square$

REMARK. The continuous functions on a compact, Hausdorff space are fairly easy to understand. But many Banach algebras are still quite subtle. For instance, there are many questions about $H^\infty(D)$ or $H^\infty(B)$ (where B is the unit ball in $\mathbb{C}^n$) for which we are nowhere near to having an answer.

5.4 THE WIENER TAUBERIAN THEOREM

Norbert Wiener was one of the most powerful analysts of his day. In the celebrated paper [WIE], he used everything but the kitchen sink to establish certain rather deep convergence results for Fourier series (these are called "Tauberian theorems" because they entail one kind of convergence implying another kind of convergence). It really put the theory of Banach algebras on the map when Gelfand was able to use his new ideas to prove one of Wiener's profound results in just a few lines. We present this proof here.

Before we state the theorem and prove it, we introduce some terminology and notation.

Let A denote the space of all complex functions f on the unit circle $\mathbb{T}$ which have absolutely convergent Fourier series. So

$$f(e^{i\theta}) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta}$$

with

$$\sum_n |c_n| < \infty . \tag{6}$$

The expression (6) is the norm on this space. Clearly A is a Banach space. In fact, in a natural fashion, A is isometrically isomorphic to ℓ^1.

We also note that A is a commutative Banach algebra under pointwise multiplication. For, if $f \in A$ with $f(e^{i\theta}) = \sum_n c_n e^{in\theta}$ and $g \in A$ with $g(e^{i\theta}) = \sum_n b_n e^{in\theta}$, then

$$f(e^{i\theta}) \cdot g(e^{i\theta}) = \sum_n \left(\sum_k c_{n-k} b_k \right) e^{in\theta} .$$

Therefore

$$\|fg\| = \sum_n \left| \sum_k c_{n-k} b_k \right| \le \sum_k |b_k| \sum_n |c_{n-k}| = \|f\| \cdot \|g\| .$$

Furthermore, the function that is identically 1 is the unit of A, and $\|1\| = 1$.

THEOREM 5.2. *Suppose that the continuous function f on the circle group has absolutely convergent Fourier series*

$$f(e^{i\theta}) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} ,$$

so that

$$\sum_n |c_n| < \infty .$$

Assume further that f vanishes nowhere. Then

$$\frac{1}{f(e^{i\theta})} = \sum_{n=-\infty}^{\infty} \gamma_n e^{in\theta}$$

with

$$\sum_n |\gamma_n| < \infty .$$

So the reciprocal of a nonvanishing function with absolutely convergent Fourier series also has absolutely convergent Fourier series.

Proof. Let A be as above. Set $f_0(e^{i\theta}) = e^{i\theta}$. Then $f_0 \in A$, $1/f_0 \in A$, and $\| f_0^n \| = 1$ for every integer n. If h is any multiplicative linear functional of A and $h(f_0) = \lambda$, then the fact that $\|h\| \leq 1$ implies that

$$|\lambda^n| = |h(f_0^n)| \leq \| f_0^n \| = 1$$

for any integer n, hence $|\lambda| = 1$. Thus, by the Gelfand-Naimark theorem, to each h there corresponds a point $e^{i\alpha}$ in the circle group such that $h(f_0) = e^{i\alpha}$. Hence

$$h(f_0^n) = e^{in\alpha} = f_0^n(e^{i\alpha})$$

for any integer n.

If f is given by

$$f(e^{i\theta}) = \sum_{n=-\infty}^{\infty} c_n e^{in\theta} ,$$

then $f = \sum_n c_n f_0^n$. This series converges in the topology of A. Since h is a continuous linear functional on A, we conclude that

$$h(f) = f(e^{i\alpha}) .$$

Our hypothesis that f vanishes at no point of the circle group then says that f is not in the kernel of any multiplicative linear functional of A. And now Theorem 5.13(c) guarantees that f is invertible in A. □

CHAPTER 6

Topological Vector Spaces

6.1 Basic Ideas

The most basic mathematical structure for functional analysis is that of a Banach space. We have learned that the richer structure of Hilbert space can lead to greater depth and insight. Certainly the most elementary structure—more primitive than either Banach space or Hilbert space, but which still yields useful results—is that of topological vector space. We shall discuss the basic ideas here.

Convexity was implicit in much of what we did with Banach and Hilbert spaces. In the current presentation, convexity will play a more explicit role. We say that a set E in a linear space is *convex* if, whenever $x, y \in E$ and $0 \le t \le 1$, then

$$(1-t)x + ty \in E\,.$$

This condition simply says that the segment connecting x and y in E also lies in E.

In this chapter and what follows, we will make use of multi-index notation. Fix attention on a Euclidean space $\mathbb{R}^n$. A *multi-index* α on $\mathbb{R}^n$ is an n-tuple of nonnegative integers: $\alpha = (a_1, a_2, \ldots, a_n)$. We let

$$x^\alpha = x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \cdots \cdot x_n^{\alpha_n}\,,$$

$$\frac{\partial^\alpha}{\partial x^\alpha} = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x^{\alpha_n}}\,.$$

Other variants of this notation will be used as needed.

DEFINITION. A *topological vector space* X is a vector space (over a field k, which is usually either $\mathbb{R}$ or $\mathbb{C}$) which is endowed with a topology so that the maps

$$\begin{aligned}(x, y) &\mapsto x + y \\ (\lambda, x) &\mapsto \lambda x\end{aligned}$$

are continuous from $X \times X$ to X or from $k \times X$ to X respectively.

We say that a topological vector space (often abbreviated TVS) is *locally convex* if there is a basis for the topology consisting of convex sets. Most any TVS that we shall encounter will be locally convex and Hausdorff.[1]

DEFINITION. Let X be a linear space. A *seminorm* on X is a function $x \mapsto \rho(x)$, from X to $[0, \infty)$, such that

(a) $\rho(x + y) \le \rho(x) + \rho(y)$ for all $x, y, \in X$,

(b) $\rho(\lambda x) = |\lambda|\rho(x)$ for all $x \in X$, $\lambda \in k$.

The difference between a seminorm and a norm is that, in the former, $\rho(x) = 0$ does not imply that $x = 0$.

We use seminorms to generate a subbasis for a topology as follows:

THEOREM 6.1. *Let $\{\rho_\alpha\}_{\alpha \in A}$ be a family of seminorms on a linear space X. If $x \in X$, $\alpha \in A$, and $\epsilon > 0$, let us define*

$$U_{x,\alpha,\epsilon} = \{y \in X : \rho_\alpha(y - x) < \epsilon\} .$$

Let $\mathcal{T}$ be the topology generated by the $U_{x,\alpha,\epsilon}$ (that is, we think of the collection of all $U_{x,\alpha,\epsilon}$ as a subbasis for the topology). We have these properties of the topology:

(i) *For each $x \in X$, the collection of finite intersections of sets $U_{x,\alpha,\epsilon}$ forms a neighborhood basis at x.*

(ii) *If $\{x_j\}_{j \in J}$ is a net[2] in X, then $x_j \to x$ if and only if $\rho_\alpha(x_j - x) \to 0$ for all $\alpha \in A$.*

(iii) *The pair $(X, \mathcal{T})$ is a locally convex TVS.*

[1] We note that any Banach or Hilbert space is automatically locally convex and Hausdorff. The basis for the topology in that case is just the metric balls.

[2] Here a *net* is a mapping from a directed set into X. The idea of net generalizes that of sequence. See [KRA5], [KRA6].

Proof. For part (i), let $x \in \cap_1^k U_{x_j,\alpha_j,\epsilon_j}$. Then let $\delta_j = \epsilon_j - \rho_\alpha(x - x_j)$. By the triangle inequality, we have $x \in \cap_1^k U_{x,\alpha_j,\delta_j} \subset \cap_1^k U_{x_j,\alpha_j,\epsilon_j}$. Thus the assertion follows from the definition of subbasis.

For part (ii) it suffices, in view of (i), to notice that $\rho_\alpha(x_j - x) \to 0$ if and only if $\{x_j\}$ is eventually in $U_{x,\alpha,\epsilon}$ for every $\epsilon > 0$.

For part (iii), we see that the continuity of the vector operations follows from the definition of continuity in terms of nets and from part (ii). Indeed, if $x_j \to x$ and $y_j \to y$, then

$$\rho_\alpha((x_j + y_j) - (x + y)) \le \rho_\alpha(x_j - x) + \rho_\alpha(y_j - y) \to 0 .$$

Hence $x_j + y_j \to x + y$. If also $\lambda_j \to \lambda$, then eventually $|\lambda_j| \le C \equiv |\lambda| + 1$. So

$$\begin{aligned}\rho_\alpha(\lambda_j x_j - \lambda x) &\le \rho_\alpha(\lambda_j(x_j - x)) + \rho_\alpha((\lambda_j - \lambda)x) \\ &\le C \cdot \rho_\alpha(x_j - x) + |\lambda_j - \lambda|\rho_\alpha(x) .\end{aligned}$$

It then follows that $\lambda_j x_j \to \lambda x$. Furthermore, the sets $U_{x,\alpha,\epsilon}$ are convex. For if $y, z \in U_{x,\alpha,\epsilon}$, then

$$\begin{aligned}\rho_\alpha(x - [ty + (1-t)z]) &\le \rho_\alpha(tx - ty) + \rho_\alpha((1-t)x + (1-t)z) \\ &< t\epsilon + (1-t)\epsilon = \epsilon .\end{aligned}$$

The local convexity of the topology thus follows from (i). □

We have

PROPOSITION 6.2. *Let X and Y be vector spaces with topologies defined, respectively, by families $\{\rho_\alpha\}_{\alpha \in A}$ and $\{\mu_\beta\}_{\beta \in B}$ of seminorms. Let $T : X \to Y$ be a linear map. Then T is continuous if and only if, for each $\beta \in B$, there exist $\alpha_1, \dots, \alpha_k \in A$ and $C > 0$ such that*

$$\mu_\beta(Tx) \le C \cdot \sum_1^k \rho_{\alpha_j}(x) .$$

REMARK. One of our main applications of topological vector spaces will be to the theory of distributions. In that context, the condition given in this proposition will be the most useful means of verifying the continuity of a linear mapping.

Proof of the Proposition. If the condition

$$\mu_\beta(Tx) \le C \cdot \sum_1^k \rho_{\alpha_j}(x) .$$

holds and if $\langle x_j \rangle$ is a net converging to $x \in X$, then by part (ii) of the last theorem we have $\rho_\alpha(x_j - x) \to 0$ for all α. Hence $\mu_\beta(Tx_j - Tx) \to 0$ for all β. Thus $Tx_j \to Tx$. We conclude that T is continuous.

Conversely, if T is continuous, then for every $\beta \in B$ there is a neighborhood U of 0 in X such that $\mu_\beta(Tx) < 1$ for $x \in U$. Part (i) of the last theorem tells us that we may assume that $U = \cap_1^k U_{x,\alpha_j,\epsilon_j}$. Let $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_k\}$. Then $\mu_\beta(Tx) < 1$ whenever $\rho_{\alpha_j}(x) < \epsilon$ for all j.

Given $x \in X$, there are now two possibilities: If $\rho_{\alpha_j}(x) > 0$ for some j, then let $y = \epsilon x / \sum_1^k \rho_{\alpha_j}(x)$. So $\rho_{\alpha_j}(y) < \epsilon$ for all j. Hence

$$\mu_\beta(Tx) = \sum_1^k \epsilon^{-1} \rho_{\alpha_j}(x) \mu_\beta(Ty) \le \epsilon^{-1} \sum_1^k \rho_{\alpha_j}(x) .$$

If instead $\rho_{\alpha_j}(x) = 0$ for all j, then $\rho_{\alpha_j}(rx) = 0$ for all j and all $r > 0$. Hence $r\mu_\beta(Tx) = \mu_\beta(T(rx)) < 1$ for all $r > 0$. Therefore $\mu_\beta(Tx) = 0$. Then $\mu_\beta(Tx) \le \epsilon^{-1} \sum_1^k \rho_{\alpha_j}(x)$ in this case as well. $\square$

Exercises for the Reader Let X be a vector space equipped with the topology induced by a family $\{\rho_\alpha\}_{\alpha \in A}$ of seminorms.

1. The linear space X is Hausdorff if and only if, for each $x \neq 0$ in X, there is an $\alpha \in A$ such that $\rho_\alpha(x) \neq 0$.

2. If X is Hausdorff and A is countable, then X is metrizable with a translation-invariant metric.[3]

6.2 FRECHET SPACES

DEFINITION. A complete Hausdorff TVS whose topology is defined by a countable family of seminorms is called a *Fréchet space*.

It is important now to look at several examples of topological vector spaces.

EXAMPLE. Consider the space of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. For each compact $K \subseteq \mathbb{R}$, define the seminorm

$$\rho_K(f) = \sup_{x \in K} |f(x)| .$$

[3] Here a metric d is *translation invariant* if $d(x, y) = d(x + z, y + z)$ for all $x, y, z \in X$.

An equivalent, and somewhat simpler, way to define the topology is with the countable family of seminorms

$$\mu_j(f) = \sup_{x \in [-j,j]} |f(x)|$$

for $j = 1, 2, 3, \ldots$.

We see that this space is complete, so it is a Fréchet space.

EXAMPLE. Consider the space of measurable functions on $\mathbb{R}$ that are locally integrable (i.e., integrable on compact sets). Define the seminorms, for $k = 1, 2, \ldots$,

$$\rho_k(f) = \int_{\{x:|x|\leq k\}} |f(x)|\, dx\,.$$

This gives a Fréchet space (just because L^1 is complete).

EXAMPLE. Fix a positive integer k and consider the C^k functions on an open set $U \subseteq \mathbb{R}^N$. For $K \subseteq U$ compact and $0 \leq j \leq k$, define seminorms

$$\mu_{j,K}(f) \equiv \sup_{x \in K, |\alpha| \leq j} \left| \frac{\partial^\alpha}{\partial x^\alpha} f(x) \right|\,.$$

Then it is not difficult to see that this is a Fréchet space.

CHAPTER 7

Distributions

7.1 Preliminary Remarks

The idea of generalized function has roots in the nineteenth century. Even more than 100 years ago, mathematicians wanted a way to say that a function satisfies a differential equation in some "weak" sense. The impetus was to develop a notion of function that is not conceived "point by point" as we usually do. This set of ideas developed further traction in the twentieth century, especially because of various questions in harmonic and functional analysis. It was in 1950 that Laurent Schwartz wrote his definitive book [SCH] enunciating the theory of distributions (or generalized functions).

The key insight for freeing the notion of function from a point-by-point realization is to think of a function as an element of a dual space. We already do this with L^p functions (which are equivalence classes, and certainly not defined point-by-point). There are different theories of distributions, depending on which space we are calculating the dual of.

Let U be an open set in $\mathbb{R}^N$. We consider $C_c^\infty(U)$ to be the collection of those functions f that are infinitely differentiable and supported in some compact subset K of U. Here the choice of K will depend on f. We equip C_c^∞ with the seminorms

$$\rho_{\alpha,K}(f) = \sup_K \left| \frac{\partial^\alpha}{\partial x^\alpha} f \right|$$

for α a multi-index and K compact in U. It is quite common in this subject to denote $C_c^\infty(U)$ by the symbol $\mathcal{D}(U)$.

Definition. The topology on $C_c^\infty(U)$ has these features:

(a) A sequence $\{\phi_j\}$ in $C_c^\infty(U)$ *converges in* C_c^∞ to ϕ if $\{\phi_j\} \subseteq C_c^\infty(K)$ for some compact set $K \subseteq U$ and $\phi_j \to \phi$ in the topology of $C_c^\infty(K)$.

That is to say, we require that $(\partial^\alpha/\partial x^\alpha)\phi_j \to (\partial^\alpha/\partial x^\alpha)\phi$ uniformly for all α.

(b) If X is a locally convex topological vector space and $T : C_c^\infty(U) \to X$ is a linear map, then T is *continuous* if T restricted to $C_c^\infty(K)$ is continuous for each compact $K \subseteq U$. That is, it is continuous if $T\phi_j \to T\phi$ whenever $\phi_j \to \phi$ in $C_c^\infty(K)$ and each $K \subseteq U$ is compact.

(c) A linear mapping $T : C_c^\infty(U) \to C_c^\infty(U')$ is *continuous* if, for each compact $K \subseteq U$, there is a compact $K' \subseteq U'$ so that $T(C_c^\infty(K)) \subseteq C_c^\infty(K')$ and T is continuous from $C_c^\infty(K)$ to $C_c^\infty(K')$.

7.2 What is a Distribution?

DEFINITION. A *distribution* on U is a continuous linear functional on $C_c^\infty(U)$. The space of all distributions on U is denoted by $\mathcal{D}'(U)$. We set $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^N)$. We impose the weak-$*$ topology on $\mathcal{D}'(U)$.

It is important to have several incisive examples of distributions. We have chosen these examples with two ideas in mind: (i) to show that ordinary functions are also distributions and (ii) to exhibit some distributions that are new "generalized functions" which are *not* ordinary functions.

EXAMPLE. Let f be a locally integrable function on $\mathbb{R}^n$. Then f may be paired with any C_c^∞ function φ by

$$\varphi \longmapsto \int f(x)\varphi(x)\,dx\,.$$

Thus f induces a distribution. Here we are taking U to be all of $\mathbb{R}^n$.

EXAMPLE. Consider the mapping

$$C_c^\infty(B) \ni \varphi \longmapsto \varphi(0)\,.$$

Here B is the unit ball of $\mathbb{R}^N$. This distribution may be represented by integration against the Dirac delta measure[1] δ.

EXAMPLE. As before, let B be the unit ball of $\mathbb{R}^N$. Consider the mapping

$$C_c^\infty(B) \ni \varphi \longmapsto \varphi(0) + \frac{\partial}{\partial x_1}\varphi(0)\,.$$

This is a distribution supported at the origin.

[1] The Dirac delta measure is the measure with total mass 1 at the origin and no mass elsewhere.

7.3 Operations on Distributions

We can come up with some more interesting examples of distributions by introducing some operations on distributions. That we now do.

We begin by noting that the pairing of a distribution α with a testing function ϕ can be denoted $\alpha(\phi)$, but it is also often denoted $\langle \alpha, \phi \rangle$.

Proposition 7.1. *Let $\{V_\alpha\}_{\alpha \in A}$ be a collection of open subsets of U, and let $V = \cup_\alpha V_\alpha$. If $\beta, \gamma \in \mathcal{D}'(U)$ and $\beta = \gamma$ on each V_α, then $\beta = \gamma$ on V.*

Proof. To say that $\beta = \gamma$ on V_α means simply that, whenever φ is C_c^∞ with compact support in V_α then $\beta(\varphi) = \gamma(\varphi)$.

Notice that, if $\varphi \in C_c^\infty(V)$ then, by compactness, there are $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that $\operatorname{supp}\varphi \subseteq \cup_{j=1}^k V_{\alpha_j}$. Choose a partition of unity $\psi_1, \psi_2, \ldots, \psi_k \in C_c^\infty$ such that $\operatorname{supp}\psi_j \subseteq V_{\alpha_j}$ for each j and $\sum_1^k \psi_j \equiv 1$ on $\operatorname{supp}\varphi$. Then

$$\langle \beta, \varphi \rangle = \sum \langle \beta, \psi_j \varphi \rangle = \sum \langle \gamma, \psi_j \varphi \rangle = \langle \gamma, \varphi \rangle . \qquad \square$$

If $\beta \in \mathcal{D}'(U)$, then there is a maximal open subset of U on which $\beta = 0$. We call the complement of this open set the *support of* β.

Example. Let B be the unit ball in $\mathbb{R}^N$ and define

$$\beta(\varphi) = \varphi(0)$$

for any $\varphi \in C_c^\infty(B)$. Then the support of β is $\{0\}$.

Example. For φ a C_c^∞ function on $\mathbb{R}$, set

$$\alpha(\varphi) = -\int_{-\infty}^0 \varphi(x)\,dx + \int_0^\infty \varphi(x)\,dx .$$

Then the support of α is $\mathbb{R}$.

Definition. Let $\gamma \in \mathcal{D}'(U)$ for some open set $U \subseteq \mathbb{R}^N$.

(a) If α is any multi-index, then we define $(\partial^\alpha/\partial x^\alpha)\gamma$ by

$$\langle (\partial^\alpha/\partial x^\alpha)\gamma, \varphi \rangle = (-1)^{|\alpha|} \langle \gamma, (\partial^\alpha/\partial x^\alpha)\varphi \rangle .$$

Clearly this definition is motivated by integration by parts.

(b) Given $\psi \in C^\infty(U)$, we define $\psi\gamma$ by

$$\langle \psi\gamma, \varphi \rangle = \langle \gamma, \psi\varphi \rangle .$$

(c) If $y \in \mathbb{R}^N$ and f is any function on $\mathbb{R}^N$ then we define the *translation* by y as $\tau_y f(x) = f(x-y)$. We set

$$\langle \tau_y \gamma, \varphi \rangle = \langle \gamma, \tau_{-y} \varphi \rangle .$$

(d) If S is an invertible linear transformation of $\mathbb{R}^N$, then we set

$$\langle \gamma \circ S, \varphi \rangle = |\det S|^{-1} \langle \gamma, \varphi \circ S^{-1} \rangle .$$

(e) If $\psi \in C_c^\infty$, then define $\widetilde{\psi}(x) = \psi(-x)$. We define the convolution of the distribution γ with ψ by

$$\gamma * \psi(x) = \langle \gamma, \tau_x \widetilde{\psi} \rangle .$$

It turns out that $\gamma * \psi$ is a C^∞ function (see Proposition 7.2 below).

(f) Let ψ be as in (e). Define $\gamma * \psi$ by

$$\langle \gamma * \psi, \varphi \rangle = \langle \gamma, \varphi * \widetilde{\psi} \rangle .$$

This definition is consistent with what we said in part **(e)**.

In what follows, we use the abbreviation ∂^α to denote $(\partial^\alpha / \partial x^\alpha)$ and ∂_j to denote $\partial / \partial x_j$. The next proposition will help to clarify the preceding definition.

PROPOSITION 7.2. *Suppose that $U \subseteq \mathbb{R}^N$ is open and $\psi \in C_c^\infty$. Let $V = \{x : x - y \in U \text{ for } y \in \operatorname{supp} \psi\}$. For $\gamma \in \mathcal{D}'(U)$ and $x \in V$, let $\gamma * \psi(x) = \langle \gamma, \tau_x \widetilde{\psi} \rangle$. Then*

(a) $\gamma * \psi \in C^\infty(V)$,

(b) $\partial^\alpha(\gamma * \psi) = [\partial^\alpha \gamma] * \psi = \gamma * [\partial^\alpha \psi]$,

(c) *For any $\varphi \in C_c^\infty(V)$, $\int (\gamma * \psi)\varphi = \langle \gamma, \varphi * \widetilde{\psi} \rangle$.*

Proof. Let $e_1, e_2, \ldots, e_N$ be the standard basis for $\mathbb{R}^N$. If $x \in V$, then there exists $t_0 > 0$ such that $x + te_j \in U$ for $|t| < t_0$. We can then easily verify that

$$t^{-1}(\tau_{x+te_j} \widetilde{\psi} - \tau_x \widetilde{\psi}) \to \tau_x \widetilde{\partial_j \psi} \text{ in } C_c^\infty(U) \text{ as } t \to 0 .$$

It follows that $\partial_j(\gamma * \psi)(x)$ exists and equals $\gamma * \partial_j \psi(x)$. By induction then, $\gamma * \psi \in C^\infty(V)$ and $\partial^\alpha(\gamma * \psi) = \gamma * \partial^\alpha \psi$. Furthermore, since $\partial^\alpha \widetilde{\psi} = (-1)^{|\alpha|} \widetilde{\partial^\alpha \psi}$ and $\partial^\alpha \tau_x = \tau_x \partial^\alpha$, we see that

$$\begin{aligned}(\partial^\alpha \gamma) * \psi(x) &= \langle \partial^\alpha \gamma, \tau_x \widetilde{\psi} \rangle = (-1)^{|\alpha|} \langle \gamma, \partial^\alpha \tau_x \widetilde{\psi} \rangle = \langle \gamma, \tau_x \widetilde{\partial^\alpha \psi} \rangle \\ &= \gamma * (\partial^\alpha \psi)(x) .\end{aligned}$$

For the next step, if $\varphi \in C_c^\infty(V)$, we have

$$\varphi * \widetilde{\psi}(x) = \int \varphi(y)\psi(y-x)\,dy = \int \varphi(y)\tau_y\widetilde{\psi}(x)\,dy\,.$$

The integrand here is continuous and supported in a compact subset of U, so the integral can be approximated by Riemann sums. That is to say, for each large $m \in \mathbb{N}$, we can approximate $\operatorname{supp}\varphi$ by a union of cubes of side length 2^{-m} centered at points $y_1^m, y_2^m, \ldots, y_{k(m)}^m \in \operatorname{supp}\varphi$. Then the corresponding Riemann sums

$$S^m \equiv 2^{-Nm}\sum_j \varphi(y_j^m)\tau_{y_j^m}\widetilde{\psi}$$

are supported in a common compact subset of U and converge uniformly to $\varphi * \widetilde{\psi}$ as $m \to \infty$. Likewise,

$$\partial^\alpha S^m = 2^{-Nm}\sum_j \varphi(y_j^m)\tau_{y_j^m}\partial^\alpha\widetilde{\psi}$$

converges uniformly to $\varphi * \partial^\alpha\widetilde{\psi} = \partial^\alpha(\phi * \widetilde{\psi})$. Hence $S^m \to \varphi * \widetilde{\psi}$ in $C_c^\infty(U)$. Thus we have

$$\begin{aligned}
\langle \gamma, \varphi * \widetilde{\psi}\rangle &= \lim_{m\to\infty}\langle \gamma, S^m\rangle \\
&= \lim_{m\to\infty} 2^{-Nm}\sum_j \varphi(y_j^m)\langle \gamma, \tau_{y_j^m}\widetilde{\psi}\rangle \\
&= \int \varphi(y)\langle \gamma, \tau_y\widetilde{\psi}\rangle\,dy \\
&= \int \varphi(y)\gamma * \psi(y)\,dy\,. \qquad \square
\end{aligned}$$

7.4 Approximation of Distributions

Typical distributions are the Dirac delta mass, and one or more derivatives of the Dirac delta mass. These are highly singular objects—as we might expect generalized functions to be. But it is a nice fact that distributions may be approximated—in a suitable topology—by smooth functions.

LEMMA 7.3. *On $\mathbb{R}^N$, suppose that $\psi \in C_c^\infty$ and that $\int \psi = 1$. Set*

$$\psi_t(x) = t^{-N}\psi(x/t)\,.$$

Let $\varphi \in C_c^\infty$. Then

$$\psi_t * \varphi \to \varphi$$

uniformly as $t \to 0$.

Proof. Let $\epsilon > 0$. Choose $\delta > 0$ small enough that $|\varphi(s) - \varphi(u)| < \epsilon$ when $|s - u| < \delta$. We calculate that

$$\begin{aligned} |\psi_t * \varphi(x) - \varphi(x)| &= \left| \int \psi_t(y)\varphi(x-y)\,dy - \varphi(x) \right| \\ &= \left| \int \psi_t(y)\,[\varphi(x-y) - \varphi(x)]\,\,dy \right| \\ &\le \int |\psi_t(y)|\epsilon\,dy \end{aligned}$$

provided that t is small enough (so that $x - y$ and x are close enough together). This last equals

$$\epsilon \int |\psi(y)|\,dy$$

by a simple change of variable. Finally this equals $C\epsilon$. □

LEMMA 7.4. *On $\mathbb{R}^N$, suppose that $\varphi \in C_c^\infty$, $\psi \in C_c^\infty$, and $\int \psi = 1$. Let*

$$\psi_t(x) = t^{-N}\psi(x/t)\,.$$

Then

(a) *Given any neighborhood U of $\operatorname{supp}\varphi$, we have $\operatorname{supp}(\varphi * \psi_t) \subseteq U$ for t sufficiently small.*

(b) *$\varphi * \psi_t \to \varphi$ in C_c^∞ as $t \to 0$.*

Proof. Suppose that $\operatorname{supp}\psi \subseteq \{x : |x| \le R\}$. Then $\operatorname{supp}(\varphi * \psi_t)$ is contained in the set of points whose distance from $\operatorname{supp}\varphi$ is at most tR. This is included in a fixed compact set if $t \le 1$ and is included in U if t is small. Further, $\partial^\alpha(\varphi * \psi_t) = (\partial^\alpha\varphi) * \psi_t \to \partial^\alpha\varphi$ uniformly as $t \to 0$ by the preceding lemma. □

PROPOSITION 7.5. *For any open set $U \subseteq \mathbb{R}^N$, $C_c^\infty(U)$ is dense in $\mathcal{D}'(U)$ in the topology of $\mathcal{D}'(U)$.*

Proof. Let $\gamma \in \mathcal{D}'(U)$. We shall approximate γ by distributions supported in compact subsets of U, and then approximate the latter by functions in $C_c^\infty(U)$.

Let $\{V_j\}$ be an increasing sequence of precompact open subsets of U whose union is U. For each j we can pick $\zeta_j \in C_c^\infty(U)$ such that $\zeta_j = 1$ on $\overline{V_j}$. Given $\varphi \in C_c^\infty(U)$, for j sufficiently large we have $\operatorname{supp}\varphi \subseteq V_j$. Therefore $\langle \gamma, \varphi \rangle = \langle \gamma, \zeta_j\varphi \rangle = \langle \zeta_j\gamma, \varphi \rangle$. Thus $\zeta_j\gamma \to \gamma$ as $j \to \infty$.

Since $\operatorname{supp}\zeta_j$ is compact, $\zeta_j\gamma$ can be regarded as a distribution on $\mathbb{R}^N$. Let ψ, ψ_t be as in the preceding lemma. Set $\widetilde{\psi}(x) = \psi(-x)$. Then $\int \widetilde{\psi} = 1$. Thus, given $\varphi \in C_c^\infty$, we have that $\varphi * \widetilde{\psi}_t \to \varphi$ in C_c^∞. But then Lemma 7.4 tells us that $(\zeta_j\gamma) * \psi_t \in C^\infty$ and $\langle(\zeta_j\gamma) * \psi_t, \varphi\rangle = \langle\zeta_j\gamma, \varphi * \widetilde{\psi}_t\rangle \to \langle\zeta_j\gamma, \varphi\rangle$. Thus $(\zeta_j\gamma) * \psi_t \to \zeta_j\gamma$ in $\mathcal{D}'$. In summary, every neighborhood of γ in $\mathcal{D}'(U)$ contains the C^∞ functions $(\zeta_j\gamma) * \psi_t$ for j large and t small.

Finally note that $\operatorname{supp}\zeta_j \subseteq V_k$ for some k. If $\operatorname{supp}\varphi \cap \overline{V}_k = \emptyset$, then for sufficiently small t we know that $\operatorname{supp}(\varphi * \widetilde{\psi}_t) \cap \overline{V}_k = \emptyset$. Hence $\langle(\zeta_j\gamma) * \psi_t, \varphi\rangle = \langle\gamma, \zeta_j(\varphi * \widetilde{\psi}_t)\rangle = 0$. Thus

$$\operatorname{supp}((\zeta_j\gamma) * \psi_t) \subseteq \overline{V}_k \subseteq U . \qquad \square$$

We conclude this discussion with a simple but entertaining example.

EXAMPLE. Let

$$h(x) = \begin{cases} 0 & \text{if} \quad x \le 0 \\ 1 & \text{if} \quad x > 0 . \end{cases}$$

We wish to calculate h'. We see that

$$\begin{aligned} \langle h', \varphi\rangle &= -\langle h, \varphi'\rangle \\ &= -\int_{-\infty}^0 0 \cdot \varphi'(x)\,dx - \int_0^\infty 1 \cdot \varphi'(x)\,dx \\ &= -0 - \Big[\varphi(\infty) - \varphi(0)\Big] \\ &= \varphi(0) . \end{aligned}$$

Thus

$$h' = \delta ,$$

where δ is the Dirac delta mass at the origin. The function h is known as the *Heaviside function*.

7.5 THE FOURIER TRANSFORM

One important operation from analysis, that we would certainly want to have working with distributions, is the Fourier transform. To make such a theory work, we would need a set of test functions (analogous to C_c^∞) that is preserved under the Fourier transform (see [KRA3]). Certainly C_c^∞ is not so preserved. The Paley-Wiener theorem tells us that the Fourier transform of a C_c^∞ function *cannot* be compactly supported.

We need another space of testing functions. What works best is the so-called *Schwartz space* $\mathcal{S}$. These are C^∞ functions f on $\mathbb{R}^N$ with the property that

$$\rho_{\alpha,\beta}(f) \equiv \sup |x^\alpha \partial^\beta f|$$

is finite for every choice of α and β. It is elementary to check that $\mathcal{S}$ is in fact preserved under the Fourier transform. If we equip $\mathcal{S}$ with the topology coming from the seminorms $\rho_{\alpha,\beta}$, then the Fourier transform is bicontinuous. Then we take the *Schwartz distributions* to be the dual space of $\mathcal{S}$. We denote this space of distributions by $\mathcal{S}'$. If $\alpha \in \mathcal{S}'$, then we define $\widehat{\alpha}$ by

$$\langle \widehat{\alpha}, \varphi \rangle \equiv \langle \alpha, \widehat{\varphi} \rangle .$$

This definition is inspired by the Plancherel formula and the Poisson summation formula.

EXAMPLE. Let us calculate the Fourier transform of the Dirac delta mass δ at the origin. We have

$$\langle \widehat{\delta}, \varphi \rangle = \langle \delta, \widehat{\varphi} \rangle = \widehat{\varphi}(0) .$$

Of course $\widehat{\varphi}(0)$ (where φ is a testing function) is nothing other than $\int \varphi \, dx$. So we see that

$$\widehat{\delta}(\varphi) = \int \varphi(x) \, dx .$$

CHAPTER 8

Spectral Theory

8.1 Background

Certainly one of the premier theorems in all of functional analysis is the spectral theorem. It says that, on Hilbert space H, any reasonable bounded linear operator can be represented as multiplication by an L^∞ function (acting on L^2). Much of the modern theory, especially the theory of normal operators, depends critically on the spectral theorem.

There are many versions of the spectral theorem, both for bounded and for unbounded operators. Here, in the spirit of simplicity, we concentrate on a basic version for bounded operators.

Some preliminary, background terminology is this:

DEFINITION. An operator $T \in \mathcal{B}(H)$ is said to be

(a) *normal* if $TT^* = T^*T$,

(b) *self-adjoint* if $T^* = T$,

(c) *unitary* if $T^*T = I = TT^*$,

(d) a *projection* if $T^2 = T$ and T is self-adjoint.

Some additional terminology: An *algebra* of sets (all of which are subsets of a given set X) is a collection of sets that is closed under finite union and complementation. A σ*-algebra* is an algebra that is closed under countable union.

DEFINITION. Let $\mathcal{M}$ be a σ-algebra on a set Ω. Let H be a Hilbert space. Then an i*-resolution* on $\mathcal{M}$ is a mapping

$$E : \mathcal{M} \to \mathcal{B}(H)$$

such that

(a) $E(\emptyset) = \mathbf{0}$, $E(\Omega) = I$,

(b) Each $E(\omega)$ is a self-adjoint projection,

(c) $E(\omega \cap \omega') = E(\omega)E(\omega')$,

(d) If $\omega \cap \omega' = \emptyset$, then $E(\omega \cup \omega') = E(\omega) + E(\omega')$,

(e) For every $x, y \in H$, the set function $E_{x,y}$ defined by

$$E_{x,y}(\omega) = \langle E(\omega)x, y\rangle$$

is a complex measure on $\mathcal{M}$.

It is quite common to take $\mathcal{M}$ to be the σ-algebra of all Borel sets in a compact or locally compact Hausdorff space. In that case we augment condition (e) to require that each $E_{x,y}$ be a *regular Borel measure*.

The properties we have enunciated for an i-resolution make the next two results nearly obvious (or see [RUD2, p. 302]):

PROPOSITION 8.1. *If E is an i-resolution and $x \in H$, then the mapping*

$$\omega \to E(\omega)x$$

is a countably additive, H-valued measure on $\mathcal{M}$.

PROPOSITION 8.2. *Let E be an i-resolution. If $\omega_j \in \mathcal{M}$ and $E(\omega_j) = 0$ for $j = 1, 2, 3, \ldots$, and if $\omega = \cup_{j=1}^{\infty}\omega_j$, then $E(\omega) = 0$.*

A useful, and intuitively appealing, preliminary result is this:

THEOREM 8.3. *If $f : H \times H \to \mathbb{C}$ is sesquilinear[1] and bounded, in the sense that*

$$M = \sup\{|f(x, y)| : \|x\| = \|y\| = 1\} < \infty\,, \tag{1}$$

then there exists a unique $S \in \mathcal{B}(H)$ that satisfies

$$f(x, y) = \langle x, Sy\rangle\,. \tag{2}$$

Furthermore, $\|S\| = M$.

[1]A quadratic form is sesquilinear if it is complex linear in the first entry and conjugate complex linear in the second entry.

Proof. Since $|f(x, y)| \leq M\|x\|\|y\|$, the mapping

$$x \mapsto f(x, y)$$

is, for each $y \in H$, a bounded linear functional on H. This functional has norm at most $M\|y\|$. Thus, to each element $y \in H$, there corresponds a unique element $Sy \in H$ such that (2) holds. Also $\|Sy\| \leq M\|y\|$.

Clearly $S : H \to H$ is additive. If $\alpha \in \mathbb{C}$, then

$$\langle x, S(\alpha y)\rangle = f(x, \alpha y) = \overline{\alpha} f(x, y) = \overline{\alpha}\langle x, Sy\rangle = \langle x, \alpha Sy\rangle$$

for all $x, y \in H$. Thus S is linear. So $S \in \mathcal{B}(H)$ and $\|S\| \leq M$.

We also know that

$$|f(x, y)| = |\langle x, Sy\rangle| \leq \|x\|\|Sy\| \leq \|x\|\|S\|\|y\| .$$

This gives the opposite inequality $M \leq \|S\|$. ☐

8.2 The Main Result

We enunciate a preliminary version of the spectral theorem.

THEOREM 8.4. *Let E be an i-resolution. Then the formula*

$$\langle(\Psi f)x, y\rangle = \int_\Omega f \, dE_{x,y} ,$$

for $x, y \in H$, defines an isometric isomorphism Ψ of the Banach algebra $L^\infty(E)$ onto a closed, normal subalgebra A of $\mathcal{B}(H)$.[2] *This isomorphism also enjoys the properties*

$$\Psi(\overline{f}) = \Psi(f)^*$$

for $f \in L^\infty(E)$ and

$$\|(\Psi f)x\|^2 = \int_\Omega |f|^2 \, dE_{x,x}$$

for $x \in H$ and $f \in L^\infty(E)$.

Finally, an operator $S \in \mathcal{B}(H)$ commutes with every $E(\omega)$ if and only if S commutes with every $\Psi(f)$.

The first formula in the theorem is frequently, just for simplicity, abbreviated as

$$\Psi(f) = \int_\Omega f \, dE .$$

[2] Here a subalgebra A if called *normal* if, whenever, $x \in A$, then $x^* \in A$.

Proof of the Theorem. The spectral theorem is so important that it is definitely worthwhile to enunciate it in several different ways. See [RUD1], [RUD2], and [DUS, v. 3] for other discussions of this result.

Let $\{\omega_1, \omega_2, \ldots, \omega_k\}$ be a partition of Ω with $\omega_j \in \mathcal{M}$. Let s be a simple function such that $s = T_j$ on ω_j. Define $\Psi(s) \in \mathcal{B}(H)$ by

$$\Psi(s) = \sum_{j=1}^{k} T_j E(\omega_j) . \tag{3}$$

Each $E(\omega_j)$ is self-adjoint, so

$$\Psi(s)^* = \sum_{j=1}^{k} \overline{T}_j E(\omega_j) = \Psi(\overline{s}) . \tag{4}$$

If $\{\omega_1', \omega_2', \ldots, \omega_m'\}$ is another partition of this sort, and if $t = \mu_j$ on ω_j', then

$$\Psi(s)\Psi(t) = \sum_{j,k} T_j \mu_k E(\omega_j) E(\omega_k') = \sum_{j,k} T_j \mu_k E(\omega_j \cap \omega_k') .$$

Since st is the simple function that equals $T_j \mu_k$ on $\omega_j \cap \omega_k'$, we see that

$$\Psi(s)\Psi(t) = \Psi(st) . \tag{5}$$

A completely analogous argument now shows that

$$\Psi(\alpha s + \beta t) = \alpha \Psi(s) + \beta \Psi(t) . \tag{6}$$

If $x, y \in H$, then (3) shows that

$$\langle (\Psi s)x, y \rangle = \sum_{j=1}^{k} T_j ((E\omega_j)x, y) = \sum_{j=1}^{k} T_j E_{x,y}(\omega_j) = \int_\Omega s \, dE_{x,y} . \tag{7}$$

By (4) and (5), we have

$$\Psi(s)^* \Psi(s) = \Psi(\overline{s})\Psi(s) = \Psi(\overline{s}s) = \Psi(|s|^2) .$$

Hence (7) gives us that

$$\|(\Psi s)x\|^2 = \langle \Psi(s)^* \Psi(s) x, x \rangle = \langle \Psi(|s|^2)x, x \rangle = \int_\Omega |s|^2 \, dE_{x,x} . \tag{8}$$

As a result,

$$\|(\Psi s)x\| \le \|s\|_\infty \|x\| . \tag{9}$$

On the other hand, if $x \in \mathcal{R}(E(\omega_j))$, then

$$(\Psi s)x = \alpha_j E(\omega_j)x = \alpha_j x \tag{10}$$

since the projections $E(\omega_j)$ have mutually orthogonal ranges. If j is chosen so that $|\alpha_j| = \|s\|_\infty$, then (9) and (10) tell us that

$$\|\Psi(s)\| = \|s\|_\infty . \tag{11}$$

Let $f \in L^\infty(E)$. There is a sequence of simple, measurable functions s_j that converges to f in the norm of $L^\infty(E)$. By (11), the corresponding operators $\Psi(s_j)$ form a Cauchy sequence in $\mathcal{B}(H)$ which is thus norm-convergent to an operator that we shall call $\Psi(f)$. One sees easily that $\Psi(f)$ does not depend on the particular choice of the $\{s_j\}$. Clearly (11) now tells us that

$$\|\Psi(f)\| = \|f\|_\infty \tag{12}$$

for $f \in L^\infty(E)$.

The first line of the theorem follows from (7) (with s_j in place of s) since each $E_{x,y}$ is a finite measure. Also the second and third lines of the theorem follow from (4) and (8). Finally, if bounded, measurable functions f and g are approximated, in the norm of $L^\infty(E)$, by simple measurable functions s and t, then we see that (5) and (6) hold with f and g in place of s and t.

In sum, Ψ is an isometric isomorphism of $L^\infty(E)$ into $\mathcal{B}(H)$. Since $L^\infty(E)$ is complete, its image $A = \Psi(L^\infty(E))$ is closed in $\mathcal{B}(H)$ because of (30).

Finally, if an operator S commutes with every $E(\omega)$, then S commutes with $\Psi(s)$ whenever s is simple. Thus the approximation process used above shows that S commutes with every member of A. $\square$

The spectral theorem that we shall prove below states that every bounded, normal operator T on a Hilbert space induces an i-resolution E on the Borel subsets of its spectrum $\sigma(T)$. Also that T can be reconstructed from E by an integral of the type considered in our last theorem.

We begin by considering algebras of operators, and then specialize down to single operators.

THEOREM 8.5. *Let A be a closed, normal subalgebra of $\mathcal{B}(H)$ that contains the identity operator I. Let $\triangle$ be the maximal ideal space of A. Then the following statements hold:*

(a) *There exists a unique i-resolution E on the Borel subsets of $\triangle$ that satisfies*

$$T = \int_{\triangle} \widehat{T}\, dE\,,$$

for every $T \in A$, where $\widehat{T}$ is the Gelfand transform of T.

(b) $E(\omega) \neq 0$ *for every nonempty open set $\omega \subseteq \triangle$.*

(c) *An operator $S \in \mathcal{B}(H)$ commutes with every $G \in A$ if and only if S commutes with every projection $E(\omega)$.*

REMARK. The formula in the statement of the theorem should be understood to mean

$$\langle Tx, y\rangle = \int_{\triangle} \widehat{T}\, dE_{x,y}\,. \tag{13}$$

Proof of the Theorem. Since $\mathcal{B}(H)$ is a B^*-algebra, our given algebra A is a commutative B^*-algebra. The Gelfand-Naimark theorem (Theorem 5.14) thus asserts that $T \to \widehat{T}$ is an isometric $*$-isomorphism of A *onto* $C(\triangle)$.

We now get a straightforward proof of the uniqueness of E. Suppose that E satisfies (13). Since $\widehat{T}$ ranges over all of $C(\triangle)$, the hypothesized regularity of the complex Borel measures $E_{x,y}$ shows that each $E_{x,y}$ is uniquely determined by (13). Use the Riesz representation theorem. Since, by definition,

$$\langle E(\omega)x, y\rangle = E_{x,y}(\omega)\,,$$

each projection $E(\omega)$ is also uniquely determined by (13).

We turn to the existence of E. If $x, y \in H$, then Theorem 5.14 (the Gelfand-Naimark theorem) shows that

$$\widehat{T} \to \langle Tx, y\rangle$$

is a bounded linear functional on $C(\triangle)$ with norm $\leq \|x\| \cdot \|y\|$ (since $\|\widehat{T}\|_\infty = \|T\|$). By the Riesz representation theorem, there is a unique regular complex Borel measure $\mu_{x,y}$ on $\triangle$ such that

$$\langle Tx, y\rangle = \int_{\triangle} \widehat{T}\, d\mu_{x,y} \tag{14}$$

for all $x, y \in H$ and $T \in A$.

When $\widehat{T}$ is real, T is self-adjoint. Thus $\langle Tx, y\rangle$ and $\langle Ty, x\rangle$ are complex conjugates of each other. As a result,

$$\mu_{x,y} = \overline{\mu}_{x,y} \tag{15}$$

for $x, y \in H$.

For fixed $T \in A$, the left-hand side of (14) is linear in x and conjugate-linear in y. The uniqueness of the measures $\mu_{x,y}$ implies then that $\mu_{x,y}(\omega)$ is, for every Borel set $\omega \subseteq \triangle$, a sesquilinear functional. Since $\|\mu_{x,y}\| \leq \|x\| \cdot \|y\|$, it follows that

$$\int_{\triangle} f \, d\mu_{x,y}$$

is a bounded sesquilinear functional on H for every bounded, Borel function f on $\triangle$. Theorem 8.3 tells us that there corresponds to each such f an operator $\Phi(f) \in \mathcal{B}(H)$ with

$$\langle (\Phi f)x, y\rangle = \int_{\triangle} f \, d\mu_{x,y} \tag{16}$$

for all $x, y \in H$. Comparison with (14) demonstrates that

$$\Phi(\widehat{T}) = T \tag{17}$$

for all $T \in A$. Hence Φ is an extension of the mapping $\widehat{T} \to T$ that takes $C(\triangle)$ onto A.

If f is real, then line (15) tells us that $\langle \Phi(f)x, y\rangle$ is the complex conjugate of $\langle \Phi(f)y, x\rangle$. This fact implies that $\Phi(f)$ is self-adjoint.

Next we shall prove that

$$\Phi(fg) = \Phi(f)\Phi(g) \tag{18}$$

for bounded Borel functions f and g. If $S, T \in A$, then $(ST)\widehat{\ } = \widehat{S}\widehat{T}$. Then (14) implies that

$$\int_{\triangle} \widehat{S}\widehat{T} \, d\mu_{x,y} = \langle STx, y\rangle = \int_{\triangle} \widehat{S} \, d\mu_{Tx,y} \, . \tag{19}$$

Since $\widehat{A} = C(\triangle)$, we conclude that

$$\widehat{T} d\mu_{x,y} = d\mu_{Tx,y}$$

for every choice of x, y, and T. The integrals (19) remain equal if $\widehat{S}$ is replaced by f. Thus

$$\begin{aligned} \int_{\triangle} f\widehat{T} \, d\mu_{x,y} &= \int_{\triangle} f \, d\mu_{Tx,y} \\ &= \langle (\Phi f)Tx, y\rangle \\ &= \langle Tx, z\rangle \\ &= \int_{\triangle} \widehat{T} \, d\mu_{x,z} \, . \end{aligned} \tag{20}$$

Here $z = \Phi(f)^* y$. The reasoning used above also shows that the first and last integrals in (20) remain equal when $\widehat{T}$ is replaced by any bounded Borel function g. Thus

$$\begin{aligned}\langle \Phi(fg)x, y\rangle &= \int_\triangle fg\, d\mu_{x,y} \\ &= \int_\triangle g\, d\mu_{x,z} \\ &= \langle \Phi(g)x, z\rangle \\ &= \langle \Phi(f)\Phi(g)x, y\rangle\,.\end{aligned}$$

That proves (18).

We can define E. If ω is a Borel subset of $\triangle$, let f be the characteristic function of ω. Put $E(\omega) = \Phi(f)$.

By (18), $E(\omega \cap \omega') = E(\omega)E(\omega')$. With $\omega' = \omega$, we see then that each $E(\omega)$ is a projection. Since $\Phi(f)$ is self-adjoint when f is real, it follows that each $E(\omega)$ is self-adjoint. Clearly $E(\emptyset) = \Phi(0) = 0$. Line (17) implies that $E(\triangle) = I$. The finite additivity of E follows from (16). And the equation

$$\langle E(\omega)x, y\rangle = \mu_{x,y}(\omega) \tag{21}$$

also follows from (16). Thus E is an i-resolution.

The proof of part (a) of the theorem is now complete, because (13) follows from (15) and (21).

For (b), suppose now that ω is open and $E(\omega) = 0$. If $T \in A$ and $\widehat{T}$ has support in ω, then the equation in part (a) implies that $T = 0$. Thus $\widehat{T} = 0$. Since $\widehat{A} = C(\triangle)$, Urysohn's lemma tells us that $\omega = \emptyset$. That establishes (b).

For (c), select $S \in \mathcal{B}(H)$ and $x, y \in H$. Put $z = S^* y$. For any $T \in A$ and any Borel set $\omega \subseteq \triangle$, we have

$$\langle STx, y\rangle = \langle Tx, z\rangle = \int_\triangle \widehat{T}\, dE_{x,z}\,, \tag{22}$$

$$\langle TSx, y\rangle = \int_\triangle \widehat{T}\, dE_{Sx,y}\,, \tag{23}$$

$$\langle SE(\omega)x, y\rangle = \langle E(\omega)x, z\rangle = E_{x,z}(\omega)\,, \tag{24}$$

$$\langle E(\omega)Sx, y\rangle = E_{Sx,y}(\omega)\,. \tag{25}$$

If $ST = TS$ for every $T \in A$, then the measures in (22) and (23) are equal, hence $SE(\omega) = E(\omega)S$. The same reasoning establishes the converse. □

The next, and final, result is our spectral theorem. We commonly refer to the E in the theorem as the *spectral decomposition* of the operator T.

THEOREM 8.6. *Let $T \in \mathcal{B}(H)$ be normal. Then there exists a unique i-resolution E on the Borel subsets of $\sigma(T)$ such that*

$$T = \int_{\sigma(T)} \lambda \, dE(\lambda) .$$

Every projection $E(\omega)$ commutes with every $S \in \mathcal{B}(H)$ which commutes with T.

REMARK. Here we see rather explicitly that any normal operator on H can be realized as multiplication by an essentially bounded function.

Proof. Let A be the smallest closed subalgebra of $\mathcal{B}(H)$ that contains I, T, and T^*. Since T is normal, the preceding theorem applies to A. By Theorem 8.5, the maximal ideal space of A can be identified with $\sigma(T)$ so that $\widehat{T}(\lambda) = \lambda$ for all $\lambda \in \sigma(T)$. The existence of E now follows from the preceding theorem.

Obversely, if E exists so that the equation in the statement of the theorem holds, then Theorem 8.5 tells us that

$$p(T, T^*) = \int_{\sigma(T)} p(\lambda, \overline{\lambda}) \, dE(\lambda) . \tag{26}$$

Here p is any polynomial in two variables with complex coefficients. By the Stone-Weierstrass theorem, these polynomials are dense in $C(\sigma(T))$. The projections $E(\omega)$ are thus uniquely determined by the integrals (26). So they are determined by T—just as in the uniqueness part of the proof of the last theorem.

If $ST = TS$, then also $ST^* = T^*S$ by the Fuglede-Putnam-Rosenblum Theorem (see [RUD2, p. 300]). So S commutes with every member of A. By part (c) of the preceding theorem, $SE(\omega) = E(\omega)S$ for every Borel set $\omega \subseteq \sigma(T)$. □

Let E be the spectral decomposition of a normal operator $T \in \mathcal{B}(H)$. Let f be a bounded, Borel function on $\sigma(T)$. We then denote the operator

$$\Psi(f) = \int_{\sigma(T)} f \, dE$$

by $f(T)$. With this notation, we can express the results of Theorems 8.4, 8.5, 8.6 as

The mapping $f \to f(T)$ *is a homomorphism of the algebra of all bounded Borel functions on* $\sigma(T)$ *into* $\mathcal{B}(H)$ *that carries the function 1 to* I*. It also carries the identity function on* $\sigma(T)$ *to* T *and it satisfies*

$$\overline{f}(T) = f(T)^*$$

with

$$\|f(T)\| \leq \sup\{|f(\lambda)| : \lambda \in \sigma(T)\}\,.$$

If $f \in C(\sigma(T))$*, then equality holds in this last display.*

CHAPTER 9

CONVEXITY

9.1 Introductory Thoughts

In this chapter we treat the classical notion of convexity. Indeed, convexity is an old idea. It occurs in some of Archimedes's treatments of the concept of arc length. But the idea of convexity was not actually formalized until the treatise [BOF] appeared in 1934. Since then it has been studied intensely, both in the classical Euclidean setting (see [KRA4]) and in the more general setting of infinite dimensions (see, for instance, [BAP], [SIM]).

DEFINITION. Let X be a topological vector space, and let $E \subseteq X$. We say that E is *convex* if, whenever $x, y \in X$, then

$$(1-t)x + ty \in E \quad \text{for all} \quad 0 \leq t \leq 1 \, .$$

EXAMPLE. Let X be any Banach space. Let $B(\mathbf{0}, 1)$ be the open ball with center $\mathbf{0}$ and radius 1. Then $B(\mathbf{0}, 1)$ is certainly convex, as the triangle inequality shows.

Let X be the Banach space ℓ^2. Let $\mathbf{1} = (1, 0, 0, \ldots)$ be an element of X. Let

$$E = B(\mathbf{0}, 1) \setminus B(\mathbf{1}, 1/2) \, .$$

Then E is *not* convex. For the points $(3/4, 5/8, 0, \ldots)$ and $(3/4, -5/8, 0, \ldots)$ both lie in E. But their midpoint (corresponding to $t = 1/2$) is $(3/4, 0, 0, \ldots)$, and that does *not* lie in E.

DEFINITION. Let X be a topological vector space and $E \subseteq X$. The *convex hull* of E is the smallest convex set in X that contains E. In other words, it is the intersection of all convex sets that contain E. See Figure 9.1.

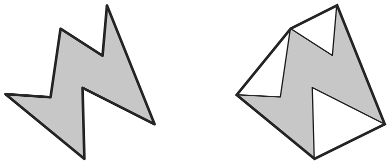

FIGURE 9.1. Convex hull.

DEFINITION. Let E be a subset of a vector space X. We say that a point $s \in E$ is an *extreme point* of E if the following condition holds: If $x, y \in E$, and if there exists a number t with $0 < t < 1$, and $tx + (1-t)y = s$, then $x = y = s$. See Figure 9.2.

We call a set $S \subseteq E$ an *extreme set* for E if, whenever $x, y \in E$, $0 < t < 1$, and $tx + (1-t)y \in S$, then $x \in S$ and $y \in S$.

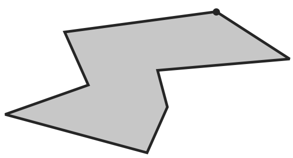

FIGURE 9.2. An extreme point.

9.2 SEPARATION THEOREMS

THEOREM 9.1. *Let A and B be disjoint, nonempty, convex sets in a topological vector space X.*

(a) *If A is open then there exist $\Lambda \in X^*$ and $\gamma \in \mathbb{R}$ such that*

$$Re\,\Lambda x < \gamma \le Re\,\Lambda y$$

for every $x \in A$ and $y \in B$.

(b) *If A is compact, B is closed, and X is locally convex, then there exist $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$, such that*

$$Re\,\Lambda x < \gamma_1 < \gamma_2 < Re\,\Lambda y$$

for every $x \in A$ and every $y \in B$.

Proof. It is sufficient to prove the result for real scalars.

For (a), fix $a_0 \in A$, $b_0 \in B$. Put $x_0 = b_0 - a_0$. Set $C = A - B + x_0$. Then C is a convex neighborhood of $\mathbf{0}$ in X. Let p be the Minkowski functional of C. The term p satisfies the hypothesis of the Hahn-Banach theorem. Since $A \cap B = \emptyset$, $x_0 \notin C$, hence $p(x_0) \geq 1$.

Define $f(tx_0) = t$ on the subspace M of X generated by x_0. If $t \geq 0$, then

$$f(tx_0) = t \leq tp(x_0) = p(tx_0)\,.$$

If $t < 0$, then $f(tx_0) < 0 \leq p(tx_0)$. Thus $f \leq p$ on M. By the Hahn-Banach theorem, f extends to a linear functional Λ on X that also satisfies $\Lambda \leq p$. In particular, $\Lambda \leq 1$ on C. Thus $\Lambda \geq -1$ on $-C$, so $|\Lambda| \leq 1$ on the neighborhood $C \cap (-C)$ of 0. We conclude that $\Lambda \in X^*$.

Let $a \in A$ and $b \in B$. Then we have

$$\Lambda a - \Lambda b + 1 = \Lambda(a - b + x_0) \leq p(a - b + x_0) < 1$$

since $\Lambda x_0 = 1$, $a-b+x_0 \in C$, and C is open. We conclude that $\Lambda a < \Lambda b$.

We conclude that $\Lambda(A)$ and $\Lambda(B)$ are disjoint, convex subsets of $\mathbb{R}$, with $\Lambda(A)$ to the left of $\Lambda(B)$. Also $\Lambda(A)$ is an open set since A is open and since every nonconstant linear functional on X is an open mapping. Let γ be the right endpoint of $\Lambda(A)$ to obtain the conclusion of part (a).

For (b), we note that there is a convex neighborhood V of $\mathbf{0}$ in X such that $(A + V) \cap B = \emptyset$. Part (a), with $A + V$ in place of A, shows that there exists $\Lambda \in X^*$ such that $\Lambda(A + V)$ and $\Lambda(B)$ are disjoint convex subsets of $\mathbb{R}$ with $\Lambda(A + V)$ open and to the left of $\Lambda(B)$. Since $\Lambda(A)$ is a compact subset of $\Lambda(A + V)$, we derive the conclusion of (b). $\square$

COROLLARY 9.2. *If X is a locally convex space then X^* separates points of X.*

Proof. If $x_1, x_2 \in X$ and $x_1 \neq x_2$, then apply (b) of the theorem with $A = \{x_1\}$ and $B = \{x_2\}$. $\square$

THEOREM 9.3. *Suppose that X is a vector space and that X' is a separating vector space (i.e., when $x_1 \neq x_2$ then there is a Λ with $\Lambda x_1 \neq \Lambda x_2$) of linear functionals on X. Then the X' topology τ' makes X into a locally convex space whose dual space is X'.*

Proof. We can see that τ' is a Hausdorff topology. The linearity of the members of X' shows that τ' is translation invariant. If $T_1, T_2, \ldots, T_k \in X'$, if $r_j > 0$, and if

$$V = \{x : |T_j x| < r_j \text{ for } 1 \leq j \leq k\}\,, \tag{1}$$

then V is convex, balanced, and $V \in \tau'$. In fact the collection of all V of the form in (1) is a local basis for τ'. Thus τ' is a locally convex topology on X.

If (1) holds, then $\frac{1}{2}V + \frac{1}{2}V = V$. This proves that addition is continuous. Suppose now that $x \in X$ and α is a scalar. Then $x \in sV$ for some $s > 0$. If $|\beta - \alpha| < r$ and $y - x \in rV$, then

$$\beta y - \alpha x = (\beta - \alpha)y + \alpha(y - x)$$

lies in V, provided that r is so small that

$$r(s + r) + |\alpha| r < 1 .$$

As a result, scalar multiplication is continuous.

We now have proved that τ' is a locally convex vector topology. Every $\Lambda \in X'$ is τ'-continuous. Conversely, suppose that Λ is a τ'-continuous linear functional on X. Then $|\Lambda x| < 1$ for all x in some set V of the form (1). We can conclude then that $\Lambda = \sum \alpha_j T_j$. Since $T_j \in X'$, and X' is a vector space, we conclude that $\Lambda \in X'$. □

PROPOSITION 9.4. *Suppose that X is a topological vector space on which X^* separates points. Let A, B be disjoint, nonempty, compact, convex sets in X. Then there is a $\Lambda \in X^*$ such that*

$$\sup_{x \in A} Re\, \Lambda x < \inf_{y \in B} Re\, \Lambda y .$$

REMARK. This result is fundamental. It enunciates a means of separating two convex sets. In the plane there would be a line separating A and B, and that would tell us right away what Λ has to be.

Proof of the Proposition. Let X_w be X equipped with the weak topology. The sets A and B are apparently compact in X_w. They are also closed in X_w just because X_w is a Hausdorff space. Since X_w is locally convex, part (b) of Theorem 9.1 can be applied to X_w. This give us a $\Lambda \in (X_w)^*$ that satisfies the displayed equation in the proposition. But we know, from Theorem 9.3, that $(X_w)^* = X^*$. □

Perhaps the most fundamental and compelling result about convexity in infinite dimensions is the following striking theorem. It is illustrated nicely in Figure 9.3.

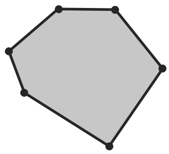

FIGURE 9.3. The Krein-Milman theorem.

9.3 The Main Result

THEOREM 9.5 (Krein-Milman). *Let X be a topological vector space in which X^* separates points. If K is a compact, convex set in X, then K is the closed, convex hull of its extreme points.*

Proof. Let $\mathcal{P}$ be the collection of all compact extreme sets of K. Since $K \in \mathcal{P}$, we see that $\mathcal{P} \neq \emptyset$. We shall use these properties of $\mathcal{P}$ (to be proved below):

(a) The intersection S of any nonempty subcollection of $\mathcal{P}$ is a member of $\mathcal{P}$ *unless* $S = \emptyset$.

(b) If $S \in \mathcal{P}$, $\Lambda \in X^*$, μ is the maximum value of $\operatorname{Re}\Lambda$ on S, and

$$S_\Lambda = \{x \in S : \operatorname{Re}\Lambda x = \mu\},$$

then $S_\Lambda \in \mathcal{P}$.

We observe that (a) is obvious.

For (b), suppose that, for $x, y \in K$ and $0 < t < 1$, we have $tx + (1-t)y = z \in S_\Lambda$. Since $z \in S$ and $S \in \mathcal{P}$, we see that $x, y \in S$. Therefore $\operatorname{Re}\Lambda x \leq \mu$, $\operatorname{Re}\Lambda y \leq \mu$. Since $\operatorname{Re}\Lambda z = \mu$ and Λ is linear, we find that $\operatorname{Re}\Lambda x = \mu = \operatorname{Re}\Lambda y$. Thus $x, y \in S_\Lambda$.

Choose some $S \in \mathcal{P}$. Let $\mathcal{P}'$ be the collection of all members of $\mathcal{P}$ that are subsets of S. Since $S \in \mathcal{P}'$, certainly $\mathcal{P}'$ is not empty. We partially order $\mathcal{P}'$ by set inclusion. Let Ω be a maximal totally ordered subcollection of $\mathcal{P}'$. Let M be the intersection of all members of Ω.

Since Ω is a collection of compact sets with the finite intersection property, $M \neq \emptyset$. By (a), $M \in \mathcal{P}'$. The maximality of Ω implies that no proper subset of M belongs to $\mathcal{P}$. We see from (b) that every $\Lambda \in X^*$ is constant on M. Since X^* separates points of X, we conclude that M has only one point. Thus M is an extreme point of K. $\square$

EXAMPLE. A closed cube in $\mathbb{R}^N$ is the closed, convex hull of its vertices.

A closed ball in $\mathbb{R}^N$ is the closed, convex hull of *all* its boundary points (that is to say, it is the closed, convex hull of the sphere).

CHAPTER 10

Fixed-Point Theorems

Of course the granddaddy of all fixed-point theorems is that of L. E. J. Brouwer, proved in the early twentieth century. It says this:

THEOREM 10.1. *Let F be a continuous mapping of the closed unit ball in $\mathbb{R}^N$ to itself. Then there is a point P in the closed unit ball such that $F(P) = P$.*

Popular expositors have fun explaining the two-dimensional version of this theorem in terms of stirring a bowl of soup with grated cheese on top.

Such frivolity tends to disguise the fact that the fixed-point theorem is a profound result of mathematical analysis, and it has important applications and consequences. In this chapter we intend to present two different versions of the fixed-point principle on Banach spaces, and to show how they can be used to derive important and substantial results in the subject.

Much of our presentation here is indebted to [PAT].

10.1 Banach's Theorem

The first infinite-dimensional fixed-point theorem is well known and is actually due to Stefan Banach.

DEFINITION. Let X be a metric space with metric ρ and $F : X \to X$ a mapping. We call F a *contraction* if there is a real number K, $0 < K < 1$, such that

$$\rho(F(x), F(y)) \leq K \cdot \rho(x, y) .$$

THEOREM 10.2. *Let X be a complete metric space with metric ρ. Let $F : X \to X$ be a contraction. Then there is a unique point $P \in X$ such that $F(P) = P$.*

EXAMPLE. Let X be the Banach space ℓ^2 and let $F : X \to X$ be given by $f(x) = (1/2)x + (1, 1/2, 1/3, \ldots)$. Then it is easily checked that f is a contraction with constant $K = 1/2$. The unique fixed point is $P = (2/1, 2/2, 2/3, 2/4, \ldots)$.

EXAMPLE. Let X be the Banach space $L^2([0, 1])$. Let F be the mapping

$$F(f) = (x + 1) \int_0^1 x \cdot f(x)\, dx\,.$$

Then we can check that

$$\begin{aligned} |F(f) - F(g)| &= \left|(x + 1) \int_0^1 x f(x)\, dx - (x + 1) \int_0^1 x g(x)\, dx\right| \\ &= \left|(x + 1) \int_0^1 x \cdot [f(x) - g(x)]\, dx\right| \\ &\leq |x + 1| \int_0^1 x^2\, dx^{1/2} \int_0^1 |f(x) - g(x)|^2\, dx^{1/2} \\ &= |x + 1| \cdot (1\sqrt{3}) \|f - g\|_{L^2}\,. \end{aligned}$$

Thus

$$\|F(f) - F(g)\|_{L^2} \leq \left(\frac{7}{9}\right)^{1/2} \cdot \|f - g\|_{L^2}\,.$$

We conclude that F has a unique fixed-point, but it is not at all clear what that fixed point is.

REMARK. The property of being a contraction may seem rather special. But our examples and applications will show that in fact contractions are not so hard to come by.

Proof of the Theorem. First observe that, if x_1, x_2 are fixed-points of F, then

$$d(x_1, x_2) = d(F(x_1), F(x_2)) \leq K d(x_1, x_2)\,,$$

from which it follows that $d(x_1, x_2) = 0$. Hence $x_1 = x_2$. That establishes the uniqueness of the fixed-point.

Choose a point $x_0 \in X$ and define

$$\begin{aligned} x_1 &= F(x_0) \\ x_2 &= F(x_1) \\ x_3 &= F(x_2) \end{aligned}$$

etcetera. We see that

$$\begin{aligned} d(x_{n+1}, x_n) &= d(F(x_n), F(x_{n-1}) \leq K \cdot d(x_n, x_{n-1}) \\ &= K \cdot d(F(x_{n-1}), F(x_{n-2})) \\ &\leq K^2 d(x_{n-1}, x_{n-2}) \\ &\leq \cdots \\ &\leq K^n d(x_1, x_0) . \end{aligned}$$

If $n, m \in \{1, 2, \ldots\}$ then

$$\begin{aligned} d(x_{n+m}, x_n) &\leq d(x_{n+m}, x_{n+m-1}) + d(x_{n+m-1}, x_{n+m-2}) \\ &\qquad + \cdots + d(x_{n+1}, x_n) \\ &\leq (K^{n+m-1} + K^{n+m-2} + \cdots + K^n) d(x_1, x_0) \\ &\leq \frac{K^n}{1-K} \cdot d(x_1, x_0) . \end{aligned}$$

We conclude that $\{x_n\}$ is a Cauchy sequence. So it has a limit point $\overline{x} \in X$. Since F is continuous (indeed, it is Lipschitz), we may conclude that

$$F(\overline{x}) = F(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = \overline{x} .$$

So $\overline{x}$ is the fixed-point that we seek. □

REMARK. It is worth noting that completeness is an essential hypothesis in the theorem. As an instance, if $X = (0, 1]$ with the usual Euclidean metric, then the mapping $F : X \to X$ given by $F(x) = x/2$ is a contraction but has no fixed-point.

It will be useful in what follows to have a couple of variants of Banach's contraction mapping fixed-point theorem.

COROLLARY 10.3. *Let X be a complete metric space and Y a topological space. Let $F : X \times Y \to X$ be a continuous function. Assume that, uniformly in Y, F is a contraction in X. That is to say, assume that*

$$d(F(x_1, y), F(x_2, y)) \leq K \cdot d(x_1, d_2)$$

for all $x_1, x_2 \in X$, $y \in Y$, and some $0 < K < 1$. Then, for each fixed $y \in Y$, the mapping $x \mapsto F(x, y)$ has a unique fixed-point $\varphi(y) \in X$. Furthermore, the function $y \mapsto \varphi(y)$ is continuous from Y to X.

Proof. The only thing we need prove is the continuity of φ. For $y, y_0 \in Y$, we have

$$\begin{aligned}
&d(\varphi(y), \varphi(y_0)) \\
&\qquad = d(F(\varphi(y), y), F(\varphi(y_0), y_0)) \\
&\qquad \leq d(F(\varphi(y), y), F(\varphi(y_0), y)) + d(F(\varphi(y_0), y), F(\varphi(y_0), y_0)) \\
&\qquad \leq K \cdot d(\varphi(y), \varphi(y_0)) + d(F(\varphi(y_0), y), F(\varphi(y_0), y_0)),
\end{aligned}$$

which implies

$$d(\varphi(y), \varphi(y_0)) \leq \frac{1}{1-K} \cdot d(F(\varphi(y_0), y), F(\varphi(y_0), y_0)).$$

Since this right-hand side goes to 0 as $y \to y_0$, we obtain the desired continuity. □

COROLLARY 10.4. *Let X be a complete metric space and let $F : X \to X$. If F^n (the composition of F with itself n times) is a contraction for some $n \geq 1$, then F has a unique fixed-point $\overline{x} \in X$.*

Proof. Let $\overline{x}$ be the unique fixed-point of F^n, given by our theorem above. Then

$$F^n(F(\overline{x})) = F(F^n(\overline{x})) = F(\overline{x}),$$

which implies that $F(\overline{x}) = \overline{x}$. Since any fixed-point of F is also clearly a fixed-point of F^n, we obtain the uniqueness automatically. □

10.2 Two Applications

Our first application of the Banach fixed-point theorem is to a proof of the implicit function theorem. There are many proofs of the implicit function theorem, but it turns out that the proof using a fixed-point theorem is the most flexible and useful (see [KRP] for the details of this assertion).

In what follows, we let $B_X(P, r)$ denote a metric ball in the Banach space X and $\overline{B}_X(P, r)$ the closed metric ball. We interpret derivatives in a Banach space in the usual Fréchet sense (see [RUD2, p. 248]).

We shall formulate the implicit function theorem as follows:

THEOREM 10.5. *Let X, Y, Z be Banach spaces. Let $U \subseteq X \times Y$ be an open set and $u_0 = (x_0, y_0) \in U$. Let $F : U \to Z$. Assume that*

(a) *F is continuous and $F(u_0) = 0$,*

(b) *$D_y F(u)$ exists for every $u = (x, y) \in U$,*

(c) *$D_y F$ is continuous at u_0 and $D_y F(u_0)$ is invertible.*

Then there exists $\alpha, \beta > 0$ for which $\overline{B}_X(x_0, \alpha) \times \overline{B}_Y(y_0, \beta) \subseteq U$ and a unique continuous function $\varphi : \overline{B}_X(x_0, \alpha) \to \overline{B}_Y(y_0, \beta)$ such that the relation

$$F(x, y) = 0 \quad \Longleftrightarrow \quad y = \varphi(x)$$

holds for all $(x, y) \in \overline{B}_X(x_0, \alpha) \times \overline{B}_Y(y_0, \beta)$.

REMARK. The proof of this result is surprisingly straightforward. Contrast, for example, with the more classical proof that appears in [RUD3].

Proof of the Theorem. We may assume, with no loss of generality, that $x_0 = 0$ and $y_0 = 0$. Define

$$\Phi(x, y) = y - \left[D_y F(0, 0)\right]^{-1} F(x, y)$$

for $(x, y) \in U$. By (a), Φ is continuous from U into Y. We see that

$$D_y \Phi(x, y) = I - \left[D_y F(0, 0) \right]^{-1} \circ D_y F(x, y) .$$

By (c), there is a $\gamma > 0$ small enough that

$$\|D_y \Phi(x, y)\| \leq \frac{1}{2}$$

for all $(x, y) \in B_X(0, \gamma) \times B_Y(0, \gamma) \subseteq U$. Thus simple estimates and the continuity of Φ imply that

$$\|\Phi(x, y_1) - \Phi((x, y_2)\|_Y \leq \frac{1}{2}\|y_1 - y_2\|_Y$$

for $\|x\|_X$, $\|y_1\|_Y$, $\|y_2\|_Y$ all less than or equal to β, which is less than γ.

Using (a), we can find $0 < \alpha < \beta$ such that

$$\|\Phi(x, 0)\|_Y \leq \frac{\beta}{2}$$

for all $\|x\|_X \leq \alpha$. Then, for $\|x\|_X \leq \alpha$ and $\|y\|_Y \leq \beta$,

$$\|\Phi(x, y)\|_Y \leq \|\Phi(x, 0)\|_Y + \|\Phi(x, y) - \Phi(x, 0)\|_Y \leq \frac{1}{2}(\beta + \|y\|_Y) \leq \beta .$$

Thus the continuous map $\Phi : \overline{B}_X(0, \alpha) \times \overline{B}_Y(0, \beta) \to \overline{B}_Y(0, \beta)$ is a contraction on $\overline{B}_Y(0, \beta)$ uniformly in $\overline{B}_X(0, \alpha)$. From our first corollary to the Banach fixed-point theorem, there exists a unique continuous function $f : \overline{B}_X(0, \alpha) \to \overline{B}_Y(0, \beta)$ such that $\Phi(x, f(x)) = f(x)$, that is, $F(x, f(x)) = 0$. $\square$

For our next application, we consider the Cauchy problem

$$\begin{cases} x'(t) & = & f(t, x(t)) \text{ for } t \in I \\ x(t_0) & = & x_0 \, . \end{cases}$$

We seek a closed interval I, with t_0 belonging to the interior of I, and also a differentiable function $x : I \to X$ (where X is a given Banach space) so that the system is satisfied. This is the standard first-order initial value problem in ordinary differential equations.

It is a familiar fact that this system is equivalent to solving the integral equation

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds$$

for $t \in I$. Our theorem is this:

THEOREM 10.6. *Assume the following hypotheses:*

(a) f *is continuous,*

(b) *The inequality*

$$\|f(t, x_1) - f(t, x_2)\|_X \le k(t)\|x_1 - x_2\|_X$$

for all $(t, x_1), (t, x_2) \in U$ *holds for some* $k(t) \in [0, \infty)$*,*

(c) $k \in L^1((t_0 - a, t_0 + a))$ *for some* $a > 0$*,*

(d) *There exist* $m \ge 0$ *and* $\overline{B}_{\mathbb{R}\times X}(u_0, s) \subseteq U$ *such that*

$$\|f(t, x)\|_X \le m$$

for all $(t, x) \in \overline{B}_{\mathbb{R}\times X}(u_0, s)$*.*

Then there exists a $\tau_0 > 0$ *such that, for any* $\tau < \tau_0$*, there is a unique solution* $x \in C^1(I_\tau, X)$ *to the Cauchy problem, with* $I_\tau = [t_0 - \tau, t_0 + \tau]$*.*

REMARK. Those familiar with the classical proof will notice here that things are much more streamlined. Many of the classical estimates are absorbed into the functional-analytic machinery.

Proof of the Theorem. Let $r = \min\{a, s\}$ and set

$$\tau_0 = \min\left\{r, \frac{r}{m}\right\} .$$

Select $\tau < \tau_0$ and consider the complete metric space $Z = \overline{B}_{C(I_r, X)}(x_0, r)$ with the metric induced by the norm of $C(I_r, X)$. Here x_0 is the constant function equal to x_0. Since $\tau < r$, if $z \in Z$ then $(t, z(t)) \in \overline{B}_{\mathbb{R}\times X}(u_0, r) \subseteq U$ for all $t \in I_r$. Thus, for $z \in Z$, define

$$F(z)(t) = x_0 + \int_{t_0}^{t} f(y, z(y))\, dy$$

for $t \in I_r$. Notice that

$$\sup_{t\in I_r} \|F(z)(t) - x_0\| \leq \sup_{t\in I_r} \left| \int_{t_0}^{t} \|f(y, z(y))\|_X\, dy \right| \leq m\tau \leq r\,.$$

We conclude that F maps Z into Z. The last task for us is to show that F^n is a contraction on Z for some $n \in \mathbb{N}$. By induction on n, we shall show that, for every $t \in I_r$,

$$\|F^n(z_1)(t) - F^n(z_2)(t)\|_X \leq \frac{1}{n!} \left| \int_{t_0}^{t} k(y)\, dy \right|^n \cdot \|z_1 - z_2\|_{C(I_r,X)}\,. \qquad (1)$$

For $n = 1$ this is clear. So assume it is true for $n-1$ with $n \geq 2$. Taking $t > t_0$ (the argument for $t < t_0$ is analogous), we see that

$$\begin{aligned}
&\|F^n(z_1)(t) - F^n(z_2)(t)\|_X \\
&\qquad = \|F(F^{n-1}(z_1))(t) - F(F^{n-1}(z_2))(t)\|_X \\
&\qquad \leq \int_{t_0}^{t} \|f(y, F^{n-1}(z_1)(y)) - f(y, F^{n-1}(z_2)(y))\|_X\, dy \\
&\qquad \leq \int_{t_0}^{t} k(y)\|F^{n-1}(z_1)(y) - F^{n-1}(z_2)(y)\|_X\, dy \\
&\qquad \leq \frac{1}{(n-1)!}\left[\int_{t_0}^{t} k(y)\left(\int_{t_0}^{y} k(w)\, dw\right)^{n-1} dy\right] \|z_1 - z_2\|_{C(I_r,X)} \\
&\qquad = \frac{1}{n!}\left(\int_{t_0}^{t} k(y)\, dy\right)^n \|z_1 - z_2\|_{C(I_r,X)}\,.
\end{aligned}$$

Thus, from (1), we get

$$\|F^n(z_1) - F^n(z_2)\|_{C(I_r,X)} \leq \frac{1}{n!}\|k\|^n_{L^1(I_r)}\|z_1 - z_2\|_{C(I_r,X)}\,,$$

and this shows that, for n big enough, F^n is a contraction. By means of our second corollary, we conclude that F has a unique fixed-point. That is clearly the unique solution to our integral equation, hence solves the Cauchy problem. $\square$

10.3 The Schauder Theorem

One of the most well known and widely used fixed-point theorems is that due to Schauder. Here we present the theorem and give a nice, modern application.

We begin with a lemma.

LEMMA 10.7. *Let K be a nonempty, compact, convex subset of a finite-dimensional real Banach space X. Then every continuous function $f : K \to K$ has a fixed-point $\overline{x} \in K$.*

Proof. Since X is homeomorphic to $\mathbb{R}^N$ for some N, we assume without loss of generality that $X = \mathbb{R}^N$. We may also assume that $K \subseteq \overline{B}^N$ (the closed unit ball). For every $x \in \overline{B}^N$, let $p(x) \in K$ be the unique point of minimum norm for the set $x - K$. Observe that $p(x) = x$ for every $x \in K$. Furthermore, p is continuous on $\overline{B}^N$. In fact, given $x_j, x \in \overline{B}^N$ with $x_j \to x$, we see that

$$\begin{aligned}\|x - p(x)\| &\leq \liminf_{x_j \to \infty} \|x - p(x_j)\| \\ &\leq \liminf \|x - x_j\| + \inf_{k \in K} \|x_j - k\| \to \|x - p(x)\|\end{aligned}$$

as $j \to \infty$. Thus $x - p(x_j)$ is a minimizing sequence as $x_j \to x$ in $x - K$. This implies the convergence of $p(x_j) \to p(x)$. Define $g(x) = f(p(x))$. Then g maps $\overline{B}^N$ continuously onto K. By the Brouwer fixed-point theorem, there is now a $\overline{x} \in K$ such that $g(\overline{x}) = \overline{x} = f(\overline{x})$. ☐

THEOREM 10.8 (Schauder-Tychonoff). *Let X be a locally convex space, $K \subseteq X$ a nonempty and convex set, and $K_0 \subseteq K$ compact. If $f : K \to K_0$ is continuous, then there exists a point $\overline{x} \in K_0$ such that $f(\overline{x}) = \overline{x}$.*

Proof. Denote by $\mathcal{B}$ the local basis for the topology of X generated by the separating family of seminorms $\mathcal{P}$ on X. Given $U \in \mathcal{B}$, we may invoke the compactness of K_0 to find $x_1, x_2, \ldots, x_k \in K_0$ such that

$$K_0 \subseteq \bigcup_{j=1}^{k} (x_j + U)\,.$$

Let $\varphi_1, \ldots, \varphi_k \in C(K_0)$ be a partition of unity for K_0 subordinate to the open cover $\{x_j + U\}$. Define

$$f_U(x) = \sum_{j=1}^{k} \varphi_j(f(x)) x_j$$

for $x \in K$. Then we have

$$f_U(K) \subseteq K_U \equiv \mathrm{co}\,(\{x_1, \ldots, x_k\}) \subseteq K\,.$$

Here co stands for "convex hull."

Lemma 10.8 yields the existence of a point $x_U \in K_U$ such that $f_U(x_U) = x_U$. Then

$$x_U - f(x_U) = f_U(x_U) - f(x_U) = \sum_{j=1}^{k} \varphi_j(f(x_U))(x_j - f(x_U)) \in U \quad (2)$$

for $\varphi_j(f(x_U)) = 0$ whenever $x_j - f(x_U) \notin U$.

Invoking again the compactness of K_0, there exists

$$\overline{x} \in \bigcap_{W \in \mathcal{B}} \overline{\{f(x_U) : U \in \mathcal{B}, U \subseteq W\}} \subseteq K_0\,. \quad (3)$$

Choose $p \in \mathcal{P}$ and $\epsilon > 0$ and let

$$V = \{x \in X : p(x) < \epsilon\} \in \mathcal{B}\,.$$

Since f is continuous on K, there is a $W \in \mathcal{B}$ with $W \subseteq V$ such that

$$f(x) - f(\overline{x}) \in V$$

whenever $x - \overline{x} \in 2W$, $x \in K$. Furthermore, by (3), there is a $U \in \mathcal{B}$, $U \subseteq W$, such that

$$\overline{x} - f(x_U) \in W \subseteq V\,. \quad (4)$$

Taking (2) and (4) together, we find that

$$x_U - \overline{x} = (x_U - f(x_U)) + (f(x_U) - \overline{x}) \in U + W \subseteq W + W = 2W\,.$$

This yields

$$f(x_U) - f(\overline{x}) \in V\,. \quad (5)$$

Lines (4) and (5) tell us that

$$p(\overline{x} - f(\overline{x})) \leq p(\overline{x} - f(x_U)) + p(f(x_U) - f(\overline{x})) < 2\epsilon\,.$$

Since p and ϵ are arbitrary, we conclude that $p(\overline{x} - f(\overline{x})) = 0$ for every $p \in \mathcal{P}$, and that implies $f(\overline{x}) = \overline{x}$. $\square$

One of the big open problems in functional analysis is to determine whether any bounded linear mapping T of a Hilbert space H has an invariant subspace. This would be a proper subspace $E \subseteq H$ such that $T(E) \subseteq E$. The result is known to be false for general Banach spaces (see [ENF]), but it is still an important open problem for Hilbert spaces.

One of the really nice positive results along these lines is the following (see [LOM]):

THEOREM 10.9 (Lomonosov). *Let X be a Banach space. Let $T \in \mathcal{B}(X)$ be a nonscalar operator commuting with a nonzero compact operator $P \in \mathcal{B}(X)$. Then T has a hyperinvariant subspace (i.e., it is invariant for all operators commuting with T).*

Proof. The argument is by contradiction. Let $\mathcal{A}$ be the algebra of operators commuting with T. It is immediate that, if T has no hyperinvariant subspace, then $\overline{\mathcal{A}x} = X$ for every $x \in X$, $x \neq 0$.

Without loss of generality, let S be an operator with $\|S\|_{\mathcal{B}(X)} \leq 1$. Choose $x_0 \in X$ such that $\|Sx_0\| > 1$ (which also implies $\|x_0\| > 1$) and set $B = \overline{B}_X(x_0, 1)$. For $x \in \overline{SB}$ (observe that this x cannot be the zero vector), there is a $T' \in \mathcal{A}$ such that $\|T'x - x_0\| < 1$. Hence every $x \in \overline{SV}$ has an open neighborhood V_x such that $T'V_x \subseteq B$ for some $T' \in \mathcal{A}$. By the compactness of $\overline{SB}$, we find a finite cover $V_1, \ldots, V_k$ and $T_1', \ldots, T_k' \in \mathcal{A}$ such that

$$T_j' V_j \subseteq B$$

for $j = 1, 2, \ldots, k$.

Let $\varphi_1, \ldots, \varphi_k \in C(\overline{SB})$ be a partition of unity for $\overline{SB}$ subordinate to the open cover $\{V_j\}$. Define, for $x \in B$,

$$f(x) = \sum_{j=1}^{k} \varphi_j(Sx) T_j' Sx \,.$$

Then f is a continuous function from B into B. Since $T_j'S$ is a compact map for every j, it is easy to see that $f(B)$ is relatively compact. Thus Schauder's fixed point theorem tells us that there is an $\overline{x} \in B$ such that $f(\overline{x}) = \overline{x}$.

Define the operator $\widetilde{T} \in \mathcal{A}$ by

$$\widetilde{T} = \sum_{j=1}^{k} \varphi_j(S\overline{x}) T_j' \,.$$

Then we have

$$\widetilde{T} S \overline{x} = \overline{x} \,.$$

But $\widetilde{T}S$ is a compact operator, thus the eigenspace F of $\widetilde{T}S$ relative to the eigenvalue 1 is finite-dimensional. Since $\widetilde{T}S$ commutes with T, we conclude that F is invariant for T. That means that T has an eigenvalue, and thus a hyperinvariant subspace, contradicting our assumption. $\square$

Table of Notation

Notation	Section	Definition
X	1.2	normed linear space
$\|x\|$	1.2	norm of x
$B(x, r)$	1.2	open ball with center x and radius r
$\overline{B}(x, r)$	1.2	closed ball with center x and radius r
X	1.2	Banach space
$C([0, 1])$	1.2	the space of continuous functions on [0, 1]
$C(X)$	1.2	the space of continuous functions on X
ℓ^1	1.2	the space of summable sequences
ℓ^p	1.2	the space of pth power summable sequences
ℓ^∞	1.2	the space of bounded sequences
$\mathbb{R}^N$	1.3	Euclidean N-space
Λ	1.4	a linear operator
X^*	1.4	the space of bounded linear functionals on X (the dual space)
$\widehat{f}(j)$	1.4	the jth Fourier coefficient of f
d	1.5	a metric
G_δ	1.5	the intersection of countably many open sets
U, V	1.6	the open unit ball in a Banach space
$\widehat{\lambda}$	1.6	extension of the linear functional λ
$\mathbb{T}$	1.7	the circle group
S_N	1.7	a partial summation operator for Fourier series
$D_N(t)$	1.7	the Dirichlet kernel
L^p	1.7	the Lebesgue space of pth power integrable functions
c_0	1.7	doubly infinite sequences that vanish at infinity

Notation	Section	Definition
G	1.7	the graph of a linear operator
$G(z, w)$	1.7	the Green's function
$P_r(\theta)$	1.7	the Poisson kernel
H	3.1	an inner product space
H	3.1	a Hilbert space
$\langle x, y\rangle$	3.1	the inner product in a Hilbert space
E	3.1	a subspace of a given space
$x^\perp$	3.2	the orthogonal space to x
$E^\perp$	3.2	the orthogonal complement of E
μ_A	3.2	the Minkowski functional of A
p, q	3.2	Hilbert space projections
λ	3.2	a bounded linear functional
$\{u_\alpha\}_{\alpha\in A}$	3.2	an orthonormal set
$\widehat{x}(\alpha)$	3.2	the αth Fourier coefficient of x
$\mathcal{B}(X, Y)$	4.2	the bounded operators from X to Y
$\mathcal{B}(X)$	4.2	the bounded operators from X to X
$\|x^*\|$	4.2	the norm of x^*
$\|L\|$	4.2	the norm of the operator L
T^*	4.2	the adjoint of the operator T
$\langle x, x^*\rangle$	4.2	the pairing of a Banach space element and a dual element
$\mathcal{N}(T)$	4.3	the null space of the operator T
$\mathcal{R}(T)$	4.3	the range of the operator T
$\sigma(T)$	4.3	the spectrum of the operator T
N	4.3	a complemented space
α	4.3	$\dim \mathcal{N}(T - \lambda I)$
β	4.3	$\dim X/\mathcal{R}(T - \lambda I)$
α^*	4.3	$\dim \mathcal{N}(T^* - \lambda I)$
β^*	4.3	$\dim X^*/\mathcal{R}(T^* - \lambda I)$
A	5.1	a Banach algebra
e	5.1	the unit element in a Banach algebra
$\sigma(x)$	5.1	the spectrum of an element of a Banach algebra

Notation	Section	Definition
φ	5.1	a multiplicative linear functional in a Banach algebra
x^{-1}	5.1	the inverse of an element in a Banach algebra
$\rho(x)$	5.1	the spectral radius of x
$\mathcal{I}, \mathcal{J}$	5.3	an ideal in a Banach algebra
$\triangle$	5.3	the maximal ideal space of a Banach algebra
$A/\mathcal{I}$	5.3	the quotient of the Banach algebra A by the ideal $\mathcal{I}$
$x \mapsto x^*$	5.3	an involution on a Banach algebra
$\widehat{x}$	5.3	the Gelfand transform
ρ, ρ_α	6.1	a seminorm
TVS	6.1	topological vector space
$U_{x,\alpha,\epsilon}$	6.1	a subbasis element for the topology on a TVS
$C_c^\infty(U)$	7.1	infinitely differentiable functions of compact support in U
$\mathcal{D}'(U)$	7.2	distributions with support in U
δ	7.2	the Dirac delta mass
$\operatorname{supp}\varphi$	7.3	the support of φ
$\langle\alpha, \phi\rangle$	7.3	pairing of a distribution with a testing function
$\partial^\alpha\varphi$	7.3	the derivative of order α of φ
$\gamma * \varphi$	7.3	convolution of a distribution with a testing function
$\widetilde{\psi}(x)$	7.3	the same as $\psi(-x)$
$t^{-N}\psi(x/t)$	7.4	a Friedrichs mollifier
$h(x)$	7.4	the Heaviside function
$\widehat{\alpha}$	7.5	the Fourier transform of a distribution
$\mathcal{M}$	8.1	a sigma algebra of sets
$E, E_{x,y}$	8.1	an i-resolution
Ψ	8.1	the isomorphism from the spectral theorem
s	9.1	an extreme point
F^n	10.2	iteration of the mapping F with itself n times

Glossary

adjoint of an operator *T* A linear operator T^* so that

$$\langle Tx, y\rangle = \langle x, T^* y\rangle.$$

algebra A collection of objects equipped with binary operations of addition and multiplication, and also with a notion of scalar multiplication.

algebra of sets A collection of sets that is closed under finite union and complementation.

Baire category theorem The result that a complete metric space cannot be written as the countable union of nowhere dense sets.

balanced set A set B in a linear space with the property that $\alpha B \subseteq B$ for any scalar α with $|\alpha| \leq 1$.

ball Given a metric space (X, d), the ball $B(P, r)$ with center P and radius r is the set $B(P, r) = \{x \in X : d(x, P) < r\}$.

Banach-Alaoglu theorem The theorem that says that the closed unit ball in the dual of a Banach space is compact in the weak-$*$ topology.

Banach algebra A Banach space that is also an algebra.

Banach fixed-point theorem See the *contraction mapping fixed-point theorem*.

Banach space A normed linear space that is complete in the topology induced by the norm.

Banach-Steinhaus theorem A theorem that characterizes the uniform boundedness of a collection of linear operators.

Bessel's inequality In a Hilbert space, the inequality

$$\sum_{\alpha \in A} |\widehat{x}(\alpha)|^2 \leq \|x\|^2 .$$

Here $\widehat{x}(\alpha)$ are the Fourier coefficients of x with respect to a given orthonormal set.

Borel sets The σ algebra generated by the open sets.

bounded linear functional A linear functional λ that satisfies an inequality of the form $|\lambda(x)| \leq C\|x\|$ for some constant $C > 0$.

bounded linear operator A linear operator L that satisfies an inequality of the form $\|Lx\|_y \leq C\|x\|_X$ for some constant $C > 0$.

Cauchy problem This is the initial value problem for an ordinary differential equation given by

$$\begin{cases} x'(t) & = & f(t, x(t)) \text{ for } t \in I \\ x(t_0) & = & x_0 \,. \end{cases}$$

Cauchy sequence A sequence $\{x_j\}$ in a normed linear space X with the property that, if $\epsilon > 0$, then there is a $K > 0$ so that if $j, k > K$, then $\|x_j - x_k\| < \epsilon$.

characteristic function Given a set S, the characteristic function χ_S is equal to 1 at points of S and equal to 0 at points of the complement of S.

circle group The group $\mathbb{T}$ that is equivalent to the reals $\mathbb{R}$ modulo 2π.

closed ball Given a metric space (X, d), the closed ball $\overline{B}(P, r)$ with center P and radius r is the set $\overline{B}(P, r) = \{x \in X : d(x, P) \leq r\}$.

closed graph theorem Logically equivalent to the open mapping principle. Characterizes continuous linear mappings in terms of closed graphs.

codimension of a subspace M of X The dimension of X/M.

compact operator A linear operator T from a Banach space X to a Banach space Y with the property that $T(U)$ has compact closure when U is the unit ball in X.

complement of a subspace M of X A subspace with the property that there is another subspace N with $M \cap N = \{\mathbf{0}\}$ and $M + N = X$.

complete orthonormal set In a Hilbert space, an orthonormal set so that, if x is orthogonal to every element of the set, then $x = \mathbf{0}$.

composition of a distribution with an invertible linear transformation If S is an invertible linear transformation of $\mathbb{R}^N$ and γ a distribution, then we set

$$\langle \gamma \circ S, \varphi \rangle = |\det S|^{-1} \langle \gamma, \varphi \circ S^{-1} \rangle \,.$$

continuous linear functional A linear functional that is continuous in the usual sense of topology.

continuous linear operator A linear operator that is continuous in the usual sense of topology.

contraction A mapping F of a metric space (X, d) to itself with the property that there is a constant K between 0 and 1 such that

$$d(F(x), F(y)) \leq Kd(x, y) .$$

contraction mapping fixed-point theorem The result that a contraction on a complete metric space has a unique fixed-point.

convex hull of a set E The smallest convex set that contains E.

convex set A set E in a linear space X such that, if $x, y \in E$, then $(1 - t)x + ty \in E$ for all $0 \leq t \leq 1$.

convolution of a distribution with a smooth function If γ is a distribution and ψ a smooth function then we set

$$\gamma * \psi(x) = \langle \gamma, \tau_x \widetilde{\psi} \rangle .$$

Here $\widetilde{\psi}(x) = \psi(-x)$ and τ is a translation operator.

derivative of a distribution If λ is a distribution, D a derivative, and φ a testing function, then we define

$$\langle D\lambda, \varphi \rangle \equiv (-1)^k \langle \lambda, D\varphi \rangle ,$$

where k is the degree of D.

Dirac delta mass A measure of mass 1 supported at the origin.

distribution A generalized function, usually specified as an element of the dual space of a space of smooth testing functions.

dual space The space X^* of bounded linear functionals on a given space X.

eigenvalue of an operator *T* A scalar λ so that $T - \lambda I$ is not one-to-one.

eigenvector of an operator *T* A vector x with the property that $Tx = \lambda x$ for some eigenvalue λ.

extreme point of a set E We say that s is an extreme point of E if, whenever $x, y \in E$, $0 < t < 1$, and $tx + (t - t)y = s$, then $x = y = s$.

extreme set for a set E We say that S is an extreme set for E if, whenever $x, y \in E$, $0 < t < 1$, and $tx + (1-t)y \in S$, then $x \in S$ and $y \in S$.

finite-dimensional space A linear space with a basis having finitely many elements.

first category A set that can be written as the countable union of nowhere dense sets.

fixed point of a mapping F A point P such that $F(P) = P$.

Fourier series For an integrable function f on the circle group $\mathbb{T}$, this is the series $\sum_j c_j e^{ij\theta}$ with $c_j = (1/(2\pi)) \int_0^{2\pi} f(t)e^{-ijt}\, dt$.

Fréchet space A complete, Hausdorff topological vector space with the property that its topology is defined by a countable family of seminorms.

Gelfand-Mazur theorem The theorem that says that if a complex Banach algebra has the property that every nonzero element is invertible then the algebra is isometrically isomorphic to the complex number field.

Gelfand-Naimark theorem A theorem that establishes an isometric isomorphism of a given Banach algebra A with the continuous functions on the maximal ideal space.

Gelfand transform The mapping that assigns to each x in a Banach algebra X a function $\widehat{x} : \triangle \to \mathbb{C}$ by way of the formula

$$\widehat{x}(h) = h(x) .$$

Green's function A function with the singularity of the fundamental solution of the Laplacian but zero boundary values.

Hahn-Banach theorem The theorem that guarantees the existence of a bounded extension of a given linear functional from a subspace to the total space.

Hausdorff maximality theorem A result that allows one to extract a maximal element from a collection. Related to Zorn's lemma.

Hilbert space A Banach space that is an inner product space.

ideal A linear subspace $\mathcal{J}$ of X such that if $y \in \mathcal{J}$ and $x \in X$ then $xy \in \mathcal{J}$.

identity element An element e in the Banach algebra X such that $ex = xe = x$ for every $x \in X$.

implicit function theorem The theorem that tells us when we can solve the equation $f(x, y) = 0$ for y in terms of x.

injective A mapping or function is injective if it is one-to-one.

inner product space A linear space equipped with a bilinear operator into the field of scalars that satisfies certain natural properties.

invertible element in a Banach algebra X An element $x \in X$ for which there exists a $y \in X$ with $xy = e$.

invertible operator A linear operator $T : X \to Y$ with the property that there is an operator $S : Y \to X$ so that $ST = TS = I$.

i**-resolution** Let $\mathcal{M}$ be a σ-algebra on a set Ω. Let H be a Hilbert space. Then an *i-resolution* on $\mathcal{M}$ is a mapping

$$E : \mathcal{M} \to \mathcal{B}(H)$$

such that

(a) $E(\emptyset) = \mathbf{0}$, $E(\Omega) = I$,

(b) Each $E(\omega)$ is a self-adjoint projection,

(c) $E(\omega \cap \omega') = E(\omega)E(\omega')$,

(d) If $\omega \cap \omega' = \emptyset$, then $E(\omega \cup \omega') = E(\omega) + E(\omega')$,

(e) For every $x, y \in H$, the set function $E_{x,y}$ defined by

$$E_{x,y}(\omega) = \langle E(\omega)x, y \rangle$$

is a complex measure on $\mathcal{M}$.

kernel of a multiplicative linear functional The set of elements that are mapped to 0.

Krein-Milman theorem The result that a compact, convex set is the convex hull of its extreme points.

linear functional A linear operator whose range is the field of scalars (usually $\mathbb{R}$ or $\mathbb{C}$).

linear operator A function L from a normed linear space X to a normed linear space Y that is linear in the sense that $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$.

locally convex space We say that X is a locally convex space if there is a basis for the topology consisting of convex sets.

Lomonosov's theorem The theorem that says that a Banach space operator that commutes with a compact operator has an invariant subspace.

maximal ideal of a Banach algebra X A proper ideal in X that is contained in no other ideal.

maximal ideal space of a Banach algebra X The collection of multiplicative linear functionals on X.

Minkowski functional A function that measures the convexity of a set.

multiplication of a distribution by a smooth function If γ is a distribution and ψ a smooth function then we set

$$\langle \psi\gamma, \varphi \rangle = \langle \gamma, \psi\varphi \rangle .$$

multiplicative linear functional on a Banach algebra X A ring homomorphism from X to the scalar field.

norm The device for measuring the lengths of vectors in a normed linear space.

normal operator A bounded operator T on the Hilbert space H such that $TT^* = T^*T$.

normed linear space A vector space that is equipped with a norm that is compatible with the algebraic structure on the space.

open mapping principle The theorem that says that the inverse of a univalent, surjective linear mapping is also continuous.

orthonormal set In a Hilbert space, a set of unit vectors that are mutually orthogonal.

parallelogram law The identity, valid in a Hilbert space, that says

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 .$$

partition of unity A collection of C_c^∞ functions that sums to 1.

perpendicular complement of an element x In an inner product space, the collection of elements that are orthogonal to x.

perpendicular complement of a set S In an inner product space, the collection of elements that are orthogonal to all elements of S.

Poisson kernel The reproducing kernel for the space of harmonic functions.

projection In a Hilbert space H, a canonical, self-adjoint, idempotent operator that maps H onto a given subspace E.

Riemann-Lebesgue lemma The result that says that the Fourier coefficients of an integrable function vanish at infinity.

Riesz-Fischer theorem The theorem that characterizes which sequences of scalars arise as the Fourier coefficients of an element of Hilbert space.

Riesz representation theorem The theorem that characterizes the dual of the space of continuous functions on a compact, Hausdorff space. Also the theorem that says that any bounded, linear functional on a Hilbert space is given by inner product with some element of the space.

scalar field The field over which our vector spaces and normed linear spaces live. This field is denoted by k, and is usually the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$.

Schauder-Tychonoff theorem A fixed-point theorem for compact, convex sets.

Schwartz distribution An element of the dual space of the Schwartz space.

Schwartz function A smooth function f on $\mathbb{R}^N$ is said to be Schwartz if

$$\rho_{\alpha,\beta}(f) \equiv \sup |x^\alpha \partial^\beta f|$$

is finite for every choice of α and β. We use the seminorms $\rho_{\alpha,\beta}$ to topologize the space of Schwartz functions.

Schwarz inequality A standard inequality that estimates the size of the inner product of two vectors in terms of the norms of those vectors:

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\| .$$

second category A set that is not of first category.

self-adjoint operator An operator T on the Hilbert space H such that $T = T^*$.

seminorm Like a norm, but without the property that $\|x\| = 0 \Rightarrow x = 0$.

sesquilinear A quadratic form that is complex linear in the first entry and conjugate complex linear in the second entry.

σ-algebra A collection of sets that is closed under countable union and complementation.

space of functions A normed linear space whose elements are functions on some common domain.

spectral radius of an element x in a Banach algebra X The radius of the spectrum of x.

spectral theorem A theorem that represents the bounded linear operators on a Hilbert space in terms of multiplication of the space of L^2 functions by L^∞ functions.

spectrum of a Banach algebra element x The set of scalars λ such that $x - \lambda e$ is not invertible.

spectrum of an operator T The collection of scalars λ so that $T - \lambda I$ is not invertible.

Stone-Weierstrass theorem The theorem that gives a sufficient condition for a subalgebra of $C(X)$ to be dense in $C(X)$.

subspace A subset of a given linear space X that is closed under addition and scalar multiplication.

support of a distribution The complement of the set where the distribution vanishes.

surjective A mapping or function is surjective if it is onto.

topological vector space A vector space that is endowed with a topology so that addition and scalar multiplication are continuous.

totally bounded set An set E with the property that, for each $\epsilon > 0$, there is a finite collection of balls of radius ϵ that covers E.

translation of a distribution If $y \in \mathbb{R}^N$ and f is any function on $\mathbb{R}^N$ then we define the *translation* by y as $\tau_y f(x) = f(x - y)$. If γ is a distribution, then we set

$$\langle \tau_y \gamma, \varphi \rangle = \langle \gamma, \tau_{-y} \varphi \rangle .$$

triangle inequality The result that estimates the norm of the sum of two vectors in terms of the norms of the individual vectors:

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| .$$

uniform boundedness principle Same as the Banach-Steinhaus theorem.

unitary operator An operator T on the Hilbert space H such that $T^*T = I = TT^*$.

von Neumann algebra The algebra of bounded operators on a Hilbert space.

weak-$*$ topology A topology on the dual X^* of a Banach space that is equivalent to pointwise convergence.

weak topology A topology on a Banach space that is induced by the dual space.

Wiener Tauberian theorem If a nonvanishing, continuous function has absolutely convergent Fourier series, then so does its reciprocal.

Zorn's lemma A logical result that enables one to extract maximal elements from a collection.

Bibliography

[BAP] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces*, 2nd ed., Reidel, New York, 1986.

[BOH] H. F. Bohenblust, Extensions of functionals on complex linear spaces *Bull. Amer. Math. Soc.* 44(1938), 91–93.

[BOF] T. Bonneson and W. Fenchel, *Theorie der konvexen Körper*, Springer-Verlag, Berlin, 1934.

[DUS] N. Dunford and J. Schwartz, *Linear Operators*, Interscience Publishers, New York, 1958–1971.

[ENF] P. Enflo, On the invariant subspace problem for Banach spaces, *Acta Math.* 158(1987), 213–313.

[FOL] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., Wiley, New York, 1999.

[KAK] S. Kakutani, Some characterizations of Euclidean space, *Japan. J. Math.* 16(1939), 93–97.

[KAN] E. Kaniuth, *A Course in Commutative Banach Algebras*, Springer, New York, 2009.

[KAT] Y. Katznelson, *Introduction to Harmonic Analysis*, John Wiley and Sons, New York, 1968.

[KEL] J. L. Kelley, Banach spaces with the extension property, *Trans. AMS* 72(1952), 323–326.

[KRA1] S. G. Krantz, *Real Analysis and Foundations*, 1st, ed., CRC Press, Boca Raton, FL, 1992.

[KRA2] ——, *Function Theory of Several Complez Variables*, 2nd ed., American Mathematical Society, Providence, RI, 2001.

[KRA3] ——, *A Panorama of Harmonic Analysis*, Mathematical Association of America, Washington, D.C., 1999.

[KRA4] ——, Convexity in real analysis, *Real Analysis Exchange*, to appear.

[KRA5] ——, *Elements of Topology: Theory and Practice*, Taylor & Francis, Boca Raton, FL, 2009.

[KRA6] ——, *A Guide to Topology*, Mathematical Association of America, Washington, D.C., 2009.

[KRA7] ——, *A Guide to Real Variables*, Mathematical Association of America, Washington, D.C., 2009.

[KRP] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem*, Birkhäuser Publishing, Boston, 2002.

[LAX] P. D. Lax, On the existence of Green's functions, *Proc. Amer. Math. Soc.* 3(1952), 526–531.

[LIM] B. V. Limaye, *Functional Analysis*, John Wiley and Sons, New York, 1981.

[LIT] J. Lindenstrauss and L. Tzafriri, On the complemented subspaces problem, *Israel J. Math.* 9(1971), 263–269.

[LOM] V. Lomonosov, Invariant subspaces of the family of operators that commute with a completely continuous operator, (Russian) *Funkcional. Anal. i Prilożen.* 7(1973), 55–56.

[PAT] V. Pata, fixed-point Theorems and Applications, manuscript, `vittorino.pata@polimi.it`.

[RES] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, New York, 1972.

[RIC] C. Rickart, *General Theory of Banach Algebras*, van Nostrand, Princeton, NJ, 1960.

[RUD1] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1966.

[RUD2] ——, *Functional Analysis*, McGraw-Hill, New York, 1973.

[RUD3] ——, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York, 1976.

[SCH] L. Schwartz, *Théorie des Distributions*, Hermann, Paris, 1950.

[SIM] B. Simon, *Convexity: An Analytic Viewpoint*, Cambridge University Press, Cambridge, 2011.

[SOB] A. Sobczyk, On the extension of linear transformations, *Trans. Amer. Math. Soc.* 55(1944), 153–169.

[WIE] N. Wiener, Tauberian theorems, *Annals of Math.* 33(1932), 1–100.

[YOS] K. Yosida, *Functional Analysis*, 6th ed., Springer-Verlag, New York, 1980.

Index

About the Author

Steven G. Krantz was born in San Francisco, California in 1951. He received the B.A. degree from the University of California at Santa Cruz in 1971 and the Ph.D. from Princeton University in 1974.

Krantz has taught at UCLA, Penn State, Princeton University, and Washington University in St. Louis. He served as Chair of the latter department for five years.

Krantz has published more than 60 books and more than 160 scholarly papers. He is the recipient of the Chauvenet Prize and the Beckenbach Book Award of the MAA. He has received the UCLA Alumni Foundation Distinguished Teaching Award and the Kemper Award. He has directed 18 Ph.D. theses and 9 Masters theses.